DOCTORAL THESIS

Embracing variability:

the use of mixed effects models in Ecology

Melina de Souza Leite

*A thesis submitted to the Institute of Biosciences the University at the University of São Paulo in fulfillment of the requirements for the degree of **Doctor in Biological Sciences**, area **Ecology***

Supervisor: Dr. Paulo Inácio de Knegt López de Prado

São Paulo

2022

TESE DE DOUTORADO

Incorporando a variabilidade: o uso de modelos de efeitos mistos em ecologia

Melina de Souza Leite

Tese apresentada ao Instituto de Biociências da Universidade de São Paulo para a obtenção de título de **Doutora em Ciências Biológicas** *na Área de* **Ecologia**

Orientador: Dr. Paulo Inácio de Knegt López de Prado

São Paulo
2022

"Essentially, all models are wrong, but some are useful"

George Box

Agradecimentos

Meus sinceros agradecimentos a todos aqueles que me acompanharam, me acolheram e me ajudaram a iniciar e finalizar essa enriquecedora etapa da minha jornada profissional e pessoal. Os agradecimentos formais e específicos de cada capítulo se encontram ao final de cada um. Aqui é o momento de me desmanchar em gratidão.

Aos mestres

Primeiramente, um enorme e especial agradecimento ao meu orientador, Paulo Inácio, também conhecido como PI (mas não confundir com meu gato, Pi). Quando entrei na USP para exercer a função de especialista em laboratório, em agosto de 2010, ninguém do departamento sabia muito para quê eu "servia". E PI me acolheu, apoiou e me incentivou a continuar no meu desenvolvimento intelectual e profissional. Lembro bem em 2011, quando ele e o prof. Alexandre Adalardo (Alê), me convidaram a fazer a disciplina de R da pós, com planos de me convidar a trabalhar com ensino e pesquisa usando o R. Esta disciplina foi um divisor de águas para mim e foi certamente a semente das ideias nos campos da ecologia e estatística que culminaram nesta tese. PI sempre foi meu grande conselheiro profissional, me ajudando em estratégias de persuasão, argumentação e principalmente a navegar pelos aspectos menos racionais nas relações interpessoais e disputas de poder no ambiente profissional. PI é um grande mestre que faz a diferença no mundo! Não só pela ciência, mas também pelo despertar da consciência, pela forma com que ele toca tão profunda e positivamente as pessoas, com confiança e esperança! Ah, como o mundo seria muito melhor se houvesse mais pessoas como o PI!

Ao Alê por me despertar interesse em ensinar ecologia, estatística e afins. Foram longos anos de monitoria de R, disciplinas de graduação e pós, eu sempre com livre acesso ao seu laboratório e convites para colaborações. Grande incentivador do aprendizado, e do compartilhamento de conhecimento. Graça a ele, tive a grande oportunidade de ir ao workshop do ForestGEO e de ter colaborações que eu jamais imaginara!

Sinceros agradecimentos ao Jean e ao seu grande laboratório-família LEPAC. Me despertou interesse em ecologia da paisagem e sempre deixou a porta aberta para me misturar à comunidade do LEPAC, onde fiz grandes amigos! Seu apoio e colaboração durante o doutorado foram cruciais!

Me orgulho muito de ter tido não apenas um, nem dois, mas TRÊS coorientadores internacionais! Imagina a sorte de ter conseguido juntar 3 mentes brilhantes orbitando meu projeto. Fico feliz por os ter convencido de que valia a pena investir o tempo deles em me orientar. Tudo começou com Marco, meu mentor no Workshop do ForestGEO (julho 2019), que ouviu meu projeto e me ajudou a moldá-lo de forma mais atraente e interessante. Depois, convencemos a Lisa e o Sean em

embarcar nesta aventura! Curiosamente, estes três deram um minicurso de modelos mistos durante o workshop. Como não os agregar ao meu doutorado?

Ao Marco com sua mente brilhante, rápida, afiada, trazendo um componente agilizado à minha tese. Me recebeu em Leiden para um estágio sanduíche de 2 meses e me ajudou a atingir a meta de terminar o manuscrito do segundo capítulo propondo um workshop com Sean e Lisa, no qual tive o privilégio de ter os 3 em dedicação exclusiva ao meu projeto por 4 dias! Incrível!

Lisa é meu grande exemplo de mulher na ciência! Uma das pesquisadoras que eu mais admiro e que tanto aprendi ao longo destes anos de convivência. Me recebeu por 1 ano na Alemanha e durante os piores momentos da pandemia foi uma ótima orientadora e companhia explorando trilhas, pedaladas e rios pela Baváia. Agradeço muito por todo apoio e suporte, por entender tão bem como uma mulher se sente neste meio e por compartilhar comigo o caminho da cientista. É um exemplo de equilíbrio entre trabalho e vida pessoal, e uma grande amiga!

Sean é um "influencer" em seus aspectos mais positivos, dinâmico, cativante, animado, sempre pronto para ajudar e discutir ciência! Me deu a oportunidade de conseguir uma bolsa do Smithsonian Institution e, apesar de eu não ter ido aos EUA visitá-lo, me fez sentir em casa em seu laboratório através das reuniões virtuais semanais. Abriu muitas oportunidades e aprendi muito com ele e seu time!

Ao meu querido comitê de acompanhamento do doutorado, formado pelo Alê e Tadeu Siqueira (UNESP Rio Claro) pelas revisões de projeto, reuniões de acompanhamento e por me orientarem no caminho da pesquisa!

À minha querida colaboradora e amiga mais que especial Lari Boesing! Por me incentivar ao doutorado e me confiar seus preciosos dados para as investidas no capítulo 1 desta tese!

Aos segundos lares

Sinceros agradecimentos aos que tornam meu local de trabalho, o departamento de ecologia, agradável e saudável. Ao Conselho do Departamento e aos chefes de departamento durante estes últimos 5-6 anos, por terem ajudado com a escala de trabalho durante o doutorado e também por receberem meus pedidos de afastamento para as viagens de estágio sanduíche, congressos e afins. Agradeço igualmente ao Instituto de Biociências e à Universidade de São Paulo pelo suporte financeiro e institucional. Às queridas e maravilhosas Vera e Shirlene por serem um exemplo de profissionalismo, eficiência e carisma! Aos meus colegas de trabalho Luís, Welington, Natália, Talita, Lenilda, Maurício, Walter, Marco, pelas boas relações e principalmente pelo apoio que recebi.

Ao PPGE por todo suporte e por ser um programa de pós-graduação progressista e de excelência, no qual tanto aprendizado, tanta troca me foi enriquecedora! À CAPES, via financiamento PROEX do PPGE, pelas diárias para viagens de trabalho em congressos e workshops! Aos professores do departamento e do PPGE pelos

ensinamentos e apoio, especialmente Adriana Martini, Camila Castanho, Glauco Machado, Paulo Guimarães, Renata Pardini. Às comissões formadas e geridas pelos alunos que é tão importante para o programa: Café Existencial, Comissão PROEX, EcoEncontros.

Aos colegas e amigos da ecologia e dos laboratórios que me acolheram durante o doutorado: LET, LAGE, LEPAC, LabTrop. É incrível como um ambiente profissional enriquecedor e estimulante faz a diferença na formação do cientista!

À Universidade de Regensburg e ao Laboratório de Ecologia Teórica. Ao professor Florian Hartig que me recebeu abertamente em seu laboratório e que me fez sentir muito bem na Alemanha! Aos passeios de caiaque, trilha na floresta com meus colegas de laboratório, Max, Johannes, Lena, Lukas, e às frutíferas reuniões nas quais aprendi muito! Ainda na Alemanha, ao Hanns Jochen Drißle, por me receber em sua linda e aconchegante casa por um ano! Ao Florian Tschernik pelas trilhas na Bavária e nos Alpes, por me mostrar tanto da natureza quando da cultura e gastronomia de seu povo, incluindo a difícil língua alemã e seus dialetos!

À Universidade de Leiden, ao Centro de Ciências Ambientais e ao laboratório do Marco por me receberem por dois meses em pleno inverno em 2021. Aos colegas de república enquanto estive na Holanda, ao Leon e seus amigos italianos pela boa comida e boas conversas!

Ao Smithsonian Institution pela bolsa de doutorado sanduíche, a qual desfrutei durante minha estadia na Alemanha. Aos colegas do Laboratório de Ecologia Quantitativa pelas ótimas reuniões e por abrirem espaço para mim!

Ao ForestGEO por me levar para seu workshop de 2019, onde tudo começou para o capítulo 2 e as frutíferas colaborações internacionais. Também pelo travel grant para passar os 3 primeiros meses na Alemanha. Agradecimentos especiais ao Stuart Davies, Caly McCarthy e Lauren Krizel.

Aos computadores espalhados por 3 continentes que realizaram a árdua tarefa de rodar quase 3 mil (foram mais, mas impossível contabilizar os descartados) modelos bayesianos para finalizar as análises do capítulo 2. Estão estes no Instituto de Biociências, USP (Darwin), na Universidade de Regensburg, no Smithsonian Institution, e na Universidade da Bayreuth. Obrigado àqueles que mantém estas máquinas poderosas funcionando!

Ao lar, doce lar!

Ao meu lar paulistano que me abriga por quase 12 anos, o Refúgio! Minhas amigas queridas que por aqui moraram e que ainda se fazem presente, especialmente à Camilinha (de Barros), Lari (Boesing) e Pati (Torres, que ainda mora!)! Não posso esquecer do meu gatinho Pi Patel, que me mata de alegria e alergia!

Aos meus parceiros de caminhada, aos amigos que vem e vão, que estão do lado, que estão guardados no peito. Aqui é a hora que eu vou esquecer de alguém, e ficar

chateada depois de me lembrar que esqueci (desculpa, minha memória é quase que inversamente proporcional ao carinho que tenho). Aqui também mora o perigo de ficar emotiva e querer ser breve por não saber lidar muito bem com os sentimentos, principalmente os de gratidão. Mas vamos lá. Às queridas amigas de jornada do doutorado, Dani Zoo (Daniela Coelho) e Juju (Júlia Barreto). À Dani agradeço muito pelas conversas profissionais e pessoais, sérias, mas também misturadas com um toque de sensualidade que só a Bahia dá a seus cidadãos! À Juju que fica difícil agradecer e dizer no que ela me ajudou, porque foi em precisamente TUDO nestes últimos 4 anos! Minha irmã gêmea mais nova, companheira de aventuras e de desventuras na cidade, na praia, na montanha! Aos amigos de agora ou de outrora, Chico, Adrian, Isa, Andrés, Bru, Gabi, e muitos outros que direta ou indiretamente me ajudaram no caminho!

À minha psicóloga, Vanessa, por seu profissionalismo que me ajudou a encontrar o equilíbrio, a saúde mental e o contentamento com o doutorado e a vida!

Ao meu companheiro de jornada, Joqui, pelos longos anos (milênios praticamente) juntos, pelo carinho, afeto, cuidado, e todo amor! Por cuidar de mim, cuidar do Pi e do nosso lar, doce lar!

À minha família querida, mamãe, Juju, Digo. Meu núcleo querido, meu abrigo em Tarituba! À história viva e ao sentimento de paz que me traz voltar sempre à origem, ao aconchego do ninho. Aos bichinhos fofos, claro, que fizeram e fazem parte da nossa família!

À Natureza em seu conceito mais abrangente (nunca achei que um dia agradeceria por isso!). Fonte de inspiração, renovação, motivação por estudá-la e vivenciá-la em suas formas mais distintas: nas florestas, campos, rios, lagos, montanhas, rochas, praias, mares, sob sol, chuva, neve, lama o que for! Que alegria me dá ser bicho do mato!

Resumo

A ecologia é uma ciência que se preocupa constantemente em testar hipóteses ou mesmo entender processos ecológicos a partir de padrões em dados, os quais são na maioria das vezes descritos com uso de ferramentas estatísticas. Modelos mistos são cada vez mais usados em ecologia devido à natureza hierárquica da maioria dos dados ecológicos e são modelos especialmente úteis em se particionar fontes de variabilidade. Modelos mistos podem fornecer respostas quantitativas que descrevem que fração de um padrão em particular é explicada por diferentes mecanismos: o "quanto" em vez de "qual". Essa mudança de perspectiva pode se muito útil em ecologia tanto para entender contribuição relativa de processos ecológicos, quanto descrever melhor padrões globais de variabilidade. Nesta tese, eu emprego e advogo pelo uso de modelos mistos em perguntas ecológicas bem distintas, mas com um arcabouço analítico comum: partição de variabilidade em modelos mistos.

No primeiro capítulo, buscamos entender como a matriz (não habitat) modula a contribuição relativa de processos de montagem de comunidades (filtro ambiental e deriva) de aves florestais com a de perda de habitat em paisagens fragmentadas. Para modelar a ocorrência das aves, usamos um modelo misto com associação entre os atributos das espécies e a cobertura florestal como efeitos fixos e diversos componentes aleatórios para medir força relativa de certos processos. Encontramos que o filtro ambiental por perda de habitat é o processo dominante e pode ser relaxado ou reforçado dependendo da qualidade da matriz, evidenciando que esta tem um forte impacto na modulação dos processos de montagem da comunidade.

No segundo capítulo, desenvolvemos e aplicamos um *framework* conceitual e analítico através da partição de variâncias de um modelo misto para atribuir a variabilidade em taxas vitais de árvores devido a variações espaciais, temporais, entre espécies e suas interações. Aplicamos o *framework* no crescimento, mortalidade e recrutamento de 21 florestas globalmente distribuídas cobrindo mais de 2,9 milhões de árvores de aproximadamente 6,5 mil espécies. A proposta de decomposição da variação das taxas vitais de árvores tem um alto potencial para identificar os mecanismos estruturantes da dinâmica florestal, pois destaca os caminhos mais promissores para pesquisas futuras, aumentando a compreensão das contribuições relativas dos grupos de mecanismos para a demografia e diversidade florestal.

Junto às ferramentas de partição de variâncias, destacamos nesta tese o grande potencial dos modelos mistos em conectar inferência estatística e ecológica, incorporando a variabilidade como fonte de informações diversas e de qualidade tanto para entender contribuição relativa de processos ecológicos, quanto descrever padrões mais detalhados em sistemas ecológicos complexos.

Palavras-chave: padrões ecológicos, processos e mecanismos ecológicos, estatística ecológica, ecologia de comunidades, ecologia de florestas, ecologia de paisagens, modelos hierárquicos/multiníveis ou de efeito misto, partição de variâncias, componentes de variação

Abstract

Ecology is a science that is constantly concerned with testing hypotheses or even understanding ecological processes from patterns in data, which are mostly described using statistical tools. Mixed-effects models are increasingly used in ecology given the hierarchical nature of most ecological data and are especially useful in partitioning sources of variability. Mixed-effects models can provide quantitative answers that describe what fraction of a particular pattern is explained by different mechanisms: the "how much" rather than the "which". This shift in perspective can prove very useful in ecology both to understand relative contribution of ecological processes, and to better describe global patterns of variability. In this thesis, I employ and advocate for the use of mixed-effects models in very different ecological questions, but with a common analytical framework: variability partitioning in mixed-effects models.

In the first chapter, we seek to understand how matrix (non-habitat) modulates the relative contribution of community assembly processes (environmental filter and drift) of forest birds due to habitat loss in fragmented landscapes. To model bird occurrence, we used a mixed-effect model with association between species attributes and forest cover as fixed effects and several random components to measure relative strength of certain processes. We found that the environmental filter by habitat loss is the dominant process and can be relaxed or strengthened depending on matrix quality, evidencing that the matrix has a strong impact on modulating community assembly processes.

In the second chapter, we develop and apply a conceptual and analytical framework through variance partitioning of a mixed-effect model to attribute variability in tree vital rates due to differences among species, space, and time, and their interactions. We apply the framework to growth, mortality, and recruitment rates of 21 globally distributed forests covering over 2.9 million trees of approximately 6,500 species. Our framework has a high potential for identifying the structuring mechanisms of forest dynamics, as it highlights the most promising avenues for future research by increasing understanding of the relative contributions of groups of mechanisms to forest demography and diversity.

Along with variance partitioning tools, we highlight in this thesis the great potential of mixed-effects models in connecting statistical and ecological inference, incorporating variability as a diverse and good source of information both to understand relative

contribution of ecological processes and to describe more detailed patterns in complex ecological systems.

Keywords: ecological patterns, ecological processes and mechanisms, ecological statistics, community ecology, forest ecology, landscape ecology, hierarchical/multilevel or mixed-effect models, variance partitioning, variance components

Sumário

Introdução autobiográfica

*"Entre o que eu estudo antes
e escrevo no projeto,
o que eu estudo depois
e executo,
o que eu quero escrever
e o que você agora lê,
fatalmente, há um abismo"[1]*

Entender como as coisas variam é um dos meus temas preferidos, já que variabilidade é tão importante para entendermos ecologia e evolução como também fonte inspiradora da estatística. A ideia dessa tese começa, então, do meu interesse em como variabilidade nos faz conectar inferência ecológica e inferência estatística. Em compreender como nós, ecólogos, inferimos e relacionamos processos ecológicos a partir de padrões observados, e como estes padrões emergem de modelos estatísticos aplicados aos dados ecológicos. Sempre me intrigou como é feita a tradução de modelos ecológicos em modelos estatísticos e como as premissas assumidas em ambos se conversam (ou não).

Este interessante debate entre inferência ecológica e estatística é bastante estudado na área da ecologia de comunidades (Vellend, 2016; Leibold & Chase, 2017). Por exemplo, quando se busca entender como as comunidades ecológicas são montadas e quais os processos ou princípios que as governam. A grande maioria dos modelos estatísticos que tentam inferir processos de montagem de comunidades, tais como seleção, deriva, dispersão, partem dos padrões em dados de abundâncias de espécies em um conjuntos de locais, as metacomunidades (Leibold *et al.*, 2004). Os padrões são geralmente variáveis que descrevem a composição (presença/ausência) e/ou abundância das espécies nos locais como variável resposta, e variáveis que expressam as possíveis causas desta estrutura como preditoras, por exemplo variáveis ambientais ou espaciais (distâncias entre locais). Nesse campo, grandes avanços teóricos e estatísticos foram conquistados entremeados por muito debate sobre vantagens e limitações de abordagens, principalmente por se tratar em sua maioria de dados observacionais (e não experimentais) e de modelos fenomenológicos (e não mecanísticos). Um dos grandes

---

[1] Paródia de um texto de/ou atribuído a Alejandro Jodorowsky: "Entre o que eu penso, o que eu quero dizer, o que eu digo e você ouve, o que você quer ouvir e o que você acha que entendeu, há um abismo".

debates que permanece em voga é o problema de se querer inferir processos ecológicos a partir de padrões nos dados, já que um padrão pode ter sido gerado por diferentes processos e um mesmo processo pode gerar diferentes padrões (Fortin & Dale, 2005; McGill *et al.*, 2007; Ulrich *et al.*, 2010; Baldridge *et al.*, 2015).

De onde parti

O meu projeto de doutorado se iniciou com uma busca em entender primeiro as conexões entre teorias de montagem de comunidades e a forma com que pesquisadores testam suas hipóteses. Os objetivos gerais eram (1) identificar, (2) comparar e testar métodos estatísticos que propunham a identificação e quantificação dos processos de montagem de comunidades, e por fim, (3) aplicar tais métodos em perguntas ecológicas usando comunidades sob diferentes condições.

Para o objetivo 1, eu mergulhei fundo na literatura e identifiquei duas grandes classes de métodos estatísticos, que serviram de arcabouço para o desenvolvimento da tese. A classe mais tradicional é baseada em métodos de ordenação multivariada que buscam identificar associações entre espécies em determinados locais e destas com variáveis ambientais e espaciais. Os livros-texto mais estudados foram Legendre & Legendre (2012), Borcard et al. (2011), Manly et al. (2017), por fornecerem um panorama geral destas técnicas com foco em comunidades ecológicas, tais como análise de correspondência canônica, escalonamento métrico multidimensional, análise de coordenadas principais (ou mesmo de componentes principais), e análise de redundância.

A partir da a análise de redundância (RDA), Borcard et al. (1992) propuseram a primeira forma de particionar a variação total na composição ou abundância das espécies da comunidade em componentes independentes de: efeitos ambientais (puro), efeitos espaciais (puro), estruturação espacial dos dados ambientais (a interseção entre espaço e ambiente), e o resíduo. Este método foi o precursor da partição de variabilidade em dados ecológicos multivariado que permitiu quantificar a contribuição relativa de grupos de variáveis explicativas (ambientais e espaciais) na estrutura das comunidades. Durante as últimas décadas, este método foi extensamente estudado e proposto como sendo capaz de separar processos baseados em nicho de processos neutros (Gilbert & Lechowicz, 2004; Diniz Filho *et al.*, 2012; Tuomisto *et al.*, 2012), ou mesmo para classificar as metacomunidades em diferentes tipos (Cottenie, 2005). Muito

dos trabalhos metodológicos desenvolvidos a partir dessa proposta focaram nas análises espaciais multi-escalas na teoria de metacomunidades (Legendre *et al.*, 2005; Dray *et al.*, 2012). Apesar do forte desenvolvimento dos métodos de ordenação entre as décadas de 1990 até 2010, a maioria das críticas recai sobre a falta de articulação entre os resultados das análises estatísticas e a inferência ecológica (Gilbert & Bennett, 2010; Tuomisto *et al.*, 2012; Vellend *et al.*, 2014).

A segunda classe de métodos que ganhou força e um rápido desenvolvimento na última década é a abordagem baseada em modelos (*model-based approach,* Warton *et al.*, 2015), principalmente com o uso de modelos lineares generalizados mistos, que neste caso são chamados de modelo conjunto de distribuição de espécies (*Joint Species Distribution Models,* JSDM). Estes modelos foram propostos como uma extensão de modelos lineares de distribuição de espécies (Ovaskainen & Soininen, 2011; Pollock *et al.*, 2014), cuja variável resposta é a abundância ou presença/ausência da espécie nos locais e as preditoras são, geralmente, variáveis ambientais e/ou espaciais. Estes modelos são também comumente usados para analisar relações atributo-ambiente (Jamil *et al.*, 2013; ter Braak, 2019), para separar processos dependentes de escala (Ovaskainen *et al.*, 2017), podendo levar em consideração atributos das espécies, autocorrelação espacial e filogenética (Warton *et al.*, 2015; Ovaskainen *et al.*, 2017; Miller *et al.*, 2018; ter Braak, 2019; Pichler & Hartig, 2021). Essa abordagem é promissora dados os recentes avanços em técnicas computacionais e permite muita flexibilidade na estrutura dos dados, possui também as vantagens das análises de diagnóstico de modelos lineares (Warton *et al.*, 2015).

Existem muitas maneiras de se conduzir estudos empíricos em ecologia de comunidades e, como vimos, o desenvolvimento e aplicação de novos métodos tem ocorrido ininterruptamente pelos últimos 50 anos (Vellend, 2016) e caracteriza-se como um grande impulso das pesquisas atuais (Anderson *et al.*, 2011; Legendre & Legendre, 2012; Warton *et al.*, 2015; Ovaskainen *et al.*, 2017). Navegar neste oceano não foi nada fácil. Desvios de rota foram tomados, mas me levaram a lugares interessantes e portos mais seguros.

Antes de avançarmos para "onde cheguei", devo mencionar um ponto de partida importante: o fato de eu trabalhar desde 2010 como ecóloga especialista em laboratório no Departamento de Ecologia da USP, o mesmo que desenvolvi esta tese. Foi uma escolha consciente envolver diferentes laboratórios na minha pesquisa de doutorado,

aproveitando as parcerias e permitindo continuidade de projetos já estabelecidos. Estas colaborações seriam particularmente importantes para o capítulo 3, trazendo temas e sistemas distintos em ecologia para um mesmo arcabouço teórico-metodológico. As parcerias estabelecidas entre Laboratório de Ecologia Teórica (o qual desenvolvi a tese), o Laboratório de Ecologia de Paisagens e Conservação (LEPAC) e o Laboratório de Ecologia de Florestas Tropicais (LabTrop) enriqueceram muito o projeto e ampliaram a rede de colaborações com pesquisadores de diversas outras instituições. A seguir, descrevo o que realmente aconteceu.

Aonde cheguei

"Não sei, só sei que foi assim!"
Chicó[2]

Dentre os objetivos originalmente propostos, o objetivo 1 de identificar os métodos estatísticos em montagem de comunidades foi parcialmente alcançado dados os estudos de imersão no tema, mas não publicado. O objetivo 2 previa comparar e testar as duas grandes classes de métodos estatísticos descritos acima, mas decidimos com meu comitê de acompanhamento[3], não focar nesse objetivo, porque suspeitávamos que outros pesquisadores poderiam já estar desenvolvendo estudos a respeito. Essa suspeita se confirmou, pois artigos sobre o tema foram publicados por diferentes grupos de pesquisa em pouco tempo (e.g., Norberg *et al.*, 2019; Roberts, 2019; ter Braak, 2019; Wilkinson *et al.*, 2019). Nesse momento, observamos um maior potencial em desenvolver as pesquisas empíricas, seguindo para a aplicação destes métodos a perguntas ecológicas concretas. Partimos então para o objetivo 3, que se desmembrou nos capítulos 1 e 2 da tese.

O primeiro capítulo segue fielmente o arcabouço teórico do projeto original descrito brevemente acima, envolvendo as teorias ecológicas de montagem de comunidades em perguntas de ecologia aplicada. Em parceria com pesquisadores do LEPAC[4], analisamos a importância da matriz (não-habitat) na modulação de processos de (des)montagem de comunidades de aves com a perda de habitat em paisagens

---

[2] O Auto da Compadecida, Ariano Suassuna
[3] O comitê era formado pelos professores Alexandre Adalardo de Oliveira (USP) e Tadeu Siqueira (UNESP-Rio Claro) e foi muito importante para me ajudar a navegar pelos caminhos da tese de forma mais assertiva e consistente.
[4] Jean Paul Metzger e Andrea Larissa Boesing, coautores do artigo do primeiro capítulo.

fragmentadas. Neste capítulo, combinamos Ecologia da Paisagem com Ecologia de Comunidades e desenvolvemos um modelo JSDM capaz de particionar a variabilidade na probabilidade de ocorrência das aves em diferentes componentes para assim entender a contribuição relativa dos diferentes processos como filtro ambiental e deriva ecológica e como essa contribuição relativa varia em função da qualidade da matriz.

Vocês verão que o segundo capítulo tomou uma rota alternativa, mas ainda assim mantém o eixo central das relações entre ecologia e estatística da proposta. No projeto original, eu propus analisar processos de montagem de comunidades de árvores em florestas tropicais, buscando principalmente quantificar a contribuição relativa dos processos de montagem ao longo de gradientes ambientais. Com ambos os capítulos, eu teria analisado comunidades bem distintas taxonomicamente (aves *versus* árvores) e ambientalmente (paisagens fragmentadas *versus* florestas contínuas pouco perturbadas), podendo investigar quais processos seriam relativamente mais importantes do que outros dependendo da comunidade estudada. Os dados seriam provenientes de duas parcelas de florestas coordenadas pelo LabTrop e que fazem parte da rede mundial de parcelas permanentes ForestGEO (Forest Global Ecology Observatory Netowork, Davies *et al.*, 2021).

Em junho de 2019, tive a oportunidade de levar o projeto para o Workshop do ForestGEO. Lá conheci meus futuros coorientadores internacionais[5], que se interessaram pelas ideias sobre uso de partição de variâncias em modelos mistos, mas que me sugeriram desenvolver os modelos com dados dinâmicos de taxas vitais de árvores, ou seja crescimento, mortalidade e recrutamento. Neste momento, comecei a ampliar o foco da minha pesquisa, imergindo no mundo das teorias e mecanismos ecológicos em dinâmica florestal. A partir daí, desenvolvemos um arcabouço conceitual e analíticos para particionar os grandes grupos de fontes de variabilidade em taxas vitais de árvores – espécie, espaço, tempo e suas interações. Apliquei este arcabouço em dados de 21 parcelas permanentes espalhadas pelo mundo de forma a entender os padrões macroecológicos das fontes de variabilidade nas taxas vitais em árvores.

---

[5] Prof. Marco Visser da Universidade de Leiden, Prof. Sean McMahon do Smithsonian Institution e Profa. Lisa Hülsmann da Universidade de Bayreuth.

Modelando a variabilidade: o maravilhoso mundo dos modelos mistos

*"Statistics is concerned with the variability
that is evident in any body of data."*
Searle et al. 2006

Percebendo rápida disseminação do uso de modelos mistos[6] atualmente na ecologia, é de se pensar que sejam uma descoberta recente no campo da estatística. Ledo engano! Searle et al. (2006) nos conta a história do desenvolvimento do interesse estatístico nos componentes de variação (ou variância) dos dados e de como surgiram os precursores de modelos de efeito aleatório e mistos. Curiosamente, a primeira formulação de um modelo de efeitos aleatórios foi em 1861, no qual o astrônomo Airy assume que diferentes noites de observação telescópica devem possuir diferentes estimativas por efeitos de "circunstância atmosférica e pessoais" (Searle *et al.*, 2006). Temos então *noite* como a primeira variável de efeitos aleatórios da história[7]! Mas foi a Análise de Variâncias (ANOVA), desenvolvida por Ronald Fisher entre as décadas de 1920 e 1940, um dos primeiros métodos a resumir a variabilidade nos dados, permitindo particionar a variância observada entre categorias de um fator ou mais fatores (variável explicativa) da variabilidade residual dos dados. Posteriormente, o desenvolvimento dos modelos mistos substituiu o uso da ANOVA como a ferramenta mais comum para análises de variância e partição de variabilidade nos dados (Wainwright *et al.*, 2007; Bolker *et al.*, 2009; Boisgontier & Cheval, 2016). Mas afinal o que são modelos mistos e quais suas utilidades na ecologia?

Modelos mistos começaram a ganhar destaque nas ciências ecológicas apenas a partir dos anos 2000 (e.g., McMahon & Diez, 2007; Bolker, 2009; Cressie *et al.*, 2009; Zuur *et al.*, 2009). Ainda assim, acredito que pouco do potencial destes modelos está sendo aplicado e, portanto, esta tese é também uma tentativa de promover o uso e estudo destes modelos na ecologia. Há quem diga que os modelos mistos deveriam ser a regra e não a exceção na aplicação de modelos lineares (e.g., McElreath, 2016) e que vieram para revolucionar a pesquisa ecológica, removendo restrições que forçaram

---

[6] Ou modelos de efeito misto, ou modelos hierárquicos ou multiníveis.
[7] Este fato me lembrou do meu artigo do mestrado (Leite *et al.*, 2013), no qual também usei noite**,** além do indivíduo, como efeito aleatório para modelar os padrões de atividade da cuíca d'água, *Chironectes minimus*. Este foi o primeiro modelo misto que construí e publiquei!

ecólogos a simplificar em demasia os modelos estatísticos e a ignorar importantes distinções entre erros de medida, erros de processo e incerteza de modelos (Cressie *et al.*, 2009). Uma das razões é que dados ecológicos são frequentemente agrupados/hierárquicos, necessitando de um modelo com estrutura aleatória para evitar pseudoreplicação (Arnqvist, 2020). Entretanto, ainda temos muito a aprender sobre o comportamento destes modelos, suas utilidades e limitações.

O que diferencia um modelo linear "simples" de um modelo misto é a distinção entre **efeitos fixos e aleatórios** na tomada de decisão de como uma variável categórica deve ser modelada. Isso depende primordialmente do objetivo do estudo, do contexto dos dados e desenho amostral, e de quais propriedades se espera obter do modelo e da variável. Uma das formas de decisão mais simples é avaliar se o interesse está nas estimativas de cada categoria (*narrow-sense inference*, McLean *et al.*, 1991), se as categorias são independentes, mutualmente exclusivas e completamente observáveis. Nesse caso, deveriam ser estimados por efeitos fixos, por exemplo, sexo ou tratamentos diferentes em experimentos. Efeito aleatórios são uma escolha geralmente quando a variabilidade entre as categorias e não as estimativas exatas de cada uma são de interesse (*broad-sense inference*), quando os níveis podem ser considerados como uma amostragem aleatória de uma população de valores, e o interesse é controlar por esta variabilidade (também podendo ser considerada pseudoreplicação caso não controlado), por exemplo blocos em experimentos ou outras unidades amostrais que agrupam muitas observações (sub-amostras). As decisões e formas de pensar um modelo com variáveis de efeitos fixos ou aleatórios e as diferenças técnicas entre eles não se esgota aqui e podem ser estudadas em Gelman & Hill (2007), Harrisson (2018), Oberpriller et al. (2022). Antes de discorrer sobre partição de variâncias em modelos mistos (a grande pérola da tese), gostaria de compartilhar um projeto realizado em paralelo à minha tese com foco, claro, em modelos mistos.

Fixo ou aleatório? O que fazer quando há poucas categorias

Quando o pesquisador decide que vai modelar uma variável categórica como um efeito aleatório, existe uma segunda pergunta a ser feita: quantas categorias essa variável tem? Ter poucas categorias significa maior incerteza ou dificuldade em se

estimar a variância entre elas. E quais as consequências para interpretação do modelo quando isso acontece? Qual o número mínimo de categorias para que uma variável seja modelada como um efeito aleatório? Essas perguntas, por simples que pareçam ainda não tinham sido investigadas a fundo, e eu me sentia frustrada toda vez que me deparava com estes questionamentos entre colegas. Na ecologia, não é difícil encontrar situações em que a variável possui poucas categorias, por exemplo em experimentos de campo no qual os blocos de experimentos são limitados.

Durante meu estágio sanduíche na Alemanha (Universidade de Regensburg), engajei dois alunos de doutorado do laboratório para realizarmos um estudo motivado justamente por estas perguntas (Oberpriller *et al.*, 2022). Nós utilizamos simulações baseadas em um contexto ecológico simples de amostragem aninhada, como se quiséssemos estudar o efeito da temperatura na altura de plantas, estabelecendo transectos altitudinais em diferentes montanhas (ver Box 1 da publicação original). A princípio, a variável montanha deveria ser modelada como efeito aleatório (*broad-sense inference*), dado que plantas de uma mesma montanha não são amostras independentes (pseudoreplicação). Nós analisamos as consequências de tratar a variável montanha com 2 a 8 níveis como efeito fixo ou aleatório em modelos corretamente especificados (ou seja, mesmo modelo usado na simulação dos dados) e modelos alternativos (sub- ou sobreparametrizado). Com poucos níveis, as estimativas de variância dos efeitos aleatórios podem ser imprecisas e enviesadas e isso pode também influenciar as estimativas dos efeitos fixos de interesse (efeito de temperatura). Calculamos as taxas de erro tipo I[8] e o poder estatístico destes modelos e encontramos que inferir a estrutura correta dos efeitos aleatórios é de grande importância para obter taxas de erro tipo I corretas. Elaboramos recomendações práticas para ecólogos na escolha das melhores estruturas do modelo dado as hipóteses e o desenho amostral do estudo. Uma destas recomendações é sempre começar com um modelo mistos com os efeitos aleatórios especificados independentemente do número de categorias e mudar para um modelo de efeitos fixos se houver problemas em estimar as variâncias dos termos aleatórios ("*singular fit*" ou variância estimada em zero). Com isso, acreditamos em escolhas mais informativas para análises baseados no desenho do estudo e natureza dos dados, e fazer inferências mais robusta para o caso de poucas categorias. Realizar este estudo foi muito

---

[8] Erro tipo I é quando não existe efeito da variável preditora sobre a resposta, neste caso, não existe efeito de temperatura nas plantas, mas o teste aponta que sim.

gratificante! É um trabalho do qual me orgulho e espero que possa contribuir para o uso e desenvolvimento de modelos mistos na ecologia.

Particionar a variabilidade: o "quanto" ao invés de "qual"

O uso de modelos cada vez mais complexos na ecologia é uma tendência que está ligada não só com a oferta das ferramentas, mas principalmente com uma mudança de mentalidade analítica, e possivelmente com o amadurecimento de teorias (Low-Décarie *et al.*, 2014). Há um bom tempo, Quinn & Dunham (1983) afirmaram ser enganosos modelar as variações observadas nos padrões e processos de comunidades atribuindo-as a apenas uma única causa, e propuseram uma visão não mutualmente exclusiva das causas de um fenômeno em termos da contribuição relativa de cada alternativa. Bolker (2007) chega à mesma conclusão ao dizer que, em vez de conceber um único experimento que determina qual explicação está certa, somos motivados a fornecer respostas quantitativas que descrevem que fração de um padrão em particular é explicada por diferentes mecanismos: o "quanto" em vez de "qual". Essa mudança de perspectiva nos leva, por exemplo, a repensar os processos ecológicos em ecologia de comunidades e como cada um contribui relativamente para a formação do padrão de abundâncias de espécies na comunidade. E as perguntas se tornam, por exemplo: *Qual a contribuição relativa de diferentes processos de montagem de comunidades após perda de habitat e fragmentação (Capítulo 1)? Qual a importância relativa de mecanismos de nicho na variabilidade de taxas vitais em árvores (Capítulo 2)?*

Para responder às perguntas de "quanto" em vez de "qual", entram em cena os componentes de variação e as formas de quantificação da contribuição relativa destes na variabilidade total dos dados. Já disse anteriormente que os primeiros métodos a se propor a partição de variâncias em ecologia de comunidades foram os métodos multivariado de RDA (Borcard et al. 1992), cujas ideias foram trazidas dos modelos lineares (p.ex. ANOVA, lembra?). Em um modelo linear, a forma mais simples para quantificar o quanto de variabilidade dos dados (variável resposta) é explicado por variáveis de interesse (variáveis explicativas) é o uso do coeficiente de determinação (R^2) que é expresso como uma proporção da variância total na resposta. Esta é uma forma crua de se assumir que existe algum tipo de variação decorrente de uma combinação de fatores desconhecidos (Bolker, 2007) que geralmente não são de interesse, os famigerados resíduos. Entretanto, R^2 é um conceito intuitivo (Nakagawa &

Schielzeth, 2013) e vai além dos testes de significância ao colocar os efeitos estudados em perspectiva da variação total nos dados (Stoffel *et al.*, 2021), e que pode significar uma quantidade biológica relevante se esta variação é representativa da população total (de Villemereuil *et al.*, 2018).

Modelos que fazem premissas explícitas sobre as causas subjacentes de variabilidade podem tanto prover mais informação sobre os processos ecológicos quanto obter mais dos dados (Bolker, 2007). Desta forma, podemos ir um pouco mais além da abordagem de partição de variâncias em modelos lineares ao especificar diferentes fontes de variabilidade e particionar a variância do que antes era tido como "não explicada" em diversos componentes. Tais componentes, por sua vez, podem refletir os diferentes processos ou grupos de processos ecológicos em questão. E aqui entram os modelos mistos e seus efeitos aleatórios, nos quais é possível atribuir uma parte da variação a componentes que não são efeitos em si, apenas fonte de variação, mudando assim a forma de entender o que "não é explicado". O histórico de uso de R^2 em modelos mistos, na verdade, começa com o termo Coeficiente de Partição de Variâncias (em inglês, Variance Partition Coefficient, VPC) para descrever a porcentagem de variação que é atribuído a um nível ou classificação particular nos dados (Browne *et al.*, 2005). Porém, dada a complexidade de modelos mistos e as relações multiníveis entre termos de efeitos fixos e aleatórios, calcular o VPC dos componentes ou mesmo o R^2 total destes modelos não é tarefa fácil. Felizmente, muitos estudos recentes têm desenvolvido métodos para a partição de variâncias em modelos mistos, especialmente para modelos com distribuição não normal[9] (Goldstein *et al.*, 2002; Browne *et al.*, 2005) e muitos na área da ecologia (Nakagawa & Schielzeth, 2013; Nakagawa *et al.*, 2017; Ives, 2019; Stoffel *et al.*, 2021).

A flexibilidade de modelagem e as técnicas de partição de variâncias (R^2 ou VPC) em modelos mistos me permitiu fazer uso criativo de variáveis categóricas que podem explicar parte da variação das comunidades de forma bem diferente do senso comum elencado nas explicações acima. O modelo utilizado no capítulo 1 é um Joint Species Distribution Models (JSDMs), que, como dito anteriormente, são uma classe de modelos mistos usados para modelar abundância e ocorrência de espécies em

---

[9] Particionar variância em modelos lineares de distribuição não normal, como Poisson ou Binomial, é um pouco mais complicado, porque a variância residual não pode ser facilmente definida nestes modelos. Assim, geralmente chama-se este tipo de R^2 de pseudo-R^2, pois, apesar de ter mesmas propriedades quanto à interpretação, são calculados de formas diferentes.

comunidades, podendo levar em consideração variáveis ambientais, atributos das espécies, autocorrelação espacial e filogenética (Warton *et al.*, 2015; Ovaskainen *et al.*, 2017; Miller *et al.*, 2018; ter Braak, 2019; Pichler & Hartig, 2021). Eu criei um modelo baseado nas características do meu estudo e dos meus objetivos, interpretando então os R^2 de cada componente de efeito aleatório e dos efeitos fixos de acordo com o que seria esperado encontrar mais variação nos dados de acordo com os processos dominantes de montagem. Por exemplo, escolhemos variáveis preditoras associadas aos processos de filtro ambiental, mas com a relação entre os atributos funcionais das aves e a o gradiente de cobertura florestal. Incorporamos também efeitos aleatórios para capturar a variação estocástica da ocorrência das espécies em manchas pequenas e isoladas de habitats remanescentes.

Já no capítulo 2, o modelo proposto é um modelo de efeitos aleatórios apenas, ou seja, não existem variáveis de efeito fixo, pois o objetivo principal é a formulação de um modelo base (*baseline*) para partição de variâncias entre grupos de componentes que sabemos criar variabilidade nos dados (*organizing principles*) de forma a elucidar padrões macroecológicos e indicar foco em testes de mecanismos e hipóteses sobre variabilidade em taxas vitais de árvores. O uso de modelos mistos, ou apenas modelos de efeito aleatório, como base para construção de modelos mais complexos tem surgido timidamente na literatura (Camac *et al.*, 2018; Solbu *et al.*, 2022), e eu acredito ser uma vantagem no entendimento primordial das fontes de variabilidade nos dados antes mesmo de testar hipóteses e mecanismos.

O compromisso com a ciência aberta

Abro este espaço para discutir um tema que vejo de extrema importância e que esteve presente durante todo o processo de formação do doutorado, o movimento da **ciência aberta**, **transparente e reprodutível**. Este movimento engloba diferentes práticas, ferramentas e códigos de conduta que visam tornar a atividade científica mais transparente, reprodutível e colaborativa, permitindo assim a reutilização, redistribuição e reprodução dos produtos científicos. O resultado é uma ciência mais íntegra, confiável e melhores forma de evitar erros, fraudes e má fé (veja Smaldino & McElreath, 2016 sobre a seleção natural da má ciência). Diversas iniciativas para promoção da ciência aberta estão surgindo recentemente na ecologia (e.g., Powers & Hampton, 2019; Figueiredo *et al.*, 2022). E não foi sem esforço que, pelo meu compromisso com a

ciência aberta, adotei diferentes práticas para tornar a pesquisa desta tese mais transparente, reprodutível e aberta.

A adoção de pequenas práticas em código reprodutível permitiu que eu pudesse publicar o *PrePrint* do primeiro capítulo da tese (Leite *et al.*, 2021), junto com os dados e códigos detalhados das análises (Leite, 2022). Os dados brutos já estavam disponíveis em repositório aberto (Boesing *et al.*, 2017) associado à publicação original (Boesing *et al.*, 2018), mostrando as vantagens da publicação de dados em novas colaborações e um exemplo de "repropósito" de dados existentes através de novas perguntas não feitas no trabalho original. A publicação do repositório de análises foi possível e facilitada com o uso de controle de versão em repositório online (GitHub) desde as primeiras análises, o uso de ferramentas computacionais (p.ex. Rmarkdown) e boas práticas de escrita de código para deixar todo o processo de análises compreensível. A publicação deste repositório de análise em um outro repositório aberto (Zenodo) com DOI citável e rastreável foi praticamente direta. Isso tornou este trabalho reprodutível computacionalmente e até mesmo didático[10]. O código foi posteriormente revisado por um grupo de trabalho (*Peer Code Review Group*) da SORTEE (Society for Open, Reliable, and Transparent Ecology and Evolutionary biology), da qual faço parte, e recebeu *feedback* sobre melhorias em relação à clareza e usabilidade, etapa esta importantíssima, porém comumente negligenciada.

O segundo capítulo da tese segue caminho similar, porém com um desafio a mais, já que muitos dos dados brutos não estão abertos publicamente por diversas razões (veja em Lima *et al.*, 2022 uma discussão importante sobre o tema). A solução que encontrei para manter a análise mais transparente possível foi disponibilizar um exemplo da modelagem com dados abertos da famosa parcela de Barro Colorado Island (Condit *et al.*, 2019). O que está público são os resultados das análises realizadas em cada floresta separadamente e as subsequentes análises com estes dados.

Pela minha experiência, publicar dados, códigos de análises e versões preliminares de manuscrito (*PrePrint*) requerem muito mais tempo para aprendizado de ferramentas, dedicação e implementação do que imaginamos. Percebo também que um ambiente estimulador e inclusivo faz toda a diferença, porque nós estudantes, quando

---

[10] Eu mesma já me beneficiei e aprendi bastante em estudar artigos com código e recursos em material complementar, entendendo muito mais a fundo quando o código está aberto e bem estruturado para a leitura humana (e não apenas da máquina).

não percebemos vantagens ou recebemos incentivos para adoção destas práticas, acabamos por abandonar qualquer iniciativa demandante e não claramente recompensadora. Porém, o número de visualizações e de interações online que obtive pelos produtos disponibilizados e o sucesso em palestra sobre o tema[11], me demostraram que o esforço vale a pena e reforça o compromisso e entusiasmo com a ciência aberta!

Estrutura da tese

O que se segue a partir daqui são:

(1) O Capítulo 1 com o texto original em inglês publicado recentemente no periódico *Journal of Applied Ecology* (Leite et al. 2022),

(2) O Capítulo 2 em inglês também escrito em formato de artigo para submissão à revista *Global Ecology and Biogeography*.

(3) A conclusão geral da tese com os aprendizados e conhecimentos gerados.

Boa leitura!

Referências

Anderson, M.J., Crist, T.O., Chase, J.M., Vellend, M., Inouye, B.D., Freestone, A.L., Sanders, N.J., Cornell, H.V., Comita, L.S., Davies, K.F., Harrison, S.P., Kraft, N.J.B., Stegen, J.C. & Swenson, N.G. (2011) Navigating the multiple meanings of β diversity: a roadmap for the practicing ecologist. *Ecology Letters*, **14**, 19–28.

Arnqvist, G. (2020) Mixed Models Offer No Freedom from Degrees of Freedom. *Trends in Ecology & Evolution*, **35**, 329–335.

Baldridge, E., Xiao, X. & White, E.P. (2015) An extensive comparison of species-abundance distribution models. *bioRxiv*, 024802.

Boesing, A.L., Nichols, E. & Metzger, J.P. (2018) Biodiversity extinction thresholds are modulated by matrix type. *Ecography*, **41**, 1520–1533.

Boesing, A.L., Nichols, E. & Metzger, J.P. (2017) Data from: Biodiversity extinction thresholds are modulated by matrix type. *Dryad, Dataset, https://doi.org/10.5061/dryad.37s4h*.

Boisgontier, M.P. & Cheval, B. (2016) The anova to mixed model transition. *Neuroscience & Biobehavioral Reviews*, **68**, 1004–1005.

Bolker, B. (2007) *Ecological Models and Data in R*, Princeton University Press, Princeton.

Bolker, B. (2009) Learning hierarchical models: advice for the rest of us. *Ecological Applications*,

---

[11] Recentemente, em julho de 2022, dei uma palestra sobre dados abertos no ciclo de seminários do PPG-Ecologia, USP. Dado o interesse e engajamento da comunidade sobre o tema, planejo seguir com estudos e diferentes forma de divulgação, por exemplo com a ideia de criar um curso sobre ciência aberta, reprodutibilidade e dados no programa.

19, 588–592.

Bolker, B.M., Brooks, M.E., Clark, C.J., Geange, S.W., Poulsen, J.R., Stevens, M.H.H. & White, J.-S.S. (2009) Generalized linear mixed models: a practical guide for ecology and evolution. *Trends in ecology & evolution*, **24**, 127–35.

Borcard, D., Gillet, F. & Legendre, P. (2011) *Numerical ecology with R,*.

Borcard, D., Legendre, P. & Drapeau, P. (1992) Partialling out the Spatial Component of Ecological Variation. *Ecology*, **73**, 1045–1055.

ter Braak, C.J.F. (2019) New robust weighted averaging- and model-based methods for assessing trait–environment relationships. *Methods in Ecology and Evolution*, **10**, 1962–1971.

Browne, W.J., Subramanian S. V., Jones K., & Goldstein H. (2005) Variance partitioning in multilevel logistic models that exhibit overdispersion. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, **168**, 599–613.

Camac, J.S., Condit, R., FitzJohn, R.G., McCalman, L., Steinberg, D., Westoby, M., Wright, S.J. & Falster, D.S. (2018) Partitioning mortality into growth-dependent and growth-independent hazards across 203 tropical tree species. *Proceedings of the National Academy of Sciences*, **115**, 12459–12464.

Condit, R.S., Perez, R., Aguilar, S., Lao, S., Foster, R. & Hubbell, S.P. (2019) *Complete data from the Barro Colorado 50-ha plot: 423617 trees, 35 years, 2019 version,*.

Cottenie, K. (2005) Integrating environmental and spatial processes in ecological community dynamics. *Ecology Letters*, **8**, 1175–1182.

Cressie, N., Calder, C. a., Clark, J.S., Ver Hoef, J.M. & Wikle, C.K. (2009) Accounting for uncertainty in ecological analysis: The strengths and limitations of hierarchical statistical modeling. *Ecological Applications*, **19**, 553–570.

Davies, S.J., Abiem, I., Abu Salim, K., Aguilar, S., Allen, D., Alonso, A., Anderson-Teixeira, K., Andrade, A., Arellano, G., Ashton, P.S., Baker, P.J., Baker, M.E., Baltzer, J.L., Basset, Y., Bissiengou, P., Bohlman, S., Bourg, N.A., Brockelman, W.Y., Bunyavejchewin, S., Burslem, D.F.R.P., Cao, M., Cárdenas, D., Chang, L.-W., Chang-Yang, C.-H., Chao, K.-J., Chao, W.-C., Chapman, H., Chen, Y.-Y., Chisholm, R.A., Chu, C., Chuyong, G., Clay, K., Comita, L.S., Condit, R., Cordell, S., Dattaraja, H.S., de Oliveira, A.A., den Ouden, J., Detto, M., Dick, C., Du, X., Duque, Á., Ediriweera, S., Ellis, E.C., Obiang, N.L.E., Esufali, S., Ewango, C.E.N., Fernando, E.S., Filip, J., Fischer, G.A., Foster, R., Giambelluca, T., Giardina, C., Gilbert, G.S., Gonzalez-Akre, E., Gunatilleke, I.A.U.N., Gunatilleke, C.V.S., Hao, Z., Hau, B.C.H., He, F., Ni, H., Howe, R.W., Hubbell, S.P., Huth, A., Inman-Narahari, F., Itoh, A., Janík, D., Jansen, P.A., Jiang, M., Johnson, D.J., Jones, F.A., Kanzaki, M., Kenfack, D., Kiratiprayoon, S., Král, K., Krizel, L., Lao, S., Larson, A.J., Li, Y., Li, X., Litton, C.M., Liu, Y., Liu, S., Lum, S.K.Y., Luskin, M.S., Lutz, J.A., Luu, H.T., Ma, K., Makana, J.-R., Malhi, Y., Martin, A., McCarthy, C., McMahon, S.M., McShea, W.J., Memiaghe, H., Mi, X., Mitre, D., Mohamad, M., Monks, L., Muller-Landau, H.C., Musili, P.M., Myers, J.A., Nathalang, A., Ngo, K.M., Norden, N., Novotny, V., O'Brien, M.J., Orwig, D., Ostertag, R., Papathanassiou, K., Parker, G.G., Pérez, R., Perfecto, I., Phillips, R.P., Pongpattananurak, N., Pretzsch, H., Ren, H., Reynolds, G., Rodriguez, L.J., Russo, S.E., Sack, L., Sang, W., Shue, J., Singh, A., Song, G.-Z.M., Sukumar, R., Sun, I.-F., Suresh, H.S., Swenson, N.G., Tan, S., Thomas, S.C., Thomas, D., Thompson, J., Turner, B.L., Uowolo, A., Uriarte, M., Valencia, R., Vandermeer, J., Vicentini, A., Visser, M., Vrska, T., Wang, X., Wang, X., Weiblen, G.D., Whitfeld, T.J.S., Wolf, A., Wright, S.J., Xu, H., Yao, T.L., Yap, S.L., Ye, W., Yu, M., Zhang, M., Zhu, D., Zhu, L., Zimmerman, J.K. & Zuleta, D. (2021) ForestGEO: Understanding forest diversity and dynamics through a global observatory network. *Biological Conservation*, **253**, 108907.

Diniz Filho, J.A.F., Rangel, T.F., Santos, T. & Mauricio Bini, L. (2012) Exploring Patterns of Interspecific Variation in Quantitative Traits Using Sequential Phylogenetic Eigenvector Regressions. *Evolution*, **66**, 1079–1090.

Dray, S., Pélissier, R., Couteron, P., Fortin, M.J., Legendre, P., Peres-Neto, P.R., Bellier, E., Bivand, R., Blanchet, F.G., De Cáceres, M., Dufour, a. B., Heegaard, E., Jombart, T., Munoz, F.,

Oksanen, J., Thioulouse, J. & Wagner, H.H. (2012) Community ecology in the age of multivariate multiscale spatial analysis. *Ecological Monographs*, **82**, 257–275.

Figueiredo, L., Scherer, C. & Cabral, J.S. (2022) A simple kit to use computational notebooks for more openness, reproducibility, and productivity in research. *PLOS Computational Biology*, **18**, e1010356.

Fortin, M.-J. & Dale, M. (2005) *Spatial analysis: a guide for ecologists*, Cambridge University Press.

Gelman, A. & Hill, J. (2007) *Data Analysis Using Regression and Multilevel/Hierarchical Models*, Cambridge University Press.

Gilbert, B. & Bennett, J.R. (2010) Partitioning variation in ecological communities: do the numbers add up? *Journal of Applied Ecology*, **47**, 1071–1082.

Gilbert, B. & Lechowicz, M.J. (2004) Neutrality, niches, and dispersal in a temperate forest understory. *Proceedings of the National Academy of Sciences of the United States of America*, **101**, 7651–7656.

Goldstein, H., Browne, W. & Rasbash, J. (2002) Partitioning Variation in Multilevel Models. *Understanding Statistics*, **1**, 223–231.

Harrison, X.A., Donaldson, L., Correa-Cano, M.E., Evans, J., Fisher, D.N., Goodwin, C.E.D., Robinson, B.S., Hodgson, D.J. & Inger, R. (2018) A brief introduction to mixed effects modelling and multi-model inference in ecology. *PeerJ*, **6**, e4794.

Ives, A.R. (2019) R2s for Correlated Data: Phylogenetic Models, LMMs, and GLMMs. *Systematic Biology*, **68**, 234–251.

Jamil, T., Ozinga, W.A., Kleyer, M. & ter Braak, C.J.F. (2013) Selecting traits that explain species–environment relationships: a generalized linear mixed model approach. *Journal of Vegetation Science*, **24**, 988–1000.

Legendre, P., Borcard, D. & Peres-Neto, P.R. (2005) Analyzing Beta Diversity: Partitioning the Spatial Variation of Community Composition Data. *Ecological Monographs*, **75**, 435–450.

Legendre, P. & Legendre, L. (2012) *Numerical ecology 3rd*, 3rd edn.

Leibold, M.A. & Chase, J.M. (2017) *Metacommunity Ecology*, Princeton University Press.

Leibold, M.A., Holyoak, M., Mouquet, N., Amarasekare, P., Chase, J.M., Hoopes, M.F., Holt, R.D., Shurin, J.B., Law, R., Tilman, D., Loreau, M. & Gonzalez, A. (2004) The metacommunity concept: a framework for multi-scale community ecology. *Ecology Letters*, **7**, 601–613.

Leite, M. de S. (2022) Data and code from: Matrix quality determines the strength of habitat loss filtering on bird communities at the landscape scale (v2.2). *Zenodo*.

Leite, M. de S., Boesing, A.L., Metzger, J.P. & Prado, P.I. (2021) Matrix quality determines the strength of habitat loss filtering on bird communities at the landscape scale.

Leite, M. de S., Queiroz, T.L., Galliez, M., de Mendonça, P.P. & Fernandez, F. a. S. (2013) Activity patterns of the water opossum Chironectes minimus in Atlantic Forest rivers of south-eastern Brazil. *Journal of Tropical Ecology*, **29**, 261–264.

Lima, R.A.F., Phillips, O.L., Duque, A., Tello, J.S., Davies, S.J., de Oliveira, A.A., Muller, S., Honorio Coronado, E.N., Vilanova, E., Cuni-Sanchez, A., Baker, T.R., Ryan, C.M., Malizia, A., Lewis, S.L., ter Steege, H., Ferreira, J., Marimon, B.S., Luu, H.T., Imani, G., Arroyo, L., Blundo, C., Kenfack, D., Sainge, M.N., Sonké, B. & Vásquez, R. (2022) Making forest data fair and open. *Nature Ecology & Evolution*, 1–3.

Low-Décarie, E., Chivers, C. & Granados, M. (2014) Rising complexity and falling explanatory power in ecology. *Frontiers in Ecology and the Environment*, 140821075159009.

Manly, B.F.J. & Navarro Alberto, J.A. (2017) *Multivariate statistical methods: a primer*, Fourth edition. CRC Press, Boca Raton.

McElreath, R. (2016) *Statistical Rethinking: A Bayesian Course with Examples in R and Stan*, CRC

Press, Boca Raton.

McGill, B.J., Etienne, R.S., Gray, J.S., Alonso, D., Anderson, M.J., Benecha, H.K., Dornelas, M., Enquist, B.J., Green, J.L., He, F., Hurlbert, A.H., Magurran, A.E., Marquet, P.A., Maurer, B.A., Ostling, A., Soykan, C.U., Ugland, K.I. & White, E.P. (2007) Species abundance distributions: moving beyond single prediction theories to integration within an ecological framework. *Ecology Letters*, **10**, 995–1015.

McLean, R.A., Sanders, W.L. & Stroup, W.W. (1991) A Unified Approach to Mixed Linear Models. *The American Statistician*, **45**, 54–64.

McMahon, S.M. & Diez, J.M. (2007) Scales of association: hierarchical linear models and the measurement of ecological systems. *Ecology Letters*, **10**, 437–452.

Miller, J.E.D., Damschen, E.I. & Ives, A.R. (2018) Functional traits and community composition: A comparison among community-weighted means, weighted correlations, and multilevel models. *Methods in Ecology and Evolution*, **10**, 415–425.

Nakagawa, S., Johnson, P.C.D. & Schielzeth, H. (2017) The coefficient of determination R^2 and intra-class correlation coefficient from generalized linear mixed-effects models revisited and expanded. *Journal of The Royal Society Interface*, **14**, 20170213.

Nakagawa, S. & Schielzeth, H. (2013) A general and simple method for obtaining R2 from generalized linear mixed-effects models. *Methods in Ecology and Evolution*, **4**, 133–142.

Norberg, A., Abrego, N., Blanchet, F.G., Adler, F.R., Anderson, B.J., Anttila, J., Araújo, M.B., Dallas, T., Dunson, D., Elith, J., Foster, S.D., Fox, R., Franklin, J., Godsoe, W., Guisan, A., O'Hara, B., Hill, N.A., Holt, R.D., Hui, F.K.C., Husby, M., Kålås, J.A., Lehikoinen, A., Luoto, M., Mod, H.K., Newell, G., Renner, I., Roslin, T., Soininen, J., Thuiller, W., Vanhatalo, J., Warton, D., White, M., Zimmermann, N.E., Gravel, D. & Ovaskainen, O. (2019) A comprehensive evaluation of predictive performance of 33 species distribution models at species and community levels. *Ecological Monographs*.

Oberpriller, J., de Souza Leite, M. & Pichler, M. (2022) Fixed or random? On the reliability of mixed-effects models for a small number of levels in grouping variables. *Ecology and Evolution*, **12**, e9062.

Ovaskainen, O. & Soininen, J. (2011) Making more out of sparse data: hierarchical modeling of species communities. *Ecology*, **92**, 289–295.

Ovaskainen, O., Tikhonov, G., Norberg, A., Guillaume Blanchet, F., Duan, L., Dunson, D., Roslin, T. & Abrego, N. (2017) How to make more out of community data? A conceptual framework and its implementation as models and software. *Ecology Letters*, **20**, 561–576.

Pichler, M. & Hartig, F. (2021) A new joint species distribution model for faster and more accurate inference of species associations from big community data. *Methods in Ecology and Evolution*, **12**, 2159–2173.

Pollock, L.J., Tingley, R., Morris, W.K., Golding, N., O'Hara, R.B., Parris, K.M., Vesk, P.A. & McCarthy, M.A. (2014) Understanding co-occurrence by modelling species simultaneously with a Joint Species Distribution Model (JSDM). *Methods in Ecology and Evolution*, **5**, 397–406.

Powers, S.M. & Hampton, S.E. (2019) Open science, reproducibility, and transparency in ecology. *Ecological Applications*, **29**, e01822.

Quinn, J.F. & Dunham, A.E. (1983) On Hypothesis Testing in Ecology and Evolution. *The American Naturalist*, **122**, 602–617.

Roberts, D.W. (2019) Comparison of distance-based and model-based ordinations. *Ecology*, **n/a**, e02908.

Searle, S.R., Casella, G. & McCulloch, C.E. (2006) *Variance Components*,.

Smaldino, P.E. & McElreath, R. (2016) The natural selection of bad science. *Royal Society Open Science*, **3**, 160384.

Solbu, E.B., van der Veen, B., Herfindal, I. & Hovstad, K.A. (2022) Analyzing dynamic species abundance distributions using generalized linear mixed models. *Ecology*, **103**, e3742.

Stoffel, M.A., Nakagawa, S. & Schielzeth, H. (2021) partR2: partitioning R2 in generalized linear mixed models. *PeerJ*, **9**, e11414.

Tuomisto, H., Ruokolainen, L. & Ruokolainen, K. (2012) Modelling niche and neutral dynamics: on the ecological interpretation of variation partitioning results. *Ecography*, **35**, 961–971.

Ulrich, W., Ollik, M. & Ugland, K.I. (2010) A meta-analysis of species-abundance distributions. *Oikos*, **119**, 1149–1155.

Vellend, M. (2016) *The theory of ecological communities (MPB-57)*, Princeton University Press.

Vellend, M., Srivastava, D.S., Anderson, K.M., Brown, C.D., Jankowski, J.E., Kleynhans, E.J., Kraft, N.J.B., Letaw, A.D., Macdonald, A.A.M., Maclean, J.E., Myers-Smith, I.H., Norris, A.R. & Xue, X. (2014) Assessing the relative importance of neutral stochasticity in ecological communities. *Oikos*, **123**, 1420–1430.

de Villemereuil, P., Morrissey, M.B., Nakagawa, S. & Schielzeth, H. (2018) Fixed-effect variance and the estimation of repeatabilities and heritabilities: issues and solutions. *Journal of Evolutionary Biology*, **31**, 621–632.

Wainwright, P.E., Leatherdale, S.T. & Dubin, J.A. (2007) Advantages of mixed effects models over traditional ANOVA models in developmental studies: A worked example in a mouse model of fetal alcohol syndrome. *Developmental Psychobiology*, **49**, 664–674.

Warton, D.I., Blanchet, F.G., O'Hara, R.B., Ovaskainen, O., Taskinen, S., Walker, S.C. & Hui, F.K.C. (2015) So Many Variables: Joint Modeling in Community Ecology. *Trends in Ecology & Evolution*, **30**, 766–779.

Wilkinson, D.P., Golding, N., Guillera-Arroita, G., Tingley, R. & McCarthy, M.A. (2019) A comparison of joint species distribution models for presence–absence data. *Methods in Ecology and Evolution*, **10**, 198–211.

Zuur, A., Ieno, E., Walker, N., Saveliev, A. & Smith, G. (2009) *Mixed effects models and extensions in ecology with R*, Springer.

Matrix quality determines the strength of habitat loss filtering on bird communities at the landscape scale

Melina de Souza Leite[1], Andrea Larissa Boesing[1,2], Jean Paul Metzger[1] &

Paulo Inácio Prado[1]

1. Department of Ecology, University of São Paulo, Rua do Matão 321, T14, São Paulo, SP, 05508-090, Brazil

2. Senckenberg Biodiversity and Climate Research Institute (SBiK-F), Georg-Voigt-Straße 14, 60325 Frankfurt am Main, Hessen, Germany

Abstract

1. Habitat loss represent a major threat to biodiversity, however, the modulation of its effects by the non-habitat matrix surrounding habitat patches is still undervalued. The landscape matrix might change community assembly in different ways. For example, low-quality matrices can accentuate environmental filtering by reducing resource availability and/or deteriorating abiotic conditions but they may also over limit dispersal of organisms and make communities more prone to ecological drift.

2. To understand how matrix quality modulates the effects of habitat loss, we quantified the relative importance of environmental filter and ecological drift in bird communities across more local (400 m buffer around sampling sites) and broader (2 km focal landscapes) gradients of habitat loss embedded in low- and high-quality matrices. We used a trait-based approach to understand habitat loss filtering effects on forest specialist and habitat generalist bird occurrences.

3. We found that low-quality matrices, composed mainly of low-productive pasturelands, increased the severity of habitat loss filtering effects for forest specialist birds, but only at the landscape scale. Bird occurrences were in general higher in high-quality matrices, i.e., more heterogeneous and with low-contrasting edges, indicating the role of the matrix quality on attenuating species extinction risks at the landscape scale, probably due to mass effect. Moreover, forest specialists presented a strong negative response to habitat loss filtering across different functional traits, while generalists presented a high variability in traits response to habitat loss.

4. *Synthesis and applications:* We raised evidence in supporting that landscape habitat loss filtering may be relaxed or reinforced depending on the quality of the matrix, evidencing that matrix quality has a strong impact in modulating community assembly processes in fragmented landscapes. In practical terms, it means that improving matrix quality may help in maintaining the high diversity of birds even without any increase in native forest cover.

Keywords: community ecology, dispersal, ecological drift, environmental filter, functional traits, landscape structure, trait-environment relationship.

Resumo

1. A perda de habitat representa uma grande ameaça à biodiversidade, entretanto, entender como seus efeitos são modulados pela matriz (não-habitat) ainda é subvalorizado. A matriz na paisagem pode mudar a montagem da comunidade de diferentes maneiras. Por exemplo, matrizes de baixa qualidade podem acentuar a filtragem ambiental de espécies por reduzir a disponibilidade de recursos e/ou deteriorar as condições abióticas, mas também podem limitar a dispersão de organismos e tornar as comunidades mais propensas à deriva ecológica.

2. Para entender como a qualidade da matriz modula os efeitos da perda de habitat, quantificamos a importância relativa do filtro ambiental e da deriva ecológica nas comunidades de aves em diferentes escalas de perda de habitat em paisagens com matrizes de baixa e alta qualidade: escala mais locais (400 m de raio em torno dos pontos de amostragem) e mais ampla (2 km de raio ao redor do centro das paisagens focais). Utilizamos uma abordagem baseada em atributos das espécies para entender os efeitos da filtragem da perda de habitat sobre as ocorrências de aves especialistas e generalistas de habitat.

3. Descobrimos que matrizes de baixa qualidade, compostas principalmente de pastagens de baixa produtividade, aumentaram a severidade dos efeitos de filtragem da perda de habitat para aves especializadas em florestas, mas apenas na escala da paisagem. As ocorrências de aves foram em geral mais elevadas em matrizes de alta qualidade, ou seja, mais heterogêneas e com bordas de baixo contraste, indicando o papel da qualidade da matriz na diminuição dos riscos de extinção de espécies na escala da paisagem, provavelmente devido ao efeito de massa (mass effect). Além disso, aves especialistas apresentaram uma forte resposta negativa à filtragem da perda de habitat através de diferentes atributos funcionais, enquanto aves generalistas apresentaram uma alta variabilidade na resposta dos atributos à perda de habitat.

4. *Síntese e aplicações*: Nossos resultados apontam que a filtragem de espécies por perda de habitat na paisagem pode ser relaxada ou reforçada dependendo da qualidade da matriz, evidenciando que a qualidade da matriz tem um forte impacto na modulação dos processos de montagem da comunidade em paisagens fragmentadas. Em termos práticos, isso significa que a melhoria da qualidade da

matriz pode ajudar a manter a alta diversidade das aves, mesmo sem qualquer aumento na cobertura florestal nativa.

Palavras-chave: ecologia de comunidades, dispersão, deriva ecológica, filtro Ambiental, estrutura da paisagem, relação atributo-ambiente.

Introduction

Anthropogenic habitat change is one of the most important drivers of biodiversity loss. The division of habitats into smaller and more isolated fragments, separated by a non-habitat matrix of human-transformed land cover, alters not only the quantity but also the quality of the habitats in the landscape (Fischer & Lindenmayer, 2007; Haddad *et al.*, 2015). Much has been studied on how habitat loss and fragmentation impact biodiversity in terms of the community and species deterministic responses (environmental filtering) and random demographic events leading to stochastic changes in community composition (ecological drift) (Baselga *et al.*, 2015; Henckel *et al.*, 2019). Nevertheless, the role of the landscape matrix in altering these deterministic and random assembly processes and thus the outcome of habitat loss on community composition is still unclear (Kupfer *et al.*, 2006; Pardini *et al.*, 2017).

Determining the matrix effects on community assembly processes in human-dominated landscapes faces two main challenges: how to disentangle community assembly processes at different spatial scales, and how to characterize the matrix context in landscapes. First, community assembly processes in acting differently at local and landscape scales may result in different metacommunity arrangements (Leibold *et al.*, 2004), which evidence the foreseen patterns of multidimensionality and scale-dependency of species richness and biodiversity changes (Chase *et al.*, 2018). One approach to disentangle community assembly processes is the use of species traits to elucidate mechanisms by which communities respond to environmental gradients (McGill *et al.*, 2006). Such trait-environment associations proved to be critical in understanding how ecological processes affect biodiversity across scales (Newbold *et al.*, 2013, 2014; Suárez-Castro *et al.*, 2018). Therefore, trait-focused approaches may show how local environments and landscapes constrain diversity (Tscharntke *et al.*, 2012), for example when the matrix influences trait diversity in habitat patches (Häkkilä *et al.*, 2017; Boesing *et al.*, 2018a).

Second, a prolific way to characterize matrix context is by its quality for a target group of organisms. Matrix quality is taxon and context-dependent, but in general, it may be assumed as the structural similarity of the matrix with the native habitat (Prevedello & Vieira, 2010). For instance, landscapes with high-quality matrices can maintain greater amounts and diversity of resources (Dunning *et al.*, 1992; Pardini *et al.*, 2009) which can be used occasionally by species living in habitat patches (spillover; Blitzer et al. 2012). High-quality matrices can also facilitate species movement across the landscape leading to higher landscape connectivity (Antongiovanni & Metzger, 2005; Fahrig, 2007). Consequently, landscapes with high-quality matrices may support more species than the ones with matrices of lower quality (Carrara *et al.*, 2015; Reider *et al.*, 2018).

The matrix may change how habitat loss and fragmentation modulate community assembly processes through its effects on (1) organismal movements (Fahrig, 2007; Watling *et al.*, 2011; Biz *et al.*, 2017), (2) the availability of supplementary or complementary resources in the landscape (Dunning *et al.*, 1992; Boesing *et al.*, 2021), and (3) abiotic edge conditions of habitat patches (Saunders et al. 1991, Pardini et al. 2009, Pfeifer et al. 2017). When movements of organisms among habitat patches are heavily limited by the matrix, dispersal may be limited and, together with habitat loss, will lead to small and functionally isolated communities, which are more prone to ecological drift (Horváth *et al.*, 2019; Siqueira *et al.*, 2020), i.e., random fluctuations of species abundances in the community leading or not to random local extinctions (Vellend, 2010). When matrix decreases resource availability in the landscape and/or accentuate unfavorable abiotic condition inside habitat patches (edge effects), habitat loss may be an even more severe environmental filter, selecting only those species able to persist under these harsher conditions (Chase, 2007). By modulating these three mechanisms, matrix quality changes how habitat loss and fragmentation alter the relative importance of the assembly processes of dispersal (potentially leading to ecological drift) and environmental filtering. Moreover, given the scale-dependent nature of ecological communities (Chase *et al.*, 2018), only a multiscale combined with a trait-based approach (Suárez-Castro *et al.*, 2018) may help to identify the relative strength of such processes across both local and landscape scales.

Bird communities are an excellent model system to investigate the effects of matrix quality on community assembly after habitat loss because it is a key factor

determining birds' movements and resource use in anthropogenic landscapes (Kennedy *et al.*, 2010; Deikumah *et al.*, 2013; Boesing *et al.*, 2018a; Barros *et al.*, 2019). Birds are especially mobile organisms, which allows them to actively search and respond to pulses of resources in ways not generally possible for other vertebrates, allowing them to leave areas in which resources are no longer sufficient and relocate to more productive locations (Whelan *et al.*, 2008). In addition, birds have large variety of responses to landscape modification: while in general habitat-specialist species are highly intolerant to human-modified environments, habitat generalists are less restrictive and more resilient to use different kinds of environments (Morante-Filho *et al.*, 2016; Boesing *et al.*, 2021).

In this study, we investigated whether and how matrix quality modulates the relative importance of assembly processes in bird communities across gradients of habitat loss. First, we hypothesize that the relative contribution of the assembly processes of habitat loss filtering and ecological drift will depend on (1) how dispersal is limited and (2) the harshness of the environmental filtering by matrices of different qualities. If bird dispersal is heavily limited by low-quality matrices, habitat loss will create small and isolated communities more prone to drift, which in turn will relatively decrease the importance of habitat loss filtering compared to landscapes with high-quality matrices. However, if the main effects of a low-quality matrix are in decreasing resource availability and/or deteriorating abiotic conditions in habitat patches, habitat loss will be an even more severe filter, increasing the relative importance of habitat loss filtering compared to landscapes with high-quality matrices. Second, given that matrix quality is an element of the landscape, we expect that the effects of habitat loss filtering on species traits will be stronger at the focal landscape scale than at the local scale – i.e. the amount of forest nearby the sampling point (hypothesis of landscape moderation of trait selection; Tscharntke et al. 2012). We compared results between forest specialist and habitat generalist birds. Since specialists are considered more sensitive to habitat loss (Pardini *et al.*, 2009; Carrara *et al.*, 2015), habitat loss filtering should be stronger and negative for them. As habitat generalists are commonly less affected by habitat loss or can even increase in abundance (Nordberg & Schwarzkopf, 2018), we predict a positive or null relationship of their traits to habitat loss.

To test our hypotheses and predictions, we modeled the occurrence of bird species in sets of landscapes from two regions with contrasting matrix qualities using

hierarchical models (Jamil *et al.*, 2013; ter Braak, 2019). These models are commonly used to analyze trait-environment associations, to disentangle scale-dependent community assembly processes (Ovaskainen *et al.*, 2017), and to sort out niche and drift effects on species abundances (Mortara, 2016). We used variance partitioning of the models (Nakagawa & Schielzeth, 2013) to compare the relative importance of processes across assemblages. Habitat loss effects across scales were investigated by both habitat loss overall effect in birds' occurrences and trait filtering effects as the variation of the effects of forest cover on species with different traits (interaction terms). We also evaluated how much (in percentage) of the species response to habitat loss is explained by their traits to understand how species traits are filtered in landscapes under different matrix contexts.

Materials and Methods

Study areas

The study was carried out in the Atlantic Forest of southeastern Brazil (Fig. 1). The entire region retains less than 30% of the original forest cover, and most of the forest is confined in small patches (< 50 ha) in different stages of regeneration (Rezende *et al.*, 2018). Our study landscapes were constrained across two regions (separated by 90 km) with similar biophysical characteristics and bird species pool, but with different agricultural matrix compositions. These matrices were denominated as high and low quality considering the perspectives or requirements of forest-dependent bird species. A higher quality matrix is associated with more heterogeneous land uses and low-contrasting edges that provide more resources and facilitate the movement of forest species across the landscape (Boesing *et al.*, 2018a). The northwest region (henceforth 'high-quality matrix') is mainly composed of a mosaic of sun-coffee, sugar cane plantations, and pastures, resulting in a more heterogeneous matrix. Coffee plantations cover around 46% of the matrix and it is usually located adjacent to forest edges, creating low-contrasting edges at the patch scale (Fig. 1A). The southeast region (henceforth 'low-quality matrix') is largely dominated by low-productive pastures and low diversity of other land use types, and so it is composed of a more homogeneous matrix. Additionally, the huge structural difference between pastures and forests creates

high-contrasting edges at the patch scale (Fig. 1B). A detailed evaluation of matrix types in the region can be found in Appendix S2 in Supporting Information.
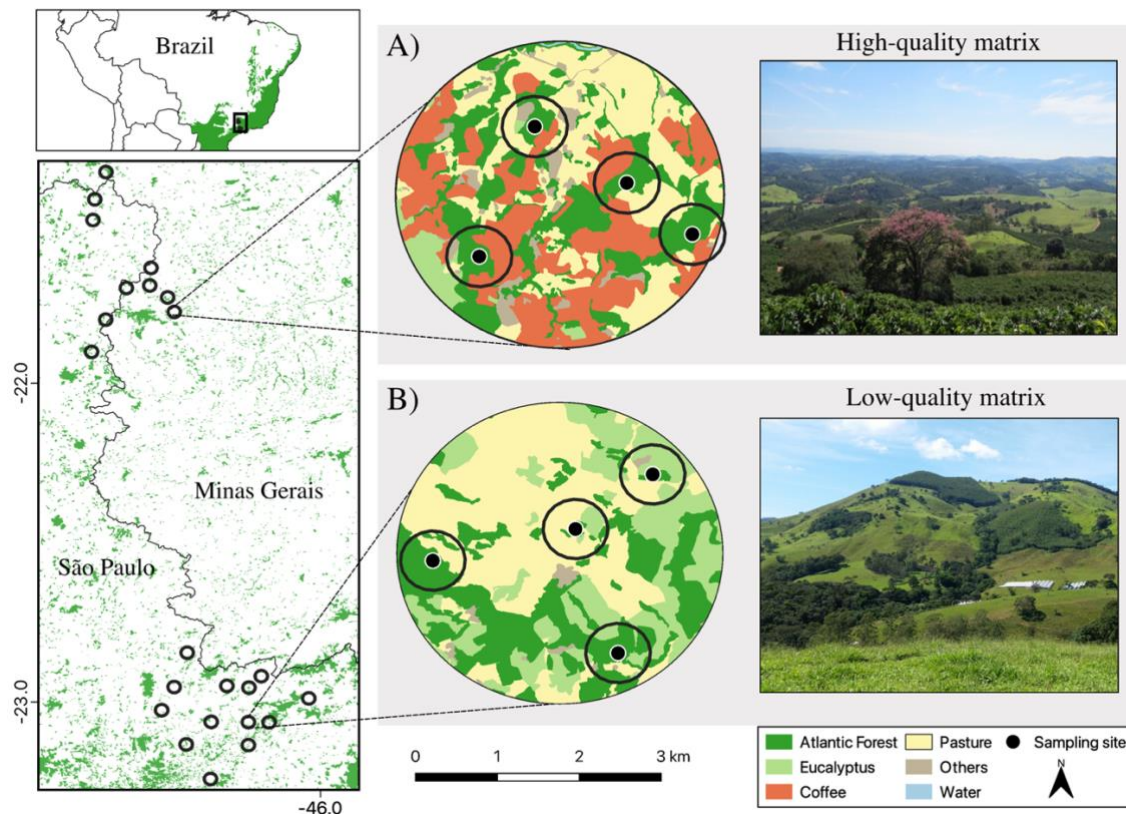

Figure 1: Location of the study area (left upper panel) within the Atlantic Forest biome (in green) with 23 studied landscapes in southeastern Brazil (left lower panel). We show two 2 km radius landscapes with 27% of native forest cover, but in (A) with a high-quality matrix – more heterogeneous with a high proportion of sun coffee plantations; and in (B) with a low-quality matrix mostly composed of low-productive pastures. In each landscape 4 local-scale sampling points of 400 m radius buffer (black dots).

We used the landscape selection approach of Pasher et al. (2013) to select independent focal landscapes that span a gradient of landscape-level forest cover, while controlling for potentially confounding factors, as altitude (ranging from 800-1300 m.a.s.l), soil type (ferric red latosol or argisol soil), presence of large highways and water reservoirs. We selected 10 and 13 focal landscapes (2 km radius, 1256 ha each) in the high-quality (ranging from 7-46% of forest cover) and low-quality matrix (12-55% of forest cover) regions, respectively. In each landscape, we placed 4 sampling sites (total 96 sites) in forest patches spaced a minimum of 800m apart (1591 ± 621 m) in a way to cover different ranges of local forest cover inside the same focal landscape. For each selected forest patch, we randomly selected a point along the forest-matrix interface and placed a sampling site within the forest patch at the end of a 100 m

transect always oriented to the center of the forest patch. The matrix type of the closest edge was always a coffee plantation in high-quality landscapes and pasture in low-quality landscapes. All selected forest patches at both regions were similar in structure and composed of intermediate to advanced second-growth forest with a well-established canopy, free from cattle disturbance in the understory and, located within private properties. See Appendix S1 and Boesing et al. (2018a) for more details about the area selection procedure.

Bird sampling and traits selection

In each sampling site, we performed a 50 m fixed radius point count and recorded all bird species detected visually or aurally during a 15 min sampling period. Studies in tropical regions indicate that 10 min point counts are sufficient to record 90% of tropical bird species (Esquivel *et al.*, 2008). Each site was visited four times, between January–April and August–November of 2014 (N=368), by the same experienced ornithologist (ALB) and both regions were sampled simultaneously (more details in Appendix S1). Permission for the data collection was obtained from the Brazilian Institute of Environment and Renewable Natural Resources (IBAMA) under SISBIO number 46697-1.

To be recorded, bird individuals needed to be perched or singing inside the radius of detection: species flying above the canopy or flying through the sample area were not recorded (Bibby *et al.*, 2000). All point counts were performed in the four hours after sunrise and in the last hour before sunset. Each point count was sampled in different times during the morning and evening periods in order to give the same chance of detection for all species (Esquivel *et al.*, 2008). To increase the detectability of more cryptic and rare species, each point count was replicated four times, between January–April and August–November of 2014 (23 landscapes, 92 sampling sites, n = 368 point counts) and both regions were sampled simultaneously. Some of the records are deposited at: https://xeno-canto.org/contributor/ETMICIBVME.

Bird traits selection followed the framework proposed by Luck et al. (2012). First, we listed the potential traits based on current knowledge on bird traits related to extinction-proneness due to habitat loss, land use, or environmental change (references in Appendix S2). Then, we selected traits based on the (1) data availability; (2) experts' knowledge (ALB); and (3) the empirical relationship among traits (correlations).

Finally, we selected five groups of response traits that are known to affect individual fitness by influencing growth, reproduction, or survival of the species on human-modified landscapes: hand-wing index (HWI), body size, nest type, diet, and foraging stratum. These traits were represented by eight operational variables described in detail in Appendix S2.

We kept separate analyses for forest specialist and habitat generalist birds (henceforth just specialists and generalists) given that habitat specialization is very important for understanding the effects of habitat loss and fragmentation (Pandit *et al.*, 2009; Kupsch *et al.*, 2019). Moreover, as both groups of species presented similar ranges of trait values on our data (Appendix S2), we also investigated if the same trait differs in response to habitat loss depending on species habitat specialization. We considered forest specialists those species that depend on forest environments for breeding and survival, and habitat generalists those that rely on different habitat types, including forests, to survive (Del Hoyo, 2014). We appealed afterwards for a bird expert revision (Scott Robinson, Univ. of Florida) to guarantee the quality of our classification. We also excluded predators to conform our assemblages to the definition of a horizontal community without trophic interactions (*sensu* Loreau, 2010) and due to the biased method of point counts for detection of such species.

Modeling

We modeled bird occurrences with hierarchical linear models commonly used to analyze trait-environment associations and disentangle scale dependency in community assembly processes (Jamil *et al.*, 2013; ter Braak, 2019). We used a binomial generalized linear mixed model (logit link function) with the number of detections of each of *n* species (four visits) in each of *m* sites as response variable (Miller *et al.*, 2018). In each species-site combination (Y_i, where *i* goes from 1 to *n* x *m* observations), we assigned predictors of the *n* species-level trait and the *m* site- and landscape-level forest cover (Miller *et al.*, 2018). Our model is described by:

$$Y_i \sim Binomial(4, p)$$

$$logit(p) = \left(\alpha + a_{spp[i]} + b_{site[i]} + c_{landscape[i]}\right) + \left(\beta_1 + d_{spp[i]}\right)forest_{cover_{site[i]}} +$$

$$\left(\beta_2 + e_{site[i]} + f_{landscape[i]}\right)trait_{spp[i]} + \beta_{12}trait_{spp[i]} \times forest_{cover_{site[i]}} + \qquad (1)$$

$$g_{site[i] \times spp[i]} + h_{landscape[i] * spp[i]}$$

where Y_i is the observed number of recorded occurrences for each of the i species-site combination (species n in site m). Fixed effects are represented in Greek and random effects in Latin letters. All random effect terms are represented by a normal distribution with mean zero and their respective estimated variances ($\sigma_a^2, \sigma_b^2, \sigma_c^2, \sigma_d^2, \sigma_e^2, \sigma_f^2, \sigma_g^2, \sigma_h^2$). Terms with random intercept and slope have additional parameters denoting the correlation between them ($\rho_{ad}, \rho_{be}, \rho_{cf}$). The model's formula syntax in the *lme4* R package (Bates *et al.*, 2015) is:

```
Y ~ trait*forest_cover + (forest_cover|species) + (trait|site) +
   (trait|landscape) + (1|landscape:species) + (1|landscape:site)
```

The fixed effects in the logit link function are the main effects of habitat loss and trait (α, β_1, β_2) and habitat loss filtering effects through traits (interaction term β_{12}). We used the percentage of forest cover at both local and landscape scales as response variables to denote habitat loss and we found no collinearity between these variables (Appendix S2). We are particularly interested in: (1) the main effects of habitat loss, i.e., how steep is the decrease in occurrence probability of all birds when habitat is lost, and (2) the interaction effect of habitat loss with traits, i.e., how species with different traits respond to habitat loss. However, these effects are not statistically easily separable from the main effects of trait in terms of variance explained (R^2, Johnson 2014). As we show below, we kept the whole structure of fixed effects in interpreting marginal R^2 as a general interpretation of habitat loss effects. Nevertheless, we also compared the estimates of habitat loss main effects among assemblages and across scales and calculated the importance of the traits in explaining habitat loss filtering (equation 2, below) to interpreting habitat loss filtering effects through traits.

Species random effects are the variation among species in their overall abundances (random intercept, $a_{spp[i]}$) and response to habitat loss (random slope, $d_{spp[i]}$) regardless of its trait values. They express the idiosyncratic response of each species to habitat loss and their differences in overall abundances. Trait-site and trait-landscape random interactions (random intercepts: $b_{site[i]}, c_{landscape[i]}$, and slopes: $e_{site[i]}, f_{landscape[i]}$) deal with the trait-mediated response to non-measured environmental gradients at local and landscape scales, respectively. This is a solution to "account for any interaction that the observed trait has with any unobserved environmental gradient" (ter Braak, 2019).

We included two additional random effects of site-species ($g_{site[i] \times spp[i]}$), and landscape-species interactions ($h_{landscape[i] \times spp[i]}$ to express that the occurrence probability of the same species may vary among sites and landscapes, regardless of the species identity, its traits, and the habitat loss gradients. This term expresses a random variation in species occurrences probabilities across sites and landscapes. The site-species interaction term is the so-called Observation Level Random Effect (OLRE), which allows for extra variance among observations. OLRE is generally used to deal with overdispersion in data modeling with the exponential family distributions (Harrison, 2015). For species occurrences, overdispersion is frequently associated with individuals' aggregations in space (e.g. Ozgul *et al.*, 2009). In this sense, at the community level, we can assume that the spatial aggregation expressed by these terms is a result of differential dispersal, which may be a limitation or excess of dispersal. Dispersal limitation of individuals among patches and landscapes may lead to ecological drift (*sensu* Vellend 2010). Although it is expected that these terms express ecological drift (Mortara, 2016), we cannot ensure that all the variances apportioned in these terms are exactly differential dispersal. However, because we have carefully handled many possible niche processes (or selection *sensu* Vellend 2010) in the model, probably, these terms do not include other processes than random variation in species occurrence probabilities. At least, if differential dispersal among patches and landscapes is an important process, it will be expressed in these two random effects and not in the other terms.

To express the **importance of traits in explaining habitat loss filtering** we calculated the proportional decrease in the variance of the species random slope for forest cover ($d_{spp[i]}$) when we include the trait-forest cover interaction term (β_{12}) in the model (Jamil *et al.*, 2013):

$$C_\beta = 1 - \frac{d_{spp[i]}(res)}{d_{spp[i]}(total)} \tag{2}$$

$d_{spp[i]}(res)$ is the variance of the random species slope from the model with forest cover and trait main effects (β_2) and the trait-environment interaction (β_{12}); $d_{spp[i]}(total)$ is the same variance term for the model without trait-environment interaction but with the main effects. This proportion explains how much of the effect of habitat loss on each species can be explained by their trait values.

We computed overall R^2 for the models by the additive variance partitioning method (Nakagawa & Schielzeth, 2013; Johnson, 2014), which is appropriate for comparing models fit to different data sets because it does not depend on sample size (Ives, 2019). We also apportioned the total variance among the model terms (marginal R^2 *sensu* Ives 2019) to compare the relative importance of processes across assemblages. We calculated the overall and marginal R^2 on the scale of the link function (logit) since it can decouple variance and mean so that the linked scale R^2 can be calculated independent of the population mean (Nakagawa & Schielzeth, 2010).

Table 1: Definitions and ecological interpretations of the statistical terms of the hierarchical linear model (equation 1). Code syntax follows lme4 R package.

Model terms	Code	Description
$\alpha; \beta_1; \beta_2; \beta_{12}$	trait*env	**Fixed effects**: main effects of **habitat loss** and traits, and **habitat loss filtering effects through traits** (interaction terms) at both local and landscape scale.
$a_{spp[i]}; d_{spp[i]}$	(env|sp)	**Random effects**: variation among species in their overall occurrences (intercept) and response to habitat loss (slope) regardless of its traits. It is the **idiosyncratic response of each species to habitat loss and their differences in overall abundances.**
$b_{site\{i\}}; e_{site}$ $c_{landscape[i]}; f_{landscape[i]}$	(trait|site) (trait|lands)	**Random effects**: interaction (random intercept and slope) between trait values and sites/landscapes. It is the possible relationship of the measured trait with any unmeasured environmental variable at the site/landscape level. **Trait-mediated response to other environmental gradients at local/landscape scale.**
$g_{spp[i]\times landscap}$	(lands:sp)	**Random effects**: overdispersion term indicating **random variation in species occurrences across landscapes**, regardless of its traits and habitat loss conditions.
$h_{spp[i]\times site[i]}$	(site:sp)	**Random effects**: overdispersion term indicating **random variation in species occurrences across sites**, regardless of its traits and habitat loss conditions.

Data analysis

We ran the model described above for each low and high-quality matrix landscapes and for forest specialists and habitat generalists separately (hereafter assemblages) to better interpret and compare overall and marginal R^2s among assemblages. We analyzed the data in two steps. First, given that ecological responses can be affected by processes acting at different spatial scales, we selected the best scale for the effect of local forest cover (Appendix S2), which was the percentage of local forest cover measured at 400 m buffer radius around each site (Fig. 1). Second, we ran

separate models for each species trait (equation 1) including local and landscape forest cover (2 km radius around the centroid of the landscapes) as predictors of habitat loss. We finally combined 4 traits with the highest explanatory power for all datasets (main diet, proportion of lower strata use, body mass, and nest type) in one model to predict the trait's response to habitat loss and to compare the marginal R^2 of each model term. The combined traits models did not show collinearity and the traits were not correlated among species (Appendix S3).

Considering that the two regions with high- and low-quality matrix landscapes may present different pools of species, we performed a robustness analysis with only the species that occurred at both regions with high- and low-quality landscapes. The results from the combined models with only the subset of species were very similar to the models with all species in each assemblage (Appendix S4), discarding any potential effect of biogeography or other sources of variation in species composition in our results. All data analysis was performed using R (version 4.0, R Core Team 2019) with *lme4* package (Bates *et al.*, 2015) for modeling, and *DHARMa* (Hartig, 2018) for quantile residuals diagnostic (Appendix S3).

Results

In our modeling approach, the R^2s of the fixed effects – trait, habitat loss, and filtering effects at both local and landscape scales – were the variance component that presented the most marked differences between assemblages (Fig. 2). It was 2.2 times higher for the specialists in the low-quality than in the high-quality matrix landscapes (22 and 10%, respectively), and it was up to 3 times higher for specialists than for both generalist assemblages (7-8%). It means that habitat loss effects alone and habitat loss filtering effects were much stronger for the specialist birds in low-quality matrix landscapes. Moreover, overall R^2 for specialists in low-quality matrix landscapes captured most of the variability in bird's occurrence (72%), followed by specialists in high-quality matrices (58%) (Fig. 2). Overall R^2 for generalists was smaller and similar across landscapes regardless of matrix quality (46-47%).

The R^2 for the terms that account for random variation in species occurrences across sites (*site:sp*), were very low (1%) for all assemblages. However, the R^2 for random variation in species occurrences at the landscape scale (*lands:sp*) was larger for

the specialists in high-quality matrices than for the specialists in low-quality matrices. This term was also larger for specialists (4-7%) than generalists (1-2%).

The terms that express the variation among species in their overall abundances and responses to habitat loss regardless of their traits (random intercept and slope: *env|sp*) presented the highest marginal R^2 for all assemblages, ranging from 32 to 40%, and being about 1.3 higher for the specialists than for the generalists. Similarly, marginal R^2 of the terms that express the effects of species traits associated with unmeasured environmental variables (*trait|site* and *trait|land*) were very low, and together they varied between 3 to 5% in all assemblages.
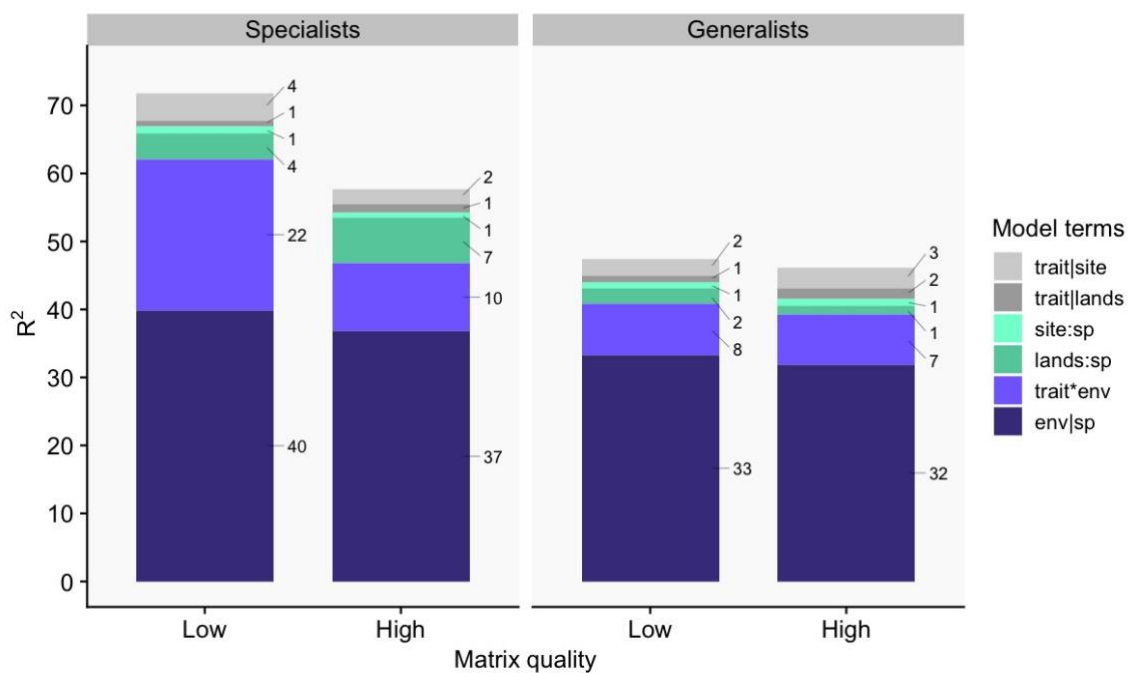

Figure 2: Overall and marginal R2s for the models with combined traits (body mass, type of nest, main diet, percentage of lower strata use) for habitat generalist and forest specialist birds in high and low-quality matrix landscapes. The number on the right side of each bar indicates the percentages for each term. See Table 1 for model terms definitions.

In comparing habitat loss effects across scales, we found that the effects of both local and landscape habitat loss in bird occurrence were stronger for the specialists in low-quality matrix landscapes (Fig. 3). While local forest cover presented similar and positive effects on bird occurrences for specialists in both high- and low-quality matrix landscapes, landscape forest cover presented a strong effect only for specialists in low-quality matrix landscapes, where this effect was larger than for local forest cover. As expected, the effects of habitat loss for generalists were generally weaker and irrelevant
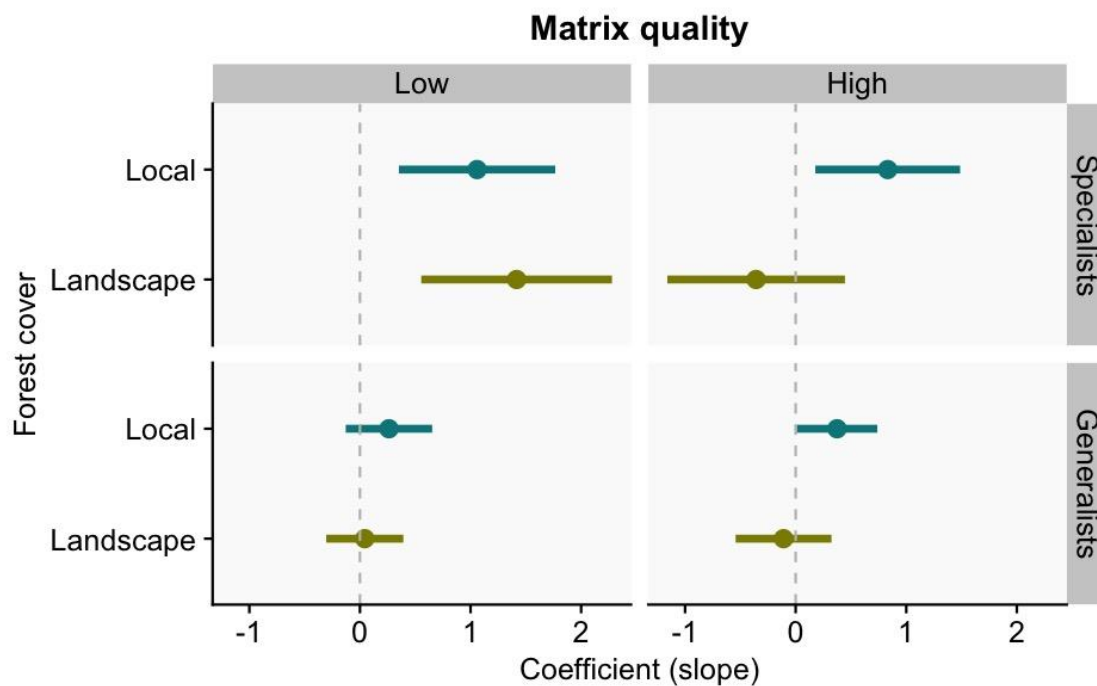
at any scale.

Figure 3: Slope coefficients (and 95% confidence intervals) of local and landscape forest cover for forest specialist and habitat generalist birds in different matrix quality landscapes. These are results for the combined traits model. See Appendix S3 for a table of all coefficients estimated for each assemblage.

The importance of species traits in explaining habitat loss filtering was higher for the assemblages in low-quality matrices (Fig. 4, combined traits model). Trait importance for the generalists in low-quality matrix landscapes explained 26% of the variability in species response to habitat loss. For the specialists, trait importance varied between 17 and 22%, and for the generalist in high-quality landscapes, it was only 14%. In general, nest type and main diet presented the highest values of trait importance for the assemblages (Fig. 4).

Bird occurrence probabilities were in general higher for the assemblage in high-quality compared to low-quality matrix landscapes (Fig. 5). For specialists, all the traits were associated with a decrease in occurrence probabilities with habitat loss. For generalists, trait-habitat loss relationships changed according to the trait. For example, generalists of closed nests responded positively to habitat loss, while those that nest in cavities responded negatively and generalists of open nests almost did not change with habitat loss (Fig. 5). Nevertheless, occurrence probabilities of frugivores, species that

build nests in cavities, and large birds decreased with habitat loss regardless of the matrix type and habitat specificity (Fig. 5).
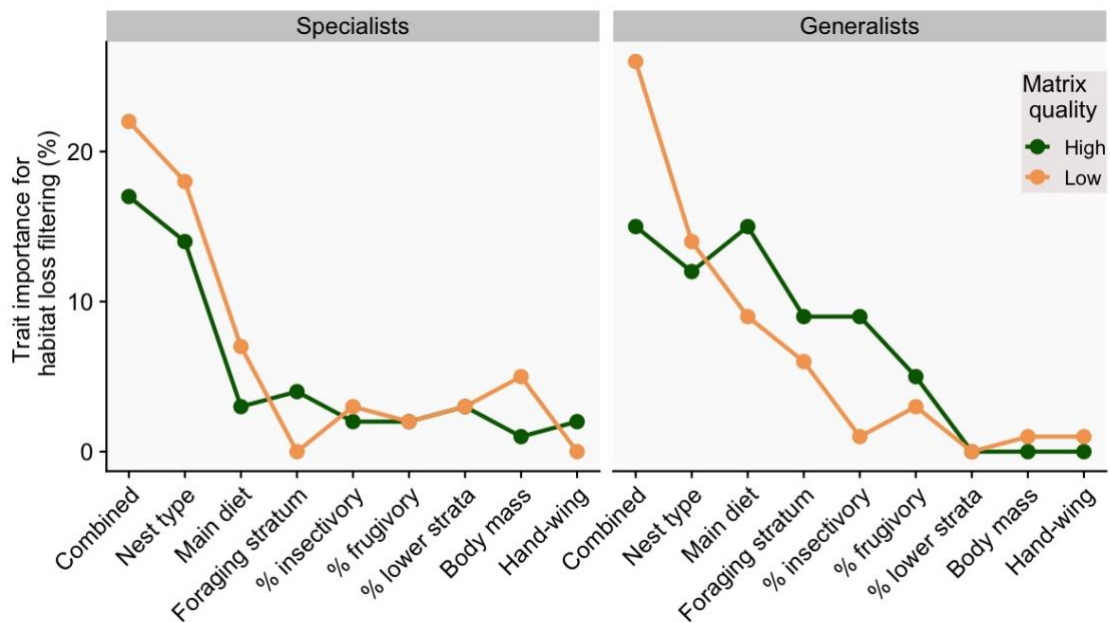

Figure 4: Importance of species traits in explaining habitat loss filtering (equation 2) for forest specialists and habitat generalist birds according to matrix quality. Trait importance was calculated for each trait in separate models and for the combined model, which includes, nest type, main diet, body mass, percentage of lower strata use, and hand-wing index.
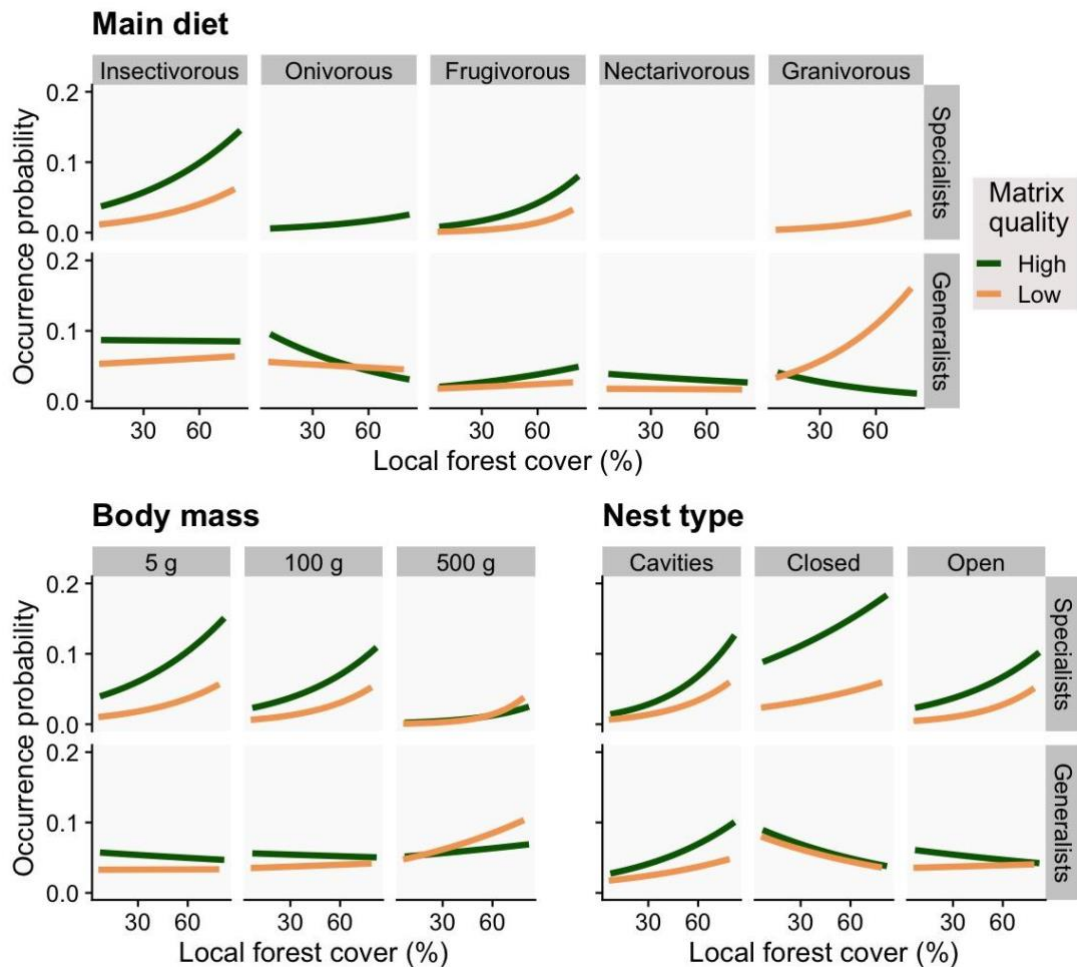
Figure 5: Predicted occurrence probability of species with local forest cover (400 m buffer; %) according to main diet, body mass, and nest type for forest specialists and habitat generalists in both high (green lines) and low-quality matrices (yellow lines). For the predictions, landscape forest cover was fixed at 30%.

Discussion

Here, we aimed to understand how the quality of the matrix surrounding habitat patches modulates the relative importance of assembly processes in bird communities across local and landscape gradients of habitat loss. We found that landscapes with low-quality matrices, composed mainly of low-productive pasturelands, increased the severity of habitat loss filtering effects for forest specialist birds (Fig. 2), with habitat loss effects at the landscape scale (Fig. 3) and habitat loss filtering effects through species traits (Fig. 4) being much stronger in those landscapes. In high quality-matrix landscapes (more heterogeneous and with low-contrasting edges), birds' occurrences were in general larger with more random variation among landscapes. Both results indicate the role of matrix quality in attenuating extinction risks in the landscape,

allowing species that would have been extinct due to habitat loss in high-quality matrices. Moreover, forest specialists presented a stronger response to habitat loss filtering, displaying a consistent negative relationship of all traits (diet, nest type, foraging stratum, and body size) with local habitat loss, while generalists presented a high variability in traits response to habitat loss (Fig. 5).

Matrix quality modulating the relative importance of habitat loss across scales

We found that matrix quality modulates habitat loss effects only for specialist birds at the landscape scale. Local habitat amount effects on birds were not dependent on the quality of the matrix. At the landscape scale, filtering processes are usually related to the selection of species according to landscape composition and configuration (Duflot *et al.*, 2014; Suárez-Castro *et al.*, 2018). Landscapes composed of low-quality matrices decrease the availability of supplementary and complementary resources in the matrix (Dunning *et al.*, 1992; Boesing *et al.*, 2021) and/or deteriorate abiotic conditions in habitat edges (Saunders *et al.*, 1991; Pardini *et al.*, 2009; Pfeifer *et al.*, 2017). These two mechanisms together with habitat loss create an even more severe environmental filter for the communities, selecting only those species able to persist under new harsh conditions (Chase, 2007). Our results for the specialists support this hypothesis of lower-quality matrices increasing severity of habitat loss filtering by showing: (1) a 2-times larger relative importance of habitat loss, traits, and filtering effects, with consequently (2) a stronger habitat loss effect, and (3) higher importance of traits in habitat loss filtering. In addition, there is evidence in our study system that the high-quality matrices, especially sun-coffee plantation matrices, are serving as a source of complementary resources for biodiversity due to the movement of organisms from forest patches to the matrix for foraging (*cross-habitat spillover*, Boesing et al. 2018b, 2021).

Additionally, landscapes composed of low-quality matrices would hinder dispersal among patches or decrease the survival rate of dispersing individuals (Fahrig, 2007; Watling *et al.*, 2011; Biz *et al.*, 2017). By that, dispersal limitation would create functionally isolated communities more prone to ecological drift (e.g., Baselga et al. 2015, Siqueira et al. 2020), which, in turn, would decrease the relative importance of species and trait filtering effects by habitat loss. However, our results pointed to the

other direction, as we found that habitat loss filtering through traits was stronger in low-quality matrix landscapes. We believe that in our study system, dispersal limitation is not an issue for the forest birds in both matrix landscapes, as the marginal R^2s for the differential dispersal (Fig. 2, *lands:sp* term) were relatively low. However, the differential dispersal term was higher in relative importance (R^2) for the specialists in the high-quality matrix, indicating that there may be not a limitation but possibly high rates of dispersal in high-quality matrices, which corroborates earlier findings (Boesing *et al.*, 2018b, 2021). High-quality matrix landscapes with higher dispersal rates, more resource availability, and milder edge effects may be preventing extinctions of isolated populations by migration of individuals from other patches (*mass effect* in metapopulation theory, Leibold et al. 2004). This mechanism would weaken the deleterious effects of habitat loss, allowing the survival of species that are not optimally suited for the new environmental conditions (Chisholm *et al.*, 2011; Leibold & Loeuille, 2015).

Habitat loss filtering depends on the quality of the matrix

The importance of species traits in explaining habitat loss filtering effects was higher for the assemblages in low-quality matrices for both generalists and specialists (Fig. 4), evidencing stronger filtering effects in low-quality matrix landscapes as discussed above. We found that specialist birds with certain traits, e.g., small-sized, insectivorous, or birds with closed nests, were more prone to extinction due to habitat loss in landscapes embedded in low-quality matrices than species with the same traits in high-quality matrices (Fig. 5). This indicates further that habitat loss has different filtering effects (stronger or weaker) over species traits depending on the quality of the matrix. Such changes might be related to effects of the matrix changing specific niches required by different species in-patches, mostly via edge-effects on both biotic and abiotic patterns (Murcia, 1995). For instance, species that make nests in cavities or build closed nests are safer against parasitism and predation (Sibly *et al.*, 2012), which are among the most impacting drivers of bird populations' decline in fragmented landscapes (Cavitt & Martin, 2002). As nest predation/parasitism often increases with edge effects (Murcia, 1995), it is stronger in landscapes with highly contrasting matrices, i.e., lower quality.

Because we were able to compare the same set of traits for habitat generalist and specialist species, we can understand better why habitat loss and fragmentation do not necessarily lead to loss of functional diversity (Boesing *et al.*, 2018a), change in functional traits (de Coster *et al.*, 2015), or even functional homogenization (Nordberg & Schwarzkopf, 2018). If specialists are replaced by generalists with similar trait values, functional differences cannot be easily observed, and thus no apparent functional differences are seen, as often reported (de Coster *et al.*, 2015; Boesing *et al.*, 2018a). However, we also found that species with some specific traits (such as nest type and body size) were always negatively impacted by habitat loss. This finding raises the point that not all habitat generalist species succeed in fragmented landscapes, and that some key functions and/or species may be indeed lost with no substitution of traits when specialists are replaced by generalists. Overall, our results reinforce the evidence that habitat loss and matrix harshness promote abiotic environmental changes inside habitat patches which make species with specific traits to be more prone to extinction than others under different matrix conditions in the landscapes.

Theoretical and practical implications

Despite a long-standing global research effort into understanding how habitat loss and fragmentation influence species loss, and at which spatial scale species loss is observed (Horváth *et al.*, 2019), there is still a considerable debate on which mechanisms are more important (Fahrig, 2017; Fletcher *et al.*, 2018). Our findings raised evidence supporting that habitat loss is a strong filtering process that often leads more vulnerable species to extinction (Pardini *et al.*, 2017), but with a considerable influence of the matrix in modulating filtering processes. We have evidenced the importance of the matrix quality in modulating habitat loss effects at the landscape scale, weakening or strengthening its severity on species filtering (Chase, 2007). Matrices of low quality accentuate habitat loss filtering in fragmented landscapes through its relatively larger effects in altering resource availability and edge effects, more than movement limitation. Consequently, the larger severity of habitat loss filtering in low-quality matrix landscapes decreases the relative importance of ecological drift on those assemblages.

Statistical quantification of the effects of dispersal and ecological drift is still an unresolved methodological problem with many caveats (Vellend *et al.*, 2014). Our

framework builds on the idea of handling possible effects of differential dispersal through terms that would take the extra variability of the data (overdispersion) produced by differences in species occurrences across sites and landscapes (Mortara, 2016). This way, we saw that dispersal limitation in low-quality matrices leading to ecological drift is not the only possible outcome after habitat loss and fragmentation. High rates of dispersal in high-quality matrix landscapes can even override environmental filtering by allowing species to occupy habitat patches where their intrinsic growth rate would be otherwise negative (Chase *et al.*, 2020). To our knowledge, this is the first empirical evidence that a not limited dispersal in high-quality matrices may result in mass effects in bird metacommunities in fragmented landscapes, although its effects are relatively small in comparison with environmental filtering.

Our binary classification of matrix quality in high and low quality is a simplification with possible caveats including for instance, the presence of other elements that could influence species' matrix perceptions. Our sampling design strived to control for matrix quality differences and other confounding factors, and series of quantitative tests were performed to reassure that such simplification of high and low quality would not neglect these other factors oversimplify the conclusions (see Appendix S2 for a detailed exploration). Despite potential limitations, such simplification proved to be very informative, as it explained an important share of the variability in birds' occurrences as differential responses to the two matrix types. In our sampling design, such appraisal would not be possible using a continuous matrix classification, which has its own merit for further studies and to respond other research questions.

In practical terms, improving matrix quality in fragmented landscapes is a key action aiming to restore and achieve more sustainable landscapes (Arroyo-Rodríguez et al., 2020; Leite et al., 2013), especially in places where conservation practices and restoration are constrained by economic and political conflicts, are costly-demanding, and requires strategic planning (Metzger *et al.*, 2021). In this sense, converting low-productive pastures into more heterogeneous environments through tree enrichment and creation of stepping-stones (Prevedello *et al.*, 2018; Silva *et al.*, 2020), or converting such pastures (even if partially) into less contrasting land uses (such as diversified crops, perennial crops and tree plantations) can increase matrix quality, minimizing the severity of habitat loss and fragmentation for forest birds (Ruffell *et al.*, 2017). In our

study areas, the conversion of abandoned and unproductive pasturelands into silviculture in the last decades was responsible for the increase in second-growth native forests (Calaboni *et al.*, 2018). We conclude that restoration and economic practices improving matrix quality may help in maintaining the high diversity of birds in the Atlantic Forest even without any increase in native forest cover.

Supporting information

APPENDIX S1: Extended methods, bird assemblages, and species trait selection

APPENDIX S2: Forest cover at local and landscape scales

APPENDIX S3: Additional models results and diagnostic

APPENDIX S4: Additional models results for species that occur at both high and low-quality matrix landscapes

Authors contributions: Melina de Souza Leite, Andrea Larissa Boesing, and Paulo Inácio Prado conceived the project. Andrea Larissa Boesing and Jean Paul Metzger delineated the sampling design. Andrea Larissa Boesing collected the data. Melina de Souza Leite and Paulo Inácio Prado analyzed the data and wrote the manuscript. All authors provided substantial revisions and comments on the manuscript and gave final approval for publication.

Conflicts of interest: We declare no conflict of interest.

Data availability statement: All data and code used in the analysis are available in Zenodo Repository http://doi.org/10.5281/zenodo.5516703 (Leite, 2022) with the published version of the GitHub repository. The original data is also available from the Dryad Digital Repository http://dx.doi.org/10.5061/dryad.37s4h (Boesing *et al.*, 2017).

References

Antongiovanni, M., & Metzger, J. P. (2005). Influence of matrix habitats on the occurrence of insectivorous bird species in Amazonian forest fragments. Biological Conservation, 122(3), 441–451. https://doi.org/10.1016/j.biocon.2004.09.005Antongiovanni, M. & Metzger, J.P. (2005) Influence of matrix habitats on the occurrence of insectivorous bird species in Amazonian forest fragments. *Biological Conservation*, **122**, 441–451.

Arroyo-Rodríguez, V., Fahrig, L., Tabarelli, M., Watling, J.I., Tischendorf, L., Benchimol, M., Cazetta, E., Faria, D., Leal, I.R., Melo, F.P.L., Morante-Filho, J.C., Santos, B.A., Arasa-Gisbert, R., Arce-Peña, N., Cervantes-López, M.J., Cudney-Valenzuela, S., Galán-Acedo, C., San-José, M., Vieira, I.C.G., Slik, J.W.F., Nowakowski, A.J. & Tscharntke, T. (2020) Designing optimal human-modified landscapes for forest biodiversity conservation. *Ecology Letters*, **23**, 1404–1420.

Barros, F.M., Peres, C.A., Pizo, M.A. & Ribeiro, M.C. (2019) Divergent flows of avian-mediated ecosystem services across forest-matrix interfaces in human-modified landscapes. *Landscape Ecology*.

Baselga, A., Bonthoux, S. & Balent, G. (2015) Temporal beta diversity of bird assemblages in agricultural landscapes: land cover change vs. stochastic processes. *PLOS ONE*, **10**, e0127913.

Bates, D., Mächler, M., Bolker, B. & Walker, S. (2015) Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, **67**, 1–48.

Bibby, C.J., Burgess, N.D., Hill, D.A. & Mustoe, S. (2000) *Bird Census Techniques*, Elsevier Science.

Biz, M., Cornelius, C. & Metzger, J.P.W. (2017) Matrix type affects movement behavior of a Neotropical understory forest bird. *Perspectives in Ecology and Conservation*, **15**, 10–17.

Blitzer, E.J., Dormann, C.F., Holzschuh, A., Klein, A.-M., Rand, T.A. & Tscharntke, T. (2012) Spillover of functionally important organisms between managed and natural habitats. *Agriculture, Ecosystems & Environment*, **146**, 34–43.

Boesing, A.L., Marques, T.S., Martinelli, L.A., Nichols, E., Siqueira, P.R., Beier, C., de Camargo, P.B. & Metzger, J.P. (2021) Conservation implications of a limited avian cross-habitat spillover in pasture lands. *Biological Conservation*, **253**, 108898.

Boesing, A.L., Nichols, E. & Metzger, J.P. (2018a) Biodiversity extinction thresholds are modulated by matrix type. *Ecography*, **41**, 1520–1533.

Boesing, A.L., Nichols, E. & Metzger, J.P. (2017) Data from: Biodiversity extinction thresholds are modulated by matrix type. *Dryad, Dataset, https://doi.org/10.5061/dryad.37s4h*.

Boesing, A.L., Nichols, E. & Metzger, J.P. (2018b) Land use type, forest cover and forest edges modulate avian cross-habitat spillover. *Journal of Applied Ecology*, **55**, 1252–1264.

ter Braak, C.J.F. (2019) New robust weighted averaging- and model-based methods for assessing trait–environment relationships. *Methods in Ecology and Evolution*, **10**, 1962–1971.

Calaboni, A., Tambosi, L.R., Igari, A.T., Farinaci, J.S., Metzger, J.P. & Uriarte, M. (2018) The forest transition in São Paulo, Brazil: historical patterns and potential drivers. *Ecology and Society*, **23**, 7.

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

Cavitt, J.E. & Martin, T.E. (2002) Effects of forest fragmentation on brood parasitism and nest predationin eastern and western landscapes. *Studies in Avian Biology*, 73–80.

Chase, J.M. (2007) Drought mediates the importance of stochastic community assembly. *Proceedings of the National Academy of Sciences*, **104**, 17430–17434.

Chase, J.M., Jeliazkov, A., Ladouceur, E. & Viana, D.S. (2020) Biodiversity conservation through the lens of metacommunity ecology. *Annals of the New York Academy of Sciences*, **1469**, 86–104.

Chase, J.M., McGill, B.J., McGlinn, D.J., May, F., Blowes, S.A., Xiao, X., Knight, T.M., Purschke, O. & Gotelli, N.J. (2018) Embracing scale-dependence to achieve a deeper understanding of biodiversity and its change across communities. *Ecology Letters*, **21**, 1737–1751.

Chisholm, C., Lindo, Z. & Gonzalez, A. (2011) Metacommunity diversity depends on connectivity and patch arrangement in heterogeneous habitat networks. *Ecography*, **34**, 415–424.

de Coster, G., Banks-Leite, C. & Metzger, J.P. (2015) Atlantic forest bird communities provide different but not fewer functions after habitat loss. *Proceedings of the Royal Society B: Biological Sciences*, **282**, 20142844.

Deikumah, J.P., McAlpine, C.A. & Maron, M. (2013) Matrix Intensification Alters Avian Functional Group Composition in Adjacent Rainforest Fragments. *PLOS ONE*, **8**, e74852.

Del Hoyo, J. (2014) *Handbook of the birds of the World alive*, Lynx Editions.

Duflot, R., Georges, R., Ernoult, A., Aviron, S. & Burel, F. (2014) Landscape heterogeneity as an ecological filter of species traits. *Acta Oecologica*, **56**, 19–26.

Dunning, J.B., Danielson, B.J. & Pulliam, H.R. (1992) Ecological processes that affect populations in complex landscapes. *Oikos*, **65**, 169.

Esquivel, M.A., Peris, S.J. & Hohenau, C.P. (2008) Influence of time of day, duration and number of counts in point count sampling of birds in an Atlantic forest of Paraguay. *Ornitologia Neotropical*, **19**, 229–242.

Fahrig, L. (2017) Ecological Responses to Habitat Fragmentation per se. *Annual Reviews of Ecology, Evolution and Systematics*, **48**.

Fahrig, L. (2007) Non-optimal animal movement in human-altered landscapes. *Functional Ecology*, **21**, 1003–1015.

Fischer, J. & Lindenmayer, D.B. (2007) Landscape modification and habitat fragmentation: a synthesis. *Global Ecology and Biogeography*, **16**, 265–280.

Fletcher, R.J., Didham, R.K., Banks-Leite, C., Barlow, J., Ewers, R.M., Rosindell, J., Holt, R.D., Gonzalez, A., Pardini, R., Damschen, E.I., Melo, F.P.L., Ries, L., Prevedello, J.A., Tscharntke, T., Laurance, W.F., Lovejoy, T. & Haddad, N.M. (2018) Is habitat fragmentation good for biodiversity? *Biological Conservation*, **226**, 9–15.

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. *Science Advances*, **1**, e1500052.

Häkkilä, M., Tortorec, E.L., Brotons, L., Rajasärkkä, A., Tornberg, R. & Mönkkönen, M. (2017) Degradation in landscape matrix has diverse impacts on diversity in protected areas. *PLOS ONE*, **12**, e0184792.

Harrison, X.A. (2015) A comparison of observation-level random effect and Beta-Binomial models for modelling overdispersion in Binomial data in ecology & evolution. *PeerJ*, **3**, e1114.

Hartig, F. (2018) DHARMa: Residual Diagnostics for Hierarchical (Multi-Level / Mixed) Regression Models.

Henckel, L., Meynard, C.N., Devictor, V., Mouquet, N. & Bretagnolle, V. (2019) On the relative importance of space and environment in farmland bird community assembly. *PLOS ONE*, **14**, e0213360.

Horváth, Z., Ptacnik, R., Vad, C.F. & Chase, J.M. (2019) Habitat loss over six decades accelerates regional and local biodiversity loss via changing landscape connectance. *Ecology Letters*, **22**, 1019–1027.

Ives, A.R. (2019) R2s for Correlated Data: Phylogenetic Models, LMMs, and GLMMs. *Systematic Biology*, **68**, 234–251.

Jamil, T., Ozinga, W.A., Kleyer, M. & ter Braak, C.J.F. (2013) Selecting traits that explain species–environment relationships: a generalized linear mixed model approach. *Journal of Vegetation Science*, **24**, 988–1000.

Johnson, P.C.D. (2014) Extension of Nakagawa & Schielzeth's R2GLMM to random slopes models. *Methods in Ecology and Evolution*, **5**, 944–946.

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.

Kupfer, J.A., Malanson, G.P. & Franklin, S.B. (2006) Not seeing the ocean for the islands: the mediating influence of matrix-based processes on forest fragmentation effects. *Global Ecology and Biogeography*, **15**, 8–20.

Kupsch, D., Vendras, E., Ocampo-Ariza, C., Batáry, P., Motombi, F.N., Bobo, K.S. & Waltert, M. (2019) High critical forest habitat thresholds of native bird communities in Afrotropical agroforestry landscapes. *Biological Conservation*, **230**, 20–28.

Leibold, M.A., Holyoak, M., Mouquet, N., Amarasekare, P., Chase, J.M., Hoopes, M.F., Holt, R.D., Shurin, J.B., Law, R., Tilman, D., Loreau, M. & Gonzalez, A. (2004) The metacommunity concept: a framework for multi-scale community ecology. *Ecology Letters*, **7**, 601–613.

Leibold, M.A. & Loeuille, N. (2015) Species sorting and patch dynamics in harlequin metacommunities affect the relative importance of environment and space. *Ecology*, **96**, 3227–3233.

Leite, M. de S. (2022) Data and code from: Matrix quality determines the strength of habitat loss filtering on bird communities at the landscape scale (v2.2). *Zenodo*.

Leite, M. de S., Tambosi, L.R., Romitelli, I. & Metzger, J.P. (2013) Landscape Ecology Perspective in Restoration Projects for Biodiversity Conservation: a Review. *Natureza & Conservação*, **11**, 108–118.

Loreau, M. (2010) *From populations to ecosystems: theoretical foundations for a new ecological synthesis*, Princeton University Press, Princeton.

Luck, G.W., Lavorel, S., McIntyre, S. & Lumb, K. (2012) Improving the application of vertebrate trait-based frameworks to the study of ecosystem services. *Journal of Animal Ecology*, **81**, 1065–1076.

McGill, B., Enquist, B., Weiher, E. & Westoby, M. (2006) Rebuilding community ecology from functional traits. *Trends in Ecology & Evolution*, **21**, 178–185.

Metzger, J.P., Villarreal-Rosas, J., Suárez-Castro, A.F., López-Cubillos, S., González-Chaves, A., Runting, R.K., Hohlenwerger, C. & Rhodes, J.R. (2021) Considering landscape-level processes in ecosystem service assessments. *Science of The Total Environment*, **796**, 149028.

Miller, J.E.D., Damschen, E.I. & Ives, A.R. (2018) Functional traits and community composition: A comparison among community-weighted means, weighted correlations, and multilevel models. *Methods in Ecology and Evolution*, **10**, 415–425.

Morante-Filho, J.C., Arroyo-Rodríguez, V. & Faria, D. (2016) Patterns and predictors of β-diversity in the fragmented Brazilian Atlantic forest: a multiscale analysis of forest specialist and generalist birds. *Journal of Animal Ecology*, **85**, 240–250.

Mortara, S.R. (2016) Fern abundance in metacommunities: linking patterns and processes with statistical models.

Murcia, C. (1995) Edge effects in fragmented forests: implications for conservation. *Trends in*

Ecology & Evolution, **10**, 58–62.

Nakagawa, S. & Schielzeth, H. (2013) A general and simple method for obtaining R2 from generalized linear mixed-effects models. *Methods in Ecology and Evolution*, **4**, 133–142.

Nakagawa, S. & Schielzeth, H. (2010) Repeatability for Gaussian and non-Gaussian data: a practical guide for biologists. *Biological Reviews*, **85**, 935–956.

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

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

Nordberg, E.J. & Schwarzkopf, L. (2018) Reduced competition may allow generalist species to benefit from habitat homogenization. *Journal of Applied Ecology*, **280**, 20122131.

Ovaskainen, O., Tikhonov, G., Norberg, A., Guillaume Blanchet, F., Duan, L., Dunson, D., Roslin, T. & Abrego, N. (2017) How to make more out of community data? A conceptual framework and its implementation as models and software. *Ecology Letters*, **20**, 561–576.

Ozgul, A., Oli, M.K., Bolker, B.M. & Perez-Heydrich, C. (2009) Upper respiratory tract disease, force of infection, and effects on survival of gopher tortoises. *Ecological Applications*, **19**, 786–798.

Pandit, S.N., Kolasa, J. & Cottenie, K. (2009) Contrasts between habitat generalists and specialists: an empirical extension to the basic metacommunity framework. *Ecology*, **90**, 2253–2262.

Pardini, R., Faria, D., Accacio, G.M., Laps, R.R., Mariano-Neto, E., Paciencia, M.L.B., Dixo, M. & Baumgarten, J. (2009) The challenge of maintaining Atlantic Forest biodiversity: A multi-taxa conservation assessment of specialist and generalist species in an agro-forestry mosaic in southern Bahia. *Biological Conservation*, **142**, 1178–1190.

Pardini, R., Nichols, E. & Püttker, T. (2017) *Biodiversity Response to Habitat Loss and Fragmentation. Reference Module in Earth Systems and Environmental Sciences*, Elsevier.

Pasher, J., Mitchell, S.W., King, D.J., Fahrig, L., Smith, A.C. & Lindsay, K.E. (2013) Optimizing landscape selection for estimating relative effects of landscape variables on ecological responses. *Landscape Ecology*, **28**, 371–383.

Pfeifer, M., Lefebvre, V., Peres, C., Banks-Leite, C., Wearn, O., Marsh, C., Butchart, S., Arroyo-Rodríguez, V., Barlow, J., Cerezo, A., Cisneros, L., D'Cruze, N., Faria, D., Hadley, A., Harris, S., Klingbeil, B., Kormann, U., Lens, L., Medina-Rangel, G., Morante-Filho, J., Olivier, P., Peters, S., Pidgeon, A., Ribeiro, D., Scherber, C., Schneider-Maunory, L., Struebig, M., Urbina-Cardona, N., Watling, J., Willig, M., Wood, E. & Ewers, R. (2017) Creation of forest edges has a global impact on forest vertebrates. *Nature*, **551**, 187–191.

Prevedello, J.A., Almeida-Gomes, M. & Lindenmayer, D.B. (2018) The importance of scattered trees for biodiversity conservation: A global meta-analysis. *Journal of Applied Ecology*, **55**, 205–214.

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.

Reider, I.J., Donnelly, M.A. & Watling, J.I. (2018) The influence of matrix quality on species richness in remnant forest. *Landscape Ecology*, **33**, 1147–1157.

Rezende, C.L., Scarano, F.R., Assad, E.D., Joly, C.A., Metzger, J.P., Strassburg, B.B.N., Tabarelli, M., Fonseca, G.A. & Mittermeier, R.A. (2018) From hotspot to hopespot: An opportunity for the Brazilian Atlantic Forest. *Perspectives in Ecology and Conservation*, **16**, 208–214.

Ruffell, J., Clout, M.N. & Didham, R.K. (2017) The matrix matters, but how should we manage it? Estimating the amount of high-quality matrix required to maintain biodiversity in fragmented landscapes. *Ecography*, **40**, 171–178.

Saunders, D.A., Hobbs, R.J. & Margules, C.R. (1991) Biological Consequences of Ecosystem Fragmentation: A Review. *Conservation Biology*, **5**, 18–32.

Sibly, R.M., Witt, C.C., Wright, N.A., Venditti, C., Jetz, W. & Brown, J.H. (2012) Energetics, lifestyle, and reproduction in birds. *Proceedings of the National Academy of Sciences*, **109**, 10937–10941.

Silva, C.M., Pereira, J.A.C., Gusmões, J.D.S.P., Mendes, B.E.P., Valente, H., Morgan, A.P., Goulart, D. & Hasui, É. (2020) Birds' gap-crossing in open matrices depends on landscape structure, tree size, and predation risk. *Perspectives in Ecology and Conservation*, **18**, 73–82.

Siqueira, T., Saito, V.S., Bini, L.M., Melo, A.S., Petsch, D.K., Landeiro, V.L., Tolonen, K.T., Jyrkänkallio-Mikkola, J., Soininen, J. & Heino, J. (2020) Community size can affect the signals of ecological drift and niche selection on biodiversity. *Ecology*, **101**, e03014.

Suárez-Castro, A.F., Simmonds, J.S., Mitchell, M.G.E., Maron, M. & Rhodes, J.R. (2018) The Scale-Dependent Role of Biological Traits in Landscape Ecology: A Review. *Current Landscape Ecology Reports*, **3**, 12–22.

Tscharntke, T., Tylianakis, J.M., Rand, T.A., Didham, R.K., Fahrig, L., Batáry, P., Bengtsson, J., Clough, Y., Crist, T.O., Dormann, C.F., Ewers, R.M., Fründ, J., Holt, R.D., Holzschuh, A., Klein, A.M., Kleijn, D., Kremen, C., Landis, D.A., Laurance, W., Lindenmayer, D., Scherber, C., Sodhi, N., Steffan-Dewenter, I., Thies, C., van der Putten, W.H. & Westphal, C. (2012) Landscape moderation of biodiversity patterns and processes - eight hypotheses. *Biological Reviews*, **87**, 661–685.

Vellend, M. (2010) Conceptual synthesis in community ecology. *The Quarterly review of biology*, **85**, 183–206.

Vellend, M., Srivastava, D.S., Anderson, K.M., Brown, C.D., Jankowski, J.E., Kleynhans, E.J., Kraft, N.J.B., Letaw, A.D., Macdonald, A.A.M., Maclean, J.E., Myers-Smith, I.H., Norris, A.R. & Xue, X. (2014) Assessing the relative importance of neutral stochasticity in ecological communities. *Oikos*, **123**, 1420–1430.

Watling, J.I., Nowakowski, A.J., Donnelly, M.A. & Orrock, J.L. (2011) Meta-analysis reveals the importance of matrix composition for animals in fragmented habitat: Connectivity meta-analysis. *Global Ecology and Biogeography*, **20**, 209–217.

Whelan, C.J., Wenny, D.G. & Marquis, R.J. (2008) Ecosystem Services Provided by Birds. *Annals of the New York Academy of Sciences*, **1134**, 25–60.

Major axes of variation in tree demography across global forests

Melina de Souza Leite[1]*, Sean M. McMahon[2], Paulo Inácio Prado[1], Stuart J. Davies[3], Alexandre Adalardo de Oliveira[1], Hannes P. De Deurwaerder[4], Salomón Aguilar[5], Kristina J. Anderson-Teixeira[5,6], Nurfarah Aqilah[7], Norman A. Bourg[8], Warren Y. Brockelman[9,10], Nicolas Castaño[11], Chia-Hao Chang-Yang[12], Yu-Yun Chen[13], George Chuyong[14], Keith Clay[15], Álvaro Duque[16], Sisira Ediriweera[17], Corneille E.N. Ewango[18], Gregory Gilbert[19], I.A.U.N. Gunatilleke[20], C.V.S. Gunatilleke[20], Robert Howe[21], Walter Huaraca Huasco[22], Akira Itoh[23,24], Daniel J. Johnson[25], David Kenfack[5], Kamil Král[26], Yao Tze Leong[27], James A. Lutz[28,29], Dr. Jean-Remy Makana[30], Yadvinder Malhi[22], William J. McShea[31], Mohizah Bt. Mohamad[32], Musalmah Nasardin[33], Anuttara Nathalang[9], Geoffrey Parker[34], Renan Parmigiani[1], Rolando Pérez[5], Richard P. Phillips[35], Pavel Šamonil[26], I-Fang Sun[36], Sylvester Tan[32], Duncan Thomas[37], Jill Thompson[38], María Uriarte[39], Amy Wolf[21], Jess Zimmerman[40], Daniel Zuleta[3], Marco D. Visser[41]**, Lisa Hülsmann[42,43,44]**

*corresponding author: melina.leite@ib.usp.br

**MDV and LH should be considered joint senior author.

[1]Department of Ecology, University of São Paulo, São Paulo, SP, Brazil, [2]Forest Global Earth Observatory, Smithsonian Environmental Research Center, Edgewater, MD, USA, [3]Forest Global Earth Observatory, Smithsonian Tropical Research Institute, Washington, DC, USA, [4]Department of Ecology and Evolutionary Biology, Princeton University,

USA, [5]Forest Global Earth Observatory, Smithsonian Tropical Research Institute, Panama, [6]Conservation Ecology Center, Smithsonian's National Zoo & Conservation Biology Institute, Front Royal, VA, 22630, USA, [7]Forest Research Institute Malaysia, Selangor, Kepong, 52110, Malaysia, [8]Conservation Ecology Center, Smithsonian's National Zoo & Conservation Biology Institute, Front Royal, VA, 22632, USA, [9]National Biobank of Thailand, National Science and Technology Development Agency, Thailand, [10]Institute of Molecular Biosciences, Mahidol University, Salaya, Phutthamonthon, Nakhon Pathom, Thailand, [11]Programa de Ecosistemas y Recursos Naturales, Instituto SINCHI, Bogotá, Colombia, [12]Department of Biological Sciences, National Sun Yat-sen University, Kaohsiung, Taiwan, [13]Department of Natural Resources and Environmental Studies, National Dong Hwa University, Hualien, Taiwan, [14]Department of Plant Science, University of Buea, POB 63 Buea, Cameroon, [15]Dept. of Ecology and Evolutionary Biology, Tulane University, New Orleans, USA, [16]Departamento de Ciencias Forestales, Universidad Nacional de Colombia Sede Medellín, Colombia, [17]Department of Science and Technology, Faculty of Applied Sciences, Uva Wellassa University, Badulla, 90000, Sri Lanka, [18]Faculté de Gestion des Ressources Naturelles Renouvelables & Faculté des Sciences, Université de Kisangani, Democratic Republic of Congo, [19]Environmental Studies Department, University of California, Santa Cruz, CA, 95064, USA, [20]Department of Botany, University of Peradeniya, Sri Lanka, [21]Dept. of Biology and Cofrin Center for Biodiversity, University of Wisconsin-Green Bay, Green Bay, WI, 54311-7001, USA, [22]Environmental Change Institute, School of Geography and the Environment, University of Oxford, Oxford, OX1 3QY, UK, [23]Osaka City University, Japan, [24]Osaka Metropolitan University, Japan, [25]School of Forest, Fisheries, and Geomatics Sciences, University of Florida, USA, [26]Department of Forest Ecology, Silva Tarouca Research Institute, Brno, 602 00, Czech Republic, [27]Forest Research Institute Malaysia, Selangor, Kepong, 52109, Malaysia, [28]Department of Wildland Resournces, Utah State University, Logan, Utah, 84322-5230, USA, [29]Ecology Center, Utah State University, Logan, Utah, 84322-5230, USA, [30]Faculte des Sciences, Universite de Kisangani, Democratic Republic of Congo, [31]Conservation Ecology Center, Smithsonian's National Zoo & Conservation Biology Institute, Front Royal, VA, 22631, USA, [32]Forest Department Sarawak, Malaysia, [33]Forest Research Institute Malaysia, Selangor, Kepong, 52111, Malaysia, [34]Smithsonian Environmental Research Center, Edgewater, MD, USA,

[35]Department of Biology, Indiana University, Bloomingtom, IN, USA, [36]Center for Interdisciplinary Research on Ecology and Sustainability, National Dong Hwa University, Hualien, Taiwan, [37]School of Biological Sciences, Washington State University, Vancouver, WA, 98686, USA, [38]UK Centre for Ecology & Hydrology, United Kingdom, [39]Department of Ecology, Evolution, and Environmental Biology, Columbia University, New York, NY, 10027, USA, [40]Department Environmental Sciences, University of Puerto Rico, Puerto Rico, [41]Institute of Environmental Sciences, Leiden University, Leiden, The Netherlands, [42]Ecosystem Analysis and Simulation Lab, University of Bayreuth, Bayreuth, Germany, [43]Theoretical Ecology, University Regensburg, Regensburg, Germany, [44]Bayreuth Center of Ecology and Environmental Research (BayCEER), University of Bayreuth, Bayreuth, Germany

O manuscrito será submetido à revista *Global Ecology and Biogeography*.

Código de análises disponíveis no repositório Zenodo:
https://doi.org/10.5281/zenodo.7250584 a partir da versão para a tese (v0.8) do repositório do GitHub https://github.com/melina-leite/varpart_treeVitalRates.

Abstract

Aim: Global forests and their structural and functional features are shaped by innumerable mechanisms that leave fingerprints in tree vital rates. Yet, although many studies have tried to quantify how specific mechanisms influence vital rates, the relative importance of these drivers across forests remains unclear. A careful assessment of how vital rates vary among species and in space and time across forests, therefore, can serve as a baseline expectation of the importance of different mechanisms in different contexts.

Location: Twenty-one forests dynamic plots worldwide.

Time period: 1981-2021

Major taxa studied: Woody plants

Methods: We developed a simple conceptual and statistical framework (variance partitioning of multilevel models) that attributes the variability in growth, survival, and recruitment to variation in species, space, and time, and interactions thereof, which we refer to as *organising principles* (OPs). The framework was applied to dynamic data from twenty-one globally distributed forest plots covering more than 2.9 million trees of approx. 6,500 species.

Results: Variability among species, i.e. the species OP, was a major source of variability in tree vital rates, explaining 28-33% of demographic variance alone, and in interaction with space 14-17%, totalling 40-43%. Models with small spatial grain (quadrat size) retained most of the spatial OP, but a large proportion of variance remained unexplained (31-55%). Temporal variability acted mostly on recruitment and mortality and in interaction with space. Variability among species declined with species richness across forests, indicating that diverse forests featured smaller interspecific differences in vital rates.

Main conclusions: Decomposing variance in vital rates into the proposed OPs has a high potential for identifying the structuring mechanisms of global forest dynamics, as it highlights the most promising avenues for future research both in terms of understanding the relative contributions of mechanisms to forest demography and diversity and for improving projections of forest ecosystems.

Introduction

Forests are an integral component of the global carbon cycle (Anderson-Teixeira et al., 2021) and are home to a majority of the terrestrial biodiversity (Pillay et al., 2022). Changes in climate and land use threaten forests but anticipating how these diverse systems might respond is challenging due to the array of mechanisms that determine forest structure and function. One approach to better understand how forests function is through the analysis of tree demography (Griffith et al., 2016): the growth, survival, and reproduction of individuals. These vital (i.e. demographic) rates combine to determine key features of forests, such as biomass stocks and fluxes (Needham et al., 2020), structural complexity (Kohyama, 1993; Iida et al., 2014), and diversity (Lasky et al., 2014). Improving our understanding of demography, for instance, may foster the development of ecological theories, for example in community ecology as coexistence and niche theory (Kohyama, 1993; Adler et al., 2010; Diez et al., 2014; Lasky et al., 2014). Moreover, demography has been identified as critical for more accurately modelling the terrestrial component in earth system models (Fisher et al., 2018) and projecting the future of the terrestrial carbon sink (Pan et al., 2011; Pugh et al., 2019). Even small changes in tree vital rates, if consistent over space or time, can affect carbon cycles (Needham et al., 2022) and thus the extent to which climate change can be mitigated by forests (Bonan, 2008; Canadell & Raupach, 2008).

Vital rates are influenced by interacting mechanisms across spatial and temporal scales creating a challenge to including demography in forest models (Weng et al., 2015). Many of these mechanisms are difficult or impossible to measure directly leading to the use of imperfect proxies (Swenson et al., 2020). There exists, however, a higher level of information that may guide studies in demographic analyses: the patterns of demographic rates themselves. The most comprehensive source of forest demography are the growth, mortality, and recruitment data derived from inventories on permanent plots (Davies et al., 2021). The three vital rates and the contextual variables ('dimensions') associated with them offer an opportunity to organise the elements of forest dynamics in ways that help infer the potential mechanisms at structuring forest dynamics. For example, by means of natural selection, species vary in their strategies to acquire and allocate resources leading to a species dimension along which phenotypes vary and, thus, also the observed vital rates of individual species (Díaz et al., 2016;

Johnson et al., 2018; Rüger et al., 2018; Needham et al., 2022). Moreover, as resource availability and stressors vary along spatial and temporal dimensions, the environmental conditions of a forest also structure its vital rates, e.g. soil and topography vary across space (Zuleta et al., 2018) and drought conditions over time (Bennett et al., 2015; Chen et al., 2019; Ovenden et al., 2021). Finally, all these dimensions have interactive effects. Functional traits vary between species and cause differential responses along spatial and temporal dimensions, for example when drought tolerant and intolerant species respond differently to a climatic event (Condit et al., 2004; Kupers et al., 2019). Gap dynamics change over both space and time, and tree responses change as forest gaps close (Wright et al., 2003). Patterns of how variability in vital rates is partitioned along key dimensions can thus reveal how important various biotic and abiotic drivers are in influencing tree demography and in extension forest dynamics.

We propose a conceptual framework that groups the mechanisms creating variation in vital rates as being related to species, space, and time. Together, these three dimensions and their interactions form seven organising principles (OPs, Table 1), which can be derived from forest inventory data. When the mechanisms that drive tree vital rates operate on unique combinations of these dimensions, quantifying the variability in vital rates that each OP describes may provide insights into the strength and importance of the associated mechanisms (Table 1). The statistical counterpart to this conceptual framework are the multilevel models (McMahon & Diez, 2007; Visser et al., 2016). These allow decomposing forest demographic data across OPs and quantifying the relative importance of each OP by estimating and partitioning the variance in each vital rate (Browne et al., 2005). In this framework, a broad assessment of the structure of variation in vital rates can be accomplished (Table 1).

We applied the framework decomposing variation in tree vital rates into the dimensions of species, space, and time as well as their interactions (OPs) for a set of twenty-one large (6 to 52 ha) forest dynamics plots from tropical to temperate forests where all stems with diameter at 1.3 m from the ground (dbh) ≥ 1 cm are identified, mapped, and measured approximately every five years (Davies et al., 2021). We then compared the relative importance of the OPs for each vital rate at each forest with the goal of identifying consistent patterns in which OPs capture variation in vital rates: **(1) among vital rates**, i.e. investigating if some OPs are more important than others for specific vital rates; **(2) across spatial scales** (grain size), given the nature of scale

dependency of ecological processes; and **(3) among forests globally** to understand how patterns may differ depending on forest diversity and structure. In answering these questions, we provide macroecological patterns of the relative importance of OPs and, thus, the first approximate assessment of their associated mechanisms in generating variation in forest demography globally. Our framework, therefore, aims to facilitate hypothesis-driven research on mechanisms by first describing the higher level patterns of vital rate variability.

Table 1: Seven organising principles (OPs) and the mechanisms that are associated with them, i.e., by creating variability of vital rates in the associated dimension *species, space* and *time* or the interaction thereof. References are example studies for the mechanisms.

Organising principles (OPs)	Related mechanisms and examples
Species	
Trees of different species have different vital rates.	**Natural selection** in response to biotic and abiotic stressors creates variation in evolutionary strategies that leads **to unique geno- and phenotypes** among individual trees manifested in different species. Species then display difference in their vital rates, as evidenced as follows: • Species have different growth forms (e.g. shrubs and trees), dispersal abilities, and regeneration strategies (Martínez-Ramos *et al.*, 2021) that are related to different allocation strategies (Rüger *et al.*, 2020), also known **as life history strategies**, leading to different **demographic niches** (Condit *et al.*, 2006) and the emergence of interspecific **demographic trade-offs**, such as growth-mortality, recruitment-mortality (Russo *et al.*, 2008), and stature-recruitment (Rüger *et al.*, 2018). • All these differences are potentially related to species functional traits (Poorter *et al.*, 2008; Adler *et al.*, 2014).
Space	
Trees in different locations (quadrats) have different vital rates.	Spatial heterogeneity created by variability in soil and topography as well as by differences in stand structure results in spatial differences of **resource availability** (nutrients, moisture, light) and **environmental stressors** (e.g. wind). In response, tree vital rates can be consistently higher in some areas than in others (Arellano, 2019): • Tree mortality may be larger on hilltops given lower water availability in soil and higher wind disturbances (Zuleta *et al.*, 2018). • Tree growth is faster and mortality larger in nutrient rich soils (Russo *et al.*, 2005; Lévesque *et al.*, 2016).
Time	
Trees during different time periods have different vital rates.	Environmental conditions are not stable in time but vary with **climate** and in response to **disturbances**, jointly affecting all species across a forest (synchronised effects): • Cyclones and other drastic climatic disturbances can kill many trees at once in a forest (Uriarte *et al.*, 2019). • Severe droughts can decrease growth and/or increase mortality directly (McDowell *et al.*, 2020) or indirectly by increasing the propensity of disease outbreaks (Negrón *et al.*, 2009). • Irregular masting events and rainfall affect growth and survival of seedlings (Martini *et al.*, 2022).
Species x space	

Trees of different species in the same location (quadrat) have different vital rates.	Due to **spatial niche** effects, species have different environmental preferences that in combination with spatial variability create certain habitats where some species perform better than others. For example: • Species adapted to low light availability have lower mortality in denser areas (Jurinitz *et al.*, 2013). • Species with more dispersive seeds recruit more in open gaps (Clark *et al.*, 2018) • Soil fertility affects species in different ways (Russo *et al.*, 2008).
Species x time	
Trees of different species during the same time period have different vital rates.	Species environmental preferences also create **temporal niche** effects that lead to asynchronous species responses to temporal variability (Fung *et al.*, 2020). For example: Species that are vulnerable to drought have higher mortality than those that are resistant or resilient (Chen *et al.*, 2019) • Species with more dispersive seeds recruit more in a favourable year (Clark *et al.*, 2018) • Species with high wood density suffer lower immediate mortality after hurricanes (Uriarte *et al.*, 2019).
Space x time	
Trees in the same location during different periods have different vital rates.	**Gap dynamics:** large tree falls open temporal gaps in the forest changing the environmental conditions of the surrounding area for a certain time (Kohyama, 1993): • Fallen trees increase immediate local mortality in the area surrounding it. • Open gaps increase light availability, allowing faster growth (Brokaw, 1987) of understory trees and recruitment (Wright *et al.*, 2003) but just during specific time periods. **Climate effects** can manifest themselves differently depending on the prevailing basic conditions in a given area. For example: • Drought events increase mortality disproportionally in valleys than on hilltops or ridges (Zuleta *et al.*, 2017). • Soil nutrients can influence growth response to drought (Lévesque *et al.*, 2016).
Species x space x time + individual	
Trees of the same species in different locations and during different time periods have different vital rates. Individual trees have different vital rates.	**Individual variation** in vital rates given genetic and phenotypic variation, spatial variation at the microscale (Schwartz *et al.*, 2020), and ontogeny. • Trees of different sizes and multi-stemmed trees have different mortality (Johnson *et al.*, 2018; Su *et al.*, 2020) and growth rates (Lu *et al.*, 2021), and their vital rates respond differently to climatic variation and disturbances. • Functional traits influence growth depending on the size of the individuals (Gibert *et al.*, 2016). • Local biotic interactions, as higher-order interactions, change individual vital rates (Li *et al.*, 2020). **Phenotypic plasticity** can change the observed vital rates of different species due to temporal and spatial environmental conditions (both biotic and abiotic). This plasticity may be seen in individual-level functional traits (Burns & Strauss, 2012).

Methods

Tree census data

We used data from twenty-one forest dynamics plots (Fig 1A) from the Forest Global Earth Observatory network (ForestGEO, Davies et al., 2021). In each plot, all stems with dbh ≥1 cm are mapped, identified, and repeatedly measured using a standardised protocol (Condit 1998). Plots used in this study range in size between 6 and 52 ha, with an inter-census measurement interval of approximately 5 years (range 3

to 10 years). All forest plots had at least 2 censuses. The forest plots cover a wide range of environmental, climatic, and edaphic conditions, with the number of species varying two orders of magnitude from 12 to 1402 (including morphospecies). In total, approximately 2.9 million trees from more than 6,500 species were repeatedly censused over periods of 3 to 40 years in more than 575 ha. For summary information on the plots and further details on how tree census data were processed see Appendix S1 in Supporting Information.

Vital rate definition and modelling

We analysed growth, mortality, and recruitment as annual (per-capita) vital rates by calculating vital rate information at the level of individual observations (sensu Kohyama et al., 2018) and fitting suitable multilevel models per forest plot and vital rate. Annual individual **growth** was calculated as dbh increment of alive trees, divided by the individual's census interval length in years, and modelled using multilevel models with a normal distribution.

Mortality rates were estimated from the status of trees - alive or dead - in each consecutive census assuming a binomial distribution (Kohyama et al., 2018). Mortality rates were annualised by using a complementary log log link function (cloglog), where the log-transformed time between individual measurements is included as an offset term (Fortin et al., 2008; Johnson et al., 2018).

Recruitment was defined as the final per-capita recruitment rate (Kohyama et al., 2018), which denotes the proportion of trees that are new recruits (i.e. not present in the previous census) and can be interpreted as the probability of an individual tree being new. Recruitment rates were estimated using the same modelling approach as for mortality, i.e. a binomial model with a cloglog link function and time interval length as an offset term. Because there is no time interval associated with individual recruits as they have not been monitored in the previous census), the time interval for recruits was calculated as the mean time interval of the survivors in the same quadrat. If there were no survivors in a specific quadrat, we used the mean time interval between the respective censuses from the entire plot.

Multilevel models

In order to quantify the variation in vital rates associated with each OP, we applied variance decomposition to multilevel models (MLMs) fitted separately for each vital rate and forest plot. MLMs are particularly useful for variance decomposition as they are able to reflect that ecological datasets contain identifiable hierarchical units or clusters (Browne et al., 2005; McMahon & Diez, 2007). MLMs can account for such interdependence by partitioning the total variance into different components of variation due to each cluster (see example in Table 1). We included *species*, spatial quadrat (*space*) and census interval (*time*) and their two-way interactions as variance components (also known as random effects) in each MLM. With that, we estimated the variance associated with each OP while respecting the hierarchical structure of the data. Following the convention of MLMs, the general structure of our models is:

$$Y_{ijkl} = \beta_0 + s_i + q_j + t_k + sq_{ij} + st_{ik} + qt_{jk} + e_{ijkl} \qquad \text{(eq. 1)}$$

Where Y_{ijkl} is the vital rate for individual observation l, from species i, in spatial quadrat j and time interval k. β_0 is the intercept. s_i, q_j, t_k are random effects for the OP species, space (quadrat) and time interval, respectively. $sq_{ij}, st_{ik}, qt_{jk}$ are random intercepts for the interactions species x space, species x time and space x time. All random effects are taken to be independent, each represented by a normal distribution with mean zero and their respective variances $\sigma_s^2, \sigma_q^2, \sigma_t^2, \sigma_{sq}^2, \sigma_{st}^2, , \sigma_{qt}^2$. The residual variance (e_{ijkl}) represents the variance explained by the three-way interaction species x space x time, and any unexplained variation among observations including non-separable measurement error and individual variation (Table 1). Residual variance is equivalent to σ_e^2 in growth models that assume a normally distributed error. For mortality and recruitment, modelled with binomial distributions, the residual variance at the link scale (i.e. linear predictor scale) is the expected variance for the binomial distribution ($\frac{\pi^2}{6} \sim 1.68$) (Nakagawa et al., 2017). We decided to not include the three-way interaction species x space x time, nor the individual variance, because of computational time limitation as these random effects would exponentially increase the number of parameters to estimate. Thus, repeated measures of the same individual are not

accounted for in the model, but we expect that, like Rüger et al (2018), the potential bias caused by repetitive tree observations is negligible.

To partition the total variance of the vital rates among the individual OPs, we calculated **variance partition coefficients (VPCs)** (Browne et al., 2005). The VPC of each OP was calculated as the proportion of its variance to the total variance of the model.

It is worth noting that we intentionally included no fixed effects in the models, in contrast to the usual statistical approach when searching for specific mechanisms, e.g., including dbh as a predictor to estimate ontogenetic effects in vital rates (Needham et al., 2022). In our framework, all mechanisms are considered through OPs, which represent the dimensions at which they generate variability. Our models provide a baseline approximation and could be extended by adding variables related to the mechanisms of interest as fixed effects (e.g., Camac et al., 2018). It may guide posterior and detailed analyses to provide further insight into how variability in vital rates arises in and is organised among OPs.

All data analyses were performed using R (R Core Team, 2022), using the R package brms (Bürkner, 2017) to build Bayesian MLMs. For all estimated parameters, we used *brms* noninformative default prior distributions. For each model, we ran three Monte Carlo Markov chains with 3,000 iterations, discarding the first 1,000 iterations and thinning with an interval of 5, resulting in 1200 posterior samples. We checked convergence of the chains using the Gelman–Rubin criterion and by visually inspecting trace plots of estimated coefficients.

Analysis framework

Organising principles among vital rates

To assess the relative importance of the OPs among vital rates, we compared the VPC results for each vital rate among the twenty-one forest plots. However, because 16 forests had too few census intervals to fit the full model (see below), i.e. less than three (Table S2), we fit a reduced version of the model (eq. 1) without the temporal OPs (dropping the variances σ_t^2, σ_{st}^2, σ_{qt}^2). The reduced models were run separately for each time interval of a forest plot, and variances were averaged across time intervals for forests with more than one census interval.

Temporal organising principles

Currently, a bottleneck of our analysis is the scarcity of data for the temporal dimension of vital rates variability. For MLMs, the estimation of the variance of a grouping variable (i.e., *time* OP) with less than four to five levels may be biased towards zero (Oberpriller et al., 2022). In our data, only five forest plots presented a reasonable number of census intervals (i.e., at least three spanning between 15 and 40 years) to be considered suitable for the full VPC analysis including temporal OPs (eq.1). We fit these MLMs to ten random subsets of 5 ha each sampled from the full forest plots, where each subset was composed of five non-overlapping quadrats of 1 ha. This subsetting procedure was necessary to restrict computational time resulting from the large number of observations, especially on the large plots that are species-rich and of high tree density (i.e., Barro Colorado Island 50 ha, Lambir 50 ha, Pasoh 50 ha, Fig 1a and Table S1.2). Variance estimates of the OPs for each forest plot were averaged across estimates of the ten subsets.

Organising principles across spatial scales

To assess how the relative importance of OP varies with spatial scale, i.e. how the choice of a specific quadrat size impacts VPCs, we divided each forest plot into non-overlapping quadrats with increasing grain size: 5x5 m (0.0025 ha), 10x10 m (0.01 ha), 20x20 m (0.04 ha), 50x50 m (0.25 ha) and 100x100 m (1 ha). Depending on the size of the plot, we trimmed the data to fit within a rectangular region with edges that were even multiples of 100 m, discarding the data outside this area (Chisholm et al., 2013). This guaranteed that each plot could be evenly divided into quadrats of 1 ha and that exactly the same area was analysed at all spatial scales. We fitted MLMs for the models without (reduced) and with temporal OPs averaged VPCs over all forest plots for each grain size and vital rate. We did not consider any spatial autocorrelation among adjacent quadrats to keep VPCs readily interpretable.

Organising principles across a global species richness gradient

Globally, species richness is one of the most distinguishing characteristics of forests and strongly correlates, for instance, with latitude (Keil & Chase, 2019), precipitation (Adler & Levine, 2007), and biome history (Wiens & Donoghue, 2004). The plots used in this analysis span two orders of magnitude in the number of species

(12 to 1402, including morphospecies) offering a unique opportunity to explore if and how sources of variability in vital rates are associated with species diversity. We therefore assessed the effect of log-transformed rarefied species richness (cf. Appendix S5) on the VPCs of *species*, *space*, *species x space* and *residual* OP using dirichlet regression from the R package DirichletReg (Maier, 2021), appropriate for response variables that are multiple categories of proportional data (Douma & Weedon, 2019).

Robustness analyses

We performed four main robustness analyses to make sure our VPCs estimates from the forest plots are robust (1) to different forest plot sizes (6 to 50 ha) for the models without temporal OPs, by subsampling and comparing VPCs of the same forest (Lambir) with both procedures (Appendix S2); (2) to the approach of computing average VPCs for the model with temporal OPs from subsampled plots (10 samples of 5 ha) (Appendix S2); (3) to changes in the modelling procedure, by including or excluding temporal OPs from the VPC analysis (Appendix S3); and (4) the presence of rare species, by excluding or including rare species, to test the influence of rarity on VPCs (Appendix S4).

VPCs estimates from all forest plots were robust to changes in plot size, and VPC estimates remained reliable after removing temporal OPs. Specifically, our main results are also robust to the presence of rare species, though excluding or regrouping rare species does result in small decreases in the *species* VPC, balanced by an increase in the *residual* and *species x space* VPC (Appendix S4).

Results

Organising principles among vital rates

When comparing the relative importance of the OPs for all twenty-one forests distributed globally, we found that, despite large differences among the plots with respect to climate, environment, species richness etc., the relative importance of the OPs was relatively similar (Fig. 1). Generally, *species* was the most important OP for explaining variance in all three vital rates, after the *residual* OP. At the smallest spatial grain (quadrat size 5x5 m), average *species* and *species x space* VPCs varied little

among vital rates, ranging from 28 to 33%, and 14 to 17%, respectively. Average *space* VPC was smaller for growth (4%), intermediate for mortality (10%) and larger for recruitment (19%). *Residual* VPCs were on average about half of the total variance for growth and mortality (55 and 47%, respectively) but smaller for recruitment (31%).
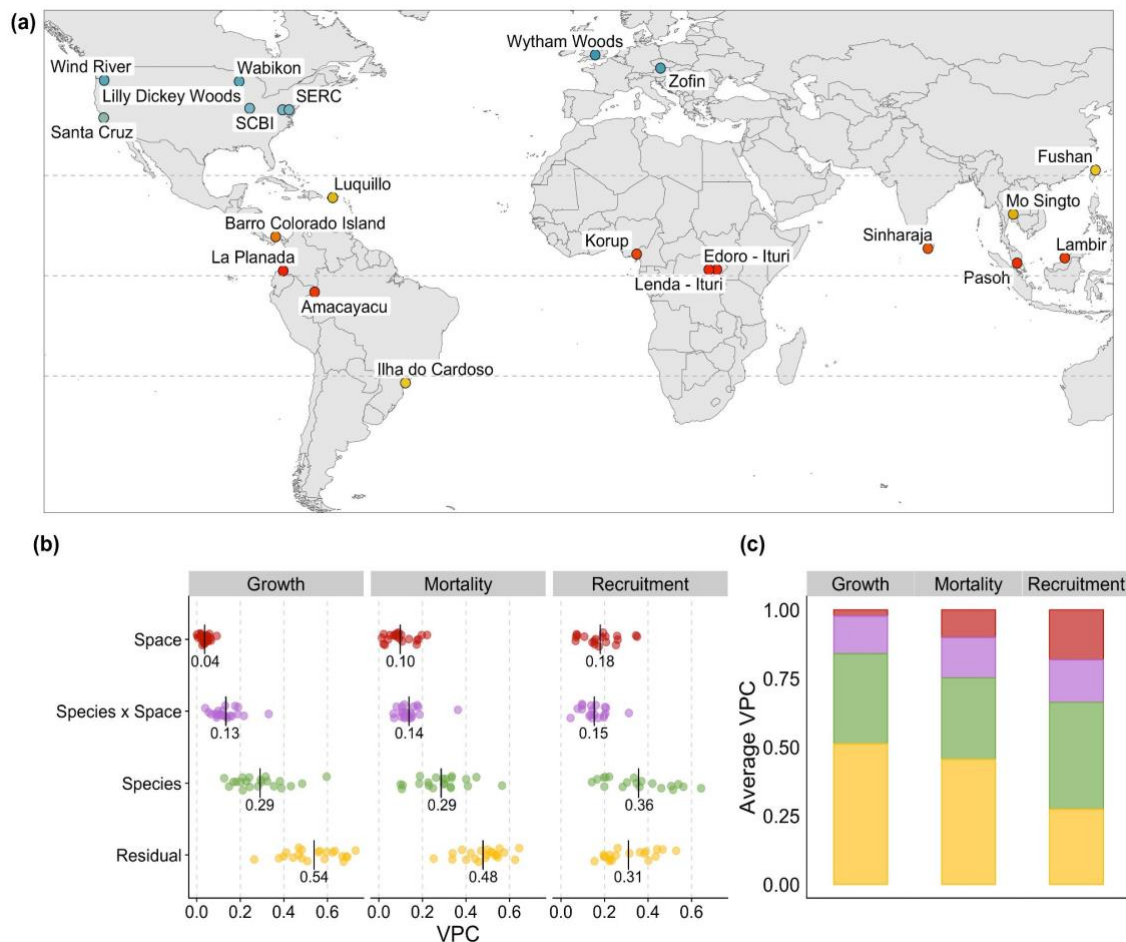

Figure 1. (a) Global distribution of the twenty-one forest plots. (b) Variance partition coefficients (VPC) of the organising principles (OPs) per vital rate - growth, mortality, and recruitment - with mean values indicated as black vertical lines and numbers. (c) Average VPCs across all plots, where colours correspond to the same OPs as in (b). Models were fitted at the 5x5 m quadrat size. Each forest plot in (a) is colored by latitude.

Temporal organising principles

When analysing demographic data from the five forest plots with more than four consecutive censuses (quadrat size 5x5 m), we found that *species* remained the most important OP to explain variance in tree vital rates, except for growth, where the *species x space* VPC was larger for four of the five plots (Fig. 2). Temporal OPs (*time*, *species x time* and *space x time*) were especially important for mortality and recruitment, where

VPCs of *space x time* (on average 10 and 15%, respectively) were larger than VPCs of *species x space* (on average 6 and 10%, respectively).
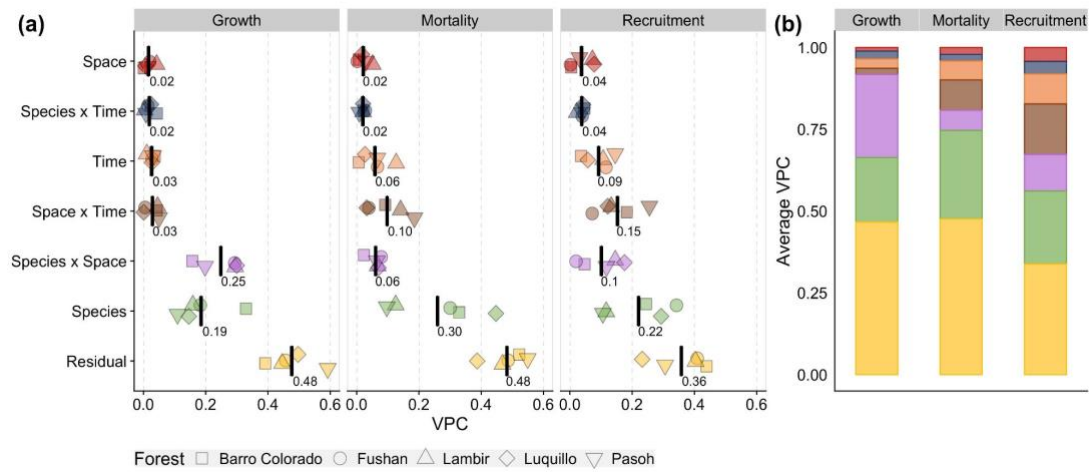

Figure 2. (a) Variance partition coefficient (VPC) of the organising principles (OPs) per vital rate - growth, mortality, and recruitment - for the five forest plots with at least four censuses. Mean VPCs across plots are presented as black lines and numbers. (b) Average VPCs across the five plots, where colours correspond to the same OPs as in (a). Models were fitted at the 5x5 m quadrat scale. See Fig. 1a for forest plot locations.

Organising principles across spatial scales

When comparing average VPCs across five spatial grain sizes, we found that the relative importance of residual variation increased with grain size for all vital rates and more accentuated for growth (Fig. 3). For instance, for the models including temporal OPs (Fig. 3b), residual variation increased from 46% at the smallest grain (quadrat size 5x5 m) to 71% at the largest grain (100x100 m). In turn, the spatial OPs - *space*, *species x space* and *space x time* - consistently decreased in relative importance with increasing spatial grain for all vital rates. The OPs *species* and *species x time* remained almost equally important across spatial grains.
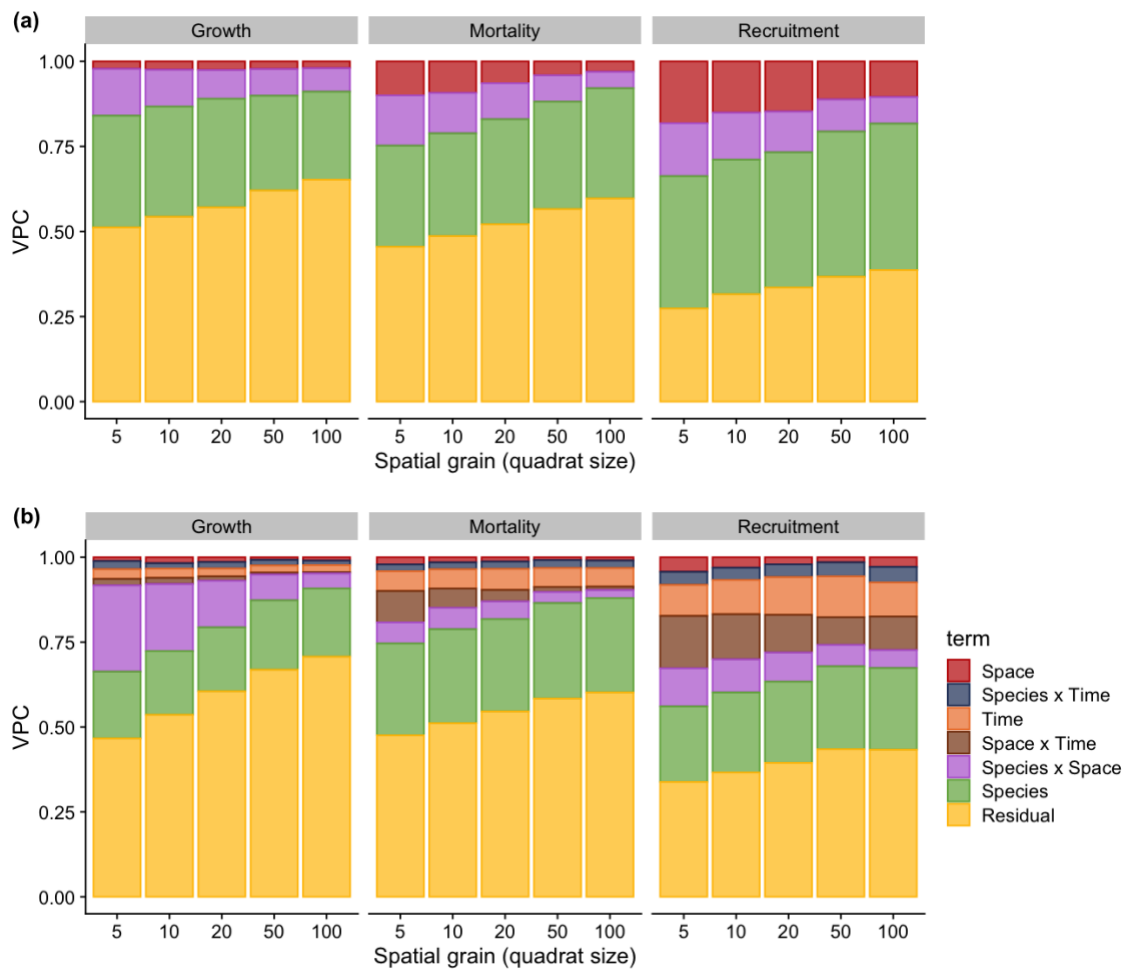
Figure 3. Average variance partition coefficients (VPCs) of each organising principle (OP) across five spatial grains from 5x5 m to 100x100 m quadrat size for the vital rates growth, mortality, and recruitment: (a) reduced models without temporal OPs for all twenty-one forests plots, and (b) full models with temporal OPs for the five (sub)tropical forest plots with enough censuses (Barro Colorado Island, Fushan, Lambir, Luquillo, and Pasoh).

Organising principles across a global species richness gradient

While the *species* OP was the most important component of vital rates variance partitioning throughout the forests, we also found that its importance decreased with species richness for recruitment and growth, but not for mortality (Fig. 4). The decrease in *species* VPC for growth and recruitment was led by a decrease in the *species* standard deviation (Fig. S5.4). This result was robust to the presence of rare species (Fig. S5.5). The other OPs showed no significant changes with species richness (Fig. 4).
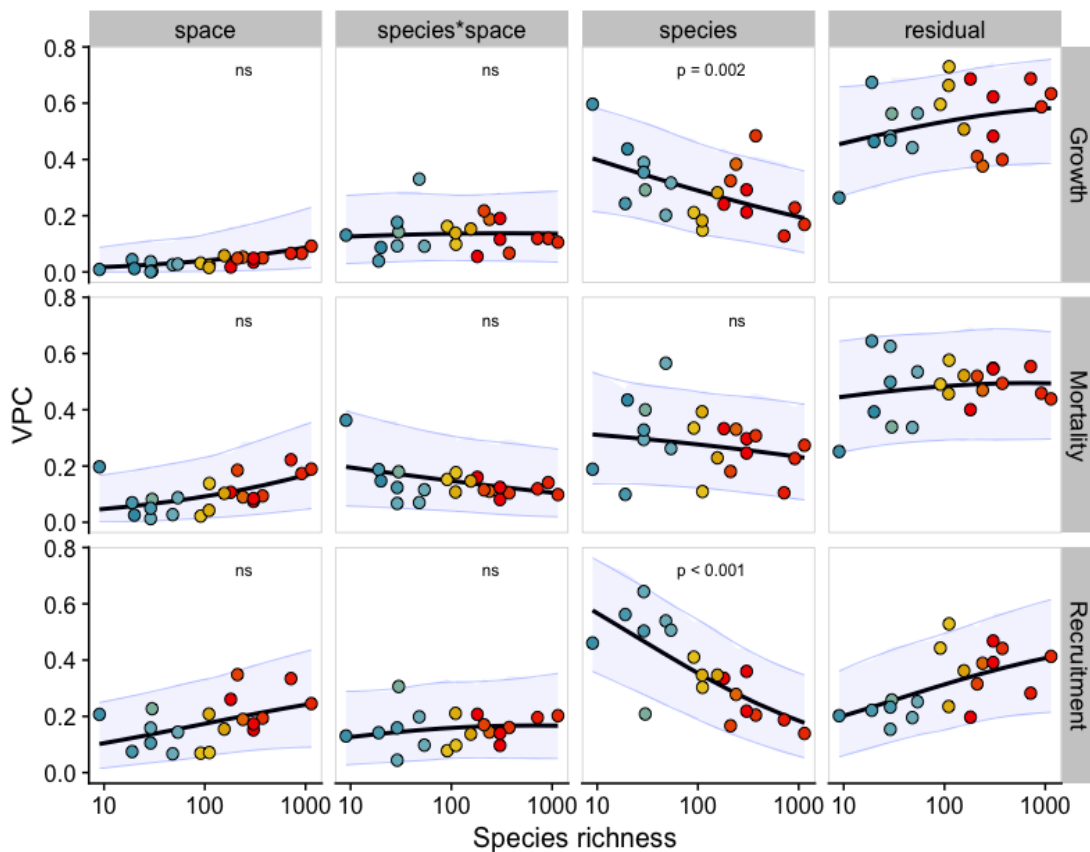
Figure 4. Variance partition coefficients (VPCc) for the organising principles (OPs) species, space, species x space and residual against rarefied species richness. OPs were estimated with the reduced model (eq. 1) without temporal OPs. Black lines are fitted relationships obtained from dirichlet regressions of VPCs against species richness; shaded blue areas are the 95% prediction intervals. p-values are shown only for the significant values after Bonferroni correction (alpha=0.016). Residual VPCs are reference categories and thus were not tested for significance. Each forest plot (dots) is colored by absolute latitude as in Fig. 1a. Species richness on the x-axis is at the logarithmic scale with base 10.

Discussion

Innumerable mechanisms operate and interact in forests and leave fingerprints of their integrated effects in tree vital rates, i.e., growth, survival, and recruitment, which together drive forest dynamics. Here, we used a simple conceptual and statistical framework to identify organising principles (OPs, Table 1) and quantify the associated variability among vital rates for more than 2.9 million trees of approx. 6,500 species in twenty-one forests across the globe. This, in turn, allows a first assessment of the relative importance of mechanisms that are underlying each OP offering a first step in narrowing down which of the mechanisms are critical for structuring global forests. In the following sections, we summarise our most striking findings, discuss some

potentially important mechanisms, and provide recommendations for an agenda of tree vital rates investigations.

Species is a major source of variability in tree vital rates

We found that *species* was the most important OP for all vital rates, explaining on average between 29 and 36% of the demographic variance across the globe (Fig. 1, reduced models). Species in interaction with space added another 13-15% variance explained, meaning that a total of 42- 51% of demographic variation can be partitioned towards species differences and species-specific responses to spatial heterogeneity (Table 1). In contrast, *space* and *time* explained relatively little variability in vital rates (Fig. 1, 2). Our results, therefore, suggest that - at least at the temporal and spatial scales covered by our datasets - spatio-temporally varying factors alone contribute less to demographic variance than evolutionary history and adaptations to the environment. Grouping individuals into species thus creates a globally important cluster of demographic variation that appears consistently most important across a wide range of forests. It is possible that this may be explained by convergent evolution, since there are globally consistent limits to evolutionary adaptations in forest environments, and forest trees can only grow or die within certain limits (Franco & Silvertown, 2004; Takada et al., 2018; Needham et al., 2022).

Our results on the importance of species support numerous ongoing research agendas. Efforts to include more realistic representation of species strategies in global vegetation models appear to be a promising route (Scheiter et al., 2013; Fisher et al., 2018; Anderegg et al., 2022), regardless of whether forest dynamics are studied in local tree neighbourhoods or larger spatial units (Fig. 3). We expect that accounting for species differences can explain up to ~36% of demographic variation, while additionally accounting for small-scale species–environment associations (Messier et al., 2010; Lasky et al., 2014) might further improve this to almost half of the variation explained. More critically, though, our work shows that there are clear limits to the improvement more realistic representations of species can bring. Programs including species strategies typically rely on functional traits (Rubio & Swenson, 2022) or demographic trade-offs (Rüger et al., 2020; Russo et al., 2021), i.e. simplifications that only partly explain more than half of among-species variation (e.g., Visser et al., 2016). Nevertheless, the global importance of species in clustering demographic variance and

its consistency across spatial scales indicates that endeavours seeking to better map species differences may have been underestimated compared to those focussing on spatial and temporal effects.

Temporal variability acts mostly on recruitment and mortality and in interaction with space

In contrast to variability among species, temporal OPs played a minor role for variability in tree vital rates, as time interval alone was responsible for only 3-7% of total variability for plots with sufficient data. Although these data probably have the most comprehensive temporal coverage of large forest areas currently available, our findings might reflect the relatively short time frame (20 to 40 years) and low temporal resolution of the inventory data (approximately 5 years). Nevertheless, variability between census intervals was detected in recruitment and to a lesser degree in mortality, but was rather unimportant for growth (Fig. 2). A possible explanation is that growth rates fluctuate within shorter periods than our 5 year census interval can capture (Dobbertin, 2005), while recruitment and mortality may exhibit several bad or good years in a row (Phillips et al., 2010; Schwartz et al., 2020).

Temporal effects were most important in interaction with space which, for instance, could indicate gap dynamics that jointly affect vital rates of most trees (Kohyama, 1993). This interpretation is consistent with the result that the *space x time* interaction OP was more important for mortality and recruitment than for growth - as mortality in gaps is known to be "spatially contagious" with falling trees killing multiple neighbours (Araujo et al., 2021), and the resulting gaps generally favour recruitment for many species (Brokaw, 1987). Additionally, some of the variability in the *space x time* OP could be the result of climatic events acting differently depending on local conditions, such as droughts that harm trees less in valleys than on ridges (Zuleta et al., 2017).

Our results on temporal OPs support a research agenda that should analyse the importance of climatic and/or temporal effects on vital rates in interaction with spatial effects. Moreover, we advocate for datasets with higher temporal resolution and longer time series, which would allow capturing larger but infrequent disturbances (Šamonil et al., 2013), thereby revealing more of the demographic importance of environmental fluctuations and temporal niches (Fung et al., 2020).

Small spatial grain variability is important

Spatial OPs were important for vital rate variability mostly in interaction with species for growth, and time for mortality and recruitment (Fig. 1 and 2), indicating the importance of spatial niches and patch dynamics (see previous section). Alone, *space* was the least important OP and only created considerable variability in models without time (Fig. 1).

Spatially acting mechanisms were best detected by dividing the plots into quadrats of 5x5 m (Fig. 3), indicating that trees interact and respond to local conditions at scales of a few metres, through mechanisms such as gap dynamics, competition, crown damage, and micro-topography (Schwartz et al., 2020). Further decreasing the spatial grain would then move below the scale of tree crowns, and begin to merely assign quadrats to single trees, here reflected by residual variance. With increasing quadrat size, less variability is explained by spatial mechanisms. Consequently, vital rates become less predictable at larger spatial grain. Nevertheless, even at the largest quadrat size of 100x100 m, spatial OPs still explained a reasonable part of the variability, with the consequence that tree species seem to distinctly respond also to environmental heterogeneity over larger areas (de Knegt et al., 2010), probably due to topography or soil nutrients (Russo et al., 2005, 2008; Zuleta et al., 2018).

Large proportion of unexplained variability in vital rates

Residual variance was consistently the dominant component of the vital rate VPCs across sites and in the temporal and spatial analyses. In multilevel models, residual variance represents the variance in the response that cannot be attributed to any of the grouping factors (here, the OPs). On one hand, this result encourages more detailed models that might include covariates that 'explain' it. For instance, both growth and mortality are known to differ across ontogeny, and thus the inclusion of tree size (as dbh) as a covariate would most likely explain some of the residual variance (Dale et al., 1985; Hülsmann et al., 2018). Moreover, functional traits at the individual level (Su et al., 2020) and structures that explicitly deal with spatial (Wiegand et al., 2017) and temporal autocorrelation may explain additional differences in individual vital rates. On the other hand, there are intrinsic limits to what can be explained by even the most detailed models, as the residual variance also includes inherent noise which by definition is unexplainable. The noise includes misattribution of species, mapping error

or measurement error (Detto et al., 2019) and chaotic behaviour known to exist in many biological systems (Beninca et al., 2015; Pearce et al., 2020). Nevertheless, individual (i.e., 'tree-level') covariates represent a promising avenue for extending our framework.

Globally, variability among species declines with species richness

Across plots, increasing species richness resulted in decreasing relative importance of the *species* OP in growth and recruitment (Fig. 4). This trend was robust to the most probable source of bias, i.e. differences in species rarity, and thus we consider that the decreasing relative importance of these OPs with species richness reflects a true ecological signal. Moreover, the decrease in the *species* VPC was determined by a decrease in the respective variance estimates, and not by an increase of variances related to the other OPs (Fig. S5.4). Similarly, Condit *et al.* (2006) found across ten tropical forests (seven in common with this study) that the range of mortality and growth rates among species decreased with higher species richness.

These results underpin that - in contrast to expectations of niche theory -the most diverse forests feature the lowest interspecific variation in demographic rates. Following the rationale of niche theory, diverse forests should have more demographic niches than low-diversity forests, as more niches allow more species to have equivalent fitness thus favouring species coexistence (Chesson, 2000). The lack of evidence for wider demographic ranges in species-rich forests (this study, Condit et al., 2006; Clark, 2010) suggests that demographic niches play a minor role for large-scale diversity patterns, hinting towards more neutral dynamics. However, coexistence is inherently high dimensional, and comparing mean species values across low dimensions (a few vital rates) only partly represents the full niche space (Clark et al. 2004). In summary, species in highly diverse forests present redundant vital rates that do not add to the diversity of demographic types, highlighting the challenges of studying and predicting changes in hyperdiverse systems.

Conclusions

As the mechanisms that influence vital rates can be grouped by the dimensions at which they operate and interact, patterns of how variance is partitioned along key dimensions can reveal how important various biotic and abiotic mechanisms are in influencing tree demography and hence forest dynamics. Here, we have shown that

variance partitioning of vital rates among key ecological dimensions, i.e. species, space and time, has a high potential to provide a first step in identifying the structuring processes of global forest dynamics. The proposed framework highlights the most promising avenues for future research both in terms of understanding the relative contributions of mechanisms to forest demography and diversity, and for predicting forest ecosystems. We hope future studies may benefit from using this approach as a conceptual and modelling baseline to narrow down which of the mechanisms are critical for structuring global forests.

Acknowledgments

Author contributions

MSL, SM, MV, LH, PIP, AOA conceived the ideas and conceptualization of the study,

MSL, SM, MV, LH conceived and designed the analyses,

MSL, LH curated the data,

MSL performed the analyses,

MSL, SM, MV, LH wrote the initial draft.

MSL, SM, MV, LH, PIP, SD, HD wrote reviewed versions of the draft.

SA, KAA, NA, NAB, WYB, NC, CHCY, YYC, GC, KC, AD, SE, CENE, GG, IAUNG, CVSG, RH, WHH, AI, DJJ, DK, KK, YTL, JAL, JRM, YM, WJMS, MBM, MN, AN,

GP, ReP, RoP, RPP, PS, IFS, ST, DT, JT, MU, AW, JZ, DZ contributed with data and provided site-specific information.

All authors contributed to the final version of the manuscript.

Competing interests

The authors declare no competing interests.

Data and code availability

The forest data that support the findings of this study are available from the ForestGEO network. For some of the sites, the data is publicly available at https://forestgeo.si.edu/explore-data. Restrictions apply, however, to the availability of the data from other sites, which were used under license for the current study, and so are not publicly available. Data are however available from the authors upon reasonable request and with permission of the principal investigators of the ForestGEO sites. Nevertheless, we provide one example of data cleaning workflow together with the code for all results and analyses on Zenodo repository (Leite, 2022) with the published version of the GitHub repository (https://doi.org/10.5281/zenodo.7250584).

References

Adler, P.B., Ellner, S.P. & Levine, J.M. (2010) Coexistence of perennial plants: an embarrassment of niches. *Ecology Letters*, **13**, 1019–1029.

Adler, P.B. & Levine, J.M. (2007) Contrasting relationships between precipitation and species richness in space and time. *Oikos*, **116**, 221–232.

Adler, P.B., Salguero-Gómez, R., Compagnoni, A., Hsu, J.S., Ray-Mukherjee, J., Mbeau-Ache, C. & Franco, M. (2014) Functional traits explain variation in plant life history strategies. *Proceedings of the National Academy of Sciences*, **111**, 740–745.

Anderegg, L.D.L., Griffith, D.M., Cavender-Bares, J., Riley, W.J., Berry, J.A., Dawson, T.E. & Still, C.J. (2022) Representing plant diversity in land models: An evolutionary approach to make "Functional Types" more functional. *Global Change Biology*, **n/a**.

Anderson-Teixeira, K.J., Herrmann, V., Morgan, R.B., Bond-Lamberty, B., Cook-Patton, S.C., Ferson, A.E., Muller-Landau, H.C. & Wang, M.M.H. (2021) Carbon cycling in mature and regrowth forests globally. *Environmental Research Letters*, **16**, 053009.

Araujo, R.F., Grubinger, S., Celes, C.H.S., Negrón-Juárez, R.I., Garcia, M., Dandois, J.P. & Muller-Landau, H.C. (2021) Strong temporal variation in treefall and branchfall rates in a tropical forest is related to extreme rainfall: results from 5 years of monthly drone data for a 50 ha plot. *Biogeosciences*, **18**, 6517–6531.

Arellano, G. (2019) Calculation of narrower confidence intervals for tree mortality rates when we know nothing but the location of the death/survival events. *Ecology and Evolution*, **9**, 9644–9653.

Benincà, E., Ballantine, B., Ellner, S.P. & Huisman, J. (2015) Species fluctuations sustained by a cyclic succession at the edge of chaos. *Proceedings of the National Academy of Sciences*, **112**, 6389–6394.

Bennett, A.C., McDowell, N.G., Allen, C.D. & Anderson-Teixeira, K.J. (2015) Larger trees suffer most during drought in forests worldwide. *Nature Plants*, **1**, 1–5.

Bonan, G.B. (2008) Forests and Climate Change: Forcings, Feedbacks, and the Climate Benefits of Forests. *Science*, **320**, 1444–1449.

Brokaw, N.V.L. (1987) Gap-Phase Regeneration of Three Pioneer Tree Species in a Tropical Forest. *The Journal of Ecology*, **75**, 9.

Browne, W.J., Subramanian, S.V., Jones, K., & H. Goldstein (2005) Variance partitioning in multilevel logistic models that exhibit overdispersion. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, **168**, 599–613.

Bürkner, P.-C. (2017) brms: An R Package for Bayesian Multilevel Models Using Stan. *Journal of Statistical Software*, **80**, 1–28.

Burns, J.H. & Strauss, S.Y. (2012) Effects of competition on phylogenetic signal and phenotypic plasticity in plant functional traits. *Ecology*, **93**, S126–S137.

Camac, J.S., Condit, R., FitzJohn, R.G., McCalman, L., Steinberg, D., Westoby, M., Wright, S.J. & Falster, D.S. (2018) Partitioning mortality into growth-dependent and growth-independent hazards across 203 tropical tree species. *Proceedings of the National Academy of Sciences*, **115**, 12459–12464.

Canadell, J.G. & Raupach, M.R. (2008) Managing Forests for Climate Change Mitigation. *Science*, **320**, 1456–1457.

Chen, Y., Uriarte, M., Wright, S.J. & Yu, S. (2019) Effects of neighborhood trait composition on tree survival differ between drought and postdrought periods. *Ecology*, **100**, e02766.

Chesson, P. (2000) Mechanisms of maintenance of species diversity. *Annual review of Ecology and Systematics*, **31**, 343–366.

Chisholm, R.A., Muller-Landau, H.C., Rahman, K.A., Bebber, D.P., Bin, Y., Bohlman, S.A., Bourg, N.A., Brinks, J., Bunyavejchewin, S., Butt, N., Cao, H., Cao, M., Cárdenas, D., Chang, L.-W., Chiang, J.-M., Chuyong, G., Condit, R., Dattaraja, H.S., Davies, S., Duque, A., Fletcher, C., Gunatilleke, N., Gunatilleke, S., Hao, Z., Harrison, R.D., Howe, R., Hsieh, C.-F., Hubbell, S.P., Itoh, A., Kenfack, D., Kiratiprayoon, S., Larson, A.J., Lian, J., Lin, D., Liu, H., Lutz, J.A., Ma, K., Malhi, Y., McMahon, S., McShea, W., Meegaskumbura, M., Razman, S.M., Morecroft, M.D., Nytch, C.J., Oliveira, A., Parker, G.G., Pulla, S., Punchi-Manage, R., Romero-Saltos, H., Sang, W., Schurman, J., Su, S.-H., Sukumar, R., Sun, I.-F., Suresh, H.S., Tan, S., Thomas, D., Thomas, S., Thompson, J., Valencia, R., Wolf, A., Yap, S., Ye, W., Yuan, Z. & Zimmerman, J.K. (2013) Scale-dependent relationships between tree species richness and ecosystem function in forests. *Journal of Ecology*, **101**, 1214–1224.

Clark, A.T., Detto, M., Muller-Landau, H.C., Schnitzer, S.A., Wright, S.J., Condit, R. & Hubbell, S.P. (2018) Functional traits of tropical trees and lianas explain spatial structure across multiple scales. *Journal of Ecology*, **106**, 795–806.

Clark, J.S. (2010) Individuals and the Variation Needed for High Species Diversity in Forest Trees. *Science*, **327**, 1129–1132.

Condit, R., Aguilar, S., Hernandez, A., Perez, R., Lao, S., Angehr, G., Hubbell, S.P. & Foster, R.B. (2004) Tropical forest dynamics across a rainfall gradient and the impact of an El Niño dry season. *Journal of Tropical Ecology*, **20**, 51–72.

Condit, R., Ashton, P., Bunyavejchewin, S., Dattaraja, H.S., Davies, S., Esufali, S., Ewango, C., Foster, R., Gunatilleke, I. a. U.N., Gunatilleke, C.V.S., Hall, P., Harms, K.E., Hart, T., Hernandez, C., Hubbell, S., Itoh, A., Kiratiprayoon, S., LaFrankie, J., Lao, S.L. de, Makana, J.-R., Noor, M.N.S., Kassim, A.R., Russo, S., Sukumar, R., Samper, C., Suresh, H.S., Tan, S., Thomas, S., Valencia, R., Vallejo, M., Villa, G. & Zillio, T. (2006) The Importance of

Demographic Niches to Tree Diversity. *Science*, **313**, 98–101.

Dale, V.H., Doyle, T.W. & Shugart, H.H. (1985) A comparison of tree growth models. *Ecological Modelling*, **29**, 145–169.

Davies, S.J., Abiem, I., Abu Salim, K., Aguilar, S., Allen, D., Alonso, A., Anderson-Teixeira, K., Andrade, A., Arellano, G., Ashton, P.S., Baker, P.J., Baker, M.E., Baltzer, J.L., Basset, Y., Bissiengou, P., Bohlman, S., Bourg, N.A., Brockelman, W.Y., Bunyavejchewin, S., Burslem, D.F.R.P., Cao, M., Cárdenas, D., Chang, L.-W., Chang-Yang, C.-H., Chao, K.-J., Chao, W.-C., Chapman, H., Chen, Y.-Y., Chisholm, R.A., Chu, C., Chuyong, G., Clay, K., Comita, L.S., Condit, R., Cordell, S., Dattaraja, H.S., de Oliveira, A.A., den Ouden, J., Detto, M., Dick, C., Du, X., Duque, Á., Ediriweera, S., Ellis, E.C., Obiang, N.L.E., Esufali, S., Ewango, C.E.N., Fernando, E.S., Filip, J., Fischer, G.A., Foster, R., Giambelluca, T., Giardina, C., Gilbert, G.S., Gonzalez-Akre, E., Gunatilleke, I.A.U.N., Gunatilleke, C.V.S., Hao, Z., Hau, B.C.H., He, F., Ni, H., Howe, R.W., Hubbell, S.P., Huth, A., Inman-Narahari, F., Itoh, A., Janík, D., Jansen, P.A., Jiang, M., Johnson, D.J., Jones, F.A., Kanzaki, M., Kenfack, D., Kiratiprayoon, S., Král, K., Krizel, L., Lao, S., Larson, A.J., Li, Y., Li, X., Litton, C.M., Liu, Y., Liu, S., Lum, S.K.Y., Luskin, M.S., Lutz, J.A., Luu, H.T., Ma, K., Makana, J.-R., Malhi, Y., Martin, A., McCarthy, C., McMahon, S.M., McShea, W.J., Memiaghe, H., Mi, X., Mitre, D., Mohamad, M., Monks, L., Muller-Landau, H.C., Musili, P.M., Myers, J.A., Nathalang, A., Ngo, K.M., Norden, N., Novotny, V., O'Brien, M.J., Orwig, D., Ostertag, R., Papathanassiou, K., Parker, G.G., Pérez, R., Perfecto, I., Phillips, R.P., Pongpattananurak, N., Pretzsch, H., Ren, H., Reynolds, G., Rodriguez, L.J., Russo, S.E., Sack, L., Sang, W., Shue, J., Singh, A., Song, G.-Z.M., Sukumar, R., Sun, I.-F., Suresh, H.S., Swenson, N.G., Tan, S., Thomas, S.C., Thomas, D., Thompson, J., Turner, B.L., Uowolo, A., Uriarte, M., Valencia, R., Vandermeer, J., Vicentini, A., Visser, M., Vrska, T., Wang, X., Wang, X., Weiblen, G.D., Whitfeld, T.J.S., Wolf, A., Wright, S.J., Xu, H., Yao, T.L., Yap, S.L., Ye, W., Yu, M., Zhang, M., Zhu, D., Zhu, L., Zimmerman, J.K. & Zuleta, D. (2021) ForestGEO: Understanding forest diversity and dynamics through a global observatory network. *Biological Conservation*, **253**, 108907.

Detto, M., Visser, M.D., Wright, S.J. & Pacala, S.W. (2019) Bias in the detection of negative density dependence in plant communities. *Ecology Letters*.

Díaz, S., Kattge, J., Cornelissen, J.H.C., Wright, I.J., Lavorel, S., Dray, S., Reu, B., Kleyer, M., Wirth, C., Prentice, I.C., Garnier, E., Bönisch, G., Westoby, M., Poorter, H., Reich, P.B., Moles, A.T., Dickie, J., Gillison, A.N., Zanne, A.E., Chave, J., Wright, S.J., Sheremet'ev, S.N., Jactel, H., Baraloto, C., Cerabolini, B., Pierce, S., Shipley, B., Kirkup, D., Casanoves, F., Joswig, J.S., Günther, A., Falczuk, V., Rüger, N., Mahecha, M.D. & Gorné, L.D. (2016) The global spectrum of plant form and function. *Nature*, **529**, 167.

Diez, J.M., Giladi, I., Warren, R. & Pulliam, H.R. (2014) Probabilistic and spatially variable niches inferred from demography. *Journal of Ecology*, **102**, 544–554.

Dobbertin, M. (2005) Tree growth as indicator of tree vitality and of tree reaction to environmental stress: a review. *European Journal of Forest Research*, **124**, 319–333.

Douma, J.C. & Weedon, J.T. (2019) Analysing continuous proportions in ecology and evolution: A practical introduction to beta and Dirichlet regression. *Methods in Ecology and Evolution*, **10**, 1412–1430.

Fisher, R.A., Koven, C.D., Anderegg, W.R.L., Christoffersen, B.O., Dietze, M.C., Farrior, C.E., Holm, J.A., Hurtt, G.C., Knox, R.G., Lawrence, P.J., Lichstein, J.W., Longo, M., Matheny, A.M., Medvigy, D., Muller-Landau, H.C., Powell, T.L., Serbin, S.P., Sato, H., Shuman, J.K., Smith, B., Trugman, A.T., Viskari, T., Verbeeck, H., Weng, E., Xu, C., Xu, X., Zhang, T. & Moorcroft, P.R. (2018) Vegetation demographics in Earth System Models: A review of progress and priorities. *Global Change Biology*, **24**, 35–54.

Fortin, M., Bédard, S., DeBlois, J. & Meunier, S. (2008) Predicting individual tree mortality in northern hardwood stands under uneven-aged management in southern Québec, Canada. *Annals of Forest Science*, **65**, 205–205.

Franco, M. & Silvertown, J. (2004) A Comparative Demography of Plants Based Upon Elasticities of Vital Rates. *Ecology*, **85**, 531–538.

Fung, T., Chisholm, R.A., Anderson-Teixeira, K., Bourg, N., Brockelman, W.Y., Bunyavejchewin, S., Chang-Yang, C.-H., Chitra-Tarak, R., Chuyong, G., Condit, R., Dattaraja, H.S., Davies, S.J., Ewango, C.E.N., Fewless, G., Fletcher, C., Gunatilleke, C.V.S., Gunatilleke, I.A.U.N., Hao, Z., Hogan, J.A., Howe, R., Hsieh, C.-F., Kenfack, D., Lin, Y., Ma, K., Makana, J.-R., McMahon, S., McShea, W.J., Mi, X., Nathalang, A., Ong, P.S., Parker, G., Rau, E.-P., Shue, J., Su, S.-H., Sukumar, R., Sun, I.-F., Suresh, H.S., Tan, S., Thomas, D., Thompson, J., Valencia, R., Vallejo, M.I., Wang, X., Wang, Y., Wijekoon, P., Wolf, A., Yap, S. & Zimmerman, J. (2020) Temporal population variability in local forest communities has mixed effects on tree species richness across a latitudinal gradient. *Ecology Letters*, **23**, 160–171.

Gibert, A., Gray, E.F., Westoby, M., Wright, I.J. & Falster, D.S. (2016) On the link between functional traits and growth rate: meta-analysis shows effects change with plant size, as predicted. *Journal of Ecology*, **104**, 1488–1503.

Griffith, A.B., Salguero-Gómez, R., Merow, C. & McMahon, S. (2016) Demography beyond the population. *Journal of Ecology*, **104**, 271–280.

Hülsmann, L., Bugmann, H., Cailleret, M. & Brang, P. (2018) How to kill a tree: empirical mortality models for 18 species and their performance in a dynamic forest model. *Ecological Applications*, **28**, 522–540.

Iida, Y., Kohyama, T.S., Swenson, N.G., Su, S.-H., Chen, C.-T., Chiang, J.-M. & Sun, I.-F. (2014) Linking functional traits and demographic rates in a subtropical tree community: the importance of size dependency. *Journal of Ecology*, **102**, 641–650.

Johnson, D.J., Needham, J., Xu, C., Massoud, E.C., Davies, S.J., Anderson-Teixeira, K.J., Bunyavejchewin, S., Chambers, J.Q., Chang-Yang, C.-H., Chiang, J.-M., Chuyong, G.B., Condit, R., Cordell, S., Fletcher, C., Giardina, C.P., Giambelluca, T.W., Gunatilleke, N., Gunatilleke, S., Hsieh, C.-F., Hubbell, S., Inman-Narahari, F., Kassim, A.R., Katabuchi, M., Kenfack, D., Litton, C.M., Lum, S., Mohamad, M., Nasardin, M., Ong, P.S., Ostertag, R., Sack, L., Swenson, N.G., Sun, I.F., Tan, S., Thomas, D.W., Thompson, J., Umaña, M.N., Uriarte, M., Valencia, R., Yap, S., Zimmerman, J., McDowell, N.G. & McMahon, S.M. (2018) Climate sensitive size-dependent survival in tropical trees. *Nature Ecology & Evolution*, **2**, 1436–1442.

Jurinitz, C.F., de Oliveira, A.A. & Bruna, E.M. (2013) Abiotic and Biotic Influences on Early-Stage Survival in Two Shade-Tolerant Tree Species in Brazil's Atlantic Forest. *Biotropica*, **45**, 728–736.

Keil, P. & Chase, J.M. (2019) Global patterns and drivers of tree diversity integrated across a continuum of spatial grains. *Nature Ecology & Evolution*, **3**, 390–399.

de Knegt, H.J., van Langevelde, F., Coughenour, M.B., Skidmore, A.K., de Boer, W.F., Heitkönig, I.M.A., Knox, N.M., Slotow, R., van der Waal, C. & Prins, H.H.T. (2010) Spatial autocorrelation and the scaling of species–environment relationships. *Ecology*, **91**, 2455–2465.

Kohyama, T. (1993) Size-Structured Tree Populations in Gap-Dynamic Forest--The Forest Architecture Hypothesis for the Stable Coexistence of Species. *The Journal of Ecology*, **81**, 131.

Kohyama, T.S., Kohyama, T.I. & Sheil, D. (2018) Definition and estimation of vital rates from repeated censuses: Choices, comparisons and bias corrections focusing on trees. *Methods in Ecology and Evolution*, **9**, 809–821.

Kupers, S.J., Wirth, C., Engelbrecht, B.M.J., Hernández, A., Condit, R., Wright, S.J. & Rüger, N. (2019) Performance of tropical forest seedlings under shade and drought: an interspecific trade-off in demographic responses. *Scientific Reports*, **9**, 18784.

Lasky, J.R., Uriarte, M., Boukili, V.K. & Chazdon, R.L. (2014) Trait-mediated assembly processes predict successional changes in community diversity of tropical forests. *Proceedings of the National Academy of Sciences*, **111**, 5616–5621.

Leite, M. de S. (2022) Data and code from: Major axes of variation in tree demography across global

forests. https://doi.org/10.5281/zenodo.7250584.

Lévesque, M., Walthert, L. & Weber, P. (2016) Soil nutrients influence growth response of temperate tree species to drought. *Journal of Ecology*, **104**, 377–387.

Li, Y., Mayfield, M.M., Wang, B., Xiao, J., Kral, K., Janik, D., Holik, J. & Chu, C. (2020) Beyond direct neighbourhood effects: higher-order interactions improve modelling and predicting tree survival and growth. *National Science Review*.

Lu, R., Qiao, Y., Wang, J., Zhu, C., Cui, E., Xu, X., He, Y., Zhao, Z., Du, Y., Yan, L., Shen, G., Yang, Q., Wang, X. & Xia, J. (2021) The U-shaped pattern of size-dependent mortality and its correlated factors in a subtropical monsoon evergreen forest. *Journal of Ecology*, **109**, 2421–2433.

Maier, M.J. (2021) DirichletReg: Dirichlet Regression.

Martínez-Ramos, M., Gallego-Mahecha, M. del M., Valverde, T., Vega, E. & Bongers, F. (2021) Demographic differentiation among pioneer tree species during secondary succession of a Neotropical rainforest. *Journal of Ecology*, **109**, 3572–3586.

Martini, F., Chang-Yang, C.-H. & Sun, I.-F. (2022) Variation in biotic interactions mediates the effects of masting and rainfall fluctuations on seedling demography in a subtropical rainforest. *Journal of Ecology*, **n/a**.

McDowell, N.G., Allen, C.D., Anderson-Teixeira, K., Aukema, B.H., Bond-Lamberty, B., Chini, L., Clark, J.S., Dietze, M., Grossiord, C., Hanbury-Brown, A., Hurtt, G.C., Jackson, R.B., Johnson, D.J., Kueppers, L., Lichstein, J.W., Ogle, K., Poulter, B., Pugh, T.A.M., Seidl, R., Turner, M.G., Uriarte, M., Walker, A.P. & Xu, C. (2020) Pervasive shifts in forest dynamics in a changing world. *Science*, **368**.

McMahon, S.M. & Diez, J.M. (2007) Scales of association: hierarchical linear models and the measurement of ecological systems. *Ecology Letters*, **10**, 437–452.

Nakagawa, S., Johnson, P.C.D. & Schielzeth, H. (2017) The coefficient of determination R2 and intra-class correlation coefficient from generalized linear mixed-effects models revisited and expanded. *Journal of The Royal Society Interface*, **14**, 20170213.

Needham, J.F., Johnson, D.J., Anderson-Teixeira, K.J., Bourg, N., Bunyavejchewin, S., Butt, N., Cao, M., Cárdenas, D., Chang-Yang, C.-H., Chen, Y.-Y., Chuyong, G., Dattaraja, H.S., Davies, S.J., Duque, A., Ewango, C.E.N., Fernando, E.S., Fisher, R., Fletcher, C.D., Foster, R., Hao, Z., Hart, T., Hsieh, C.-F., Hubbell, S.P., Itoh, A., Kenfack, D., Koven, C.D., Larson, A.J., Lutz, J.A., McShea, W., Makana, J.-R., Malhi, Y., Marthews, T., Bt. Mohamad, M., Morecroft, M.D., Norden, N., Parker, G., Shringi, A., Sukumar, R., Suresh, H.S., Sun, I.-F., Tan, S., Thomas, D.W., Thompson, J., Uriarte, M., Valencia, R., Yao, T.L., Yap, S.L., Yuan, Z., Yuehua, H., Zimmerman, J.K., Zuleta, D. & McMahon, S.M. (2022) Demographic composition, not demographic diversity, predicts biomass and turnover across temperate and tropical forests. *Global Change Biology*, **n/a**.

Needham, J.F., Johnson, D.J., Chang-Yang, C.-H., Anderson-Teixeira, K.J., Bourg, N.A., Bunyavejchewin, S., Butt, N., Cao, M., Càrdenas, D., Chen, Y.-Y., Chuyong, G., Davies, S.J., Duque, A., Ewango, C.E.N., Fernando, E.S., Fisher, R.A., Fletcher, C.D., Foster, R., Hao, Z., Hart, T., Hsieh, C.-F., Hubbell, S.P., Kenfack, D., Koven, C.D., Larson, A.J., Lutz, J.A., Makana, J.-R., Malhi, Y., Marthews, T.R., McShea, W., Mohamad, M.B., Morecroft, M.D., Norden, N., Parker, G., Shringi, A., Sukumar, R., Sun, I.-F., Suresh, H., Tan, S., Thomas, D.W., Thompson, J., Uriarte, M., Valencia, R., Yao, T.-L., Yap, S.L., Yuan, Z., Yuehua, H., Zimmerman, J.K., Zuleta, D. & McMahon, S.M. (2020) Demographic functional types determine carbon dynamics of global forests. 13.

Negrón, J.F., McMillin, J.D., Anhold, J.A. & Coulson, D. (2009) Bark beetle-caused mortality in a drought-affected ponderosa pine landscape in Arizona, USA. *Forest Ecology and Management*, **257**, 1353–1362.

Oberpriller, J., de Souza Leite, M. & Pichler, M. (2022) Fixed or random? On the reliability of mixed-effects models for a small number of levels in grouping variables. *Ecology and*

Evolution, **12**, e9062.

Ovenden, T.S., Perks, M.P., Clarke, T., Mencuccini, M. & Jump, A.S. (2021) Life after recovery: Increased resolution of forest resilience assessment sheds new light on post-drought compensatory growth and recovery dynamics. *Journal of Ecology*, 1365-2745.13576.

Pan, Y., Birdsey, R.A., Fang, J., Houghton, R., Kauppi, P.E., Kurz, W.A., Phillips, O.L., Shvidenko, A., Lewis, S.L., Canadell, J.G., Ciais, P., Jackson, R.B., Pacala, S.W., McGuire, A.D., Piao, S., Rautiainen, A., Sitch, S. & Hayes, D. (2011) A Large and Persistent Carbon Sink in the World's Forests. *Science*, **333**, 988–993.

Pearce, M.T., Agarwala, A. & Fisher, D.S. (2020) Stabilization of extensive fine-scale diversity by ecologically driven spatiotemporal chaos. *Proceedings of the National Academy of Sciences*, **117**, 14572–14583.

Phillips, O.L., van der Heijden, G., Lewis, S.L., López-González, G., Aragão, L.E.O.C., Lloyd, J., Malhi, Y., Monteagudo, A., Almeida, S., Dávila, E.A., Amaral, I., Andelman, S., Andrade, A., Arroyo, L., Aymard, G., Baker, T.R., Blanc, L., Bonal, D., de Oliveira, Á.C.A., Chao, K.-J., Cardozo, N.D., da Costa, L., Feldpausch, T.R., Fisher, J.B., Fyllas, N.M., Freitas, M.A., Galbraith, D., Gloor, E., Higuchi, N., Honorio, E., Jiménez, E., Keeling, H., Killeen, T.J., Lovett, J.C., Meir, P., Mendoza, C., Morel, A., Vargas, P.N., Patiño, S., Peh, K.S.-H., Cruz, A.P., Prieto, A., Quesada, C.A., Ramírez, F., Ramírez, H., Rudas, A., Salamão, R., Schwarz, M., Silva, J., Silveira, M., Ferry Slik, J.W., Sonké, B., Thomas, A.S., Stropp, J., Taplin, J.R.D., Vásquez, R. & Vilanova, E. (2010) Drought–mortality relationships for tropical forests. *New Phytologist*, **187**, 631–646.

Pillay, R., Venter, M., Aragon-Osejo, J., González-del-Pliego, P., Hansen, A.J., Watson, J.E. & Venter, O. (2022) Tropical forests are home to over half of the world's vertebrate species. *Frontiers in Ecology and the Environment*, **20**, 10–15.

Poorter, L., Wright, S.J., Paz, H., Ackerly, D.D., Condit, R., Ibarra-Manríquez, G., Harms, K.E., Licona, J.C., Martínez-Ramos, M., Mazer, S.J., Muller-Landau, H.C., Peña-Claros, M., Webb, C.O. & Wright, I.J. (2008) ARE FUNCTIONAL TRAITS GOOD PREDICTORS OF DEMOGRAPHIC RATES? EVIDENCE FROM FIVE NEOTROPICAL FORESTS. *Ecology*, **89**, 1908–1920.

Pugh, T.A.M., Lindeskog, M., Smith, B., Poulter, B., Arneth, A., Haverd, V. & Calle, L. (2019) Role of forest regrowth in global carbon sink dynamics. *Proceedings of the National Academy of Sciences*, **116**, 4382–4387.

R Core Team (2022) R: A language and environment for statistical computing.

Rubio, V.E. & Swenson, N.G. (2022) Functional groups, determinism and the dynamics of a tropical forest. *Journal of Ecology*, **110**, 185–196.

Rüger, N., Comita, L.S., Condit, R., Purves, D., Rosenbaum, B., Visser, M.D., Wright, S.J. & Wirth, C. (2018) Beyond the fast–slow continuum: demographic dimensions structuring a tropical tree community. *Ecology Letters*, **21**, 1075–1084.

Rüger, N., Condit, R., Dent, D.H., DeWalt, S.J., Hubbell, S.P., Lichstein, J.W., Lopez, O.R., Wirth, C. & Farrior, C.E. (2020) Demographic trade-offs predict tropical forest dynamics. *Science*, **368**, 165–168.

Russo, S.E., Brown, P., Tan, S. & Davies, S.J. (2008) Interspecific demographic trade-offs and soil-related habitat associations of tree species along resource gradients. *Journal of Ecology*, **96**, 192–203.

Russo, S.E., Davies, S.J., King, D.A. & Tan, S. (2005) Soil-related performance variation and distributions of tree species in a Bornean rain forest. *Journal of Ecology*, **93**, 879–889.

Russo, S.E., McMahon, S.M., Detto, M., Ledder, G., Wright, S.J., Condit, R.S., Davies, S.J., Ashton, P.S., Bunyavejchewin, S., Chang-Yang, C.-H., Ediriweera, S., Ewango, C.E.N., Fletcher, C., Foster, R.B., Gunatilleke, C.V.S., Gunatilleke, I.A.U.N., Hart, T., Hsieh, C.-F., Hubbell, S.P., Itoh, A., Kassim, A.R., Leong, Y.T., Lin, Y.C., Makana, J.-R., Mohamad, M.B., Ong, P., Sugiyama, A., Sun, I.-F., Tan, S., Thompson, J., Yamakura, T., Yap, S.L. & Zimmerman, J.K.

(2021) The interspecific growth–mortality trade-off is not a general framework for tropical forest community structure. *Nature Ecology & Evolution*, **5**, 174–183.

Šamonil, P., Doleželová, P., Vašíčková, I., Adam, D., Valtera, M., Král, K., Janík, D. & Šebková, B. (2013) Individual-based approach to the detection of disturbance history through spatial scales in a natural beech-dominated forest. *Journal of Vegetation Science*, **24**, 1167–1184.

Scheiter, S., Langan, L. & Higgins, S.I. (2013) Next-generation dynamic global vegetation models: learning from community ecology. *New Phytologist*, **198**, 957–969.

Schwartz, N.B., Feng, X., Muscarella, R., Swenson, N.G., Umaña, M.N., Zimmerman, J.K. & Uriarte, M. (2020) Topography and Traits Modulate Tree Performance and Drought Response in a Tropical Forest. *Frontiers in Forests and Global Change*, **3**, 136.

Su, S.-H., Guan, B.T., Chang-Yang, C.-H., Sun, I.-F., Wang, H.-H. & Hsieh, C.-F. (2020) Multi-stemming and size enhance survival of dominant tree species in a frequently typhoon-disturbed forest. *Journal of Vegetation Science*, **31**, 429–439.

Swenson, N.G., Worthy, S.J., Eubanks, D., Iida, Y., Monks, L., Petprakob, K., Rubio, V.E., Staiger, K. & Zambrano, J. (2020) A Reframing of Trait–Demographic Rate Analyses for Ecology and Evolutionary Biology. *International Journal of Plant Sciences*, **181**, 33–43.

Takada, T., Kawai, Y. & Salguero-Gómez, R. (2018) A cautionary note on elasticity analyses in a ternary plot using randomly generated population matrices. *Population Ecology*, **60**, 37–47.

Uriarte, M., Thompson, J. & Zimmerman, J.K. (2019) Hurricane María tripled stem breaks and doubled tree mortality relative to other major storms. *Nature Communications*, **10**, 1362.

Visser, M.D., Bruijning, M., Wright, S.J., Muller-Landau, H.C., Jongejans, E., Comita, L.S. & de Kroon, H. (2016) Functional traits as predictors of vital rates across the life cycle of tropical trees. *Functional Ecology*, **30**, 168–180.

Weng, E.S., Malyshev, S., Lichstein, J.W., Farrior, C.E., Dybzinski, R., Zhang, T., Shevliakova, E. & Pacala, S.W. (2015) Scaling from individual trees to forests in an Earth system modeling framework using a mathematically tractable model of height-structured competition. *Biogeosciences*, **12**, 2655–2694.

Wiegand, T., Uriarte, M., Kraft, N.J.B., Shen, G., Wang, X. & He, F. (2017) Spatially Explicit Metrics of Species Diversity, Functional Diversity, and Phylogenetic Diversity: Insights into Plant Community Assembly Processes. *Annual Review of Ecology, Evolution, and Systematics*, **48**, 329–351.

Wiens, J.J. & Donoghue, M.J. (2004) Historical biogeography, ecology and species richness. *Trends in Ecology & Evolution*, **19**, 639–644.

Wright, S.J., Muller-Landau, H.C., Condit, R. & Hubbell, S.P. (2003) Gap-Dependent Recruitment, Realized Vital Rates, and Size Distributions of Tropical Trees. *Ecology*, **84**, 3174–3185.

Zuleta, D., Duque, A., Cardenas, D., Muller-Landau, H.C. & Davies, S.J. (2017) Drought-induced mortality patterns and rapid biomass recovery in a terra firme forest in the Colombian Amazon. *Ecology*, **98**, 2538–2546.

Zuleta, D., Russo, S.E., Barona, A., Barreto-Silva, J.S., Cardenas, D., Castaño, N., Davies, S.J., Detto, M., Sua, S., Turner, B.L. & Duque, A. (2018) Importance of topography for tree species habitat distributions in a terra firme forest in the Colombian Amazon. *Plant and Soil*, 1–17.

Conclusão geral

*"Caminho se conhece andando
Então de vez em quando
é bom se perder
Perdido fica perguntando
vai só procurando
e acha sem saber"*

Chico César

Nesta tese, pude desenvolver dois estudos ecológicos que fazem uso da informação contida nos componentes de variação em comunidades ecológicas, captados por modelos mistos, para responder a perguntas bem distintas, tanto em termos de padrões observados e processos interpretados.

No Capítulo 1, estudamos ecologia de comunidades em paisagens altamente fragmentadas por atividades humanas agrícolas. Buscamos entender com a matriz (não habitat) modula a contribuição relativa de processos de montagem de comunidades (filtro ambiental e deriva) de aves florestais com a de perda de habitat em paisagens fragmentadas. Para isso, usamos um modelo misto com associação entre os atributos das espécies e a cobertura florestal como efeitos fixos e diversos componentes aleatórios para medir força relativa de certos processos. Neste modelo, que é uma modificação de modelos de distribuição de espécies conjuntos (JSDM em inglês), escolhemos os termos de efeitos fixo e aleatórios cautelosamente para interpretar cada componente do modelo como processos ecológicos distintos. Particionamos, então, a variância destes componentes para medir a importância relativa destes na ocorrência das espécies. Encontramos que o filtro ambiental por perda de habitat é o processo dominante e pode ser relaxado ou reforçado dependendo da qualidade da matriz, evidenciando que esta tem um forte impacto na modulação dos processos de montagem da comunidade.

No Capítulo 2, estávamos interessados em estudar dinâmica florestal, mais especificamente como as características funcionais e estruturais das florestas são moldadas pelos inúmeros mecanismos que deixam marcas nas taxas vitais das árvores. Embora muitos estudos tenham tentado quantificar como mecanismos específicos influenciam as taxas vitais (a pergunta "qual"), a importância relativa desses mecanismos (o "quanto") nas florestas permanece pouco clara. Nós desenvolvemos um

framework conceitual e analítico, usando partição de variâncias de um modelo misto, para atribuir a variabilidade no crescimento, sobrevivência e recrutamento à variação entre espécies, espacial, temporal, e suas interações como uma forma de desembaraçar os diferentes mecanismos que influenciam as taxas vitais das árvores. Aplicamos o *framework* nos dados de crescimento, mortalidade e recrutamento de 21 florestas globalmente distribuídas cobrindo mais de 2,9 milhões de árvores de aproximadamente 6,5 mil espécies. Encontramos que a variabilidade entre espécies foi a maior fonte de variabilidade em todas as taxas vitais, mas que esta variabilidade diminui com a riqueza de espécies, indicando que florestas diversas apresentam as menores diferenças interespecíficas entre taxas vitais. Nossa proposta de decomposição da variação das taxas vitais de árvores tem um alto potencial para identificar os mecanismos estruturantes da dinâmica florestal, pois destaca os caminhos mais promissores para pesquisas futuras, aumentando a compreensão das contribuições relativas dos mecanismos para a demografia e diversidade florestal.

Pudemos observar neste estudo parte do potencial dos modelos mistos em conectar inferência estatística e ecológica, incorporando a variabilidade como fonte de informações diversas e de qualidade sobre sistemas ecológicos complexos. As técnicas de partição de variância em modelos mistos têm se desenvolvido bastante e se mostrando cada vez mais confiáveis, assim, temos acesso à mais informação sobre os dados que de certa forma seria vista como "resíduo". Vimos a importância de se aplicar um modelo base para primeiro entender as fontes de variação dos dados para depois construir o modelo a partir das hipóteses e predições. Outra tendência que se adequa bem aos avanços da estatística ecológica é o método de inferência bayesiano para lidar com toda a complexidade e flexibilidade dos modelos mistos junto com o aumento da disponibilidade de ferramentas computacionais para facilitar a construção destes modelos.

Esta tese não é, nem de longe, um retrato fiel de todo o desenvolvimento, das oportunidades e formas de "fazer ciência" encontradas pelo meu processo intelectual durante o doutorado. Mas acredito que ainda assim reflete de certa forma os meus interesses principais, o meu olhar sobre a ecologia e o uso da estatística em conjunto com as oportunidades e colaborações com colegas e mentores. Eu sempre me impressiono com o quanto de conhecimento é capaz de se condensar em informação compacta. O quanto é preciso saber, aprender, desenvolver internamente para chegar a

algum trabalho concluído, sendo este geralmente algumas dezenas de páginas tamanho A4 em formato PDF. E como diz meu mestre e orientador, Paulo Inácio Prado, a tese é apenas um retrato do processo, o trabalho intelectual, científico não acaba agora, não acaba neste texto. E como disse o cantor e compositor B Negão: **O processo é lento!**

Anexos 1 – Material suplementar Capítulo 1

Todo o material suplementar (apêndices) apresentado é idêntico em teor ao material suplementar publicado no artigo original, também disponível em:

https://doi.org/10.1111/1365-2664.14275

Appendix S1: Study sites, bird assemblages, and trait selections

Melina Leite

2022-08-13

Complementary information about study areas and bird sampling

For the placement of the 23 focal landscapes with different forest covers, we also evaluated if forest cover varied within radii of 1, 2, or 3 km based on each landscape centroid. Landscape-level forest cover did not vary more than 5% within those radii. It shows that we are avoiding the influence of any outside larger patch on colonization processes (Pasher et al. 2013).

Bird assemblages

We restricted our analysis to forest specialist and habitat generalist species, excluding open area species eventually detected inside smaller forest patches. We recorded 180 bird species in both regions: 93 specialists and 87 generalists. **Specialist species richness** was higher in the low-quality than in the high-quality matrix region (82 and 68 species respectively), with 57 species common to both regions. However, the median of specialist species richness per landscape and site in the high-quality matrix was higher than in the low-quality matrix (Table S1.1). There were 11 specialist species (16% of total richness) in the high-quality matrix detected only once (singletons), and 15 singletons (18% of total richness) in the low-quality matrix.

Generalist species richness was quite similar between regions, with 77 and 74 species in the high and low-quality matrix, respectively (64 common species). Accordingly, generalist species richness per landscape and site did not vary much between high-quality and low-quality matrices (Table S1,1). There were 9 singleton generalist species

in the high-quality matrix (12% of total richness) and 12 in the low-quality matrix (16% of total richness). For more details in biodiversity metrics of both regions see Boesing et al. (2018).

Table S1.1. Species richness for the assemblages in the high and low-quality matrix landscapes for landscape and local spatial scales.

		Landscape	Local
Assemblage	**Total richness**	**Median (min – max)**	**Median (min – max)**
Specialists	93		
Low-quality matrix	82	32 (17 – 61)	16 (4 – 42)
High-quality matrix	68	40 (29 – 48)	24 (12 – 40)
Generalists	87		
Low-quality matrix	74	38 (34 – 44)	19 (10 – 31)
High-quality matrix	77	44.5 (38 – 49)	22.5 (15 – 31)
Total	180		

Beta diversity (*sensu* Baselga et al. 2010) between assemblages in high- and low-quality matrix regions was very small (Fig S1.1).
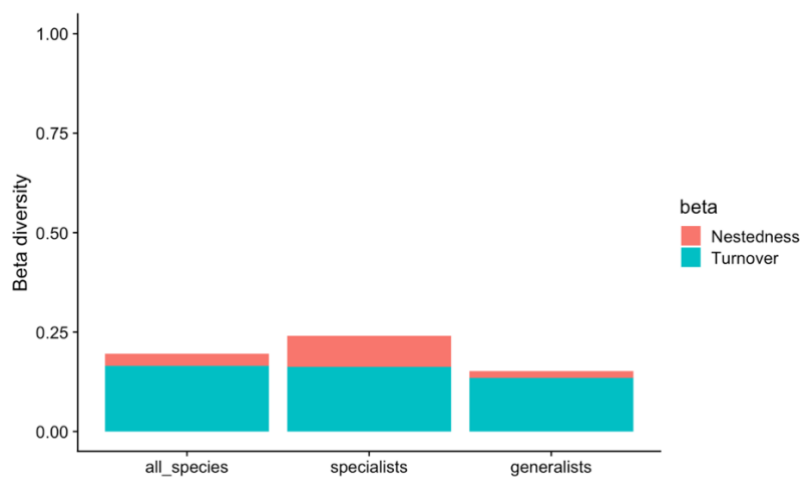

Figure S1.1. Beta diversity between high- and low-quality matrix regions partitioned in Nestedness and Turnover (*sensu* Baselga 2010), calculated with all species and separately for forest specialists and generalists. Analysis done using bepart R package (Baselga et al. 2022).

It is worth noting that our bird species list contains few species that were not expected to occur given the current knowledge of their distributional range based on field observations, especially in the South of Minas Gerais. The region is located in an ecotone which might explain some species associated to humid forests occurring further

into the Atlantic forest distribution range. We also highlight that the region is undersampled since it is an agricultural dominated region poorly attractive for ornithologists and bird watchers. Thus, many other interesting species might occur undetectable in the region. Fortunately, we have vouchers (audio records) of some of the species recorded deposited on Xeno-canto website (https://xeno-canto.org/contributor/ETMICIBVME), including some of the unexpected species, as *Grallaria varia, Anabazenops fuscus, Chamaeza campanisona*. Thus, it may be possible that our records could contribute to expanding the distributional range of *Anabazenops fuscus, Chamaeza campanisona, Grallaria varia, Drymophila rubricollis,* and *Drymophila squamata*. It is important to highlight that the detection of these species was in general rare, which indicates that they may have low population.

Even though some of the species mentioned above are known as 'unlikely' in the region, an unpublished study on niche species modeling of birds in the Brazilian Atlantic Forest shows that our study area (both regions) is suitable for many of these unexpected species (Acosta et al. *in prep*). In this study, the current suitability of the Atlantic Forest biome for each bird species was obtained using species distribution modeling (SDM), which relies on species' records and on climate variables to build models of geographic distribution. We show in Figure S1.2 habitat suitability maps for seven species of interesting findings in the study area: *Anabazenops fuscus, Chamaeza meruloides, Chamaeza campanisona, Drymophila ferruginea, Drymophila malura, Grallaria varia, Odontophorus capueira.*
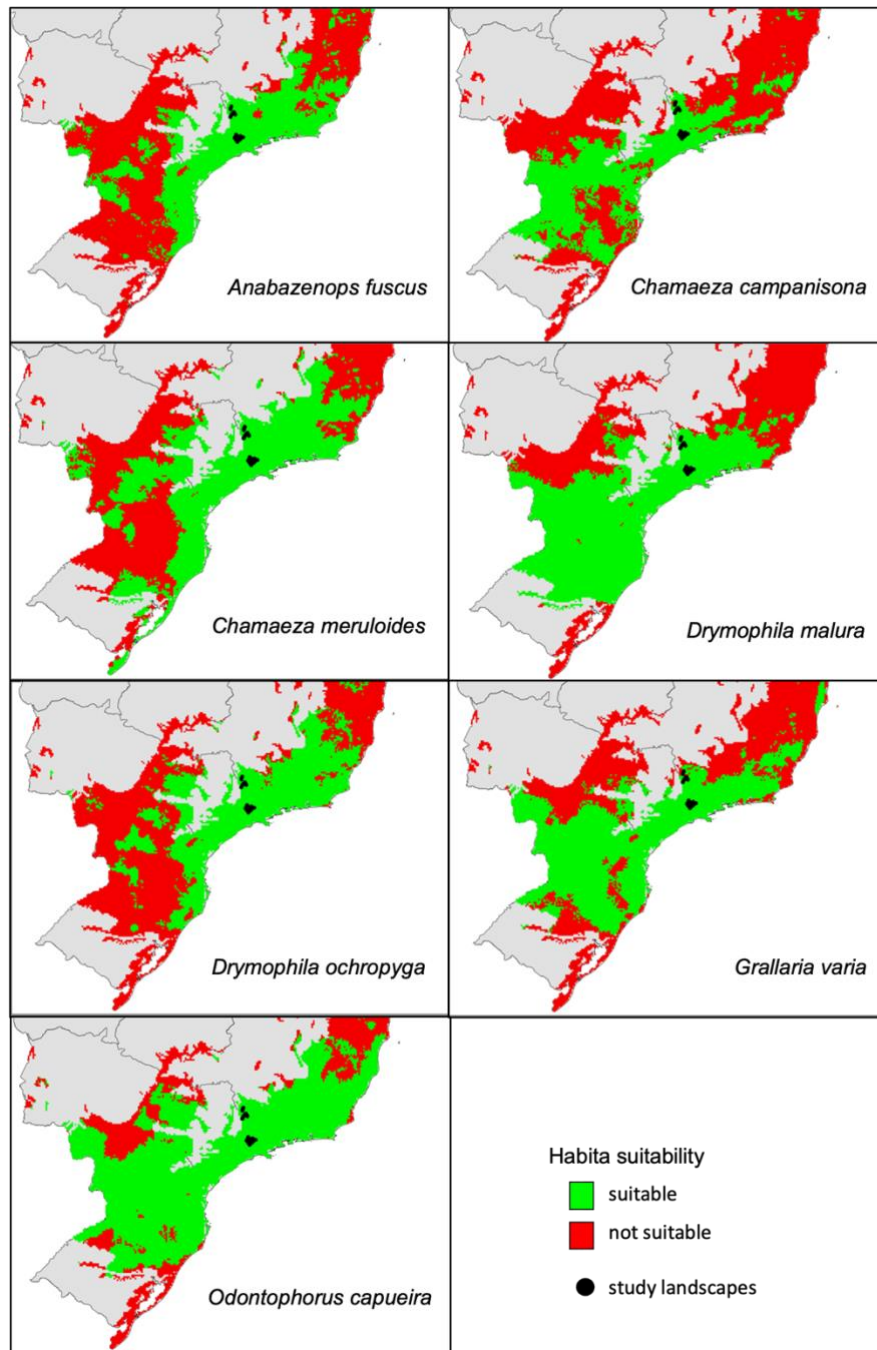
Figure S1.2. Habitat suitability for some species recorded in our field surveys, with focus in the south and southeastern area of the Brazilian Atlantic Forest. Black points are the sampling sites of our study. Source Acosta et al. *in prep.*

Reference

Acosta, A.L.; Giannini, T.C.; De Coster, G.; Banks-Leite, C. Metzger, J.P. in prep. How climate change can affect the environmental services provided by birds in the Atlantic Forest.

Species traits selection

Bird traits selection and the associated hypothesis of how species may respond to habitat loss according to its traits are summarized in Table S1.2 and described in detail in the following paragraphs.

Table S1.2. Bird traits with the information of the operational variables, the hypothesis of species habitat loss filtering according to trait values, and the sources of data acquisition.

Trait	Operational variable	Hypothesis	Data source
Body size	Continuous. Log of mean body mass for both sexes in grams.	Abundance of large species decreases more intensely with habitat loss.	Ramirez et al. 2008; Rodrigues et al. 2019
Nest type	Categorical. Nest in cavities; open or semi-open nest; closed nest.	The abundance of species with open/semi-open nests decreases more intensely with habitat loss than for species with other nest types.	Del Hoyo et al. 2014
Diet	Categorical main diet: omnivorous, frugivorous, nectarivorous, insectivorous, granivorous. Continuous. Percentage of fruits in the diet. Continuous. Percentage of invertebrates in the diet.	The abundance of frugivorous and insectivorous species decreases more intensely with habitat loss than for species with other diets.	Sick 1997, Del Hoyo et al. 2014; Wilman et al. 2014
Foraging stratum	Categorical. Main foraging stratum: ground and/or understory, midstory and/or canopy, and all strata. Continuous. Percentage of lower strata (ground and understory) use.	The abundance of species in the ground and understory strata decreases more intensely than species using midstory, canopy, or all strata.	Sick 1997, Del Hoyo et al. 2014; Wilman et al. 2014
Hand-wing index	Continuous. Index for wing shape. Mean of Hand-wing index for both sexes.	Abundance of species with larger hand-wing index (larger dispersal ability) decreases less intensely with habitat loss.	Sheard et al. 2020

Body size is one of the key attributes of vertebrates in respect of extinction risk, reproduction, and dispersal (Owens and Bennett 2000, Ripple et al. 2017). In birds, species with large mean body sizes are often considered more vulnerable to extinction

given their low population densities, large home ranges, slow growth rates, high energetic requirements, and high sensitivity to anthropogenic overexploitation (Cardillo et al. 2005). The risk of local extinction in altered and smaller habitats correlates with mean body size (e.g. Barbaro and Halder 2009, Flynn et al. 2009, Newbold et al. 2013, Bregman et al. 2016, Bovo et al. 2018). However, sometimes this trait is not a good predictor of community changes (e.g. Tscharntke et al. 2008, Angert et al. 2011), probably because large species may also benefit from having higher mobility (Tscharntke et al. 2012). We compiled information on avian body mass (mean values from males and females) in Boesing et al. (2018), which followed Ramirez et al. (2008) and Rodrigues et al. (2019). Body mass was log-transformed before modeling.

Nest type is often associated with reproduction effort and is most likely to affect recruitment (Bennett and Owens 2002). For example, species that make nests in cavities have higher growth rates (Bellier et al. 2018) probably because it is a safer nest against parasitism and predation (Sibly et al. 2012). Nest predation and parasitism are among the most impacting factors of bird populations' decline in fragmented landscapes (Cavitt and Martin 2002). In addition, the lack of suitable nesting habitats in disturbed environments can have a strong effect on the reproductive success of certain bird species, such as those from Picidae and Psittacidae families, which require old or dead trees to build their nests (Sick 1997). We assigned the species to 3 nesting categories: closed, cavity, and open/semi-open. We collapsed open and semi-open nest types due to the low proportion of specialist species with open nests in our data. Nest type information was collected in Del Hoyo et al. (2014).

Habitat loss, fragmentation, and land-use change affect the structure of the habitat by altering differently the provision of food for birds. For example, nectarivorous, frugivorous, and insectivorous species seem to be more sensitive to habitat loss and fragmentation than omnivorous and granivorous (Sekercioglu et al. 2004, Newbold et al. 2013, Bovo et al. 2018, Chatterjee and Basu 2018). We assigned species to five **main diet** categories according to information available in the literature (Sick 1997, Del Hoyo 2014): omnivorous, insectivorous, frugivorous, nectarivorous, granivorous. Because of the special relationship found for frugivorous and insectivorous species in land-use change (Newbold et al. 2014), we also assigned the percentage of each of these components in species' diet as trait variables. The percentages of fruits and insects in the species' diet were extracted from the EltonTrait database (Wilman et al.

2014). For the analysis, we excluded the single granivorous forest specialist species in the high-quality matrix region and two omnivorous forest specialist species in the low-quality matrix region because of issues during model fit.

Foraging stratum is of most importance for birds in fragmented landscapes. Ground and understory species are more prone to extinction (Laurance and Gomez 2005), mostly because of higher dispersal limitation and avoidance of open areas (gaps, matrix, and forest edges). We assigned each species to 3 foraging strata categories: ground-understory, midstory-canopy, and all strata. We also used the percentage of use of lower foraging strata (ground and understory) as an alternative operational variable extracted from EltonTrait database (Wilman et al. 2014).

Hand-wing index is an estimate of wing shape widely adopted as a proxy for dispersal ability in birds (Sheard et al. 2020). Species with larger dispersal abilities are expected to suffer less with habitat loss and fragmentation because of higher connectivity among habitat patches. We compiled hand-wing index values from (Sheard et al. 2020).

Comparing traits between forest specialists and habitat generalists

All traits were compared between forest specialists and habitat generalists using graphical and multivariate analysis to ensure that the selected traits are comparable between groups, i.e., the range of numeric trait values between generalist and specialist species are around the same range (minimum and maximum) and that there was not a single trait that could completely separate them. We found that hand-wing index presented larger maximum values and that nectarivorous birds were only generalists and insectivorous were more common among specialists.

Below, we present the comparisons of trait values between specialists and generalists. Continuous variables were Z-score scaled and are presented in Figure S1.3, Categorical variables are summarized in Table S1.4.
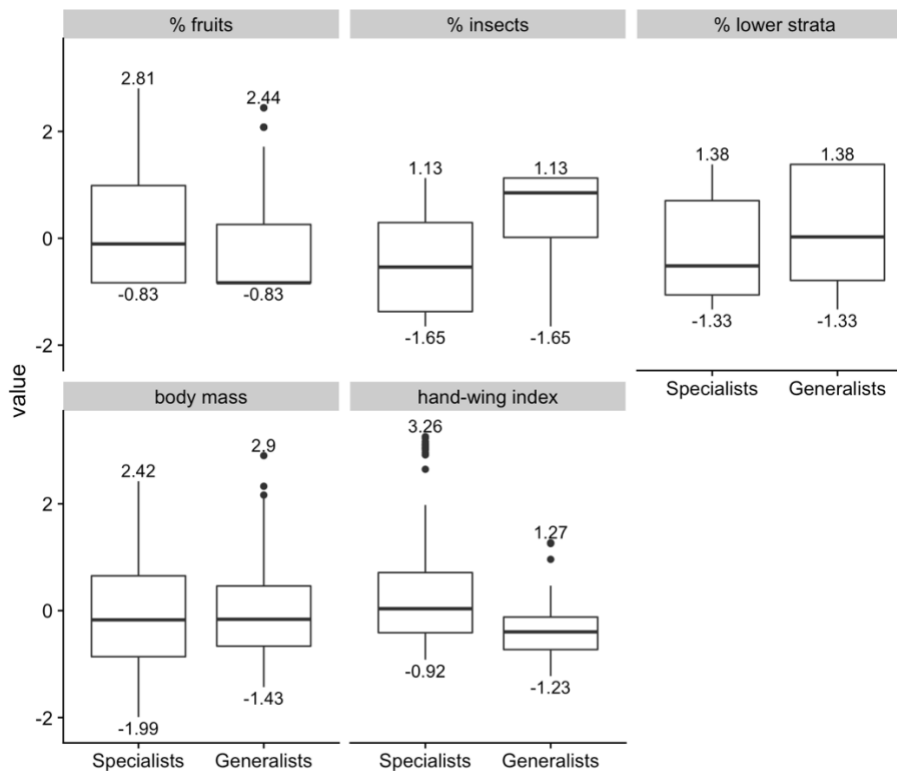
Figure S1.3. Boxplots of values for the traits measured as continuous variables for specialist and generalist birds. Values presented are maximum and minimum values in each group.

Table S1.4. Percentage of the species in each trait category for specialists and generalists. Numbers inside brackets are the number of species.

Traits	Specialists	Generalists
Nest type		
Cavities	25% (23)	22% (19)
Closed	27% (25)	21% (18)
Open/semi-open	48% (45)	57% (50)
Main diet		
Frugivorous	17% (16)	23% (20)
Granivorous	2% (2)	6% (5)
Insectivorous	77% (72)	40% (35)
Nectarivorous	0% (0)	15% (13)
Onivorous	3% (3)	16% (14)
Foraging stratum		
All	5% (5)	20% (17)
Ground/Understory	58% (54)	32% (28)
Midstory/Canopy	37% (34)	48% (42)

References on bird traits and environmental change

List the consulted references of bird traits.

Angert, A. L., Crozier, L. G., Rissler, L. J., Gilman, S. E., Tewksbury, J. J., & Chunco, A. J. (2011). Do species' traits predict recent shifts at expanding range edges?: Traits and range shifts. Ecology Letters, 14(7), 677–689. https://doi.org/10.1111/j.1461-0248.2011.01620.x

Barbaro, L., & Halder, I. V. (2009). Linking bird, carabid beetle and butterfly life-history traits to habitat fragmentation in mosaic landscapes. Ecography, 32(2), 321–333. https://doi.org/10.1111/j.1600-0587.2008.05546.x

Barros, F. M. de. (2017). Species composition, ecological functions and ecosystem services by birds across forest-matrix interfaces in tropical disturbed landscapes. UNESP.

Bellier, E., Kéry, M., & Schaub, M. (2018). Relationships between vital rates and ecological traits in an avian community. Journal of Animal Ecology, 87(4), 1172–1181. https://doi.org/10.1111/1365-2656.12826

Bennett, P. M., & Owens, I. P. F. (2002). Evolutionary Ecology of Birds—Life histories, Mating Systems and Extinction. Oxford University Press. https://kar.kent.ac.uk/7528/

Boesing, A. L., Nichols, E., & Metzger, J. P. (2018). Biodiversity extinction thresholds are modulated by matrix type. Ecography, 41(9), 1520–1533. https://doi.org/10.1111/ecog.03365

Bovo, A. A. A., Ferraz, K. M. P. M. B., Magioli, M., Alexandrino, E. R., Hasui, É., Ribeiro, M. C., & Tobias, J. A. (2018). Habitat fragmentation narrows the distribution of avian functional traits associated with seed dispersal in tropical forest. Perspectives in Ecology and Conservation, 16(2), 90–96. https://doi.org/10.1016/j.pecon.2018.03.004

Bregman, T. P., Lees, A. C., MacGregor, H. E. A., Darski, B., Moura, N. G. de, Aleixo, A., Barlow, J., & Tobias, J. A. (2016). Using avian functional traits to assess the impact of land-cover change on ecosystem processes linked to resilience in tropical forests. Proceedings of the Royal Society B: Biological Sciences, 283(1844), 20161289. https://doi.org/10.1098/rspb.2016.1289

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

Cardillo, M., Mace, G. M., Jones, K. E., Bielby, J., Bininda-Emonds, O. R. P., Sechrest, W., Orme, C. D. L., & Purvis, A. (2005). Multiple causes of high extinction risk in large mammal species. Science, 309(5738), 1239–1241. https://doi.org/10.1126/science.1116030

Cavitt, J. E., & Martin, T. E. (2002). Effects of forest fragmentation on brood parasitism and nest predationin eastern and western landscapes. Studies in Avian Biology, 25, 73–80.

Chatterjee, S., & Basu, P. (2018). Food preferences determine habitat selection at multiple scales: Implication for bird conservation in tropical forests. Animal Conservation, 21(4), 332–342. https://doi.org/10.1111/acv.12397

Cleary, D. F. R., Boyle, T. J. B., Setyawati, T., Anggraeni, C. D., Loon, E. E. V., & Menken, S. B. J. (2007). Bird species and traits associated with logged and unlogged forest in borneo. Ecological Applications, 17(4), 1184–1197. https://doi.org/10.1890/05-0878

Cormont, A., Vos, C., van Turnhout, C., Foppen, R., & ter Braak, C. (2011). Using life-history traits to explain bird population responses to changing weather variability. Climate Research, 49(1), 59–71. https://doi.org/10.3354/cr01007

de Coster, G., Banks-Leite, C., & Metzger, J. P. (2015). Atlantic forest bird communities provide

different but not fewer functions after habitat loss. Proceedings of the Royal Society B: Biological Sciences, 282(1811), 20142844. https://doi.org/10.1098/rspb.2014.2844

Del Hoyo, J. (2014). Handbook of the birds of the World alive. Lynx Editions.

Flynn, D. F. B., Gogol-Prokurat, M., Nogeire, T., Molinari, N., Richers, B. T., Lin, B. B., Simpson, N., Mayfield, M. M., & DeClerck, F. (2009). Loss of functional diversity under land use intensification across multiple taxa. Ecology Letters, 12(1), 22–33. https://doi.org/10.1111/j.1461-0248.2008.01255.x

Kupsch, D., Vendras, E., Ocampo-Ariza, C., Batáry, P., Motombi, F. N., Bobo, K. S., & Waltert, M. (2019). High critical forest habitat thresholds of native bird communities in Afrotropical agroforestry landscapes. Biological Conservation, 230, 20–28. https://doi.org/10.1016/j.biocon.2018.12.001

Laurance, S. G. W., & Gomez, M. S. (2005). Clearing width and movements of understory rainforest birds. Biotropica, 37(1), 149–152. https://doi.org/10.1111/j.1744-7429.2005.04099.x

Luck, G. W., Carter, A., & Smallbone, L. (2013). Changes in Bird Functional Diversity across Multiple Land Uses: Interpretations of Functional Redundancy Depend on Functional Group Identity. PLOS ONE, 8(5), e63671. https://doi.org/10.1371/journal.pone.0063671

Luck, G. W., Lavorel, S., McIntyre, S., & Lumb, K. (2012). Improving the application of vertebrate trait-based frameworks to the study of ecosystem services. Journal of Animal Ecology, 81(5), 1065–1076. https://doi.org/10.1111/j.1365-2656.2012.01974.x

Macchi, L., Baumann, M., Bluhm, H., Baker, M., Levers, C., Grau, H. R., & Kuemmerle, T. (2019). Thresholds in forest bird communities along woody vegetation gradients in the South American Dry Chaco. Journal of Applied Ecology, 56(3), 629–639. https://doi.org/10.1111/1365-2664.13342

Martin, A. E., Desrochers, A., & Fahrig, L. (2017). Homogenization of dispersal ability across bird species in response to landscape change. Oikos, 126(7), 996–1003. https://doi.org/10.1111/oik.03859

Martin, C. A., & Proulx, R. (2016). Habitat geometry, a step toward general bird community assembly rules in mature forests. Forest Ecology and Management, 361, 163–169. https://doi.org/10.1016/j.foreco.2015.11.019

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(6), e0128923. https://doi.org/10.1371/journal.pone.0128923

Newbold, T., Hudson, L. N., Phillips, H. R. P., Hill, S. L. L., Contu, S., Lysenko, I., Blandon, A., Butchart, S. H. M., Booth, H. L., Day, J., De Palma, A., Harrison, M. L. K., Kirkpatrick, L., Pynegar, E., Robinson, A., Simpson, J., Mace, G. M., Scharlemann, J. P. W., & 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(1792), 20141371. https://doi.org/10.1098/rspb.2014.1371

Newbold, T., Scharlemann, J. P. W., Butchart, S. H. M., Şekercioğlu, Ç. H., Alkemade, R., Booth, H., & Purves, D. W. (2013). Ecological traits affect the response of tropical forest bird species to land-use intensity. Proceedings of the Royal Society B: Biological Sciences, 280(1750), 20122131. https://doi.org/10.1098/rspb.2012.2131

Owens, I. P. F., & Bennett, P. M. (2000). Ecological basis of extinction risk in birds: Habitat loss versus human persecution and introduced predators. Proceedings of the National Academy of Sciences, 97(22), 12144–12148. https://doi.org/10.1073/pnas.200223397

Ramirez, L., Diniz-Filho, J. A. F., & Hawkins, B. A. (2008). Partitioning phylogenetic and adaptive components of the geographical body-size pattern of New World birds. Global Ecology and Biogeography, 17(1), 100–110. https://doi.org/10.1111/j.1466-8238.2007.00346.x

Ripple, W. J., Wolf, C., Newsome, T. M., Hoffmann, M., Wirsing, A. J., & McCauley, D. J. (2017). Extinction risk is most acute for the world's largest and smallest vertebrates. Proceedings of

the National Academy of Sciences, 114(40), 10678–10683.
https://doi.org/10.1073/pnas.1702078114

Rodrigues, R. C., Hasui, É., Assis, J. C., Pena, J. C. C., Muylaert, R. L., Tonetti, V. R., Martello, F., Regolin, A. L., Costa, T. V. V. da, Pichorim, M., Carrano, E., Lopes, L. E., Vasconcelos, M. F. de, Fontana, C. S., Roos, A. L., Gonçalves, F., Banks-Leite, C., Cavarzere, V., Efe, M. A., … Ribeiro, M. C. (2019). ATLANTIC BIRD TRAITS: A dataset of bird morphological traits from the Atlantic forests of South America. Ecology, 100(6), e02647. https://doi.org/10.1002/ecy.2647

Şekercioğlu, Ç. H., Daily, G. C., & Ehrlich, P. R. (2004). Ecosystem consequences of bird declines. Proceedings of the National Academy of Sciences, 101(52), 18042–18047. https://doi.org/10.1073/pnas.0408049101

Sheard, C., Neate-Clegg, M. H. C., Alioravainen, N., Jones, S. E. I., Vincent, C., MacGregor, H. E. A., Bregman, T. P., Claramunt, S., & Tobias, J. A. (2020). Ecological drivers of global gradients in avian dispersal inferred from wing morphology. Nature Communications, 11(1), 2463. https://doi.org/10.1038/s41467-020-16313-6

Sibly, R. M., Witt, C. C., Wright, N. A., Venditti, C., Jetz, W., & Brown, J. H. (2012). Energetics, lifestyle, and reproduction in birds. Proceedings of the National Academy of Sciences, 109(27), 10937–10941. https://doi.org/10.1073/pnas.1206512109

Sick, H. (1997). Ornitologia brasileira. Editora Nova Fronteira. https://books.google.com.br/books?id=-RuGRAAACAAJ

Smith, Y. C. E., Smith, D. A. E., Seymour, C. L., Thébault, E., & Veen, F. J. F. van. (2015). Response of avian diversity to habitat modification can be predicted from life-history traits and ecological attributes. Landscape Ecology, 30(7), 1225–1239. https://doi.org/10.1007/s10980-015-0172-x

Tscharntke, T., Sekercioglu, C. H., Dietsch, T. V., Sodhi, N. S., Hoehn, P., & Tylianakis, J. M. (2008). Landscape constraints on functional diversity of birds and insects in tropical agroecosystems. Ecology, 89(4), 944–951. https://doi.org/10.1890/07-0455.1

Tscharntke, T., Tylianakis, J. M., Rand, T. A., Didham, R. K., Fahrig, L., Batáry, P., Bengtsson, J., Clough, Y., Crist, T. O., Dormann, C. F., Ewers, R. M., Fründ, J., Holt, R. D., Holzschuh, A., Klein, A. M., Kleijn, D., Kremen, C., Landis, D. A., Laurance, W., … Westphal, C. (2012). Landscape moderation of biodiversity patterns and processes—Eight hypotheses. Biological Reviews, 87(3), 661–685. https://doi.org/10.1111/j.1469-185X.2011.00216.x

Vance, M. D., Fahrig, L., & Flather, C. H. (2003). Effect of reproductive rate on minimum habitat requirements of forest-breeding birds. Ecology, 84(10), 2643–2653. https://doi.org/10.1890/02-0159

White, H. J., Montgomery, I. W., & Lennon, J. J. (2018). Contribution of local rarity and climatic suitability to local extinction and colonization varies with species traits. Journal of Animal Ecology, 87(6), 1560–1572. https://doi.org/10.1111/1365-2656.12881

Wilman, H., Belmaker, J., Simpson, J., Rosa, C. de la, Rivadeneira, M. M., & Jetz, W. (2014). EltonTraits 1.0: Species-level foraging attributes of the world's birds and mammals. Ecology, 95(7), 2027–2027. https://doi.org/10.1890/13-1917.1

Other references

Baselga, A. (2010). Partitioning the turnover and nestedness components of beta diversity. Global Ecology and Biogeography, 19(1), 134–143. https://doi.org/10.1111/j.1466-8238.2009.00490.x

Baselga A, David Orme, Sebastien Villeger, Julien De Bortoli, Fabien Leprieur and Maxime Logez (2022). betapart: Partitioning Beta Diversity into Turnover and Nestedness Components. R package version 1.5.6. https://CRAN.R-project.org/package=betapart

Appendix S2: Landscapes land use and forest cover at different scales

Melina Leite

2022-08-13

This appendix is a description of land uses in the focal landscapes together with the description of forest cover at both the local and the focal landscape scales. We present the baseline models for the selection of the best scale for the local forest cover variable for each dataset. For the local scale, we measured the percentage of forest cover within buffers of 400, 600 and 800 m around each sampling site. For the focaal landscape forest cover, we used the 2 km buffer around the landscape centroid.

1. Land use composition in landscapes

To clarify the differences in landscapes between regions and to show that these differences are in accordance with our categorical classification of high- and low-quality matrices for birds, we show in Figure S2.1 the composition of the main land use types per landscapes, and in Figure S2.2 and Table S2.1 the comparisons of land use types among the high- and low-quality regions.

We tested for differences in Shannon and Simpson diversity indexes using the percentage of land use types of the landscapes (Figure S2.3). We found larger diversity of land use types in highquality matrix landscapes for both diversity indices, that is, more matrix heterogeneity in high-quality matrix landscapes.

We used Principal Coordinate Analysis (PCoA) to show the clear separation in land uses of landscapes from both regions (Figure S2.4). PCoA is more adequate than PCA given the nature of the data (sum up 100%, compositional data). We used gower distance tranformation to create a distance matrix among landscapes. Among the main land uses, landscapes in the low-quality region are more associated with larger proportions of pasture, eucalyptus plantations and urban areas to a lower extent, while landscapes in the high-quality region are more associated with larger proportions of coffee, sugar cane and lower proportions of pasture to a lower extend. The PCoA analysis also shows that forest cover variation is similar between regions and are not

relevant in separating landscapes from both regions,i.e., the amount of forest cover among landscapes follows a similar gradient in both regions.

Although landscapes in the high-quality matrix region do have some proportion of pasture (mean 35.5% ± 10.2), this remains lower than the proportion of pasture in the low-quality matrix region (mean 46.9% ± 11.1). Moreover, as it can be seen in PCoA results (Figure S2.4), high-quality matrix landscapes are more negatively related to proportion of pasture when compared together with low-quality matrix landscapes.

Eucalyptus tree plantations represent the only arboreal matrix element in all 13 low-quality matrix landscapes and compose an average of 18% ± 8.7 of matrix cover, while high-quality landscapes have on average 2% ± 3.2 of eucalyptus in matrix cover but it is still present in 7 of the 10 landscapes. Although eucalyptus trees may, in principle, provide less edge effects as pastures and coffee plantations, it doesn't necessarily mean it is a high-quality matrix for birds. (Barros *et al.* (2019)) concluded that eucalyptus plantations at the same region of our low-quality matrix landscapes were matrices of lower quality because of intensive management including biocidal suppression of native understory vegetation. Such understory vegetation suppression results not only in resource poor environments, but in very simplified stratification with a limited amount of microhabitats required specially for understory and terrestrial species, allowing only a subset of more generalist species capable of using such areas (Jacoboski, Mendonça-Lima & Hartz (2016)). Moreover, eucalyptus plantations are also less perennial elements in the landscape with cycles of clearcut around 6-8 years (Rodrigues *et al.* (2019)). It means that eucalyptus plantations may not be necessarily equally high-quality matrices as initially supposed (Barros *et al.* (2019)), especially when compared with coffee plantations, which not only present a low-contrast physical structure, but also provide a variety of resources for different species yearlong, evidenced by high-rates of spillover movements from forest to them (Boesing, Nichols & Metzger (2018)).

Also note the high negative correlation among percentage of pasture and forest cover in Figure S2.5.
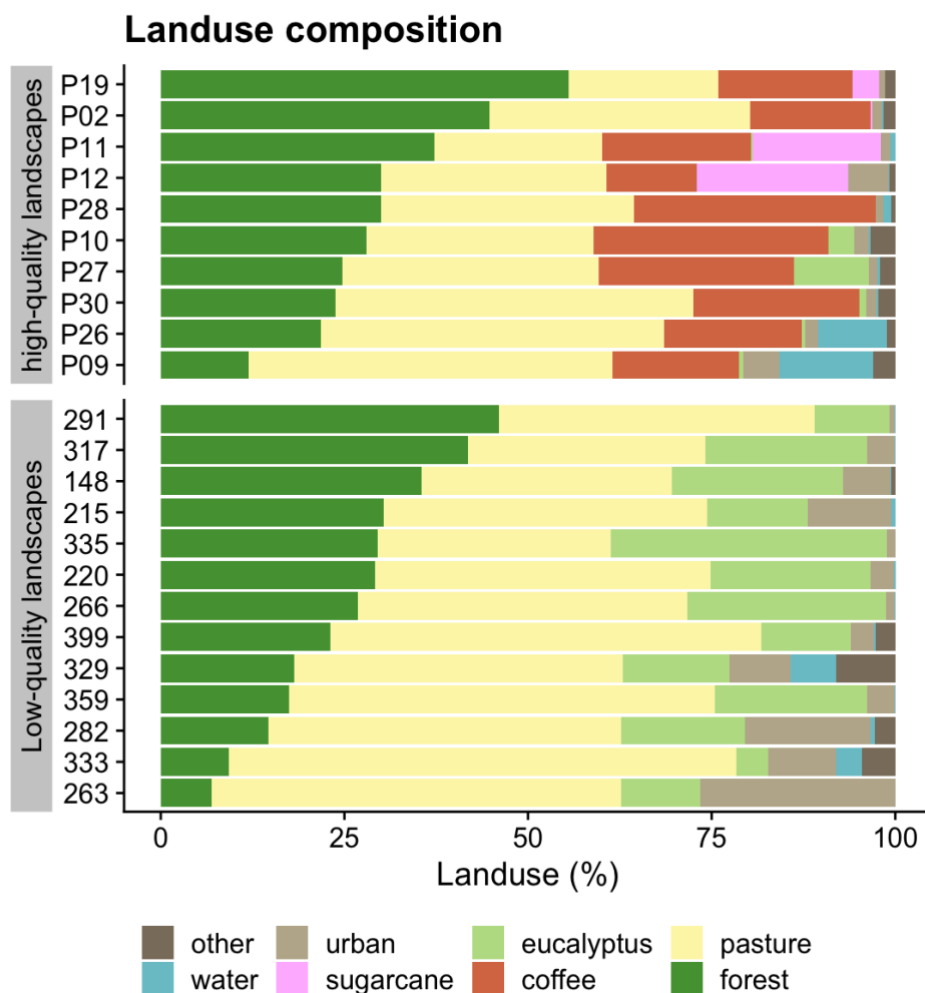
Landuse composition

Figure S2.1: Percentage of the 8 main land use types per landscape in each region.

Table S2.1: Summary table of the land use percentages for landscapes in both regions of high- and low-quality matrix landscapes.

matrix	landuse	min	mean	sd	median	max
high_quality	forest	12.0	30.8	12.4	29.0	55.5
high_quality	pasture	20.4	35.5	10.2	34.7	49.5
high_quality	coffee	12.4	21.7	6.8	19.5	32.9
high_quality	eucalyptus	0.0	1.6	3.2	0.4	10.1
high_quality	sugarcane	0.0	4.2	7.9	0.0	20.6
high_quality	urban	0.8	2.1	1.7	1.4	5.5
high_quality	water	0.0	2.5	4.6	0.3	12.7
high_quality	other	0.0	1.6	1.1	1.5	3.4
low_quality	forest	6.9	25.3	11.9	26.8	46.0
low_quality	pasture	31.7	46.9	11.1	44.8	69.1
low_quality	coffee	0.0	0.0	0.0	0.0	0.0
low_quality	eucalyptus	4.3	18.1	8.7	16.9	37.7
low_quality	sugarcane	0.0	0.0	0.0	0.0	0.0
low_quality	urban	0.7	7.4	7.5	3.8	26.6

| low_quality | water | 0.0 | 0.9 | 1.9 | 0.1 | 6.2 |
| low_quality | other | 0.0 | 1.4 | 2.5 | 0.0 | 8.0 |
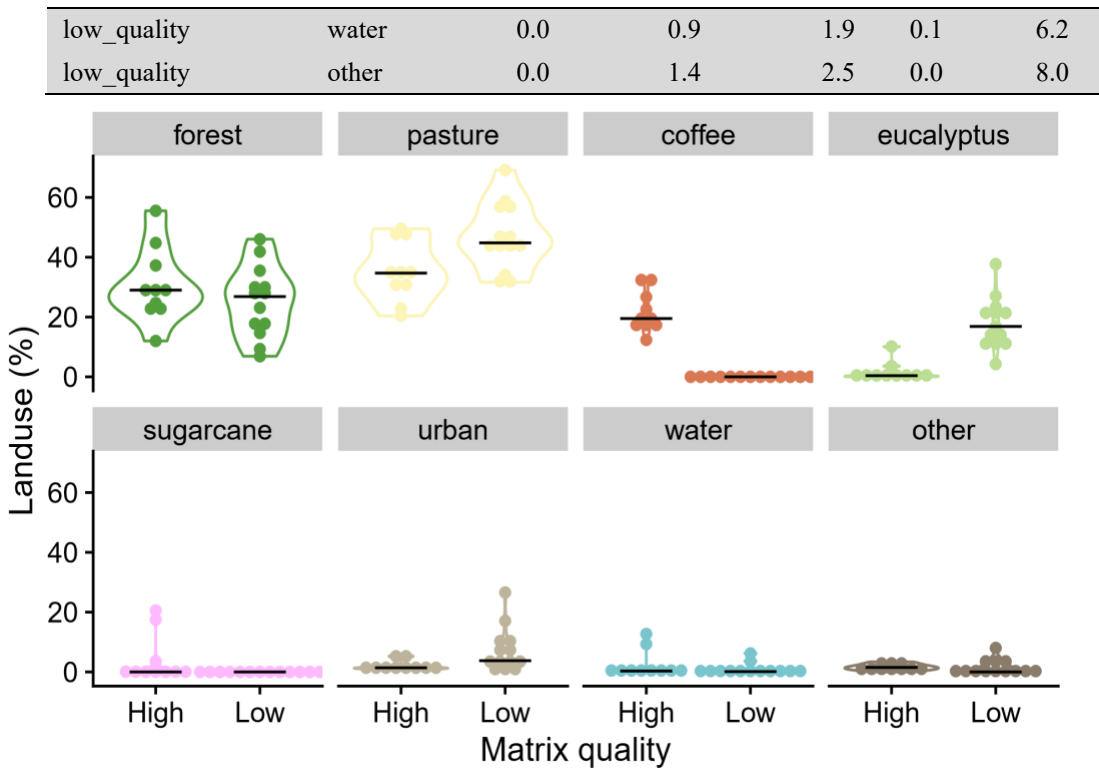
Figure S2.2: Proportions of landuse types among high- and low-quality landscapes
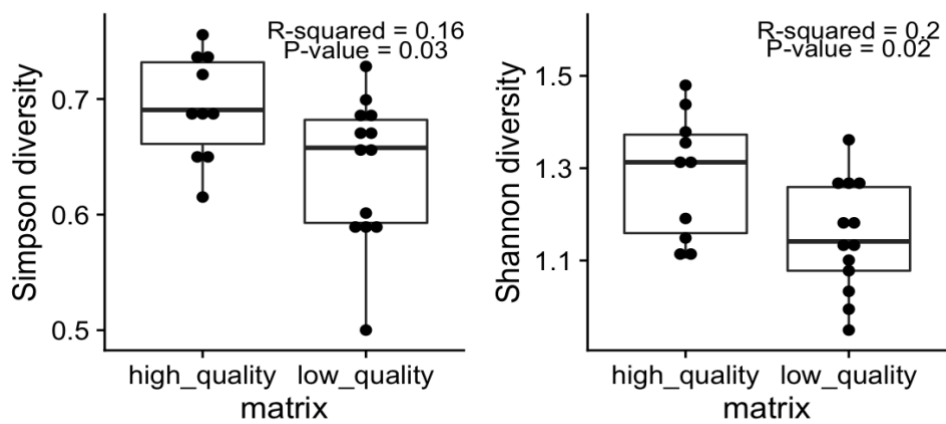

Figure S2.3: Simpson and Shannon diversity indexes for land uses in landscapes from both high- and low-quality matrix regions.
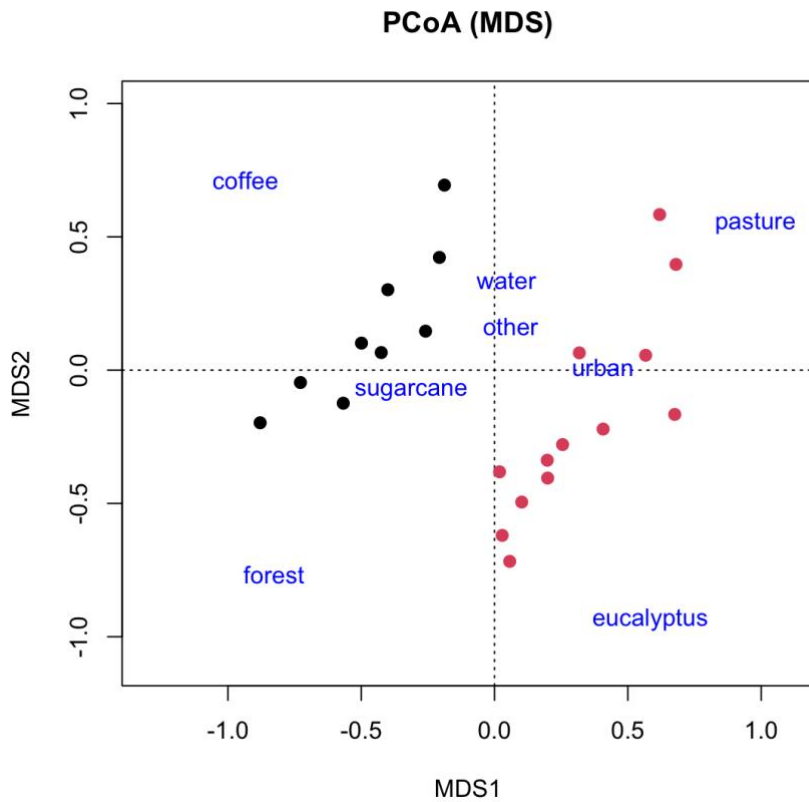
**Figure S2.4:** First 2 axes of the PCoA results for the composition of land use types among landscapes from the high-quality matrix (black dots) and low-quality matrix (red dots) regions.
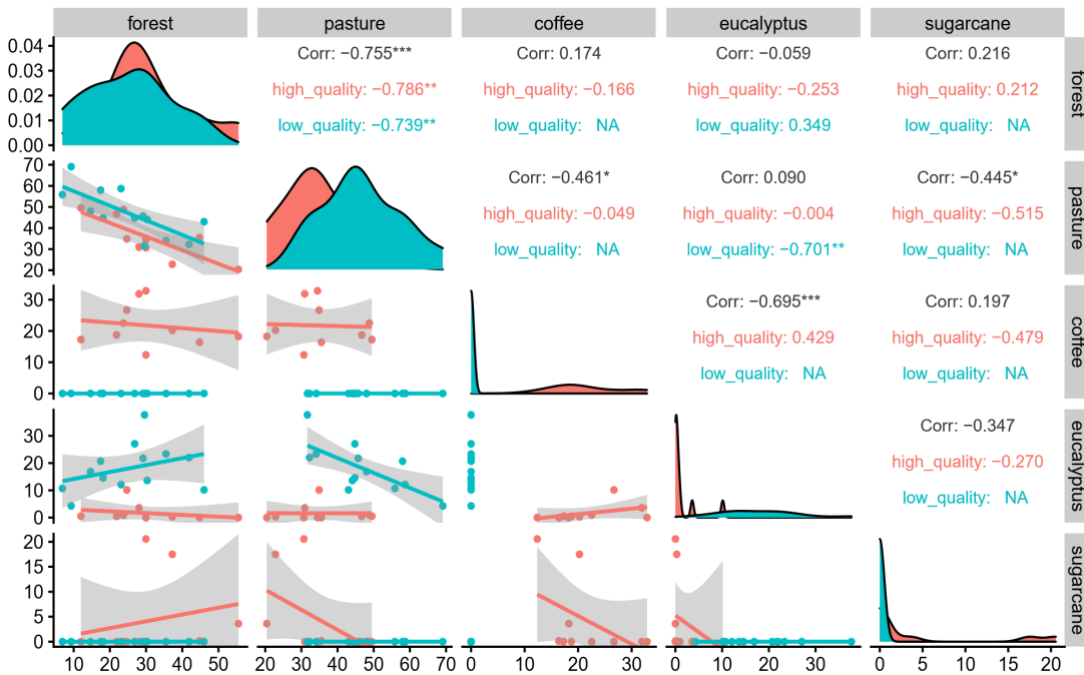
Figure S2.5: Correlation among percentages of landuse types.

2. Relationships among forest cover variables

We calculated Pearson correlation coefficients for forest cover variables in each matrix quality region (Figure S2.6). Also, we ploted the range of local forest cover (400 m) within the landscapes to see how local forest cover varies among landscapes in both regions (Figure S2.7).
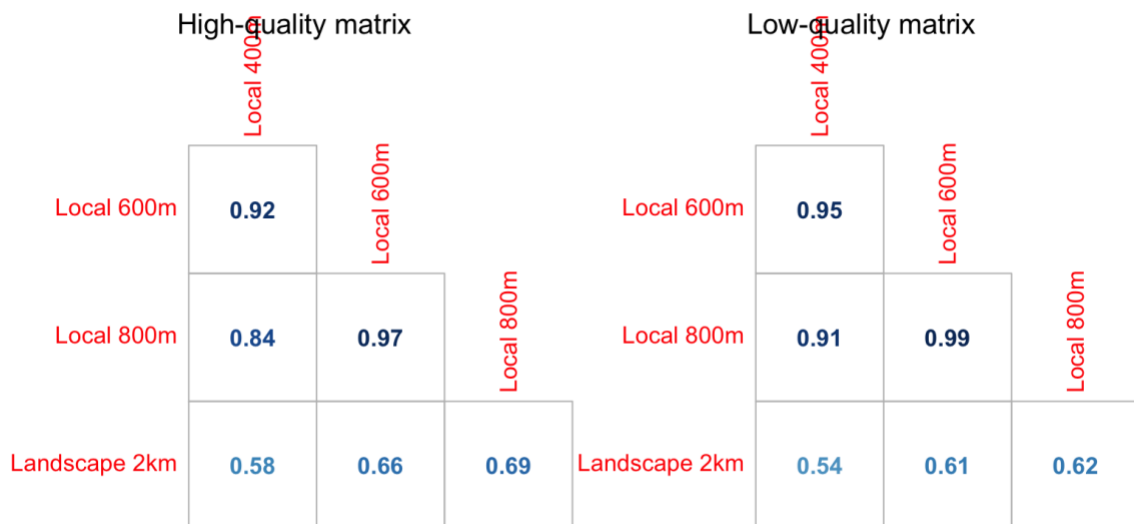
Figure S2.3: Simpson and Shannon diversity indexes for land uses in landscapes from both high- and low-quality matrix regions.
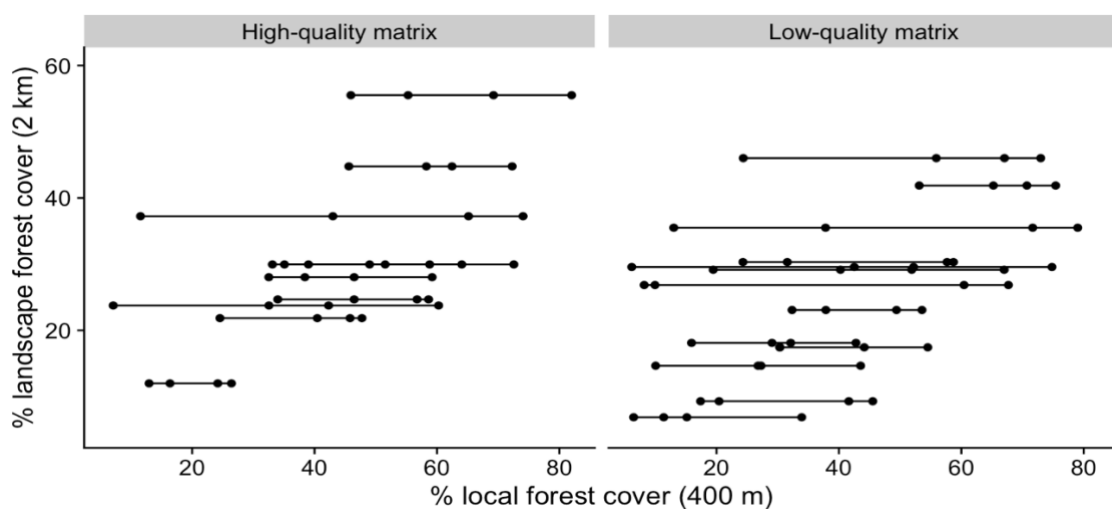

Figure S2.7: Range of local forest cover (buffer 400 m) within each landscape (buffer 2 km) for both landscapes with different matrix quality. Each line represents a landscape and the dots area the local forest cover for each sampling site.

3. Scale of effects for local forest cover

We ran different models with each local forest cover variable and selected the scale of effect using AIC model selection and the R^2 of the models. The models follow the specification presented in the paper (Modeling section), except that here we did not include trait variables, i.e., we only modeled the occurrence of species according to forest cover.

We used lme4 package (Bates *et al.* (2015)) to perform a GLMM with binomial (proportion) distribution. An example of the code for each assemblage is as follows:

```
model <- glmer(cbind(occor, n.visit-occor) ~ local.cover +
(local.cover|sp) + (1|landscape:sp) + (1|site:sp) + (1|landscape) +
(1|site), family=binomial, data=high.spe)
```

In Figure S2.8, we present the occurrence probability predicted for the models with different local forest cover scales for all the assemblages. Predictions were quite similar and decreased with forest cover for the specialists, especially in the low-quality matrix region, and increased or remained flat for the generalists.

Table S2.2: Overal and marginal r-squares and model comparisons with Akaike Information Criterion (AIC) for models with different local forest cover scales as predictor for the specialist and generalists species in both regions with different matrix qualities. For the terms see Table 1 (main text). dAIC is the difference in Akaike Information Criterion to the best model; df are the degrees of freedom.

| Model	Total	fixed		env.sp	lands.sp	site.sp	lands	site	AIC dAIC	df
Forest specialist species										
Low-quality matrix										
400m	64.3	7.5		42.9	6.2	3.3	1.0	3.4	0.00	9
600m	63.9	6.2		43.1	6.2	3.4	0.9	4.1	10.03	9
800m	63.9	6.0		43.0	6.3	3.5	1.0	4.1	18.08	9
High-quality matrix										
400m	56.6	1.3		44.3	7.5	1.8	0.7	1.1	0.00	9
600m	56.8	1.0		44.2	7.5	1.8	0.9	1.3	8.57	9
800m	56.6	0.6		44.3	7.4	1.9	0.9	1.4	12.40	9
Forest generalist species										
Low-quality matrix										
400m	46.6	0.1		39.7	3.5	2.5	0.0	0.9	0.00	9
600m	46.6	0.0		39.1	3.6	3.0	0.0	0.9	20.27	9
800m	46.5	0.0		39.0	3.6	3.0	0.0	0.9	22.54	9
High-quality matrix										
400m	44.1	0.0		37.0	2.2	3.6	0.6	0.7	11.24	9
600m	44.3	0.0		37.5	2.0	3.4	0.6	0.7	2.45	9
800m	44.3	0.1		37.5	2.0	3.5	0.6	0.7	0.00	9
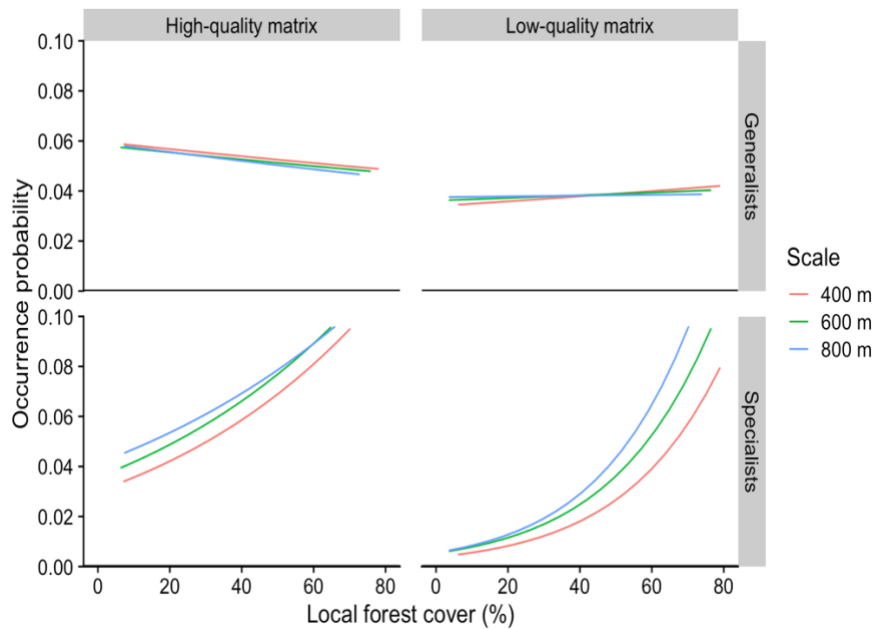
Figure S2.8: Predictions of the models with different local forest cover scales (lines) for specialists and generalists in both regions.

Residual correlations among species

We evaluated the residuals by Kendall correlations among species and among sites for the 400 m models using the predictions for site:sp random effects (Observation Level Random Effect), following the code provided by Miller, Damschen & Ives (2018). Codes for the species names are presented in the dataset available.

Below we show models residual correlation plots for the specialists in the high-quality matrix landscape. All the other assemblages presented similar results.

- Range of species correlations: -0.41, 0.46.
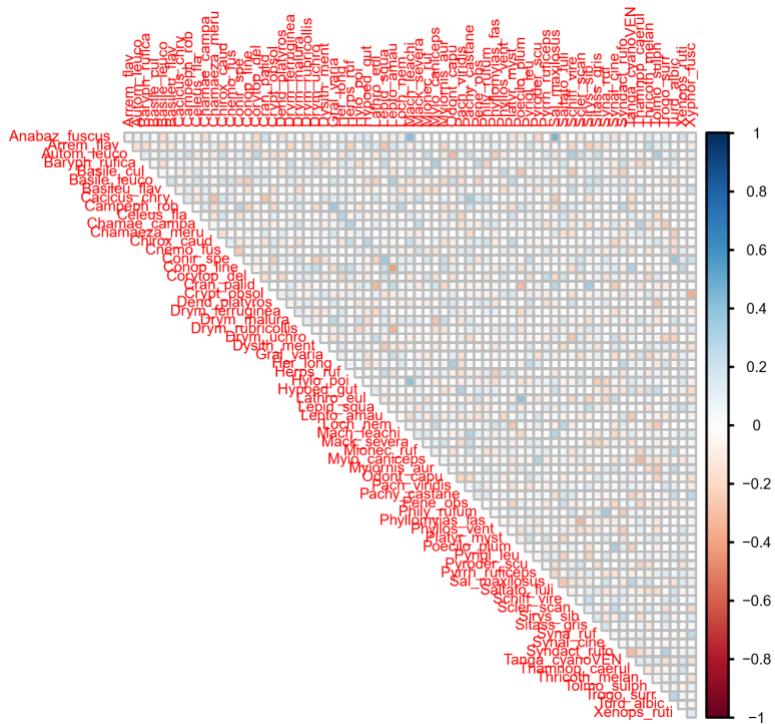- Range of sites correlations: -0.25, 0.31.

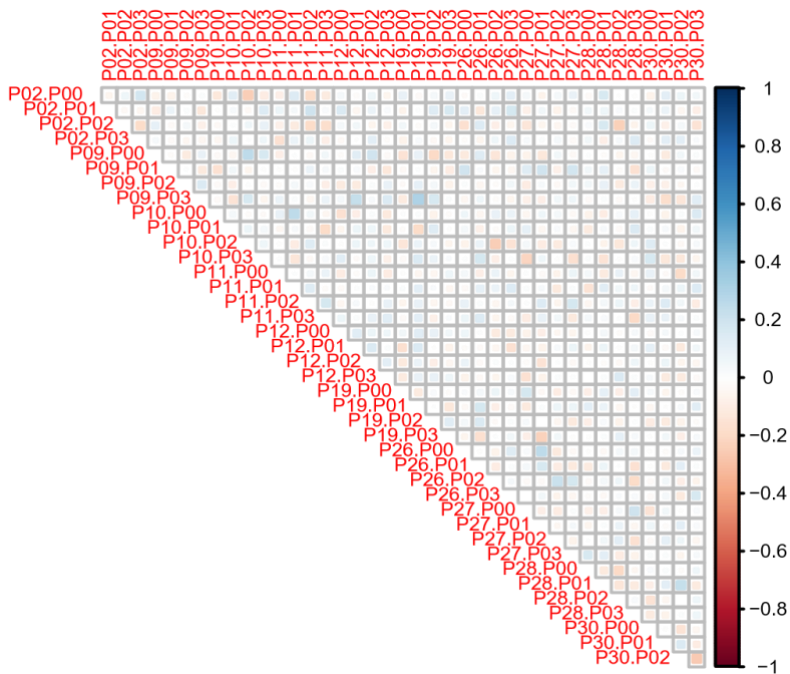**Figure S2.9:** Species residual Kendall correlations for the specialist species in the coffee matrix region.

Figure S2.10: Sites residual Kendall correlations for the specialist species in the coffee matrix region.
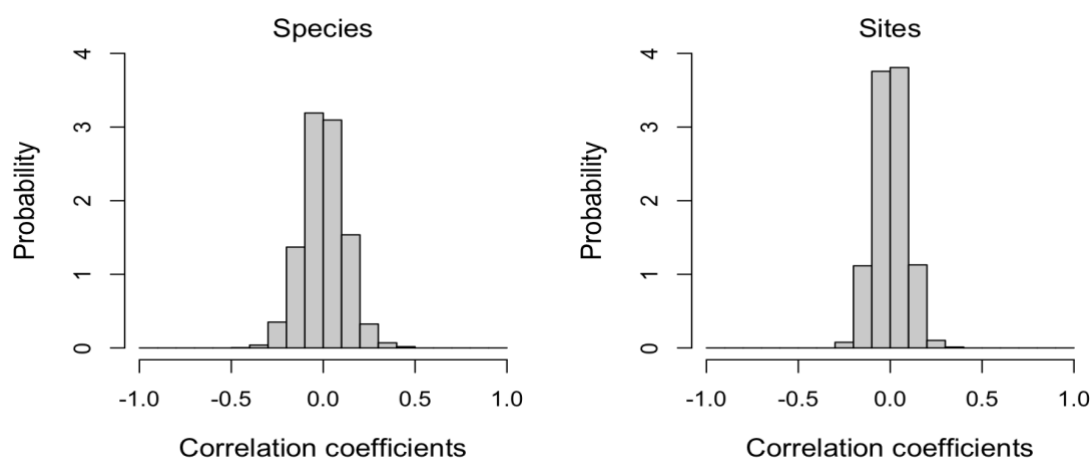
Figure S2.11: Histograms of the residual Kendall correlations for the specialists species in the coffee matrix region.

4. Including landscape forest cover

After selecting the local forest cover of 400 m radius buffer around each site for all datasets, we included the landscape forest cover (2 km radius buffer around the centroid of the landscape) in the model.

The R syntax example of this model area as follows:

```
model <- glmer(cbind(occor, n.visit-occor) ~ local.400 + landscape.2k
+ (local.400 + landscape.2k |sp) + (1|landscape:sp) + (1|site:sp) +
(1|landscape) + (1|site), family=binomial, data=high.spe)
```

Table S2.3: Variance Inflation Factor index for the variables of local forest cover and landscape forest cover.

	Local	Landscape
Specialists		
Coffee	1.26	1.26
Pasture	1.04	1.04
Generalists		
Coffe	1.13	1.13
Pasture	1.15	1.15

Before analysing results, we evaluated possible colinearity between local and landscape forest cover using the Variance Inflation Factor with the code provided by John Lefcheck (https://jonlefcheck. net/2012/12/28/dealing-with-multicollinearity-

using-variance-inflation-factors/). With VIF we found no evidence of collinearity between the forest cover scales (Table S2.3).

Predictions of the models are present in Figure S2.12. It is important to notice the differences in 20 and 40% landscape forest cover predictions for the specialists in the low-quality matrix.
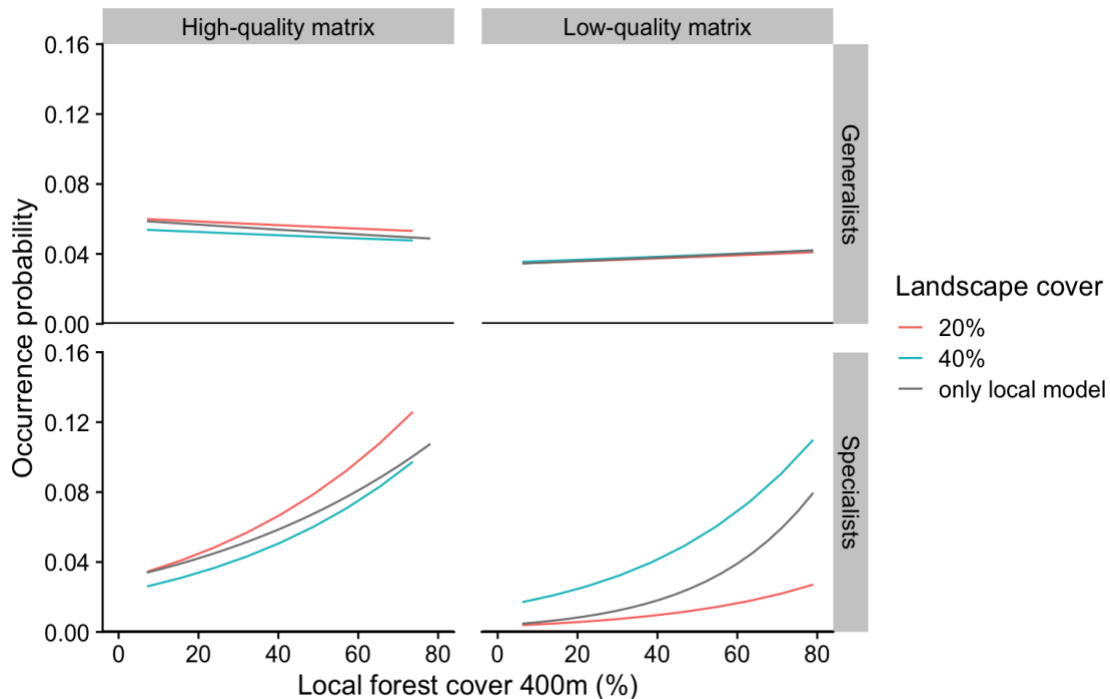

Figure S2.12: Predictions of the models without (gray lines) and with landscape forest cover scales (20 percent cover in red and 40 percent cover in blue lines) for specialists and generalists in both regions.

References

Barros, F.M., Martello, F., Peres, C.A., Pizo, M.A. & Ribeiro, M.C. (2019). Matrix type and landscape attributes modulate avian taxonomic and functional spillover across habitat boundaries in the Brazilian Atlantic Forest. Oikos 128, 1600–1612.

Bates, D., Mächler, M., Bolker, B. & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. Journal of Statistical Software 67, 1–48.

Boesing, A.L., Nichols, E. & Metzger, J.P. (2018). Biodiversity extinction thresholds are modulated by matrix type. Ecography 41, 1520–1533.

Jacoboski, L.I., Mendonça-Lima, A. de & Hartz, S.M. (2016). Structure of bird communities in eucalyptus plantations: Nestedness as a pattern of species distribution. Brazilian Journal of Biology 76, 583–591.

Miller, J.E.D., Damschen, E.I. & Ives, A.R. (2018). Functional traits and community composition: A comparison among community-weighted means, weighted correlations, and multilevel models. Methods in Ecology and Evolution 0.

Rodrigues, C.B., Taniwaki, R.H., Lane, P., Lima, W. de P. & Ferraz, S.F. de B. (2019). Eucalyptus Short-Rotation Management Effects on Nutrient and Sediments in Subtropical Streams.

Forests 10, 519.

Appendix S3: Additional models results

Melina Leite

2022-08-13

Models with traits

Specification of the models: We used lme4 package to perform a GLMM with binomial (proportion) distribution. An example of the code for each dataset are as follows:

```
mhigh.spe <- glmer(cbind(occor, n.visit-occor) ~

    forest_site400*lbody_size + forest_land*lbody_size +

    forest_site400*nest + forest_land*nest +

    forest_site400*diet + forest_land*diet +

    forest_site400*lower_stratum + forest_land*lower_stratum +

    (forest_site400 + forest_land|sp) +

    (1|landscape:sp) + (1|site:sp) +

    (lbody_size + nest + diet + lower_stratum|landscape) +

    (lbody_size + nest + diet + lower_stratum|site),

    family=binomial, data=high.spe, nAGQ = 1,

    control = glmerControl(optimizer = "bobyqa", optCtrl =
list(maxfun = 500000)))
```

We ran separate models for each assemblage and trait. Afterwards, we ran one model with the combination of the traits body mass, diet, nest type and % of lower strata use. Table S3.1 shows the marginal R^2 of all models terms.

Models coeficients

Tables S3.2, S3.3, S3.4, and S3.5 show the coefficients for each model.

Table S3.1: Overall and marginal R-squared of trait models in each dataset. For the marginal R-squared terms see Table 2 (main text).

Model	Total	trait*env	env|sp	lands:sp	site:sp	trait|lands	trait|site
Specialists							
High quality							
Combined	57.6	10.0	36.8	6.7	0.7	1.3	2.2
body mass	56.8	7.7	38.9	6.7	1.6	0.8	1.1
nest type	56.8	4.6	41.9	6.8	1.5	0.6	1.4
main diet	57.4	4.5	42.5	6.5	1.6	1.0	1.3
% frugivory	56.8	4.4	42.4	6.8	1.6	0.5	1.1
% insetivory	56.7	3.8	42.9	6.8	1.6	0.5	1.1
% lower strata	56.7	1.5	45.0	6.8	1.4	0.7	1.4
foraging stratum	56.7	2.0	44.5	6.5	1.5	0.8	1.3
hand-wing	56.6	1.6	44.9	6.8	1.2	0.5	1.6
Low quality							
Combined	71.9	22.2	39.9	3.8	1.0	0.9	4.0
body mass	70.5	15.8	45.6	3.8	2.4	0.1	2.7
nest type	71.0	17.3	44.5	3.9	2.2	0.0	3.1
main diet	71.4	18.4	43.8	3.7	2.5	0.4	2.5
% frugivory	71.3	17.2	45.1	3.8	2.6	0.0	2.6
% insetivory	71.3	18.4	44.0	3.8	2.6	0.1	2.5
% lower strata	71.1	15.5	46.2	3.8	1.9	0.4	3.2
foraging stratum	70.7	17.0	44.5	3.9	2.3	0.3	2.6
hand-wing	70.9	14.0	47.6	3.8	2.4	0.1	2.9
Generalists							
High quality							
Combined	46.2	7.4	31.9	1.3	0.9	1.6	3.1
body mass	44.7	0.2	37.8	1.6	2.9	0.6	1.5
nest type	44.6	0.6	37.4	1.7	3.1	0.6	1.1
main diet	45.8	5.2	33.7	1.5	1.7	1.1	2.7
% frugivory	44.5	0.5	37.4	1.5	3.6	0.7	0.8
% insetivory	44.4	3.2	34.8	1.2	3.7	0.9	0.7
% lower strata	44.5	1.3	36.7	1.5	3.6	0.7	0.7
foraging stratum	44.4	1.9	35.8	1.7	3.2	0.7	1.0
hand-wing	44.4	3.0	34.9	1.6	3.5	0.6	0.8
Low quality Combined	47.4	7.5	33.3	2.3	0.9	1.0	2.4
body mass	47.1	0.8	39.7	2.7	2.2	0.4	1.2
nest type	46.7	1.0	39.3	3.1	2.1	0.0	1.1
main diet	47.1	5.3	35.2	2.8	1.9	0.5	1.5
% frugivory	47.0	0.4	40.0	2.9	2.4	0.3	0.9
% insetivory	46.7	1.6	38.7	3.0	2.4	0.2	0.9
% lower strata	47.0	0.9	39.5	3.1	2.1	0.1	1.3
foraging stratum	47.0	1.2	39.2	2.9	2.4	0.1	1.1
hand-wing	46.7	2.5	37.8	3.0	2.4	0.1	0.9

Table S3.2: Fixed effects coefficients for the model of specialists in high-quality matrix landscapes.

effect	term	estimate	std.error	statistic	p.value
fixed	(Intercept)	-2.72	0.98	-2.78	0.01

effect	term	estimate	std.error	statistic	p.value
fixed	forest_site400	0.83	0.33	2.49	0.01
fixed	lbody_size	-0.62	0.28	-2.22	0.03
fixed	nestclosed	0.32	0.67	0.48	0.63
fixed	nestopen_semi	-0.28	0.56	-0.50	0.62
fixed	dietinsectivorous	0.07	0.83	0.09	0.93
fixed	dietonivorous	-1.00	1.25	-0.81	0.42
fixed	lower_stratum	0.18	0.23	0.78	0.44
fixed	forest_land	-0.36	0.41	-0.87	0.38
fixed	forest_site400:lbody_size	-0.03	0.09	-0.30	0.76
fixed	forest_site400:nestclosed	-0.38	0.19	-1.98	0.05
fixed	forest_site400:nestopen_semi	-0.24	0.17	-1.36	0.17
fixed	forest_site400:dietinsectivorous	-0.28	0.29	-0.97	0.33
fixed	forest_site400:dietonivorous	-0.35	0.44	-0.80	0.42
fixed	forest_site400:lower_stratum	0.01	0.07	0.10	0.92
fixed	lbody_size:forest_land	0.00	0.10	0.04	0.97
fixed	nestclosed:forest_land	0.30	0.22	1.35	0.18
fixed	nestopen_semi:forest_land	0.31	0.19	1.61	0.11
fixed	dietinsectivorous:forest_land	-0.04	0.33	-0.11	0.91
fixed	dietonivorous:forest_land	0.45	0.47	0.95	0.34
fixed	lower_stratum:forest_land	0.00	0.08	-0.03	0.98

Table S3.3: Fixed effects coefficients for the model of specialists in low-quality matrix landscapes.

effect	term	estimate	std.error	statistic	p.value
fixed	(Intercept)	-5.46	1.03	-5.33	0.00
fixed	forest_site400	1.06	0.36	2.94	0.00
fixed	lbody_size	-0.21	0.32	-0.67	0.50
fixed	nestclosed	0.70	0.78	0.90	0.37
fixed	nestopen_semi	-0.12	0.69	-0.18	0.86
fixed	dietgranivorous	0.13	1.82	0.07	0.95
fixed	dietinsectivorous	1.27	0.82	1.54	0.12
fixed	lower_stratum	0.36	0.28	1.28	0.20
fixed	forest_land	1.42	0.44	3.22	0.00
fixed	forest_site400:lbody_size	0.05	0.09	0.51	0.61
fixed	forest_site400:nestclosed	-0.39	0.22	-1.78	0.07
fixed	forest_site400:nestopen_semi	-0.06	0.20	-0.29	0.77
fixed	forest_site400:dietgranivorous	-0.41	0.61	-0.67	0.50
fixed	forest_site400:dietinsectivorous	-0.40	0.28	-1.45	0.15
fixed	forest_site400:lower_stratum	0.02	0.09	0.18	0.86
fixed	lbody_size:forest_land	0.01	0.12	0.04	0.96
fixed	nestclosed:forest_land	-0.59	0.28	-2.10	0.04
fixed	nestopen_semi:forest_land	-0.16	0.27	-0.62	0.54
fixed	dietgranivorous:forest_land	-0.03	0.73	-0.04	0.96
fixed	dietinsectivorous:forest_land	-0.38	0.34	-1.10	0.27
fixed	lower_stratum:forest_land	-0.15	0.12	-1.24	0.22

Table S3.4: Fixed effects coefficients for the model of generalists in high-quality matrix landscapes.

effect	term	estimate	std.error	statistic	p.value
fixed	(Intercept)	-3.45	0.55	-6.29	0.00
fixed	forest_site400	0.37	0.19	2.01	0.04

effect	term				
fixed	lbody_size	0.08	0.27	0.29	0.77
fixed	nestclosed	0.20	0.61	0.32	0.75
fixed	nestopen_semi	0.20	0.53	0.38	0.71
fixed	dietgranivorous	-0.91	0.86	-1.06	0.29
fixed	dietinsectivorous	1.00	0.50	2.00	0.05
fixed	dietnectarivorous	-0.12	0.75	-0.16	0.88
fixed	dietonivorous	0.14	0.67	0.21	0.83
fixed	lower_stratum	0.36	0.17	2.06	0.04
fixed	forest_land	-0.11	0.22	-0.50	0.62
fixed	forest_site400:lbody_size	0.05	0.09	0.61	0.54
fixed	forest_site400:nestclosed	-0.43	0.20	-2.12	0.03
fixed	forest_site400:nestopen_semi	-0.22	0.17	-1.28	0.20
fixed	forest_site400:dietgranivorous	-0.61	0.35	-1.75	0.08
fixed	forest_site400:dietinsectivorous	-0.13	0.18	-0.75	0.46
fixed	forest_site400:dietnectarivorous	-0.16	0.27	-0.61	0.54
fixed	forest_site400:dietonivorous	-0.50	0.25	-2.04	0.04
fixed	forest_site400:lower_stratum	0.04	0.05	0.71	0.48
fixed	lbody_size:forest_land	-0.01	0.10	-0.09	0.93
fixed	nestclosed:forest_land	0.18	0.21	0.88	0.38
fixed	nestopen_semi:forest_land	0.07	0.17	0.38	0.71
fixed	dietgranivorous:forest_land	0.68	0.36	1.87	0.06
fixed	dietinsectivorous:forest_land	-0.13	0.19	-0.65	0.51
fixed	dietnectarivorous:forest_land	-0.11	0.28	-0.40	0.69
fixed	dietonivorous:forest_land	0.06	0.27	0.22	0.83
fixed	lower_stratum:forest_land	-0.03	0.06	-0.58	0.56

Table S3.5: Fixed effects coefficients for the model of generalists in low-quality matrix landscapes.

effect	term	estimate	std.error	statistic	p.value
fixed	(Intercept)	-4.23	0.49	-8.66	0.00
fixed	forest_site400	0.26	0.20	1.32	0.19
fixed	lbody_size	0.22	0.26	0.83	0.41
fixed	nestclosed	0.79	0.59	1.33	0.18
fixed	nestopen_semi	0.55	0.50	1.10	0.27
fixed	dietgranivorous	0.87	1.12	0.78	0.44
fixed	dietinsectivorous	1.06	0.48	2.20	0.03
fixed	dietnectarivorous	-0.19	0.76	-0.25	0.80
fixed	dietonivorous	0.39	0.64	0.62	0.54
fixed	lower_stratum	0.24	0.18	1.33	0.18
fixed	forest_land	0.04	0.18	0.25	0.81
fixed	forest_site400:lbody_size	0.03	0.10	0.32	0.75
fixed	forest_site400:nestclosed	-0.52	0.21	-2.45	0.01
fixed	forest_site400:nestopen_semi	-0.26	0.17	-1.49	0.14
fixed	forest_site400:dietgranivorous	0.55	0.39	1.41	0.16
fixed	forest_site400:dietinsectivorous	0.02	0.19	0.12	0.90
fixed	forest_site400:dietnectarivorous	0.09	0.30	0.30	0.77
fixed	forest_site400:dietonivorous	-0.05	0.23	-0.21	0.84
fixed	forest_site400:lower_stratum	-0.10	0.06	-1.62	0.11
fixed	lbody_size:forest_land	0.00	0.09	-0.05	0.96
fixed	nestclosed:forest_land	0.27	0.18	1.46	0.14
fixed	nestopen_semi:forest_land	0.20	0.15	1.37	0.17
fixed	dietgranivorous:forest_land	-0.48	0.34	-1.42	0.15

fixed	dietinsectivorous:forest_land	-0.21	0.17	-1.23	0.22
fixed	dietnectarivorous:forest_land	-0.26	0.28	-0.94	0.35
fixed	dietonivorous:forest_land	-0.25	0.20	-1.27	0.21
fixed	lower_stratum:forest_land	0.07	0.06	1.17	0.24

Models diagnostic

Variance Inflation Factor of the model parameters for each dataset in Table S3.6.

Table S3.6: Variance Inflation Factor index for combined traits models in each dataset.

	Specialists		Generalists	
parameter	High-quality	Low-quality	High-quality	Low-quality
forest.local	1.24	1.04	1.17	1.17
body_mass	1.23	1.08	1.15	1.15
nest_closed	1.78	1.43	2.37	2.37
nest_open_semi	2.15	1.99	1.96	1.96
diet_insectivorous	1.83	1.94	2.16	2.16
diet_onivorous	1.83	1.49	2.04	2.04
lower_strata	1.37	NA	2.22	2.22
diet_granivorous	1.18	1.19	1.15	1.15
forest.landscape	NA	1.16	1.20	1.20
diet_nectarivorous	NA	NA	2.46	2.46

Example of the residual diagnostic of the model with the combined traits (main diet, body mass, nest type and % of lower strata use) for the forest specialists in high-quality matrix landscapes. The models' diagnostics for the other assemblages were all similar and can be checked in the original repository of the analysis.

Residual correlations among species and sites

Below we present the Kendall correlations for the residuals among species and sites for the models using the predictions for site:sp random effect (Observation Level Random Effect). For the residual correlations we followed the code provided by Miller, Damschen & Ives (2018).

- Range of species correlations: -0.4, 0.43.
- Range of sites correlations: -0.3, 0.27.
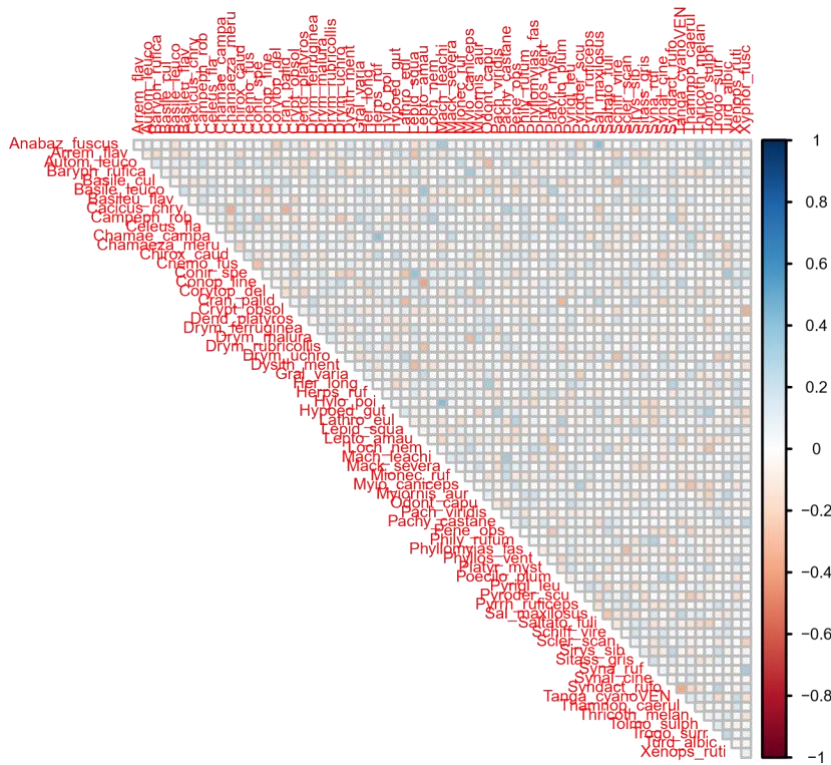
Figure S3.1: Species residual Kendall correlations for the specialist species in high-quality matrix landscapes.
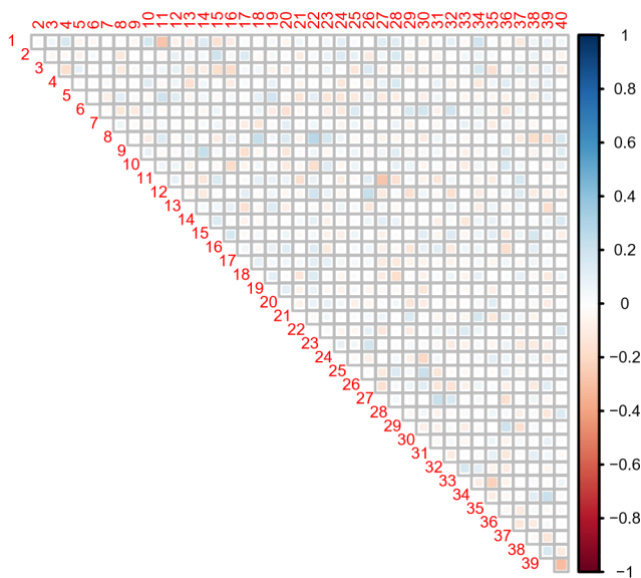
Figure S3.2: Sites residual Kendall correlations for the specialist species in high-quality matrix landscapes.
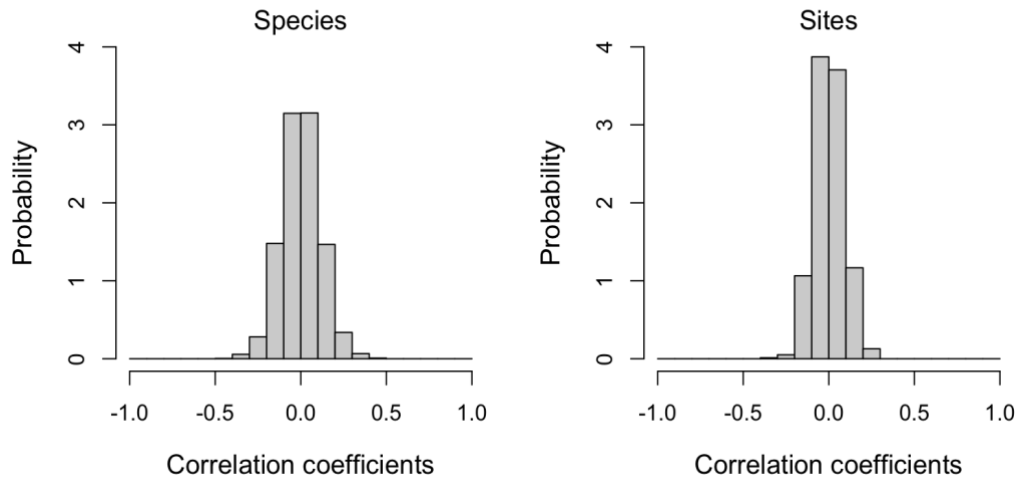
Figure S3.3: Histograms of the residual Kendall correlations for the specialists species in high-quality matrix landscapes.

Residual diagnostic

We used DHARMa package (Hartig (2018)) for the diagnostic of quantile residuals.
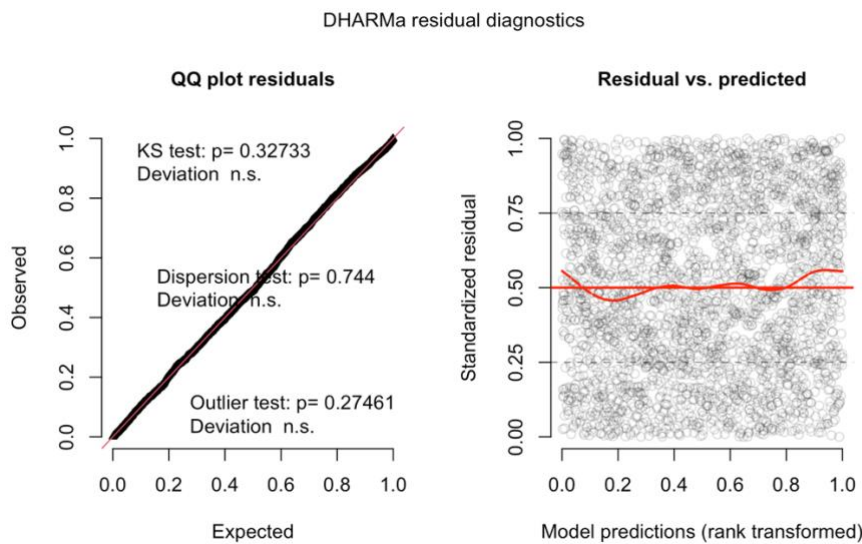
Figure S3.4: Plots for model diagnostic form DHARMa package.

Residuals against predictors:
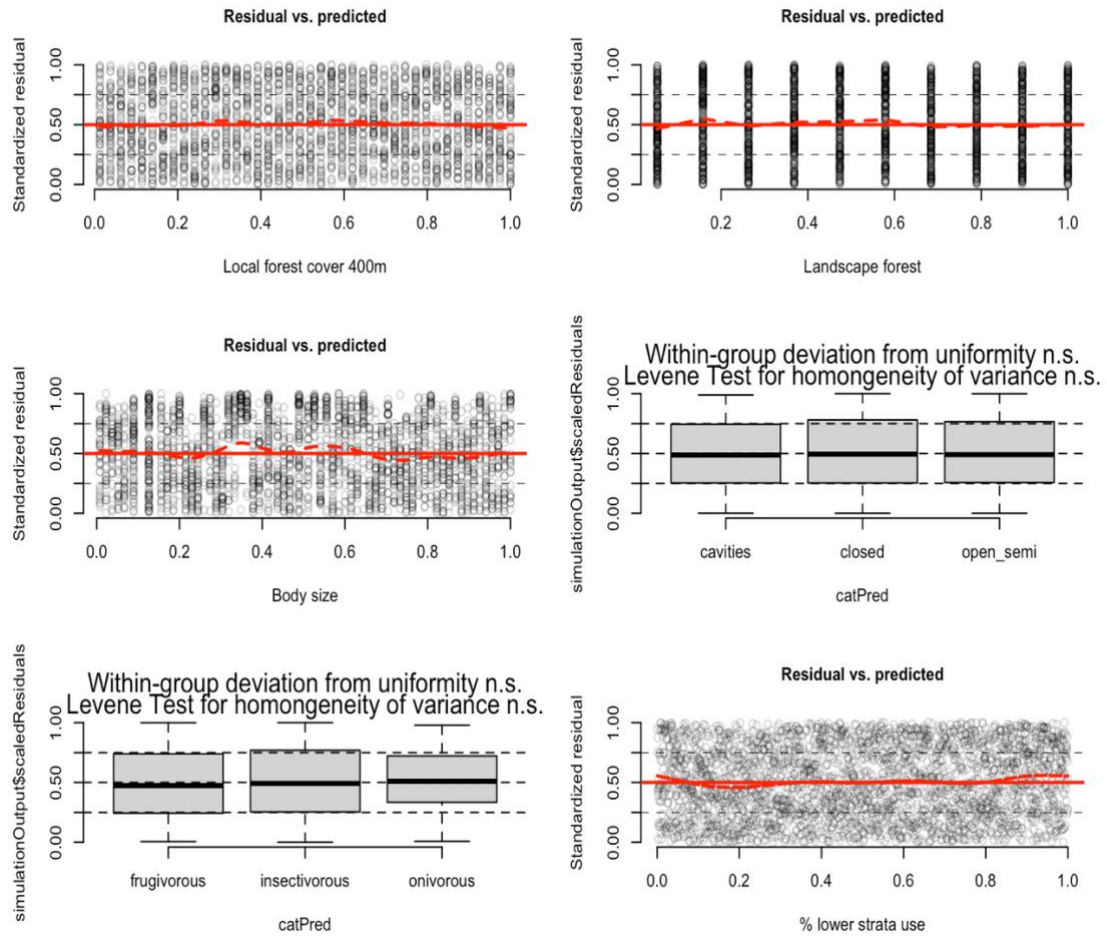
Figure S3.5: More for model diagnostic form DHARMa package.

Predictions for each species local forest cover
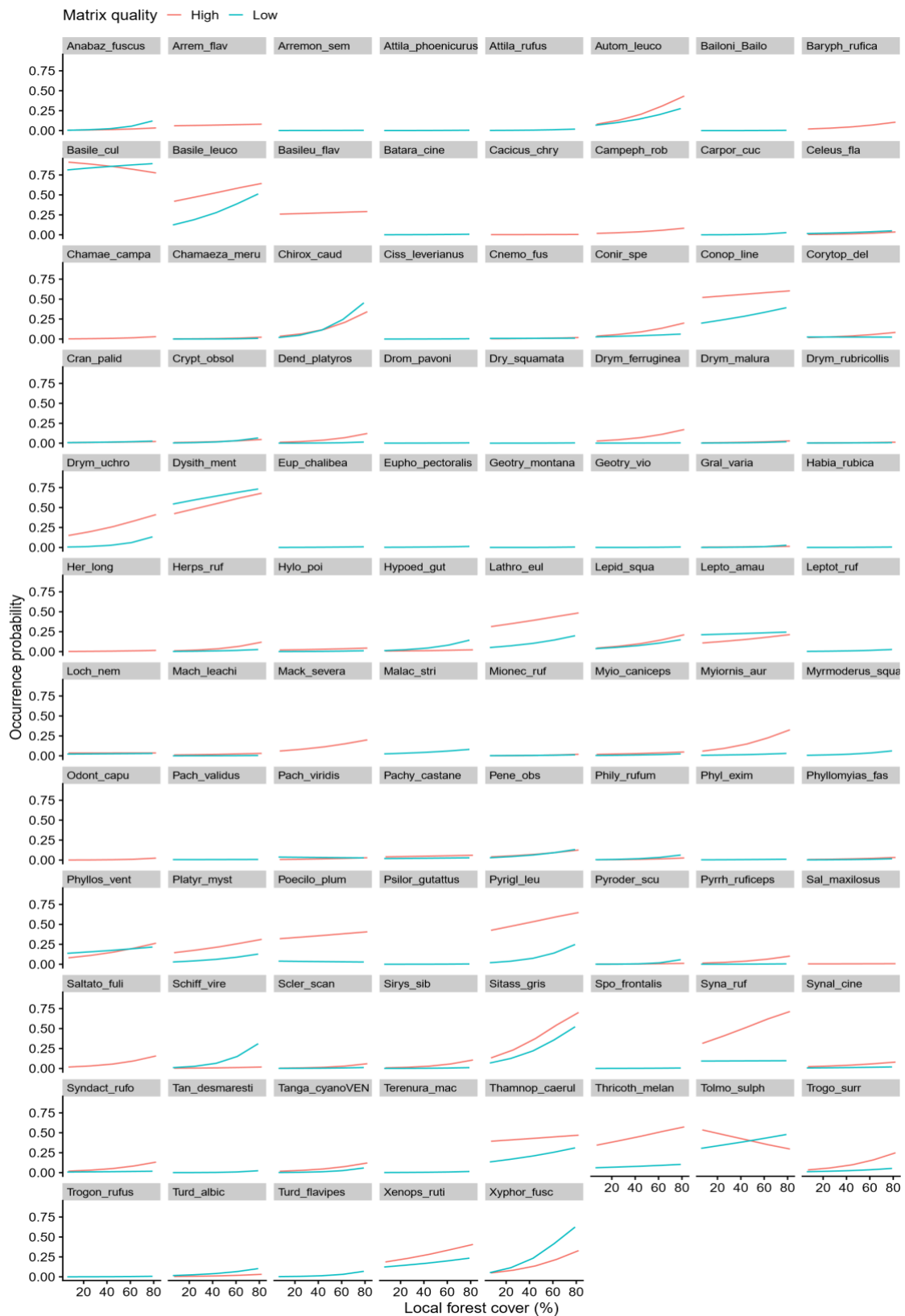
Figure S3.6: Forest specialist birds. Landscape forest cover was fixed in 30%.
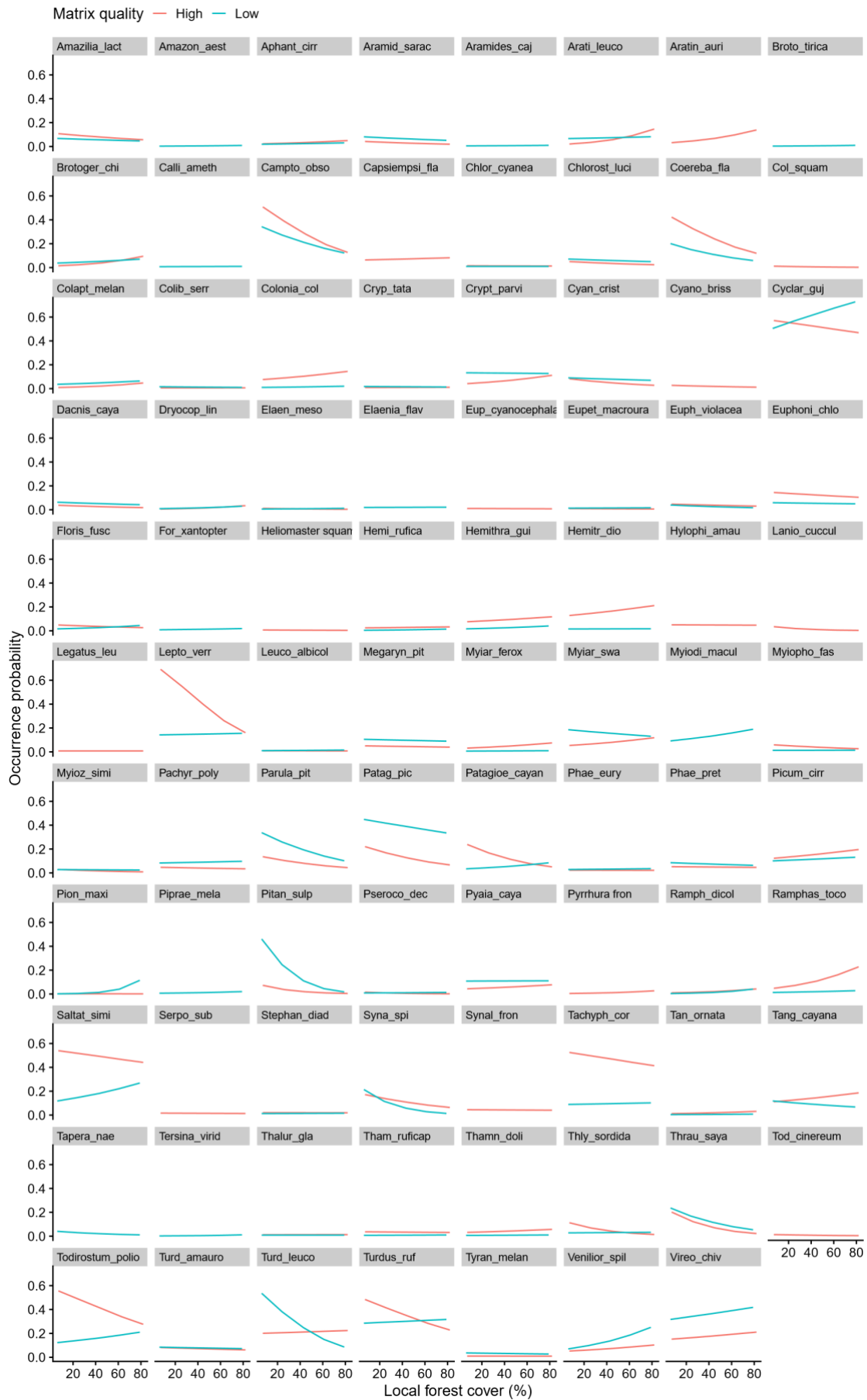
Figure S3.7: Forest specialist birds. Landscape forest cover was fixed in 30%.

References

Hartig, F. (2018). DHARMa: Residual Diagnostics for Hierarchical (Multi-Level / Mixed) Regression Models.

Miller, J.E.D., Damschen, E.I. & Ives, A.R. (2018). Functional traits and community composition: A comparison among community-weighted means, weighted correlations, and multilevel models. Methods in Ecology and Evolution 0.

Appendix S4: Robustness analysis with species that occur at both high and low-quality landscapes

Melina Leite

2022-08-13

Given that bird species pool are slightly different in both regions (see Appendix 1), the results we found comparing both high- and low-quality landscapes could be confounded by groups os species that respond differently but doesn't occur at both regions. In order to check for the robustness of our results, we performed separate analysis with only those species that occur in both regions, **57 forest specialists** (61.3% of all specialists) and **64 forest generalists** (73.6% of all generalists), and compared them with the results presented in the main text with all species.

The analyses were done with the combined trait models, the same presented at Figure 2 and 3 in the main text.

We found that even with the same groups of species analyzed, we get very similar results for the differences between high- and low-quality matrix landscapes, which reinforce that the drawned inferences regarding habitat loss responses remains the same regardless of the exclusion of 'unique' species in a given a region.

Datasets

Separating only species that occur at both regions.

Table S4.1: Checking the number of species in all 4 assemblages datasets.

assemblage	N_of_species
High-quality specialists	57
Low-quality specialists	57
High-quality generalists	64
Low-quality generalists	64

Local and landscape forest cover effects

Comparing effect sizes and 95% IC for fixed effects forest cover a local and landscape scale for each dataset: the original and the subsets of the same species (1). We see that estimates almost didn't change with the dataset.
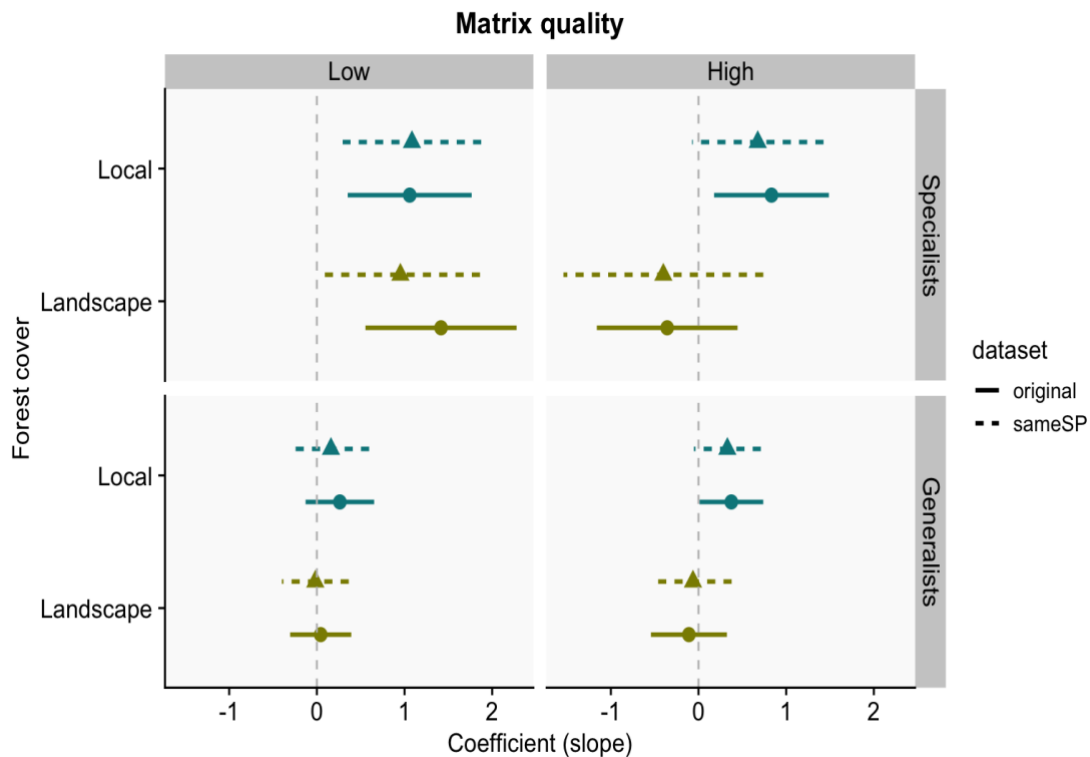
Figure S4.1: Comparing effect sizes (and 95 percent confidence intervals) for forest cover at local and landscape scale for both datasets: original with all species (triangles and continuous lines) and the subset of species that occur at both regions (circles and dashed lines).

R2 models

Comparing partial R^2 of the models' terms in both datasets, the original and the subsets of the same species (2). Qualitatively, we see that R^2 almost didn't change with the dataset, expect for specialists in low-quality matrix landscapes where fixed effects (trait*env) decreased from 22% to 17%.
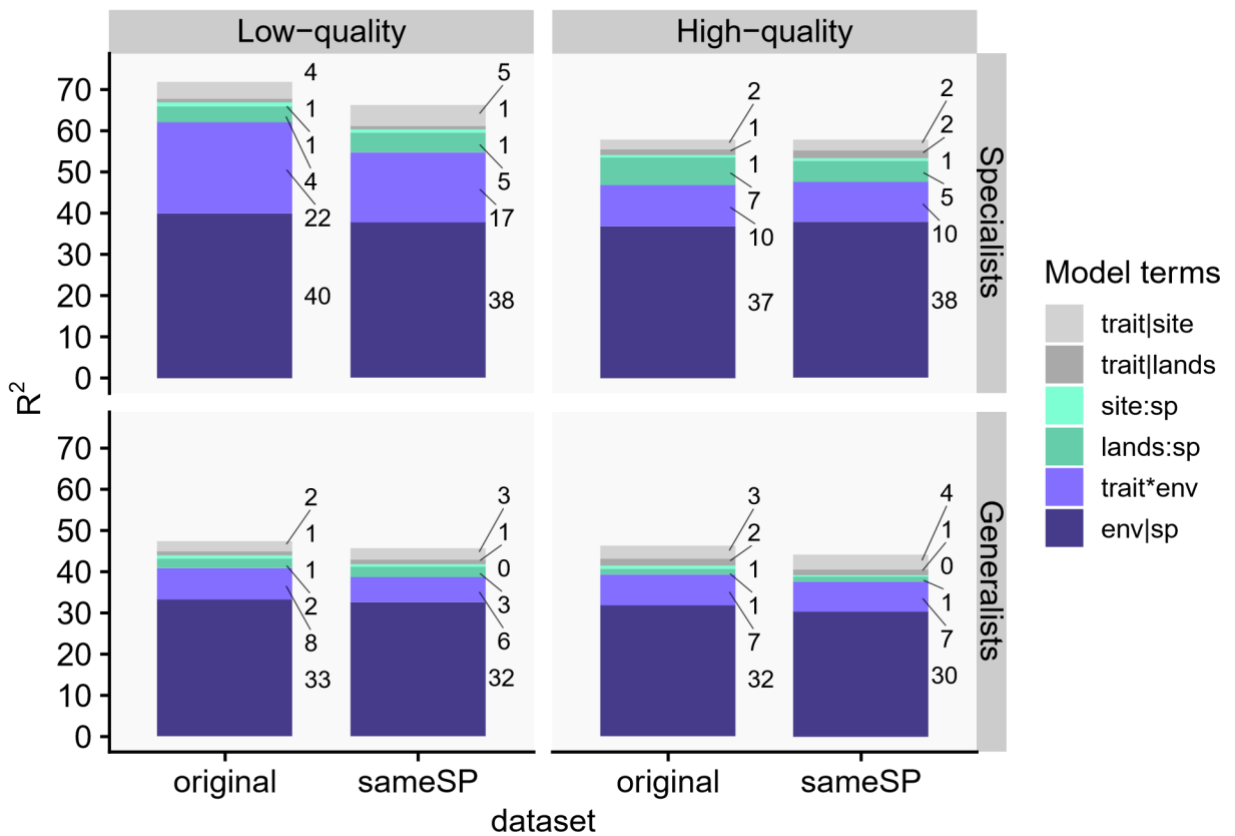
Figure S4.2: Comparing partial R-squares for models terms in both datasets: original with all species (left) and the subset of species that occur at both regions (right).

Predictions for each species local forest cover

Predictions of species occurrence probability with forest cover at the local scale for the species that occur at both regions with models fitted for both datasets: the original and the subsets of the same species for forest specialists (3) and habitat generalists (4). Landscape forest cover was fixed in 30%.
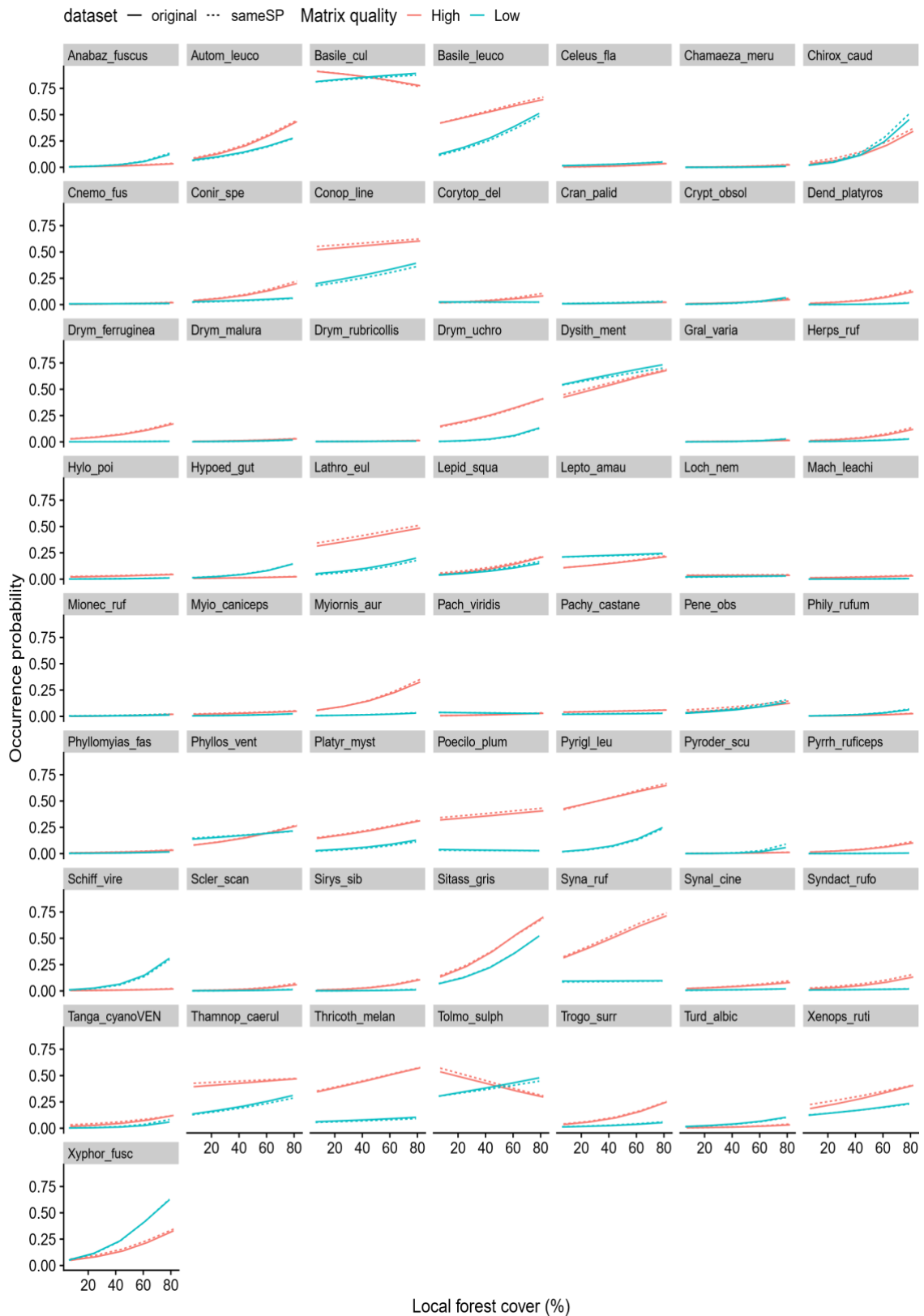
Figure S4.3: Predictions for forest specialist birds.
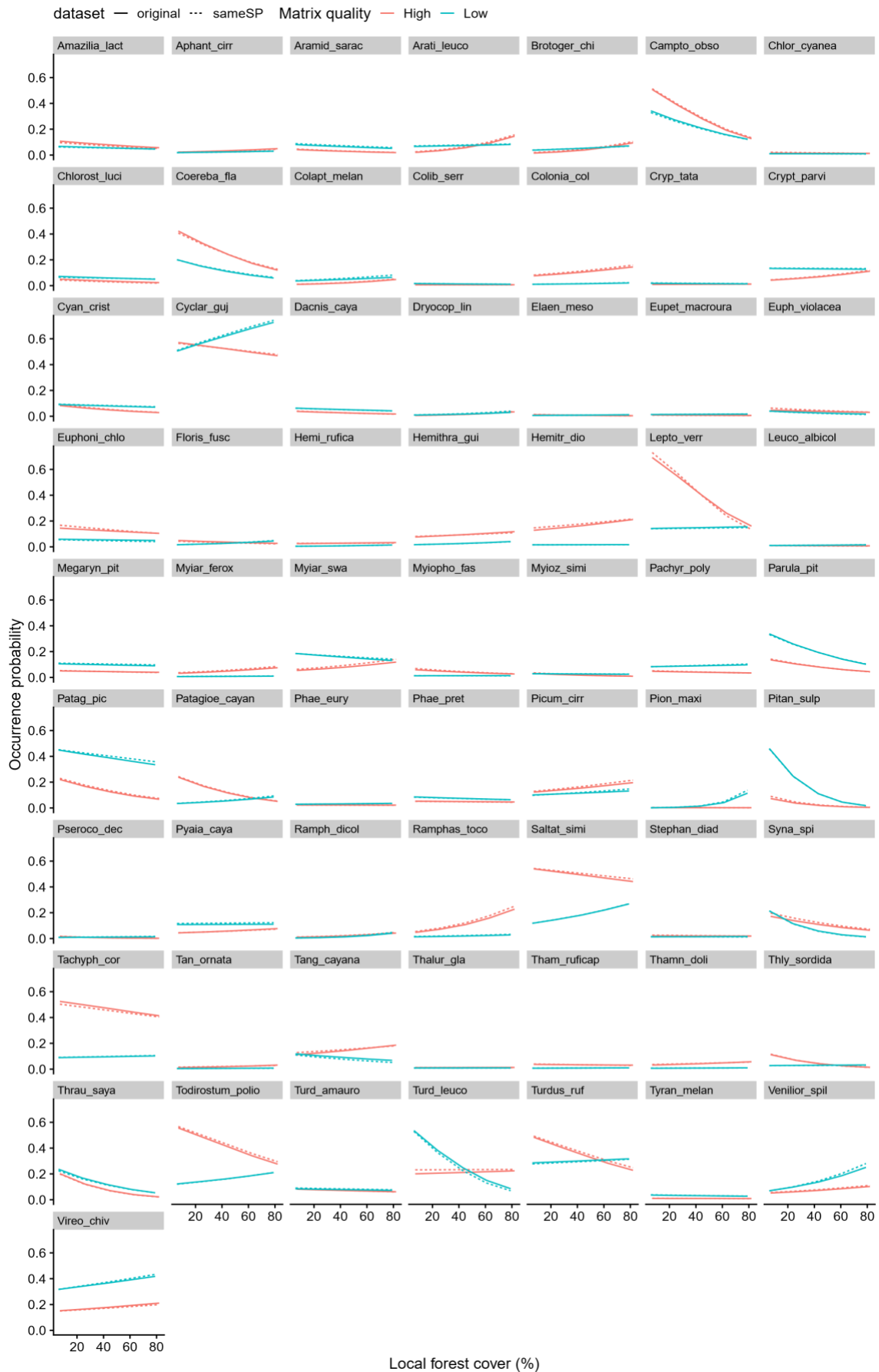
Figure S4.4: Predictions for habitat generalist birds.

Anexos 2 – material suplementar Capítulo 2

Appendix S1- Forest census data

Forest plots data preparation

In every forest plot dataset, each observation is an individual tree. For trees with more than one stem, the individual is considered alive if at least one of the stems is alive and dead if all the stems are dead. The diameter at breast height (DBH) for trees with multiple stems was calculated based on the sum of the basal area of the stems, considering the diameter of a circle.

We excluded ferns and palms from all analyses due to their non-standard growth form, and lack of secondary growth. We also excluded trees without information on x and y plot coordinates, species name, status (alive, dead or recruit), and date of measurement. Although we excluded individuals with unknown or unidentified species name, we kept morphospecies classification when existing in the data. For growth analysis, we excluded individuals with different heights of measurements of DBH in consecutive censuses and trees with growth rates more than four standard deviations from the mean, as they are likely measurement errors (Rüger et al. 2011; Condit et al. 2017).

Environmental, climatic and vegetational information of each ForestGEO plot used are in Table S1.1 and the summary information of the data used in the analysis are in Table S1.2.

References

Rüger, N., Berger, U., Hubbell, S.P., Vieilledent, G. & Condit, R. (2011) Growth Strategies of Tropical Tree Species: Disentangling Light and Size Effects. *PLoS ONE*, **6**, e25330.

Condit, R., Pérez, R., Lao, S., Aguilar, S. & Hubbell, S.P. (2017) Demographic trends and climate over 35 years in the Barro Colorado 50 ha plot. *Forest Ecosystems*, **4**, 17.

TABLE S1.1: Information of the ForestGEO plots. Data on environmental and climatic variables are from Anderson-Teixeira et al. (2015). Abbreviations and units in columns: Lat latitude, Long longitude, Köppen Climate classification zone*; MAT mean annual temperature in °C; MAP mean annual precipitation in mm/year; PET annual potential evapotranspiration in mm/day; Dominant soil Classification**; Dominant vegetation type*** Natural disturbance regime****.

Forest Plot	Latitudinal zone	Country	Lat	Long	Elevation m (min-max)	Köppen	MAT	MAP	PET	Dominant soil	Dominant vegetation	Natural disturbances
Amacayacu	Tropical	Colombia	-3.81	-70.27	89-111	Af	25.8	3216	1010	Ult	BE	FI; W; In
Barro Colorado Island	Tropical	Panama	9.15	-79.85	120-160	Am	27.1	2551	1311	Ox	BdD; BE	D; W
Fushan	Subtropical	Taiwan	24.76	121.56	600-733	Cfa	18.2	4271	1085	Ult; In	BE	H
Ilha do Cardoso	Subtropical	Brazil	-25.096	-47.9573	3-8	Cfa	22.4	2100	-	S	BE	-
Ituri - Edoro	Tropical	Democratic Republic of Congo	1.44	28.583	700-850	Af	24.3	1682	1168	Ox	BE	W; A
Ituri - Lenda	Tropical	Democratic Republic of Congo	1.44	28.583	700-850	Af	24.3	1682	1168	Ox	BE	W; A
Korup	Tropical	Cameroon	5.07	8.85	150-240	Am	26.6	5272	1050	Ult; Ox	BE	W
Lambir	Tropical	Malaysia	4.19	114.02	104-244	Af	26.6	2664	1114	Ult	BE	L; D
Lilly Dickey Woods	Temperate	USA	39.24	-86.22	230-303	Cfa	11.6	1203	981	In; Ult; Alf	BcD	W; D; Ic
La Planada	Tropical	Colombia	1.16	-77.99	1796-1840	Cfb	19	4087	-	An	BE	W
Luquillo	Tropical	Puerto Rico, USA	18.33	-65.82	333-428	Am	22.8	3548	1219	Ox; Ult	BE	H; L
Mo Singto	Tropical	Thailand	14.43	101.35	725-815	Aw	23.5	2100	1300	NA	BE; BdD	W
Pasoh	Tropical	Malaysia	2.98	102.31	70-90	Af	27.9	1788	1120	Ult	BE	W
Smithsonian Conservation Biology Institute	Temperate	USA	38.89	-78.15	273-338	Cfa	12.9	1001	1003	Alf	BcD	W, Ic

Smithsonian Environmental Research Center	Temperate	USA	38.89	-76.56	6-10	Cfa	13.2	1068	1111	Ult; In; En	BcD	H; W
Sinharaja	Tropical	Sri Lanka	6.40	80.40	424-575	Af	22.5	5016	1384	Ult	BE	W
University of California Santa Cruz	Temperate	USA	37.01	-122.08	314-332	Csb	14.8	778	-	Mo	BcD	W, Ic
Wabikon	Temperate	USA	45.55	-88.79	488-514	Dfb	4.2	805	-	Alf	BdC	W
Wind River	Temperate	USA	45.82	-121.96	352-385	Csb	9.2	2495	770	An	NE	Fi; W; In
Wytham Woods	Temperate	United Kingdom	51.77	-1.34	104-163	Cfb	10	717	637	E	BcD	-
Zofin	Temperate	Czech Republic	48.66	14.71	735-825	Cfb	6.2	866	-	S; In; Hi	BdC; NE	W; In

*Af: Tropical with significant precipitation year-round; Am: Tropical monsoon; Aw: Tropical wet and dry; Csb subtropical/mid-latitude climate with dry summers (a.k.a.: Warm-summer Mediterranean); Cfa: Humid subtropical/mid-latitude climate with significant precipitation year-round; Cfb: Oceanic with significant precipitation year-round; Dfb: Humid Continental with significant precipitation year-round.

** Alf, Alfisols; An, Andisols; E, Entisoils; Ge, Gelisols; Hi, Histosols; In, Inceptisols; Ox, Oxisols; Ult, Ultisols; S, Spodosols; Ve, Vertisols.

*** BE, broadleaf evergreen; BdD, broadleaf drought deciduous; BcD, broadleaf cold deciduous; NE, needleleaf evergreen.

****A, animal activity (destructive); D, Drought; E, Erosion; Fi, Fire; Fl, flood; H, hurricane/typhoon; Ic, Ice storms; Insect outbreaks; L, landslides; PT, permafrost thaw; W, wind storms (local).

TABLE S1.2. Summary information of the data used in the analysis. Number of species may include morphospecies.

ID	Forest Plot	Original plot size (ha)	Trimmed plot size (ha)	First census (year)	Last census (year)	Number of census intervals	Total Number of species			Total number of observations		
							Growth	Mortality	Recruitment	Growth	Mortality	Recruitment
ama	Amacayacu	25	25	2007	2017	1[+]	1156	1269	-	76580	105357	-
bci	Barro Colorado Island	50	50	1981	2016	7*	305	313	316	1310125	1558414	1620133
fus	Fushan	25	25	2003	2019	3	105	107	105	267550	325882	341664
idc	Ilha do Cardoso	10.24	9	2009	2019	1	116	117	130	17529	19081	25306
edo	Ituri - Edoro	20	20	1994	2007	2	388	412	417	280412	313806	305282
len	Ituri - Lenda	20	20	1994	2007	2	382	396	399	234991	264914	259444
kor	Korup	50	50	1997	2010	1	449	468	461	280703	327121	321582
lam	Lambir	52	50	1991	2009	3	1362	1402	1376	915971	1073643	1065051
ldw	Lilly Dickey Woods	25	25	2012	2017	1	33	36	33	20596	26496	23059
lpl	La Planada	25	25	1997	2003	1	203	205	225	71713	89251	89359
luq	Luquillo	16	15	1990	2016	3	134	146	145	106355	166558	156824
mos	Mo Singto	31	30	2003	2017	2	266	272	275	227085	273510	394809
pas	Pasoh	50	50	1986	2011	5*	880	891	891	1357235	1555377	1588440
scbi	Smithsonian Conservation Biology Institute	25.6	24	2008	2013	1	57	66	60	24265	29022	33812
serc	Smithsonian Environmental Research Center	16	16	2008	2014	1	65	71	70	19834	23200	24156
sin	Sinharaja	25	25	1993	2008	2	231	234	231	355214	399749	381212

ucsc	University of California Santa Cruz	16	6	2006	2020	2	30	31	30	12873	15753	14634
wab	Wabikon	25.6	24	2008	2018	2	33	36	34	77167	92258	84133
wfdp	Wind River	25.6	24	2010	2016	1	24	26	25	22979	25354	24420
wyw	Wytham Woods	18	18	2008	2021	4[+]	25	25	-	54198	57925	-
zof	Zofin	25	25	2012	2017	1	11	11	12	57445	58344	72764

*for growth rates, Barro Colorado Island had 6 and Pasoh 4 census intervals due to problems with DBH measurements in the first census.

[+] Recruitment rates for Amacayacu and Wytham Woods could not be analyzed. Given that, Wytham Woods was not analyzed with temporal data.

Forest plots acknowledgments and references

Amacayacu

The 25-ha Long-Term Ecological Research Project of Amacayacu is a collaborative project of the Instituto Amazónico de Investigaciones Científicas Sinchi and the Universidad Nacional de Colombia Sede Medellín, in partnership with the Unidad de Manejo Especial de Parques Naturales Nacionales and the Forest Global Earth Observatory of the Smithsonian Tropical Research Institute (ForestGEO). The Amacayacu Forest Dynamics Plot is part of ForestGEO, a global network of large-scale demographic tree plots. We acknowledge the Director and staff of the Amacayacu National Park for supporting and maintaining the project in this National Park.

Barro Colorado Island

The BCI forest dynamics research project was made possible by National Science Foundation grants to Stephen P. Hubbell: DEB-0640386, DEB-0425651, DEB-0346488, DEB-0129874, DEB-00753102, DEB-9909347, DEB-9615226, DEB-9615226, DEB-9405933, DEB-9221033, DEB-9100058, DEB-8906869, DEB-8605042, DEB-8206992, DEB-7922197, support from the Forest Global Earth Observatory, the Smithsonian Tropical Research Institute, the John D. and Catherine T. MacArthur Foundation, the Mellon Foundation, the Small World Institute Fund, and numerous private individuals, and through the hard work of over 100 people from 10 countries over the past three decades. The plot project is part of the Forest Global Earth Observatory (ForestGEO), a global network of large-scale demographic tree plots.

References

Condit R., Perez, R., Aguilar, S., Lao, S., Foster, R., Hubbell, S.P. 2019. Complete data from the Barro Colorado 50-ha plot: 423617 trees, 35 years, 2019 version.

Condit, R. 1998. Tropical Forest Census Plots. Springer-Verlag and R. G. Landes Company, Berlin, Germany, and Georgetown, Texas.

Hubbell, S.P., R.B. Foster, S.T. O'Brien, K.E. Harms, R. Condit, B. Wechsler, S.J. Wright, and S. Loo de Lao. 1999. Light gap disturbances, recruitment limitation, and tree diversity in a neotropical forest. Science 283: 554-557.

Condit R., Perez, R., Aguilar, S., Lao, S., Foster, R., Hubbell, S.P. 2019. BCI 50-ha plot taxonomy, 2019 version.

Fushan

The Fushan Forest Dynamics plot (FDP) is supported by the Taiwan Forestry Bureau, the Taiwan Forestry Research Institute and the Ministry of Science and Technology of Taiwan. We would like to express our gratitude to all field technicians and students who helped with the implementation and recensus of the Fushan FDP. We also thank the Fushan Research Center staff for providing logistic support.

Ilha do Cardoso

The 10.24ha Ilha do Cardoso Forest Dynamics Plot was established and has been supported by São Paulo Research Foundation (FAPESP grants #1999/09635-0 and #2017/11979-9), University of São Paulo (USP), Forest Global Earth Observatory and the Smithsonian Tropical Research Institute. The Principal Investigator (AAdeO) received fellowship grant from Brazil National Council for Scientific and Technological Development (CNPq #313829/2021-7) and would like to express his gratitude to many field workers, research technicians, and staff who helped during decades, working at field and USP, to make this data available.

References

Oliveira, A.A. de, Vicentini, A., Chave, J., Castanho, C. de T., Davies, S.J., Martini, A.M.Z., Lima, R.A.F., Ribeiro, R.R., Iribar, A., and Souza, V.C. (2014). Habitat specialization and phylogenetic structure of tree species in a coastal Brazilian white-sand forest. Journal of Plant Ecology 7, 134–144. 10.1093/jpe/rtt073.

Lima, R.A.F. de, Oliveira, A.A. de, Martini, A.M.Z., Sampaio, D., Souza, V.C., and Rodrigues, R.R. (2011). Structure, diversity, and spatial patterns in a permanent plot of a high Restinga forest in Southeastern Brazil. Acta Bot. Bras. 25, 633–645. 10.1590/S0102-33062011000300017.

Ituri - Edoro and Lenda

The Ituri 40-ha plot program is a collaborative project between the Centre de Formation et de Recherche en Conservation Forestière, the Wildlife Conservation Society – DRC through his conservation project in the Okapi forest Reserve, in partnership with the Forest Global Earth Observatory (ForestGEO). The Ituri plots are financially supported by the Wildlife Conservation Society, the Frank Levinson Family Foundation, and ForestGEO. The Institut Congolais pour la Conservation de la Nature graciously provided the research permit.

References

Makana J-R., Hart T.B., Liengola I., Ewango C.N.E., Hart J.A. & Condit R. (2004). Ituri Forest

Dynamics Plots, Democratic Republic of Congo. In Losos E.C. and Leigh, E.G. Jr. eds. Forest Diversity and Dynamism: Findings from a Large-Scale Plot Network. 506-516. University of Chicago Press, Chicago.

Korup

The 50-ha is a collaborative project of the University of Buea, Cameroon, and the World Wide Fund for Nature, Cameroon Program in partnership with the Forest Global Earth Observatory of the Smithsonian Tropical Research Institute (ForestGEO). Funding for the first census was provided by the International Cooperative Biodiversity Group (a consortium of the NIH, the NSF, and the USDA), with supplemental funding by the Central Africa Regional Program for the Environment (a program of USAID). Funding for the second census was provided by the Frank Levinson Family Foundation. Permission to conduct the field program in Cameroon is provided by the Ministry of Environment and Forests and the Ministry of Scientific Research and Innovation.

References

Chuyong G.B., Condit, R., Kenfack, D., Losos, E., Sainge, M., Songwe, N.C., and Thomas, D.W. (2004). Korup Forest Dynamics Plot, Cameroon. In Losos E.C. and Leigh, E.G. Jr. eds. Forest Diversity and Dynamism: Findings from a Large-Scale Plot Network. . 506-516. University of Chicago Press, Chicago.

Thomas, D.W., Kenfack, D., Chuyong, G.B., Sainge N.M., Losos E.C., Condit R.S., Songwe N.C. (2003). Tree Species of Southwestern Cameroon: Tree distribution maps, diameter tables and species documentation of the 50-ha Korup Forest Dynamics Plot. Center for Tropical Forest Science, Washington, D.C.

Kenfack D., Thomas D.W., Chuyong G.B., & Condit R. (2007). Rarity and abundance in a diverse African forest: Biodiversity Conservation 16: 2045 – 2074.

Lambir

The 52-ha Long-Term Ecological Research Project is a collaborative project of the Forest Department of Sarawak, Malaysia, the Forest Global Earth Observatory (ForestGEO) , the Arnold Arboretum of Harvard University, USA (under NSF awards DEB-9107247 and DEB-9629601), and Osaka City, Ehime & Kyoto Universities, Japan (under MEXT/JSPS KAKENHI grants 09NP0901, 22H02388, and JST/JICA-SATREPS PUBS). The Lambir Forest Dynamics Plot is part of ForestGEO, a global network of large-scale demographic tree plots. We acknowledge the Sarawak Forest Department for supporting and maintaining the project in Lambir Hills National Park.

References

Lee, H.S., P.S. Ashton, T. Yamakura, S. Tan, S.J. Davies, A. Itoh, T. Ohkubo & J.V. LaFrankie. (2002). The 52-Hectare Forest Dynamics Plot at Lambir Hills, Sarawak, Malaysia: Tree

Distribution Maps, Diameter Tables and Species Documentation.Sarawak Forest Department. 621 pp. Lee Ming Press, Kuching, Sarawak, Malaysia.

Lee, H., Davies, S. J., LaFrankie, J. V., Tan, S., Itoh, A., Yamakura, T., and Ashton, P. S. 2002. Floristic and structural diversity of 52 hectares of mixed dipterocarp forest in Lambir Hills National Park, Sarawak, Malaysia. Journal of Tropical Forest Science, 14: 379-400.

La Planada

Lilly Dickey Woods

References

Johnson, D.J., Clay, K. & Phillips, R.P. Mycorrhizal associations and the spatial structure of an old-growth forest community. Oecologia 186, 195–204 (2018). https://doi.org/10.1007/s00442-017-3987-0

Luquillo

part of the Smithsonian Institution Forest Global Earth Observatory, a worldwide network of large, long-term forest dynamics plots.

References

Zimmerman, Jess K., Liza S. Comita, Jill Thompson, María Uriarte, and Nicholas Brokaw. 2010. Patch dynamics and community metastability of a subtropical forest: compound effects of natural disturbance and human land use. Landscape Ecology 25: 1099-1111.

Mo Singto

Pasoh

References

Manokaran N & LaFrankie JV. 1990. Stand structure of Pasoh Forest Reserve, A lowland rainforest in Peninsular Malaysia. Journal of Tropical Forest Science 3: 14-24.

Kochumen KM, LaFrankie JV, Manokaran N. 1990. Floristic composition of Pasoh Forest Reserve, a lowland forest in Peninsular Malaysia. Journal of Tropical Forest Science 3: 1-13.

Manokaran N, Abd Rahman K, Azman H, Quah ES, Chong PF. 1992. Short-term population dynamics of dipterocarp trees in a lowland rainforest in Peninsular Malaysia. Journal of Tropical Forest Science 5: 97-112.

Smithsonian Conservation Biology Institute - SCBI

References

Bourg, N.A., W.J. McShea, J.R. Thompson, J.C. McGarvey, and X. Shen. 2013. Initial census, woody seedling, seed rain, and stand structure data for the SCBI SIGEO Large Forest Dynamics Plot. Ecology 94(9): 2111-2112. http://dx.doi.org/10.1890/13-0010.1

Smithsonian Environmental Research Center - SERC

Data on SERC Dynamic Forest plot was provided by Geoffrey Parker on October 8, 2020. These data were gathered as part of forest ecology studies at the Smithsonian Environmental Research Center (SERC). SERC is a participant in the Smithsonian Institution Forest Global Earth Observatory (ForestGEO) network.

Sinharaja

The 25-ha Long-Term Ecological Research Project at Sinharaja World Heritage Site is a collaborative project of the Uva Wellassa University, University of Peradeniya, the Forest Global Earth Observatory (ForestGEO) of the Smithsonian Tropical Research Institute, with supplementary funding received from the John D. and Catherine T. Macarthur Foundation, the National Institute for Environmental Science, Japan, and the Helmholtz Centre for Environmental Research-UFZ, Germany, for past censuses. The PIs gratefully acknowledge the Forest Department, Uva Wellassa University, and the Post-Graduate Institute of Science at the University of Peradeniya, Sri Lanka for supporting this project, and the local field and lab staff who tirelessly contributed in the repeated censuses of this plot.

University of California Santa Cruz - UCSC

The UCSC Forest Ecology Research Plot was made possible by National Science Foundation grants to Gregory S. Gilbert (DEB-0515520, DEB-084259, and DEB-1655896), by the Pepper-Giberson Chair Fund, the Robert Headley Presidential Chair for Integral Ecology and Environmental Justice, the UCSC Campus Natural Reserve, the University of California, the ForestGEO global network, and by the hard work of hundreds of UCSC students. We acknowledge that the land on which this research was conducted is the unceded territory of the Awaswas-speaking Uypi Tribe. The Amah Mutsun Tribal Band, comprised of the descendants of indigenous people taken to missions Santa Cruz and San Juan Bautista during Spanish colonization of the Central Coast, is today working hard to restore traditional stewardship practices on these lands and heal from historical trauma.

References

Gilbert, G.S., E. Howard, B. Ayala-Orozco, M. Bonilla-Moheno, J. Cummings, S. Langridge, I.M. Parker, J. Pasari, D. Schweizer, S. Swope. 2010. Beyond the tropics: forest structure in a temperate forest mapped plot. Journal of Vegetation Science 21: 388-405.

Wabikon

The Wabikon Lake Forest Dynamics Plot, located in the Chequamegon-Nicolet National Forest of northern Wisconsin, is part of the Smithsonian Institution's ForestGEO network. Tree censuses at the site have been supported by the 1923 Fund, the Smithsonian Tropical Research Institute, and the Cofrin Center for Biodiversity at the University of Wisconsin-Green Bay. More than 50 scientists and student assistants contributed to the first two plot censuses. We are particularly grateful for the leadership of Gary Fewless, Steve Dhein, Kathryn Corio, Juniper Sundance, Cindy Burtley, Curt Rollman, Mike Stiefvater, Kim McKeefry, and U.S. Forest Service collaborators Linda Parker and Steve Janke.

Wind River

The Wind River Forest Dynamics Plot is a collaborative project of Utah State University, the Utah Agricultural Experiment Station and the USDA Forest Service Pacific Northwest Research Station. Funding was provided by the Center for Tropical Forest Science of the Smithsonian Tropical Research Institute, Utah State University, and the Utah State Agricultural Experiment Station. We acknowledge the Gifford Pinchot National Forest and the Wind River Field Station for providing logistical support, and the students, volunteers and staff individually listed at http://wfdp.org for data collection.

References

Lutz, J. A, A. J. Larson, J. A. Freund, M. E. Swanson, and K. J. Bible. 2013. The ecological importance of large-diameter trees to forest structural heterogeneity. PLOS ONE 8(12): e82784.

Lutz, J. A., A. J. Larson, T. J. Furniss, J. A. Freund, M. E. Swanson, D. C. Donato, K. J. Bible, J. Chen, and J. F. Franklin. 2014. Spatially non-random tree mortality and ingrowth maintain equilibrium pattern in an old-growth *Pseudotsuga-Tsuga* forest. *Ecology* 95(8): 2047-2054.

Wytham Woods

The 18-ha Long-Term Forest Monitoring Plot is a collaborative project between the University of Oxford, the Centre for Ecology and Hydrology, and the Smithsonian

References

Butt, N., Campbell, G., Malhi, Y., Morecroft, M., Fenn, K., Thomas, M. (2009) Initial results from establishment of a long-term broadleaf monitoring plot at Wytham Woods, Oxford, UK. University of Oxford Report.

Zofin

References

Janík D., Král K., Adam D., Hort L., Šamonil P., Unar P., Vrška T., McMahon S.M., 2016. Tree spatial patterns of Fagus sylvatica expansion over 37 years. Forest Ecology and Management 375: 134–145.

Janík D., Vrška T., Hort L., Unar P., Král K., 2018. Where have all the tree diameters grown? Patterns in Fagus sylvatica L. diameter growth on their run to the upper canopy. Ecosphere 9 (12): 1-19.

Šamonil P**.,** Daněk P., Lutz J.A., Anderson-Teixeira, K., Jaroš J., Phillips J.D., Rousová A., Adam D., Larson A.J., Kašpar J., Janik D., Vašíčková I., Gonzáles-Akre E., Egli M. 2022. Biogeomorphic role of tree death in temperate forests: disturbance and species specific processes from local to continental scale. Ecosystems. 10.1007/s10021-022-00755-8.

Appendix S2- Robustness analysis for subsampling forest plots

Comparing forest plots of different sizes

Because forest plots differed in size (from 6 to 50 ha), we tested if differences in plot size could bias the VPC estimates. For that, we compared VPC results of the reduced model (no temporal OPs) from Lambir entire plot with the average VPCs from 10 subsets of 5 ha each. Lambir is a suitable plot for such a comparison because it has one of the largest species richness (> 1000) and it is a large plot size (50 ha).We randomly subsampled Lambir data 10 times to the size of 5 ha and ran reduced models for growth at the 5x5 m quadrat scale to build distributions of the estimates and compared mean values with the estimates of the results for the 50 ha analysis.

Such comparisons also evaluate if the approach of repeatedly fitting the vital rate models to smaller subsamples of the plot, which was necessary to run the models with temporal OPs, is robust. For this purpose, we ran the VPC analysis (eq.1) for the entire Fushan (25 ha) and Luquillo (15 ha) forest plots, at the 100x100 m quadrat size for each vital rate and compared VPC estimates against the distribution of VPC estimates obtained from 10 subsampled datasets at the same grain size.

For both robustness analyses, using subsamples of the entire plot only marginally changed the estimates and, thus, we conclude that (1) the results among forests with different plot sizes for models without temporal OPs can be compared, and (2) the average VPCs among subsampled datasets are reliable VPC estimates for models with temporal OPs.

Models without temporal OPs - Lambir forest plot

The 5 ha subsamples presented on average 86% of the original number of species (1033 of 1311) and 10% of the original number of trees (mean 32,007 of 313,544). Mean VPCs from models using the subsamples of 5 ha were very similar to VPC values calculated from the entire 50 ha forest data (Figure S2.1). The differences between mean values from the subsamples and the entire forest was negligible for *space* and *species x space* OPs and around 0.02 for *species* and *residual* OPs.
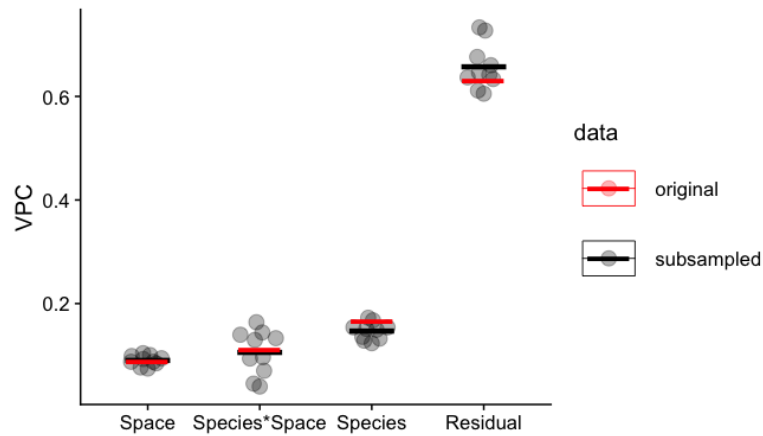
Figure S2.1: Comparison of the results for Variance Partition Coeficient (VPC) from growth reduced models (without temporal OPs) at the 5 x 5 m quadrat scale for Lambir forest plot. Red bars indicate the VPC for the entire 50 ha forest plot dataset and black bars are the mean VPC from 10 subsamples of 5 ha (gray dots).

Models with temporal OPs - Fushan and Luquillo forest plots

For the models with temporal OPs, we did the opposite: as the computational time for running models to a huge amount of data are restrictive, we evaluated if the subsampling analysis of 5 ha each plot was able to get the same results if we were doing the analysis with the entire plot. The 5 ha subsamples retained around 85% of the number of species and between 20 and 34% of the number of observations. The vast majority of VPC estimates was very similar between datasets (Figure S2.2). The largest differences (up to 0.03) were found for *species* OP in recruitment for both forest plots, for *time* OP in recruitment for Fushan, and for *species* OP in mortality for Luquillo.
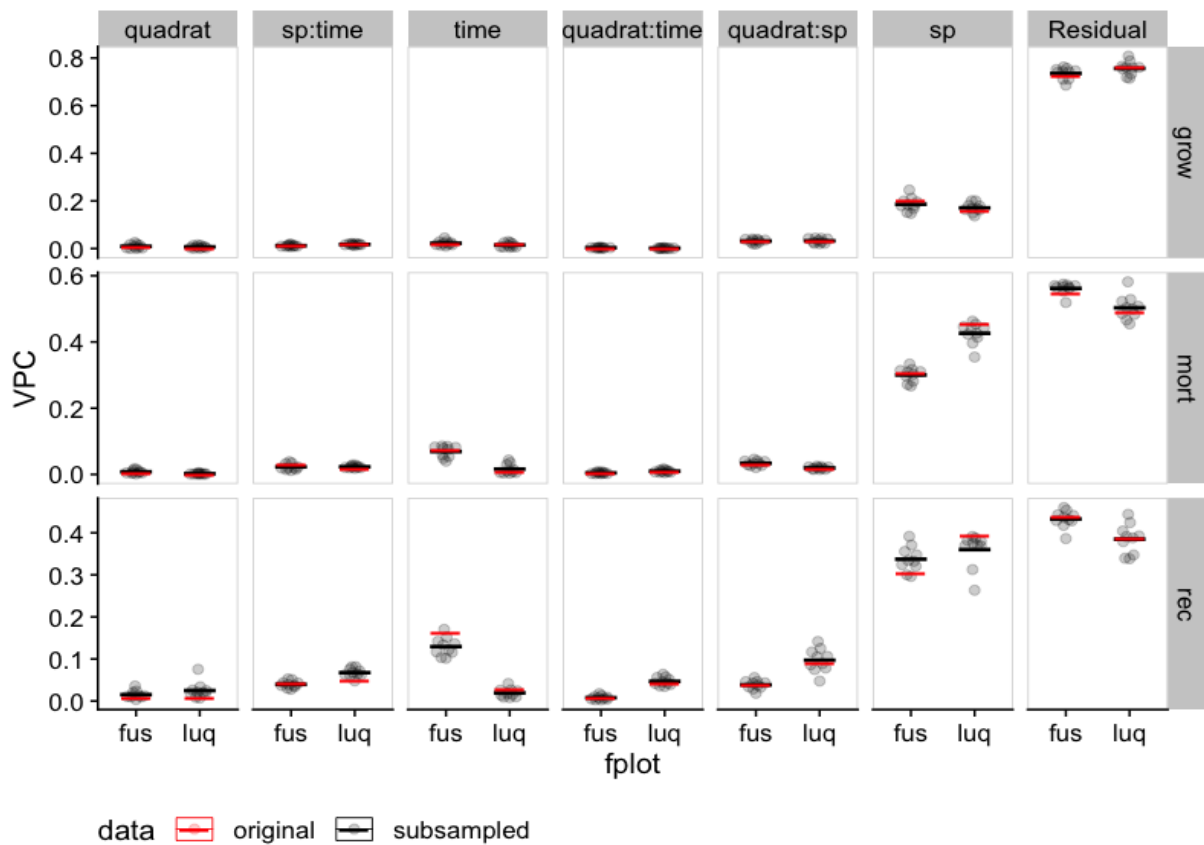
Figure S2.2: Comparison of the results for Variance Partition Coeficient (VPC) from the time models at the 100 x 100 m quadrat scale for Fushan (fus) and Luquillo (luq) forest plots. Red bars indicate the VPC for the entire 50 ha forest plot dataset and black bars are the mean VPC from 10 subsamples of 5 ha (gray dots).

Appendix S3- Comparing standard deviations of models with and without temporal organising principles

Given the low number of forest plots with four or more censuses, we could apply the main model of Equation 1 (main text) to only five forest plots (Table S1.1). However, forest plots with just one census interval can also be a reliable source of information for comparing *species*, *space*, and *species x space* organising principles. Therefore, we applied a reduced model setup without the temporal organising principles - *time, species x time*, and *space x time* - for all 21 forest plots. For forest plots with more than one census interval, we averaged the variance estimates across intervals and calculated VPCs.

Using the formula syntax of *brms* R package (Bürkner, 2017), the complete model from equation 1 (hereafter **time models**) is written as:

```
Y ~ 1 + (1|species) + (1|space) + (1|time) + (1|species:space) +
(1|species:time) + (1|space:time)
```

while the reduced model without temporal terms (hereafter **no-time models**) is:

```
Y ~ 1 + (1|species) + (1|space) + (1|species:space)
```

To understand the effects of omitting the temporal terms for the remaining standard deviations and thus the reliability of the reduced model setup, we ran both models for the five forest plots where enough census intervals were available and compared the standard deviations (SD) of the model terms. Below, we show and discuss these comparisons at the 5x5 m quadrat spatial scale.

As expected, the total standard deviation in time models was always larger (Figure S3.1) and residual standard deviations for growth (normal distribution) also did not change. Notice that the residual standard deviations for mortality and recruitment are anyway fixed to the theoretical standard deviation of binomial models with complementary log-log link function (Nakagawa *et al.*, 2017). Standard deviations for species organising principle did not change between time and no-time models, except for recruitment where it was slightly larger for some forest plots in no-time models (Figure S3.2). *Space* and *species x space* standard deviations were, in general, smaller in time models for all vital rates, except for growth, where *space* standard deviations were equal or a bit larger for time models (Figure S3.2). It may be the case that *space* and *species x space* organising principles in no-time models are also incorporating temporal variability in vital rates, especially through gap dynamics effects on tree recruitment, mortality and growth. In the lack of temporal organising principles, the footprint of fallen-tree gap formation remains in *space* and, at a lower extent, *species x space* organising principles.
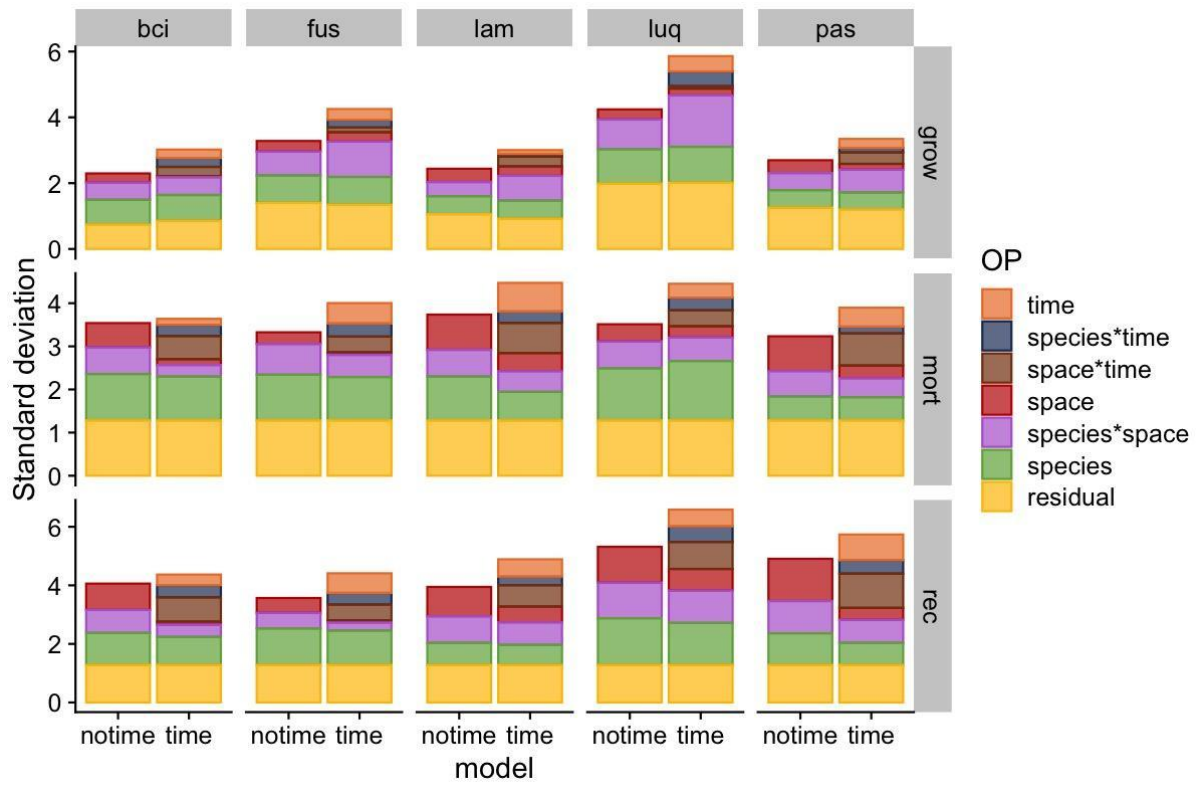
Figure S3.1. Standard deviations of models with (time models) and without temporal (no-time models) organising principles for the five forest plots with more than four census intervals: Barro Colorado Island (bci), Fushan (fus), Lambir (lam), Luquillo (luq), and Pasoh (pas).

Standard deviations

Figure S3.2. Comparing standard deviation of *species, space*, and *species x space* organising principles for models with (time models) and without temporal organising principles (no-time models) for the five forest plots with more than four census intervals: Barro Colorado Island (bci), Fushan (fus), Lambir (lam), Luquillo (luq), and Pasoh (pas). Dotted diagonal lines indicate the 1:1 threshold.
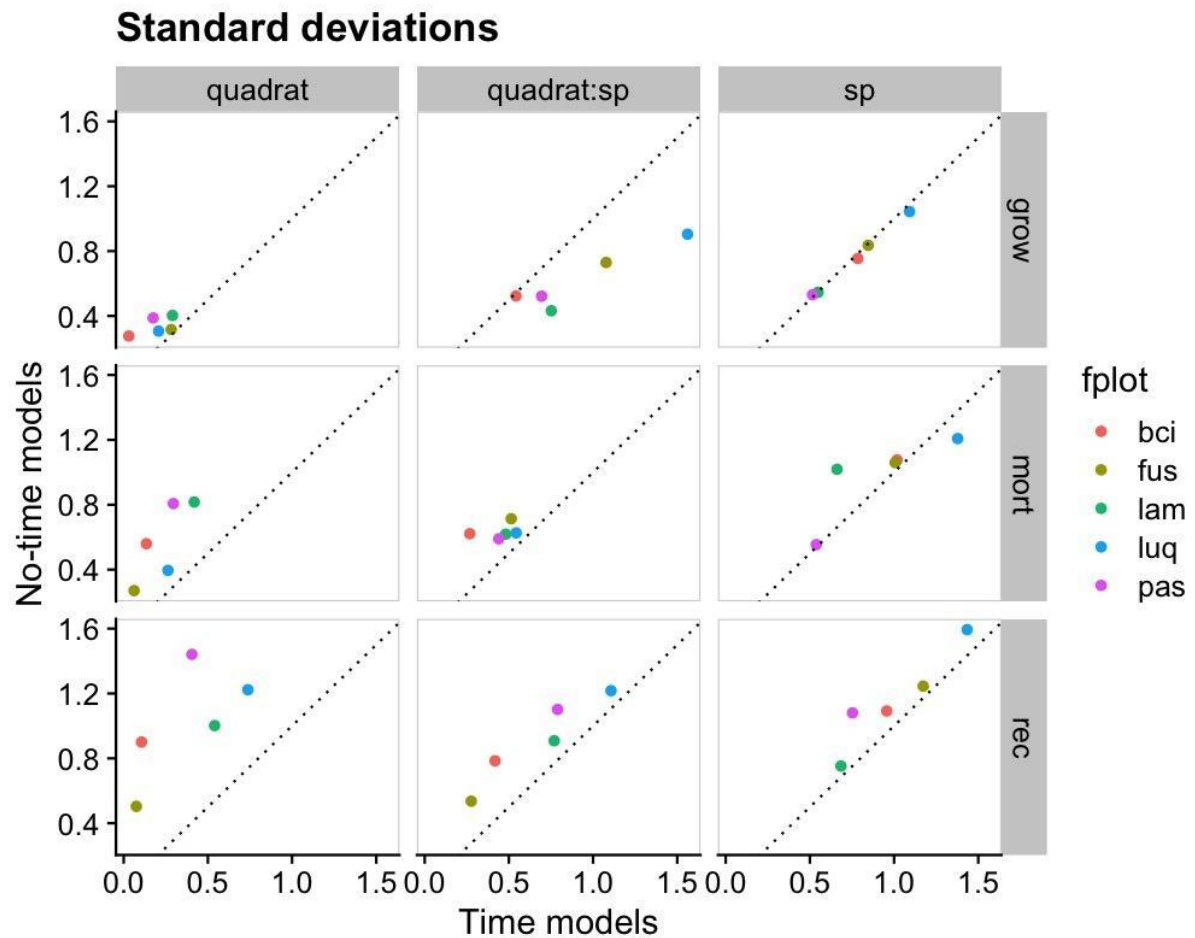
References

Bürkner, P.-C. (2017) brms: An R Package for Bayesian Multilevel Models Using Stan. Journal of Statistical Software, 80, 1–28.

Nakagawa, S., Johnson, P.C.D. & Schielzeth, H. (2017) The coefficient of determination R2 and intra-class correlation coefficient from generalized linear mixed-effects models revisited and expanded. Journal of The Royal Society Interface, 14, 20170213.

Appendix S4- Robustness analysis for the role of rare species

The presence of rare species in a forest plot can influence VPC analyses in the following ways: (1) in MLMs, rare species may 'shrink' variance estimates towards the

population mean because of the small number of observations, and (2) rare species may increase species standard deviations, either by sampling artefact (Condit et al., 2006) or because rare species represent vital rate strategies distinct from the more common species (Umaña et al., 2017). We thus assessed the extent to which rare species affect the VPC results rerunning the VPC analysis (1) without rare species and (2) with rare species grouped as a single 'species'. We used the FuzzyQ clustering algorithm in the FuzzyQ R package (Balbuena et al., 2021) to estimate the probability of each species to be common or rare based on species abundance and occupancy in 50x50 m quadrats. The method has the advantage of allowing comparisons among plots of different sizes. Both procedures showed similar results, with a small decrease in the *species* VPC, balanced by an increase in the *residual* and *species x space* VPC when excluding or regrouping rare species. We conclude that our main results are robust to the presence of rare species in the datasets.

Classifying rare species

In order to understand the possible influence of rare species in the variance partition components (VPC), we first classified species as rare or common in each forest plot and census interval data, using the 50x50 m quadrat as sampling units. To describe and compare rarity patterns in all forest, we estimated the number of species, number of individuals, and density of rare and common species. For the forests plots with more than 1 census interval, we averaged the estimates across time intervals. Summary results for the number of species and trees are shown in Table S4.1.

Across all plots and censuses, the number of tree was highly skewed to common species while species richness was skewed towards rare species (Fig. S4.1). Common species richness ranged from 20% to 56% of species richness, while it comprised from 86% to 99% of the trees. The average density (number of trees/ha) that formed the cutoff for the rare species classification was 1.39 trees/ha, (Table S4.2).

Table S4.1. Number and percentages of species and trees classified as common or rare per forest plot. See Table S1.1 and S1.2 for forest plot names and information. For forest plots with more than 1 census interval, we average the values across intervals.

	Species richness				Number of trees			
	Common		Rare		Common		rare	
Forest	N	%	N	%	N	%	N	%

ama	456	35	839	65	98878	90	11578	10
bci	135	42	185	58	334275	96	12477	4
edo	122	29	298	71	168253	96	7004	4
fus	55	50	55	50	143813	98	2376	2
idc	79	56	61	44	47740	96	1969	4
kor	186	40	282	60	335392	94	22608	6
lam	539	38	869	62	376389	86	61135	14
lwd	14	38	23	62	28004	96	1257	4
len	110	28	291	72	145785	97	4482	3
lpl	108	45	132	55	120516	96	4735	4
luq	53	35	97	65	67147	96	2838	4
mos	94	34	183	66	150534	95	7994	5
pas	375	44	474	56	360659	92	31816	8
scbi	25	35	47	65	38366	97	1248	3
serc	20	27	55	73	27031	96	1205	4
sin	118	50	118	50	207356	94	12196	6
ucsc	10	32	21	68	8738	96	394	4
wab	11	28	28	72	49562	93	3558	7
wfdp	8	31	18	69	28948	97	831	3
wyw	5	20	20	80	19458	96	740	4
zof	3	23	10	77	75368	99	381	1

Table S4.2. Summary values of density of trees (N/ha) classified as common or rare for the 21 forest plots. SD - standard deviation, Quant - quantiles.

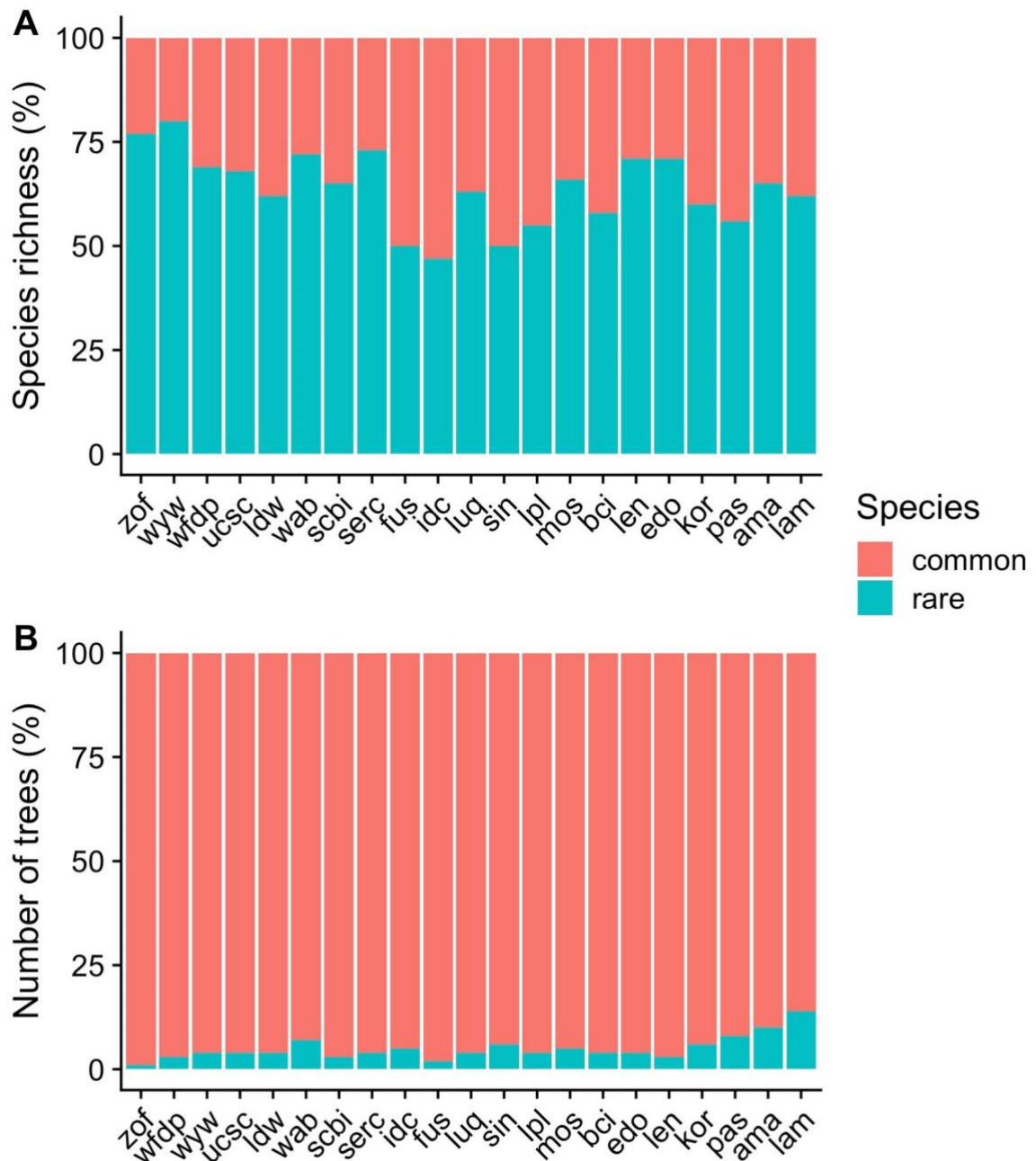|	Min	Max	Mean	SD	Median	Quant 90	Quant 95
common	1.32	3840.80	35.28	128.05	11.28	63.86	109.94
rare	0.02	297.32	1.39	5.14	0.55	2.86	4.39

Figure S4.1. Percentages of rare and common species and number of trees per forest plot.

We performed a generalised additive model (Pedersen et al., 2019) to evaluate the relationship between rarity (proportion of rare species) and number of species. We found that the percentage of rare species tended to decrease with the number of species only for forests with less than 100 species: from 77% with 10 species to 63% with 90 species. For plots with 100 or more species, the percentage of rare species varied from 62 to 60% (Figure S4.2).
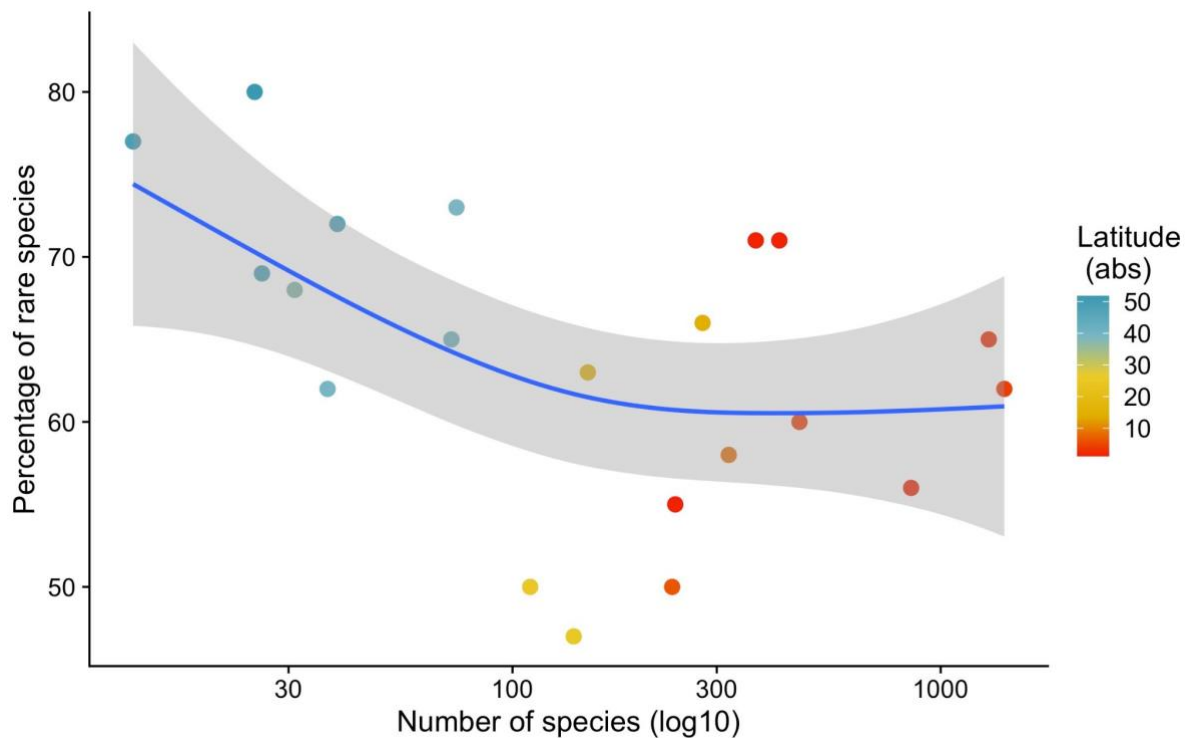
Figure S4.2. Percentage of rare species against the number of species (log10 scale) for 21 worldwide distributed forest plots (excluding Zofin given its low number of species). Blue line and grey area are the fitted results and confidence intervals for a generalised additive model showing a decrease in the percentage of rare species with the number of species but only for forests with less than 100 species. Model's adjusted $R^2 = 0.30$. Each forest plot is coloured by the latitude in absolute values.

Excluding or regrouping rare species in forest data

Rare species can influence our variance partition results in two ways. Given the small number of trees, the estimates for rare species in species random effects is done by borrowing information from the other species (assuming a normal distribution of the random effects). It may make rare species estimates vary narrowly around the mean rate, because of the stronger shrinkage effect (Gelman & Pardoe, 2006). In this case, excluding rare species could widen the distribution of the standard deviation of species random effect. On the other hand, more species increase the chance of estimating more extreme values, as rare species may also present very distinct vital rates (Umaña et al., 2017), which may narrow the distribution of species random effects if they are excluded.

We used two procedures to deal with rare species: (1) excluding rare species from the dataset, which means excluding a proportion of the number of observations in the data, and (2) renaming the rare species the dataset into one generic species name,

which does not change the total number of observations in the data. We applied both procedures for the 5 plots with more than 4 censuses using the main model in equation 1 (Fig S4.3) and also for all forest plots with the reduced model without temporal organizing principle (Fig S4.4).

Both procedures to deal with rare species presented very similar results, and they showed a decrease in *species* VPC, balanced with an increase in the *residual* and *species x space* VPC when excluding or regrouping rare species.

For the models with temporal organising principles (Fig S4.2), we classified the species in the whole dataset as rare and common based on the classification in each census interval (previous section) as some species may temporally vary in abundance/occupancy. We classified a species as rare if it was rare in half or more than half of the time intervals. The results showed that the largest absolute VPC differences appeared in *species* VPC, decreasing on average from 0.03 (recruitment) to 0.05 (mortality), which corresponds to an average of 11% relative decrease in standard deviation. *Residual* VPC increased on average between 0.01 (recruitment) and 0.06 (mortality). Although *space, time, space x time* and *species x time* standard deviations changed relatively between 3 and 24%, these differences in terms of absolute VPC were negligible (between 0,001 and 0.01). We did not find any tendency for the relative differences in VPC being related to the proportion of rare species in the data (Fig S4.2b).

Table S4.3: Average absolute differences in VPC and relative differences in standard deviation between models with rare species and models (1) excluding or (2) regrouping rare species for growth, mortality, and recruitment vital rates. Data used here were the 5 forests with more than 4 census intervals.

Organizing Principle	Growth				Mortality				Recruitment			
	Exclude rare		Regroup rare		Exclude rare		Regroup rare		Exclude rare		Regroup rare	
	VPC	%SD	VPC	%SD	VPC	%SD	VPC	%SD	VPC	%SD	VPC	%SD
space	0.002	8.8	0.002	13.3	0.001	7.9	0.003	7.9	0.001	23.3	0.002	24.3
species x time	-0.007	-13.1	-0.008	-23.6	-0.001	-5.2	-0.001	-6.7	-0.004	-5.5	-0.006	-10.1
time	0.001	4.2	0.002	10.1	0.008	3	0.008	3.3	0.008	5.3	0.011	5.4
space x time	0.005	-4	0.001	-6.9	0.008	0.6	0.002	-2.9	0.004	1.1	0.013	1.8
species x space	0.016	2.3	-0.006	0.4	0.002	-0.2	0.004	2.7	0.008	2.9	-0.009	-3.8
species	-0.040	-11.4	-0.050	-12.3	-0.049	-12.5	-0.048	-12.6	-0.025	-8.6	-0.026	-10.6

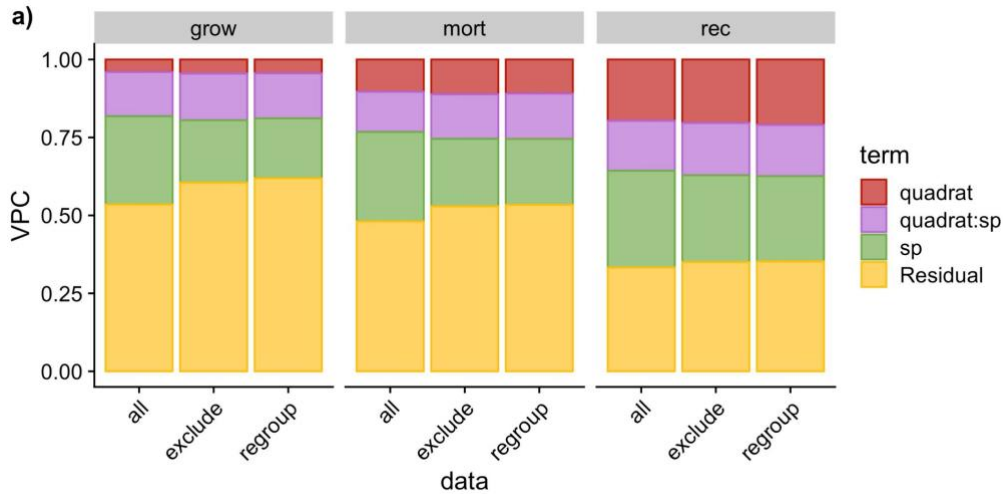| residual	0.023	1.4	0.059	8.2	0.030	-	0.033	-	0.008	0	0.016	-
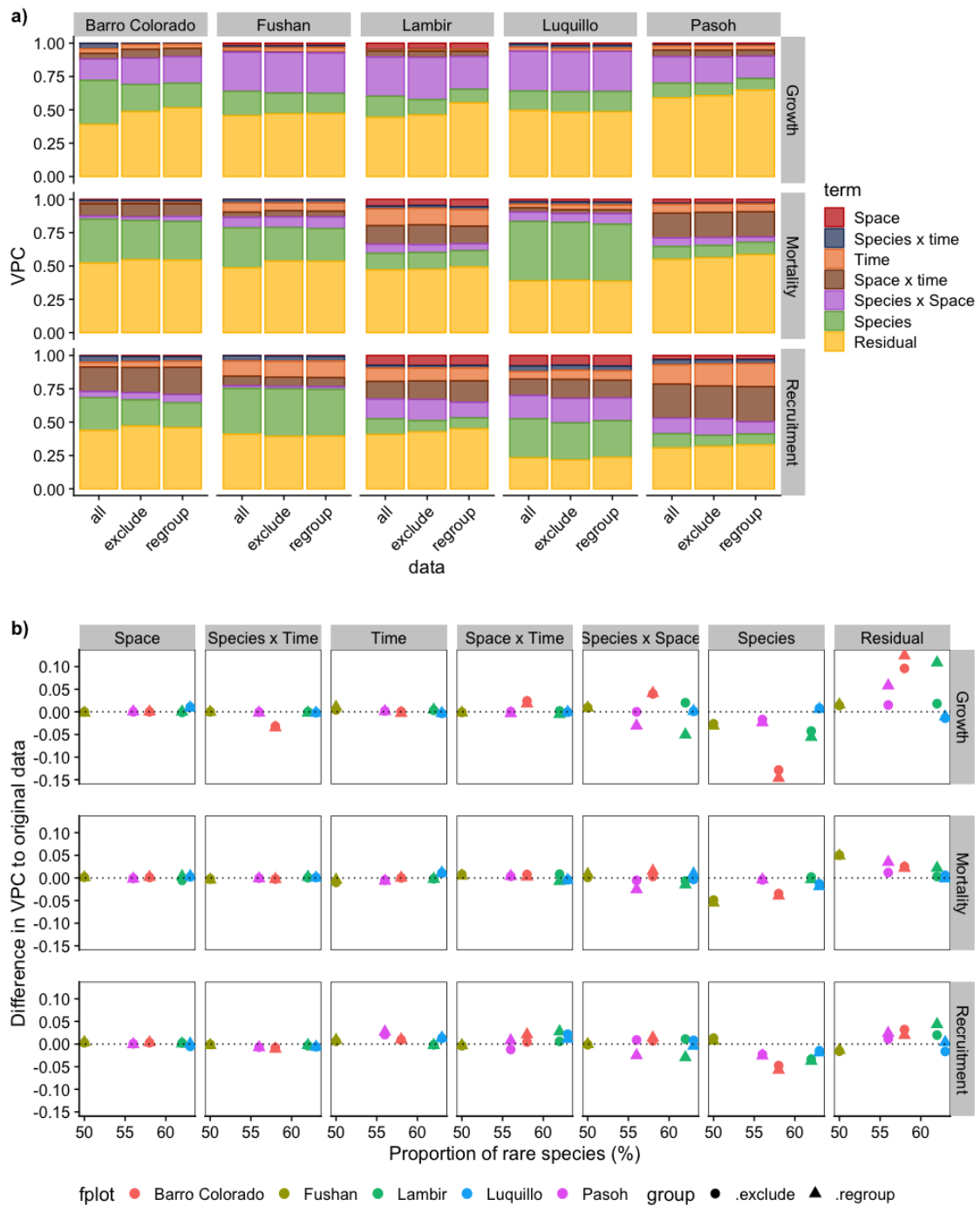
Figures S4.3. (a) Comparing Variance Partitioning Components for models with temporal organizing principles among models with all species data included (all), excluding rare (exclude) or regrouping rare species into one generic species label (regroup). (b) Differences in VPC from original data (all species) to the models excluding rare species (circles) or regrouping rare species (triangles). Results here are for the models with the 5x5 m quadrat scale.

For the models without temporal organizing principles (21 forest plots), there was an average of 10 to 15% relative decrease in *species* standard deviation (Table

S4.4), while the absolute VPC decreased between 0.03 and 0.09 (Fig S4.3). This decrease was balanced mainly by an increase in *residual* VPCs around 0.02 to 0.07, while for *space* and *space x species* the absolute differences in VPC were very small and on average smaller than 0.01.

Table S4.4: Average absolute differences in VPC and relative differences in standar deviations between models with rare species and models excluding or regrouping rare species for growth, mortality and recruitment. Analysis applied to the reduced model without temporal organizing principles for the 21 forest plots.

| Organizing Principle	Growth				Mortality				Recruitment			
	Exclude rare		Regroup rare		Exclude rare		Regroup rare		Exclude rare		Regroup rare	
	VPC	%SD	VPC	%SD	VPC	%SD	VPC	%SD	VPC	%SD	VPC	%SD
space	0.004	11	0.004	8.7	0.008	0.0	0.006	-2.6	0.006	-0.6	0.013	0.5
species x space	0.008	8.5	0.028	4.0	0.014	-1.0	0.016	-0.6	0.007	0.3	0.004	0.8
species	-0.080	-11.8	-0.091	-12.7	-0.070	-14.9	-0.075	-16.4	-0.032	-9.3	-0.047	-10.1
residual	0.070	13.2	0.084	13.9	0.048	0.0	0.053	0.0	0.017	0.0	0.019	0.0
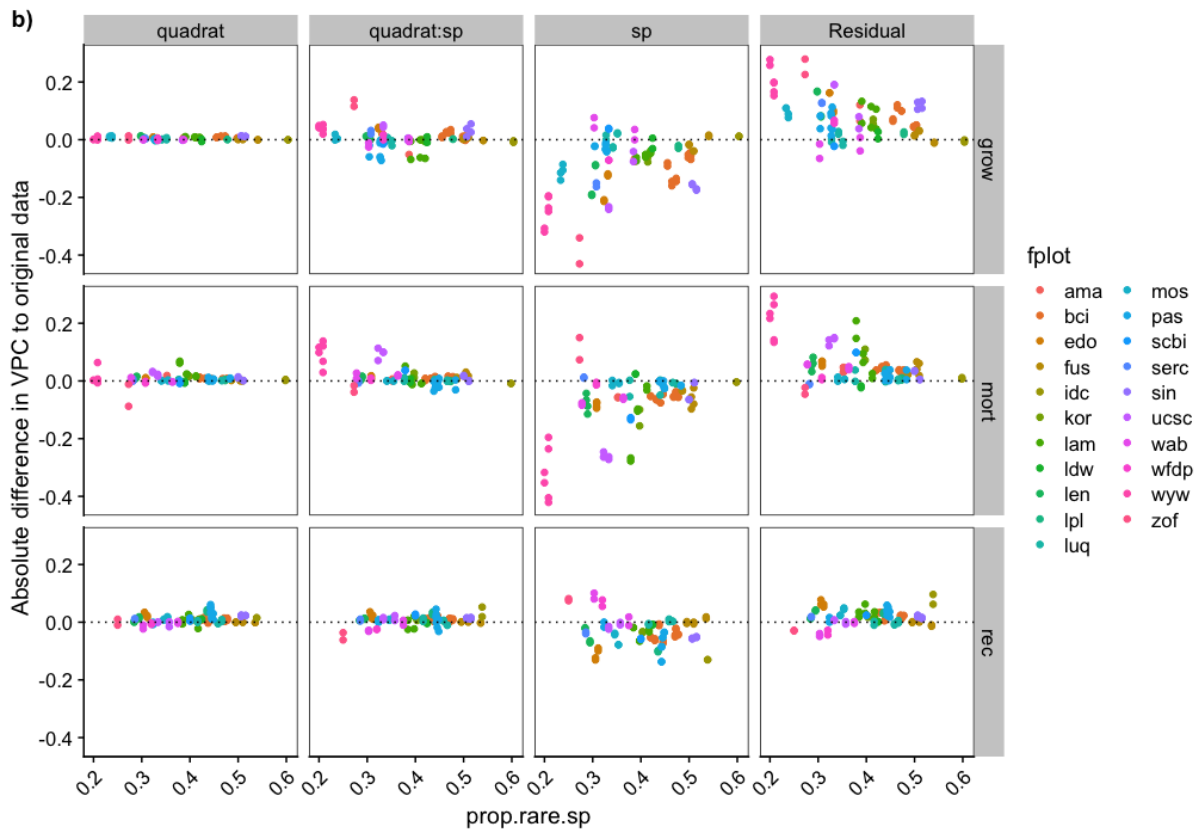
Figure S4.4. (a) Comparing Variance Partitioning Components for models without temporal organizing principles (no time models) among models with all species data included (all), excluding rare (exclude) or regrouping rare species into one generic species label (regroup). **(b)** Differences in VPC from original data (all species) to the models excluding rare species (circles) or regrouping rare species (triangles). Results here are for the models with the 5x5 m quadrat scale.

References

Balbuena, J.A., Monlleó-Borrull, C., Llopis-Belenguer, C., Blasco-Costa, I., Sarabeev, V.L. & Morand, S. (2021) Fuzzy quantification of common and rare species in ecological communities (FuzzyQ). Methods in Ecology and Evolution, 12, 1070–1079.

Condit, R., Ashton, P., Bunyavejchewin, S., Dattaraja, H.S., Davies, S., Esufali, S., Ewango, C., Foster, R., Gunatilleke, I. a. U.N., Gunatilleke, C.V.S., Hall, P., Harms, K.E., Hart, T., Hernandez, C., Hubbell, S., Itoh, A., Kiratiprayoon, S., LaFrankie, J., Lao, S.L. de, Makana, J.-R., Noor, M.N.S., Kassim, A.R., Russo, S., Sukumar, R., Samper, C., Suresh, H.S., Tan, S., Thomas, S., Valencia, R., Vallejo, M., Villa, G. & Zillio, T. (2006) The Importance of Demographic Niches to Tree Diversity. Science, 313, 98–101.

Gelman, A. & Pardoe, I. (2006) Bayesian Measures of Explained Variance and Pooling in Multilevel (Hierarchical) Models. Technometrics, 48, 241–251.

Pedersen, E.J., Miller, D.L., Simpson, G.L. & Ross, N. (2019) Hierarchical generalized additive models in ecology: an introduction with mgcv. PEERJ, 7.

Umaña, M.N., Mi, X., Cao, M., Enquist, B.J., Hao, Z., Howe, R., Iida, Y., Johnson, D., Lin, L., Liu, X., Ma, K., Sun, I.-F., Thompson, J., Uriarte, M., Wang, X., Wolf, A., Yang, J., Zimmerman, J.K. & Swenson, N.G. (2017) The role of functional uniqueness and spatial aggregation in explaining rarity in trees: Umaña et al. Global Ecology and Biogeography, 26, 777–786.

Appendix S5- Additional results comparing global forests

Species richness rarefaction

We calculated rarefied species richness to compare forest plots across the globe, given that forest plots vary from 6 to 50 ha in size. Rarefaction was constructed based on sampling increment and each sample was a quadrat of 20x20 m quadrat size. We used the R packages BiodiversityR (Kindt & Coe, 2005) and vegan (Oksanen et al., 2020), following Gotelli & Collwel (2001) suggestions for rarefaction curve construction and used the species richness estimated at the smallest plot size (Fig S5.1).
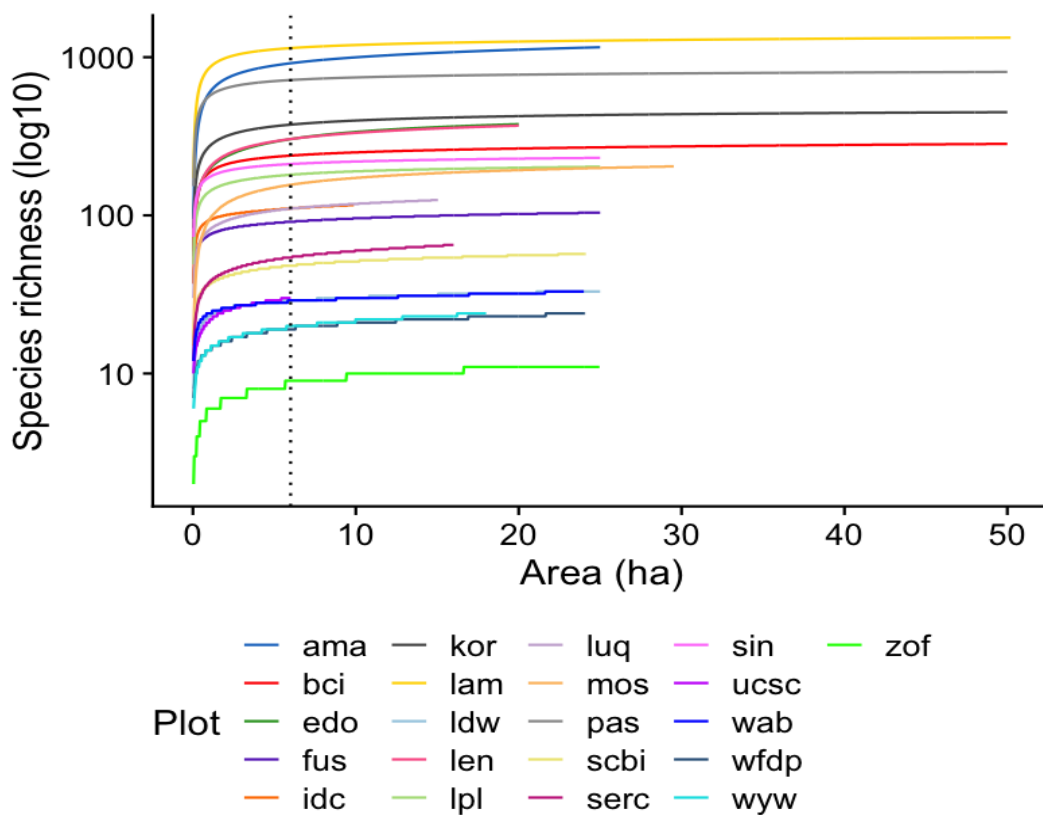
Figure S5.1. Rarefaction curves of species accumulations for the 21 forest plots. Vertical dotted line indicates 6 ha area, which is the smallest forest plot area. See Table S1.1 for forest plots abbreviations.

We compare rarefied species richness with other species richness measures, latitude, tree density and metrics of variation in density and richness, with a principal component analysis (Fig. S5.2). All species richness variables were highly correlated and presented the largest contribution to PCA axis 1, which summarised 75.7% of the

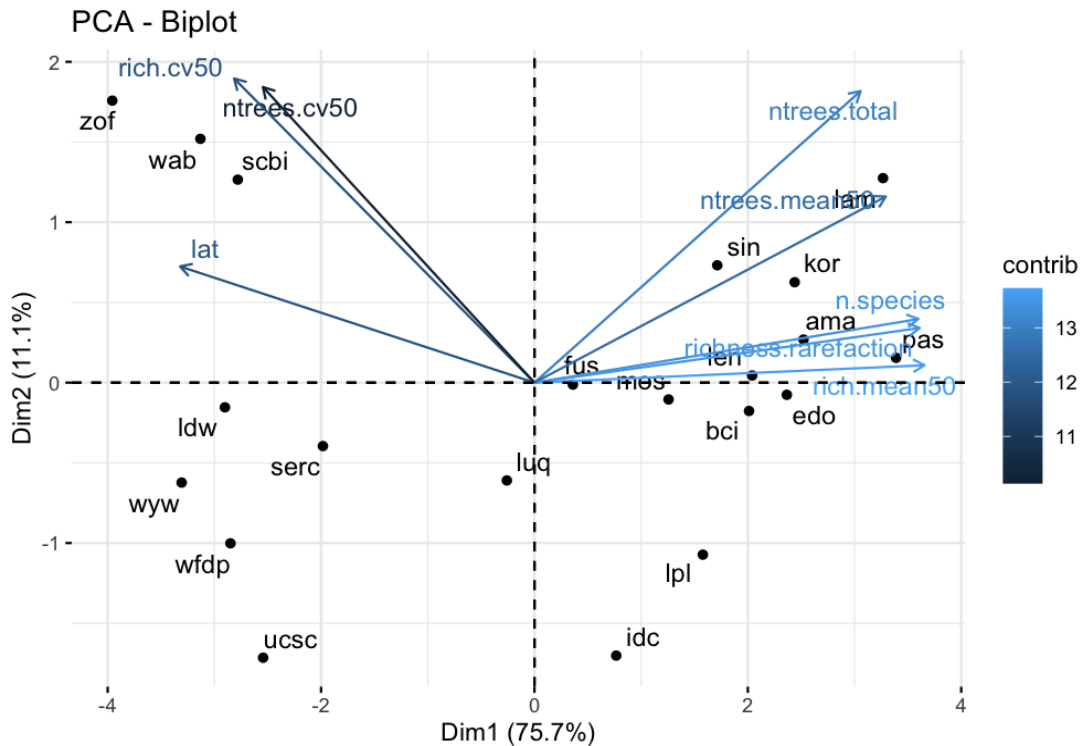variation among plots. We, therefore, used the rarified species richness to compare forest plots.

Figure S5.2. Principal Component Analysis of the 8 variables used to compare forest plots. All species richness variables were highly correlated and presented the largest contribution to Axis 1 (light blue colours in legend), which summarised 75.7% of the variation among plots. Other variables: latitude (lat), total number of species (n.species), mean number of species (rich.mean50) and coefficient of variation (rich.cv50) in 50 x 50 m quadrat size, total number of trees (ntrees.total), mean number of trees (ntrees.mean50) and coefficient of variation (ntrees.cv50) in 50x50m quadrat size. Richness and tree density variables were log-transformed. See Table S1.1 for forest plots abbreviations.

Standard deviation of organising principles across forests

One of the reasons the patterns shown in Figure 4 (main text) - decrease in species VPC with increase in forest species richness - is the increase in standard deviation of other organising principles. Here, we investigated how the overall standard deviation of forest vital rates and each specific organising principle varies in relation to rarefied species richness.

We found that overall standard deviations only decrease for recruitment (FigS5.3) and that decrease is led mainly by the decrease in species standard deviations

(Fig. S5.4), which reveals the reason species VPC also decreases with forest species richness.
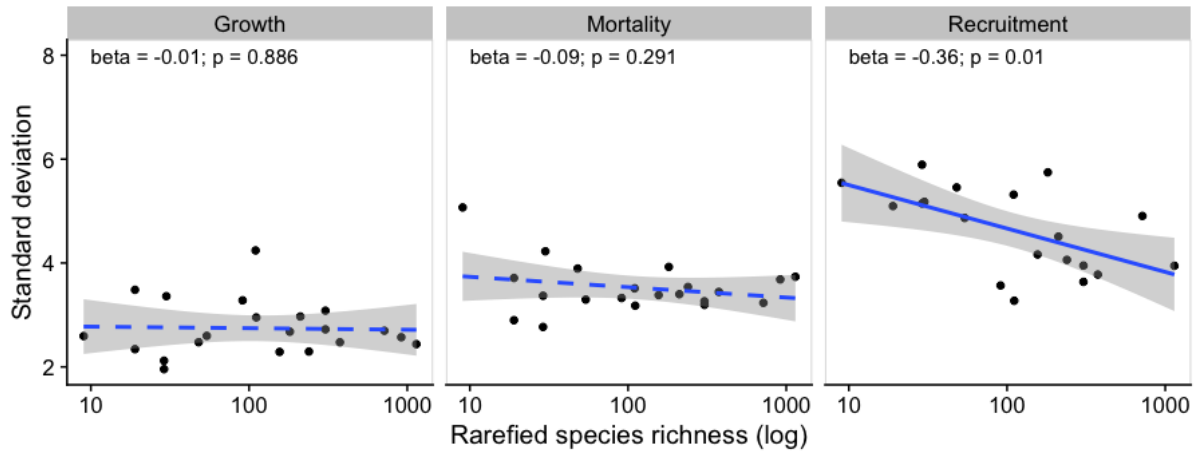

Fig S5.3. Overall standard deviation across forest rarefied species richness for the 2 forest plots. Betas (slopes) and their significance were estimated by a linear model with normal distribution.
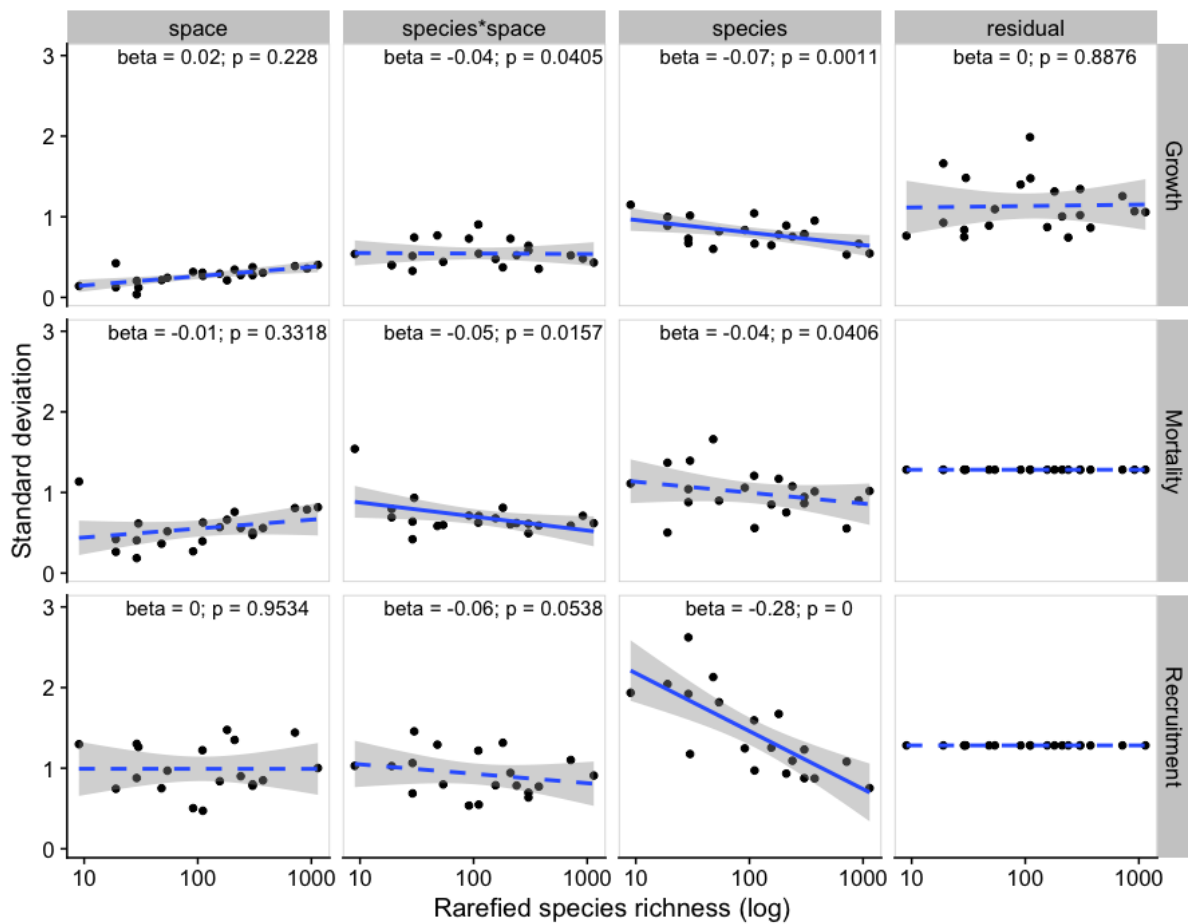

Fig S5.4. Standard deviations for each organising principle and vital rates across forest rarefied species richness for the 21 forest plots. Blue dashed lines indicate non significant betas (slopes) for fitted linear models (normal distributions) and blue solid

lines indicate a significant decrease in species standard deviation with increased species richness for recruitment. Multiple tests Bonferroni alpha-level correction 0.01666.

Dirichlet regression models excluding rare species

Results were qualitatively similar to the models with all species included (Figure 4 in the main text).
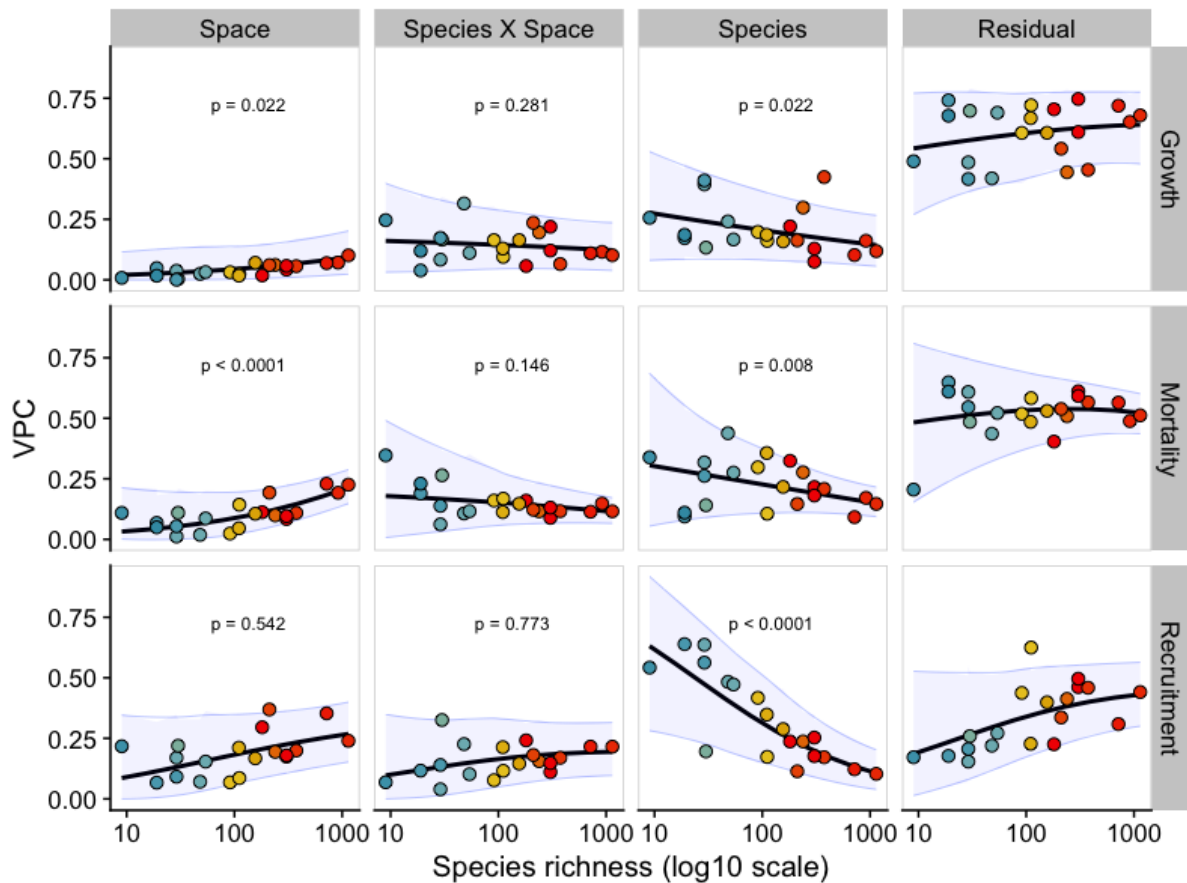

Figure S5.5. Dirichlet regression models for the relationship between organising principles VPCs and rarefied species richness applied to forest data excluding rare species. P-values should be compared with alpha after Bonferroni multiple tests correction (alpha=0.016).

References

Gotelli, N.J. & Colwell, R.K. (2001) Quantifying biodiversity: procedures and pitfalls in the measurement and comparison of species richness. Ecology Letters, 4, 379–391.

Kindt, R. & Coe, R. (2005) Tree diversity analysis: a manual and software for common statistical methods for ecological and biodiversity studies, World Agroforestry Centre, Nairobi, Kenya.

Oksanen, J., Blanchet, F.G., Friendly, M., Kindt, R., Legendre, P., McGlinn, D., Minchin, P.R., O'Hara, R.B., Simpson, G.L., Solymos, P., Stevens, M.H.H., Szoecs, E. & Wagner, H. (2020) vegan: Community Ecology Package.