



**PROGRAMA DE PÓS-GRADUAÇÃO EM ECOLOGIA, EVOLUÇÃO E
BIODIVERSIDADE**

**EFEITOS DA ESTOCASTICIDADE DEMOGRÁFICA SOBRE A DINÂMICA DE
COMUNIDADES E METACOMUNIDADES**

CRISTINA MARIANA JACOBI

**Rio Claro – SP
2024**



**PROGRAMA DE PÓS-GRADUAÇÃO EM ECOLOGIA, EVOLUÇÃO E
BIODIVERSIDADE**

**EFEITOS DA ESTOCASTICIDADE DEMOGRÁFICA SOBRE A DINÂMICA DE
COMUNIDADES E METACOMUNIDADES**

CRISTINA MARIANA JACOBI

Tese apresentada ao Instituto de Biociências do Câmpus de Rio Claro, Universidade Estadual Paulista, como parte dos requisitos para obtenção do título de Doutora em Ecologia, Evolução e Biodiversidade.

Orientador: Dr. Tadeu Siqueira

**Rio Claro – SP
2024**

J16e

Jacobi, Cristina Mariana

Efeitos da estocasticidade demográfica sobre a dinâmica de comunidades e metacomunidades / Cristina Mariana Jacobi. -- Rio Claro, 2024

93 f. : il., tabs., mapas

Tese (doutorado) - Universidade Estadual Paulista (UNESP), Instituto de Biociências, Rio Claro

Orientador: Tadeu de Siqueira Barros

1. estocasticidade demográfica. 2. metacomunidades. 3. diversidade beta. 4. variabilidade temporal. 5. variabilidade ambiental.

I. Título.

Sistema de geração automática de fichas catalográficas da Unesp. Biblioteca da Universidade Estadual Paulista (UNESP), Instituto de Biociências, Rio Claro. Dados fornecidos pelo autor(a).

Essa ficha não pode ser modificada.

CERTIFICADO DE APROVAÇÃO

TÍTULO DA TESE: Efeitos da estocasticidade demográfica sobre a dinâmica de comunidades e metacomunidades

AUTORA: CRISTINA MARIANA JACOBI

ORIENTADOR: TADEU DE SIQUEIRA BARROS

Aprovada como parte das exigências para obtenção do Título de Doutora em Ecologia, Evolução e Biodiversidade, área: Biodiversidade pela Comissão Examinadora:

Prof. Dr. TADEU DE SIQUEIRA BARROS (Participação Virtual)
Departamento de Biodiversidade / UNESP - Instituto de Biociências de Rio Claro - SP

Prof. Dr. GABRIEL NAKAMURA DE SOUZA (Participação Virtual)
Departamento de Ecologia (Pesquisador Pós-doc) / USP - Universidade de São Paulo - SP

Professora Assistente DANIELLE KATHARINE PETSCH (Participação Virtual)
Departamento de Ciências Biológicas / UNESP - Faculdade de Ciências e Letras de Assis / SP

Rio Claro, 23 de agosto de 2024

AGRADECIMENTOS

Primeiramente, agradeço ao meu orientador, Tadeu Siqueira. Cada correção, cada crítica, cada conselho que recebi ao longo destes anos me guiaram não apenas na construção desta tese, mas também na minha formação como pessoa e pesquisadora. Apesar de ter sido uma orientação à distância, Tadeu sempre foi um orientador presente, sendo super profissional e empático. Sou muito grata por todos os seus ensinamentos!

Agradeço também à professora e grande pesquisadora Marina Corrêa Côrtes pelo tempo dedicado à minha orientação durante um ano no estágio de docência. Nossas conversas e seus ensinamentos foram sempre muito enriquecedores;

aos meus amigos, familiares e ao meu marido que me acompanharam e me incentivaram ao longo desta jornada, me ajudando a não pirar nos momentos mais desafiadores;

a um revisor ou revisora anônimo da revista *Proceedings of the Royal Society B*, cuja revisão foi essencial ao desenvolvimento da minha tese;

ao Programa de Pós-Graduação em Ecologia, Evolução e Biodiversidade e à Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES). O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001.

RESUMO

Compreender os processos e mecanismos ecológicos que moldam a variação biológica em diferentes níveis, desde genes até a biosfera, é um dos maiores e mais persistentes desafios da Ecologia. Embora grande parte das pesquisas se concentre em compreender como a composição de espécies é moldada por relações determinísticas (ex. tolerância aos ambientes poluídos), também é crucial considerar o papel do acaso na dinâmica das comunidades. Nesta tese, investigo como uma faceta do acaso – a estocasticidade demográfica – influencia a variação espacial e temporal de comunidades e metacomunidades. Eu analisei extensas bases de dados sobre variáveis ambientais e abundância de peixes em riachos ao longo do tempo, utilizando duas abordagens: modelagem matemática baseada em princípios que regem a montagem de comunidades e informada por dados observados; e modelagem estatística baseada em modelos lineares generalizados e informada por modelos matemáticos. No primeiro capítulo, testei a hipótese de que a composição de espécies em comunidades formadas por populações pequenas é mais afetada pela deriva ecológica e menos influenciada pela seleção ambiental do que comunidades compostas por populações grandes. Relacionei uma medida de diversidade beta com o tamanho de 541 comunidades distribuídas entre 32 metacomunidades. Depois, eu medi a força da relação entre a composição de espécies e o ambiente dentro dessas metacomunidades. Encontrei que a composição de espécies variava mais entre as comunidades conforme reduziam em tamanho e que a relação comunidade-ambiente enfraquecia com a diminuição do tamanho. No segundo capítulo, investiguei como a estocasticidade demográfica se propaga para níveis mais altos de organização biológica no tempo. Selecionei 39 metacomunidades compostas no total por 468 comunidades, amostradas entre 1981 e 2019. Medi a variabilidade temporal da composição de espécies em escala local e regional e relacionei com o tamanho das comunidades, a extensão e o número de amostras das séries temporais, a riqueza de espécies e com a variação temporal de variáveis ambientais. O tamanho das comunidades foi o principal fator determinante da variabilidade temporal da composição de espécies em escala local, com comunidades menores exibindo maior variação ao longo do tempo. Em escala regional, não encontrei relação estatística entre a variabilidade temporal na composição de espécies e o tamanho das comunidades. Entretanto, os dados exibiram uma tendência negativa podendo indicar que uma relação poderia surgir com mais amostras. Em suma, os resultados corroboram a hipótese de que a composição de espécies em comunidades menores é mais variável no espaço e tempo e avançam em relação a estudos anteriores demonstrando que além de variar mais, a composição de espécies em comunidades menores é menos relacionada ao ambiente (mais imprevisível). Além disso, os resultados indicam que a influência da estocasticidade demográfica sobre a dinâmica temporal de comunidades e metacomunidades depende da escala, sendo mais forte na escala local. É possível inferir que o tamanho da comunidade influencia a importância dos processos de deriva e seleção e que medidas de conservação que visem o aumento do tamanho das comunidades podem evitar variações imprevisíveis na abundância das espécies e reduzir o risco de extinções locais.

Palavras-chave: composição de espécies; deriva ecológica; diversidade beta; estocasticidade demográfica; tamanho da comunidade; variabilidade temporal.

ABSTRACT

Understanding the ecological processes and mechanisms that shape biological variation at different levels, from genes to the biosphere, is one of the greatest and most enduring challenges in Ecology. While much research focuses on understanding how species composition is shaped by deterministic relationships (e.g., tolerance to polluted environments), it is also crucial to consider the role of chance in community dynamics. In this thesis, I investigate how one facet of chance - demographic stochasticity - affects the spatial and temporal variation of communities and metacommunities. I analyzed extensive datasets on environmental variables and fish abundance in streams over time, using two approaches: mathematical modeling based on principles governing community assembly and informed by observed data; and statistical modeling based on generalized linear models and informed by mathematical models. In the first chapter, I tested the hypothesis that the species composition of communities formed by small populations is more affected by ecological drift and less influenced by environmental selection than communities composed of large populations. I related a measure of beta diversity to the size of 541 communities distributed among 32 metacommunities. Then, I measured the strength of the relationship between species composition and the environment within these metacommunities. I found that species composition varied more among communities as they decreased in size and that the community-environment relationship weakened with decreasing size. In the second chapter, I investigated how demographic stochasticity propagates to higher levels of biological organization over time. I selected 39 metacommunities composed of a total of 468 communities, sampled between 1981 and 2019. I measured the temporal variability of species composition at local and regional scales and related it to community size, the extent and number of samples of the time series, species richness, and the temporal variation of environmental variables. Community size was the main determinant of the temporal variability of species composition at the local scale, with smaller communities exhibiting greater variation over time. At the regional scale, I did not find a statistical relationship between the temporal variability in species composition and community size. However, the data exhibited a negative trend, suggesting that a relationship could emerge with more samples. In summary, the results support the hypothesis that species composition in smaller communities is more variable in space and time. They advance previous studies by demonstrating that, in addition to being more variable, species composition in smaller communities is less related to the environment, making it more unpredictable. Furthermore, the results indicate that the influence of demographic stochasticity on the temporal dynamics of communities and metacommunities depends on the scale, being stronger at the local scale. It can be inferred that community size influences the importance of drift and selection processes and that conservation measures aimed at increasing community size can avoid unpredictable variations in species abundance and reduce the risk of local extinctions.

Keywords: species composition; ecological drift; beta diversity; demographic stochasticity; community size; temporal variability.

SUMÁRIO

APRESENTAÇÃO	9
1 INTRODUÇÃO GERAL	14
1.1 Referências	22
2 CAPÍTULO I - High compositional dissimilarity among small communities is decoupled from environmental variation	25
2.1 Introduction.....	27
2.2 Material and Methods	29
2.3 Results.....	33
2.4 Discussion.....	35
2.5 Data availability statement.....	39
2.6 References.....	39
2.7 Supporting Information.....	44
2.7.1 Material and methods	44
2.7.3 Results.....	47
2.7.4 Supporting information references	57
3 CAPÍTULO II - Unveiling the temporal signatures of demographic stochasticity from populations to metacommunities	60
3.1 Introduction.....	62
3.2 Material and Methods	64
3.3 Results.....	68
3.4 Discussion.....	71
3.5 References.....	75
3.6 Supplementary Information: Appendix S1	82
3.6.1 Supplementary Information Text.....	82
3.6.2 Supplementary Figures.....	85
3.6.3 Supplementary Table.....	89
3.6.4 Supporting information references	91
CONSIDERAÇÕES FINAIS	92
ANEXO I	94

APRESENTAÇÃO

Esta tese apresenta a seguinte estrutura: uma introdução geral em que apresento um breve histórico e as bases teóricas sobre a influência da estocasticidade demográfica sobre padrões de composição de espécies em comunidades e metacomunidades; dois capítulos em formato de artigo científico; e uma conclusão geral. O artigo do primeiro capítulo foi publicado na revista *Oikos* e o artigo do segundo capítulo está em processo de análise na revista *Ecology*.

Nesta seção também compartilho algumas experiências obtidas durante o processo de doutorado. Faço isso porque um doutorado não se resume apenas aos resultados obtidos, mas sim a todo o processo de aprendizagem. Além de nos aprofundarmos em conceitos teóricos, análises quantitativas e outros aspectos da atividade científica, lidamos com diversas situações e uma miscelânea de sentimentos, que apesar de não se transformarem em artigos científicos, também fazem parte do processo de construção de uma doutora.

Após terminar a minha graduação em Ciências Biológicas na Universidade Federal do Rio Grande do Sul (UFRGS), me mudei para Manaus para fazer mestrado no Instituto Nacional de Pesquisas da Amazônia (INPA). Durante esse tempo, eu tive grandes experiências, incluindo um intercâmbio e muitas atividades de campo relacionadas a ecologia de populações e comunidades. Entretanto, ao terminar o mestrado eu sentia falta de um aprofundamento maior em abordagens quantitativas aplicadas ao estudo da biodiversidade. Motivada por esse desejo, em meados de 2020, comecei a procurar por uma possível orientação com esse perfil e encontrei o Tadeu.

Pouco tempo depois das primeiras conversas com o Tadeu, fui aceita no Programa de Pós-Graduação e no início de setembro do mesmo ano, consegui uma bolsa de doutorado da Capes. Entretanto, estávamos em período de pandemia (Covid-19). As universidades estavam fechadas e não havia nenhuma previsão de retorno às atividades presenciais. Decidimos propor um projeto baseado na análise de dados já coletados de peixes. Dessa forma, além de poder aprender a extrair informações de bases globais eu continuaria trabalhando com o grupo de

organismos que eu tinha maior familiaridade. Na minha opinião, essa decisão foi acertada, pois a UNESP só retomou as atividades presenciais em abril de 2022.

Após algumas reuniões e leituras, eu havia escrito um projeto mais voltado para conservação e que não ia tanto de encontro da ecologia de metacomunidades. Então, meu orientador sugeriu que eu escrevesse o primeiro capítulo do projeto aprofundando as ideias de um trabalho recente dele. Assim, ele me auxiliaria bastante a definir as ideias e análises do primeiro capítulo e esse auxílio reduziria gradativamente na escrita do segundo e terceiro capítulos. Com base nisso, escrevi o projeto composto por três capítulos: no primeiro, eu investigaria os efeitos da estocasticidade demográfica sobre a composição de espécies; no segundo, propus entender como a substituição de espécies varia entre diferentes categorias tróficas e como a heterogeneidade ambiental pode influenciar essa variação. No terceiro capítulo, propus investigar a existência de um gradiente latitudinal de substituição de espécies de peixes no tempo. Entretanto, na prática, o projeto inicial acabou sendo bastante alterado.

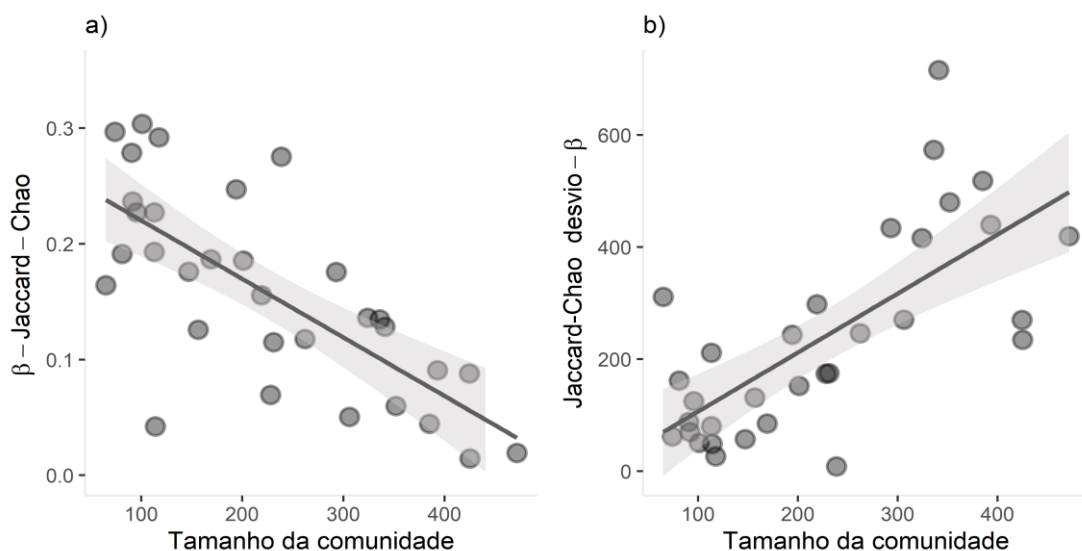
Devido a grande quantidade de informações e detalhes presentes no tema do primeiro capítulo, acabei dividindo-o em dois capítulos distintos: um para investigar o papel do acaso em gerar padrões de diversidade no espaço e outro no tempo. Outro ponto importante, que afetou o andamento da tese, foi a descoberta de um problema metodológico no meu primeiro capítulo após já tê-lo submetido a uma revista (explicarei isso em mais detalhes ainda nesta seção). Além disso, ao conhecer melhor a base de dados que utilizei, percebi que, ao contrário do que eu esperava, eu não teria dados robustos para desenvolver a ideia do gradiente latitudinal. Por fim, após finalizar o segundo capítulo da tese, optei por tentar desenvolver um terceiro capítulo que se relacionasse mais com os dois anteriores. Comecei a investigar se encontraria relações entre o tamanho do ecossistema e a força de processos determinísticos e estocásticos usando dados de peixes e macroinvertebrados. No entanto, consegui realizar apenas algumas análises preliminares, que não foram suficientes para compor um terceiro capítulo para a tese.

Em meados de abril de 2022, eu havia finalizado o artigo referente ao primeiro capítulo da tese e submetido à revista *Proceedings of the Royal Society B*. Enquanto aguardava a resposta da revista organizei e apresentei minha

qualificação. Em seguida, iniciei a seleção e análise de dados para o segundo capítulo. Após quase três meses de espera, a resposta da revista havia finalmente chegado. O revisor ou revisora havia feito um excelente trabalho e encontrado um problema nas minhas análises, problema que invalidava meus resultados!

Resumidamente, para avaliar os efeitos da estocasticidade demográfica sobre a variação espacial na composição de espécies eu havia feito duas relações: primeiro, (1) eu havia relacionado duas métricas de diversidade beta (β -Shannon e β -Jaccard-Chao) com o tamanho das comunidades e havia encontrado uma relação negativa (Figura 1a), indicando que a composição de espécies entre comunidades menores variava mais no espaço. Depois, (2) eu demonstrava que existia uma relação positiva entre o desvio beta (diferença da diversidade beta observada nas comunidades reais e a diversidade beta média de 1000 simulações geradas em um modelo nulo) e o tamanho das comunidades (Figura 1b), sugerindo que a composição de espécies das comunidades menores era mais parecida a uma composição de espécies aleatória. Esse resultado indicava uma maior atuação de processos estocásticos em comunidades menores, enquanto comunidades maiores, por apresentarem menor similaridade com uma composição aleatória, estariam sob maior influência da seleção de espécies.

Figura 1 - Exemplo de parte dos resultados encontrados na primeira versão do Capítulo 1. a) Relação negativa entre a diversidade beta e o tamanho das comunidades. b) Relação positiva entre o desvio beta e o tamanho das comunidades.



Fonte: Elaborada pela autora.

Entretanto, a pessoa que revisou meu trabalho demonstrou que as relações que eu havia encontrado podiam ser geradas apenas por processos probabilísticos sem necessidade de invocar a seleção. Para demonstrar isso o revisor simulou as mesmas relações com comunidades governadas apenas pela dinâmica neutra (sem nenhuma influência da seleção) e que variavam apenas em relação ao seu tamanho. A relação 1 podia ser encontrada apenas porque comunidades menores têm menor probabilidade de compartilhar espécies do que comunidades grandes devido a efeitos de amostragem. Além disso, como a agregação intraespecífica (tendência de indivíduos da mesma espécie se agruparem em um local) é em média maior em comunidades maiores, o desvio beta tende a aumentar com o tamanho das comunidades, gerando a relação 2. Diante disso, eu não podia sugerir que comunidades maiores seriam governadas pela seleção ambiental em contraste as comunidades menores.

Todo o tempo que eu havia dedicado para desenvolver o primeiro capítulo tinha ido por água abaixo. Nesse momento foi fundamental a orientação do Tadeu, que me fez entender que eu não deveria me preocupar com o tempo que havia passado e sim tentar usar as simulações mostradas pelo revisor para encontrar alguma métrica que pudesse ser mais adequada a minha pergunta. Afinal, o fato de o meu método de análise não ser adequado, não significava que as hipóteses e expectativas estivessem erradas. Naquele momento, eu estava diante de um grande desafio: descobrir quais métodos poderiam ser usados para testar as hipóteses que tínhamos interesse e que não seriam afetadas pelo tamanho das comunidades. Comecei a ler artigos e testar várias métricas de diversidade usando simulações baseadas em processos (isso está detalhado nos artigos e por isso não entrarei em detalhes aqui), até que encontrei uma medida de diversidade beta mais adequada para responder minhas perguntas. Agradeço o grande trabalho daquele revisor ou revisora, pois isso mudou minha tese. Afinal, é assim que a ciência avança, encontrando erros e corrigindo-os. Após eu analisar e escrever uma nova versão do Capítulo 1, ele foi aceito para publicação na revista científica Oikos! Todo este trabalho foi ainda essencial para o desenvolvimento do segundo capítulo, em que também tive que realizar simulações baseadas em processos.

Já em 2023, além de estar desenvolvendo o segundo capítulo, cursei a última disciplina obrigatória que me faltava, pois já tinha feito as demais no formato online, e comecei a fazer o estágio de docência. O estágio teve a duração de dois

semestres, sendo que o primeiro estágio foi com a disciplina de Seminários Gerais e o segundo com Ecologia de Micro-organismos. Tive a sorte de escolher uma ótima professora, a Marina, para ser minha supervisora. Aprendi muito com ela e tive experiências muito enriquecedoras na docência.

Paralelamente ao desenvolvimento do doutorado, também realizei outras atividades que considero experiências muito importantes. Por exemplo, atuei como revisora de um manuscrito submetido à *Journal of Fish Biology*, o que me fez refletir bastante sobre o processo de revisão de artigos. Também apresentei os resultados do Capítulo I em um congresso e realizei algumas parcerias de trabalho com antigos orientadores. Estas parcerias resultaram na escrita de dois artigos científicos que já estão publicados e na escrita do capítulo de um livro que será publicado ainda este ano (Anexo I). Enfim, acredito que toda a bagagem de experiências e aprendizados que a dinâmica do doutorado proporcionou se traduzem no processo de virar doutora.

1 INTRODUÇÃO GERAL

A diversidade biológica, ou biodiversidade, é representada por uma incrível variedade e quantidade de formas de vida. Toda essa diversidade se expressa em diferentes níveis, desde a variação genética entre indivíduos da mesma espécie até a variação no número de espécies diferentes em locais específicos, entre locais numa mesma região, entre regiões num mesmo bioma, entre biomas num mesmo continente, e entre diferentes regiões no globo, dando origem a padrões bastante complexos. Os padrões de composição de espécies não são estáticos, pois podem variar de um local para outro no espaço, ou se modificar ao longo do tempo. Isso porque os padrões de diversidade variam de acordo com diversos fatores, por exemplo: a escala em que são observados (Levin, 1992), as diferentes interações entre espécies e indivíduos (Leibold *et al.*, 2004; Zhang *et al.*, 2018), gradientes ambientais como altitude e latitude (Rohde, 1992; Hillebrand, 2004), e variações ambientais, como as mudanças climáticas (Pech *et al.*, 2017). Embora muitos padrões já tenham sido descritos e estudados em comunidades biológicas, ainda há muito a ser descoberto sobre como eles são gerados, mantidos e como podem mudar no futuro. Compreender os processos e mecanismos por trás dos padrões de composição de espécies é um dos maiores e mais persistentes desafios da Ecologia.

Entre o final do século passado e o início do século XXI, duas teorias distintas haviam sido propostas para explicar a estrutura e composição de comunidades biológicas: a Teoria de Nicho (Hutchinson, 1959; MacArthur; Levins, 1967) e a Teoria Neutra (Hubbell, 2001). Apesar de ambas terem sido propostas para explicar os padrões de diversidade de espécies, elas apresentavam visões opostas e mutuamente exclusivas sobre a organização das comunidades. De acordo com a perspectiva da Teoria de Nicho, a composição de espécies pode ser determinada pela coexistência de espécies com diferentes nichos ecológicos. A definição de nicho ecológico foi apresentada pela primeira vez por Grinnell em 1917, que dizia que o nicho engloba tudo que condiciona a existência de uma espécie em um local, incluindo fatores abióticos e interações com outros membros da comunidade. Em 1927, Elton definiu o nicho de uma espécie como o seu papel funcional e o seu impacto no ambiente. Hutchinson, em 1957, propôs um avanço

importante ao definir o nicho como um espaço multidimensional, com cada dimensão representando um fator ambiental relevante para a sobrevivência da espécie. Dessa forma, ao comparar os nichos de diferentes espécies seria possível quantificar a sobreposição entre eles, ajudando a entender como a competição por recursos influencia a coexistência ou a exclusão de espécies. Entretanto, como grande parte da Teoria de Nicho era baseada em modelos de competição, então Chase e Leibold (2003) apresentaram uma visão de nicho mais pluralista que reconhecia a importância de múltiplos mecanismos ecológicos (como predação, mutualismo, perturbações ambientais) em explicar a coexistência de espécies. Segundo os autores, o nicho de uma espécie reflete as condições ambientais mínimas necessárias para uma espécie se manter e como ela impacta essas condições ambientais (Chase; Leibold, 2003). A coexistência pode então ser mediada em função de: mecanismos equalizadores, que reduzem as diferenças de aptidão entre as espécies, tornando-as mais parecidas em termos de capacidade de sobreviver e reproduzir, como quando um predador ataca preferencialmente uma espécie mais abundante, reduzindo sua população e permitindo que espécies menos comuns se proliferem, e; mecanismos estabilizadores, que reduzem a sobreposição de nicho, tornando o nicho das espécies mais distinto (Chesson, 2000). Um exemplo é a partição temporal de padrões de atividade entre predadores de topo e mesocarnívoros a fim de evitar a competição e encontros diretos entre as espécies (Schuette *et al.* 2013).

Em contrapartida, a Teoria Neutra assume que todos os indivíduos em um determinado nível trófico são ecologicamente equivalentes e a dinâmica das espécies é gerada através de flutuações aleatórias na capacidade de sobreviver, reproduzir e se dispersar. Dessa forma, a composição de espécies muda ao longo do tempo de acordo com a aleatoriedade de eventos de nascimentos e mortes – um mecanismo conhecido como estocasticidade demográfica (Lande, 1993; Lande; Engen; Saether, 2003; Melbourne, 2012). A estocasticidade demográfica pode gerar flutuações aleatórias porque os eventos demográficos são probabilísticos e discretos (Melbourne, 2012). Ser probabilístico significa que cada evento demográfico tem uma probabilidade de ocorrer para cada indivíduo de uma população. Ou seja, cada indivíduo tem uma probabilidade específica de morrer ou sobreviver. Ainda que as taxas de natalidade e mortalidade da população sejam conhecidas, o nascimento ou a morte de um indivíduo não é totalmente previsível. Isso ocorre porque

características intrínsecas e extrínsecas ao indivíduo, como diferenças de fecundidade, de tamanho corporal, idade, ciclo de vida, estocasticidade ambiental, influenciam a sua aptidão e introduzem um elemento aleatório em cada evento demográfico (Melbourne, 2012; Lee; Saether; Engen, 2011). Essa imprevisibilidade somada ao fato de o resultado de cada evento demográfico ser discreto e independente (um indivíduo nasce ou não, morre ou não) causa variação nos resultados individuais que podem se acumular e levar a flutuações aleatórias no tamanho da população. Essas flutuações são maiores quando a população é pequena porque poucos eventos combinados com alguma alta probabilidade são suficientes para afastar o tamanho da população do valor esperado. Dessa forma, a estocasticidade demográfica aumenta em importância com a redução do tamanho da população, causando maior variação no seu tamanho, sua estrutura e sua distribuição ao longo do tempo, consequentemente aumentando o risco de extinção de espécies (Lande, 1993; Pimm; Jones; Diamond, 1988; Reed; Hobbs, 2004).

Uma analogia que ajuda a compreender o conceito de que eventos aleatórios afetam as populações de maneira diferente de acordo com o seu tamanho, é o de imaginar o lançamento de uma moeda. Cada lançamento tem uma probabilidade de 50% de resultar em cara ou coroa, mas não é possível prever com certeza o resultado de cada jogada individual. Além disso, a chance de obter cara ou coroa em qualquer lançamento é independente dos lançamentos anteriores, assim como os eventos demográficos de cada indivíduo numa comunidade biológica. Ao jogarmos uma moeda 100 vezes, mesmo que seja possível obter uma sequência de caras ou coroas no início, eventualmente, o número total de caras e coroas será próximo ao valor esperado de 50 cada. No entanto, se lançarmos essa moeda apenas 10 vezes, a influência de qualquer cara ou coroa individual se torna muito maior. Uma sequência de coroas ou caras consecutivas pode distorcer significativamente o resultado esperado (50% de cada) em apenas 10 lançamentos.

A ideia de que eventos aleatórios poderiam explicar a variação na composição de espécies em comunidades gerou um intenso debate na Ecologia. Além de apresentar poucos pressupostos, a Teoria Neutra apresentou explicações consistentes para alguns padrões ecológicos, como a distribuição de abundância das espécies (*species abundance distribution* – SAD; Hubbell, 2001). Essa distribuição, caracterizada por poucas espécies abundantes e muitas raras, pode ser explicada por uma maior influência da deriva na abundância das espécies raras na

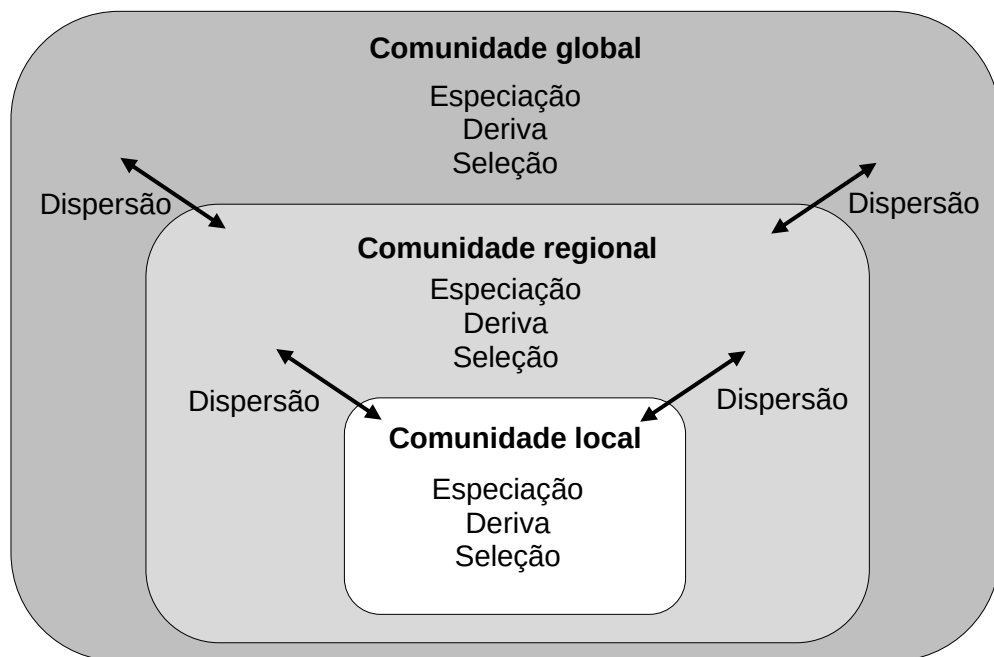
comunidade, e menor nas espécies abundantes. Entretanto, o fato de a Teoria Neutra assumir que os padrões de diversidade seriam gerados apenas por processos aleatórios foi considerado extremo por muitos pesquisadores e recebeu diversas críticas por não levar em consideração as diferenças de nicho entre as espécies. Logo, surgiram estudos demonstrando que modelos incluindo diferenças de nicho também podiam prever a distribuição de abundância das espécies, além de também explicarem outros padrões e serem importantes em modelos preditivos de distribuição de espécies (Chave *et al.*, 2002; McGill, 2003; McGill *et al.*, 2006). Toda essa discussão levou a diversas comparações entre as teorias e a busca por uma explicação abrangente sobre a estrutura e a dinâmica das comunidades levou à conclusão de que tanto os processos determinísticos (baseados no nicho) quanto os estocásticos (baseados no acaso) atuam em conjunto para explicar os padrões de diversidade (Matthews; Whittaker, 2014).

Essa visão integrada contribuiu para o surgimento de novas teorias, combinando elementos de ambas as perspectivas em diferentes escalas espaciais (Munoz; Huneman, 2016). Um marco nesse processo foi o desenvolvimento da Teoria de Metacomunidades, que define uma metacomunidade como um conjunto de comunidades locais interligadas pela dispersão e pelas interações entre diversas espécies (Leibold *et al.*, 2004). Essa nova perspectiva, além de considerar a atuação de processos determinísticos e estocásticos em gerar padrões locais de diversidade, também leva em consideração o papel desses processos em moldar padrões regionais de composição de espécies. Processos determinísticos e estocásticos passam a ser vistos como duas partes complementares ao longo de um continuum de forças ecológicas que moldam a estrutura das comunidades, com a estocasticidade assumindo um papel fundamental na geração de padrões ecológicos (Gravel *et al.*, 2006).

Utilizando a abordagem de metacomunidades, em 2010, Vellend propôs uma maneira de organizar e unificar conceitualmente a ecologia de comunidades. Ele sugere que, de forma geral, os padrões de composição de espécies são influenciados por apenas quatro tipos de processos ecológicos: a seleção de espécies, a deriva ecológica, a especiação e a dispersão. Aqui a deriva ecológica representa a força de processos estocásticos, enquanto a seleção de espécies representa basicamente processos baseados na teoria de nicho. As espécies são adicionadas às comunidades através de especiação e dispersão, e as abundâncias

relativas destas espécies são então moldadas pela deriva e pela seleção, bem como pela dispersão contínua que impulsiona a dinâmica das comunidades sob diferentes escalas (Figura 1; Vellend, 2010).

Figura 1 - Representação gráfica da teoria de comunidades ecológicas através de escalas espaciais (local, regional e global). Comunidade local: conjunto de espécies que coexistem em uma área específica; Comunidade regional: conjunto de comunidades locais; Comunidade global: conjunto de comunidades regionais.

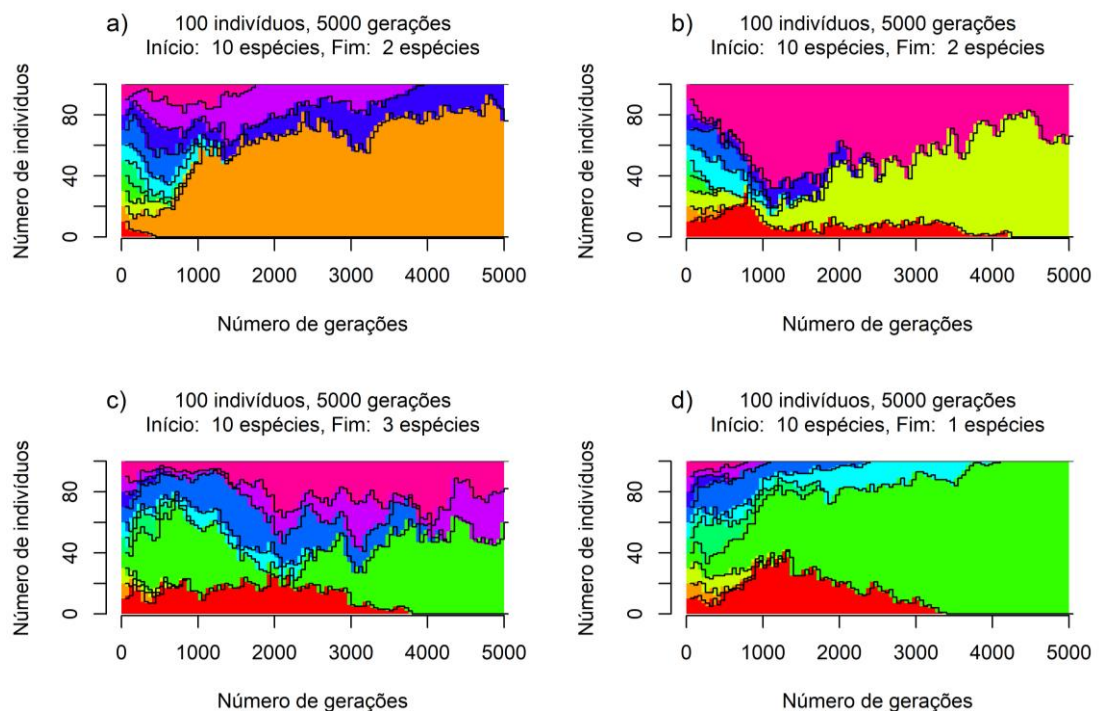


Fonte: Elaborada pela autora (modificada de Vellend, 2016, página 67).

Mas como especificamente a deriva ecológica pode afetar a composição de espécies em comunidades e metacomunidades? Se considerarmos a dinâmica de uma comunidade apenas sob o efeito da deriva ecológica, sem considerar a influência dos outros processos ecológicos, a abundância das espécies deve variar de forma aleatória, independentemente de suas identidades. Eventualmente, uma das espécies irá dominar a comunidade, enquanto a abundância de todas as outras espécies oscilarão até a extinção local (Hubbell, 2001). A probabilidade que cada espécie possui em atingir essa monodominância é igual a sua frequência inicial, e a taxa na qual isso é alcançado é negativamente relacionada ao tamanho da comunidade (Hubbell, 2001; Vellend, 2010). Para exemplificar, imagine uma comunidade formada por 100 indivíduos, divididos entre 10 espécies. Considere que

aleatoriamente, a cada geração, um indivíduo da comunidade morre e outro indivíduo nasce e ocupe o seu lugar. Sabendo que as espécies têm uma proporção igual de indivíduos no início da simulação e que não existem diferenças de fitness entre os indivíduos, a probabilidade de extinção é igual para todas as espécies. Ao realizar diferentes simulações da dinâmica dessa comunidade, é possível observar que, em cada simulação, a abundância das espécies varia aleatoriamente ao longo das gerações (Figura 2). Além disso, as espécies que crescem ou reduzem em abundância em uma simulação não são necessariamente as mesmas que crescem ou reduzem em outra simulação (Figura 2). Também podemos observar que quanto menor a frequência de uma espécie na comunidade, maior é a probabilidade dessa espécie ser extinta ao longo das gerações (Figura 2). Na quarta simulação, a monodominância foi atingida, situação em que a espécie de cor verde passa a representar 100% da abundância dos indivíduos na comunidade (Figura 2 – d).

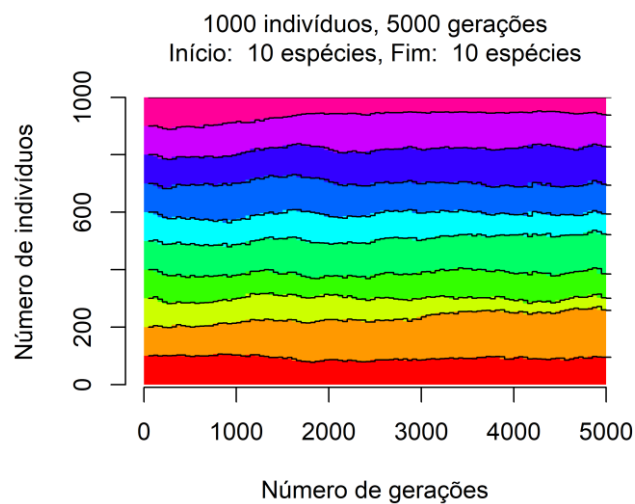
Figura 2 - Simulações da deriva aleatória da abundância de 10 espécies dentro de uma comunidade composta por 100 indivíduos ao longo de 5000 mil gerações. Cada cor representa a abundância de indivíduos por espécie.



Fonte: Elaborado pela autora (Modificado de Zelený, D. Analysis of community ecology data in R. 2017. Disponível em: https://www.davidzeleny.net/anadat-r/doku.php/en:customized_functions:random.drift).

Em virtude da natureza aleatória da deriva ecológica, não há como prever o número de gerações que levaram ao cenário de monodominância. Entretanto, quanto mais indivíduos a comunidade tiver, maior é a quantidade de gerações necessárias para que a monodominância ocorra (Hubbell, 2001). Em uma simulação similar a descrita no parágrafo anterior, mas com 1000 indivíduos, podemos observar que a abundância das espécies oscilou menos ao longo do tempo e em 5000 mil gerações nenhuma espécie foi extinta (Figura 3).

Figura 3 - Deriva aleatória da abundância de 10 espécies dentro de uma comunidade composta por 1000 indivíduos ao longo de 5000 mil gerações.



Fonte: Elaborada pela autora (Modificado de Zelený, D. Analysis of community ecology data in R. 2017. Disponível em: https://www.davidzeleny.net/anadat-r/doku.php/en:customized_functions:random.drift).

Diante disso, é possível prever que declínios no tamanho das comunidades devem aumentar a importância da deriva ecológica. Segundo Vellend (2016), comunidades menores devem apresentar menor diversidade local, maior diversidade beta (variação da composição entre comunidades) e uma relação mais fraca entre a composição de espécies e o ambiente. A composição de espécies deve variar entre comunidades ao longo do espaço, pois diferentes espécies poderão dominar diferentes comunidades locais em um determinado período, mesmo que as espécies dominantes não sejam as mais favorecidas pela seleção (enfraquecendo as relações composição-ambiente). Por esta razão, uma variação espacial na estrutura da composição de espécies entre comunidades pode ser

gerada na ausência de heterogeneidade de habitats ou de diferenças de nicho entre espécies. Da mesma forma, também é esperado que a similaridade de uma comunidade diminua ao longo do tempo, pois espécies se extinguirão e serão substituídas por imigrantes de outras espécies.

Embora a teoria sugira que a deriva ecológica possa ter um papel importante na montagem de comunidades, a maior parte das pesquisas que abordam esse assunto se limitam ao estudo de comunidades teóricas e experimentais bastante simplificadas, existindo poucas evidências empíricas de seus efeitos (Orrock; Fletcher, 2005; Orrock; Watling, 2010; Gilbert; Levine, 2017; Legault; Fox; Melbourne, 2019; Siqueira *et al.*, 2020). Testar empiricamente os efeitos da deriva ecológica é bastante desafiador, uma vez que isolar completamente os seus efeitos de outros processos ecológicos é extremamente difícil, ou até mesmo impossível.

Em um cenário preocupante de declínios populacionais e altas taxas de extinção de espécies (McCallum, 2015; He *et al.*, 2019; Almond; Grooten, Petersen, 2020), torna-se crucial compreender como e quando processos estocásticos são importantes em moldar os padrões de diversidade biológica. Populações de espécies de água doce, por exemplo, estão declinando rapidamente devido a atividades antrópicas, como a degradação do habitat, poluição da água, sobrepesca, introdução de espécies exóticas, fragmentação e mudanças climáticas (Dudgeon *et al.*, 2006; Reid *et al.*, 2019). Dessa forma, compreender como processos estocásticos afetam a biodiversidade é fundamental para entender as respostas das comunidades às mudanças globais e para o desenvolvimento de medidas de conservação eficazes, capazes de frear e reverter a perda de biodiversidade. O aumento da qualidade e quantidade de dados ecológicos, aliado ao desenvolvimento de ferramentas analíticas robustas, nos oferece um caminho promissor para alcançar esse objetivo.

Nesta tese, eu investiguei os efeitos da deriva ecológica sob diferentes níveis biológicos e escalas espaciais, analisando como a estocasticidade demográfica influencia a dinâmica de comunidades e metacomunidades de peixes de riacho na natureza. A estrutura da tese consiste em dois capítulos, que foram desenvolvidos com o apoio do meu orientador. No primeiro capítulo, tivemos como objetivo complementar os achados de Siqueira *et al.* (2020), utilizando dados mais abrangentes e considerando diferentes métricas de diversidade beta. Abordamos a

ideia de que comunidades compostas por menos indivíduos são mais afetadas pela estocasticidade demográfica, resultando em maior variação na composição de espécies entre comunidades (diversidade beta) e menor relação da composição de espécies com o ambiente. Para isso, utilizando dados empíricos coletados em diferentes continentes, avaliamos como a diversidade beta e a força da relação comunidade-ambiente variam com o tamanho das comunidades. No segundo capítulo, nós procuramos entender como a propagação de flutuações aleatórias causadas pela estocasticidade demográfica para níveis biológicos maiores (comunidades e metacomunidades) pode afetar a dinâmica temporal da biodiversidade nas escalas local e regional.

1.1 Referências

ALMOND, R. E. A.; GROOTEN, M.; PETERSEN, T. Bending the curve of biodiversity loss. **Living Planet Report**. Gland: WWF, 2020.

CHASE, J. M.; LEIBOLD, M. A. Ecological Niches. Linking Classical and Contemporary Approaches. 1. ed. Chicago: **University of Chicago Press**, 2003.

CHAVE, J.; MULLER-LANDAU, H. C.; LEVIN, S. A. Comparing classical community models: theoretical consequences for patterns of diversity. **The American Naturalist**, v. 159, n. 1, p. 1–23, 2002.

CHESSON, P. Mechanisms of Maintenance of Species Diversity. **Annual Review of Ecology and Systematics**, v. 31, n. 1, p. 343–366, 2000.

DUDGEON, D. *et al.* Freshwater biodiversity: importance, threats, status and conservation challenges. **Biological Reviews**, v. 81, n. 2, p. 163–182, 2006.

GILBERT, B.; LEVINE, J. M. Ecological drift and the distribution of species diversity. **Proceedings of the Royal Society B: Biological Sciences**, v. 284, n. 1855, p. 20170507, 2017.

GRAVEL, D.; CANHAM, C. D.; BEAUDET, M.; MESSIER, C. Reconciling niche and neutrality: the continuum hypothesis. **Ecology Letters**, 9: 399–409, 2006.

HE, F. *et al.* The global decline of freshwater megafauna. **Global Change Biology**, v. 25, n. 11, p. 3883–3892, 2019.

HILLEBRAND, H. On the Generality of the Latitudinal Diversity Gradient. **The American Naturalist**, v. 163, n. 2, p. 192–211, 2004.

HUBBELL, S. P. **The unified neutral theory of biodiversity and biogeography**. Princeton: Princeton University Press, 2001.

HUTCHINSON, G. E. Homage to Santa Rosalia or Why Are There So Many Kinds of Animals? **The American Naturalist**, v. 93, n. 870, p. 145–159, 1959.

LANDE, R. Risks of Population Extinction from Demographic and Environmental Stochasticity and Random Catastrophes. **The American Naturalist**, v. 142, n. 6, p. 911–927, 1993.

LANDE, R.; ENGEN, S.; SAETHER, B.-E. Stochastic Population Dynamics in Ecology and Conservation. [s.l.] **Oxford University Press**, 2003.

LEE, A. M.; SAETHER, B.-E.; ENGEN, S. Demographic Stochasticity, Allee Effects, and Extinction: The Influence of Mating System and Sex Ratio. **The American Naturalist**, v. 177, n. 3, p. 301–313, 2011.

LEGAULT, G.; FOX, J. W.; MELBOURNE, B. A. Demographic stochasticity alters expected outcomes in experimental and simulated non-neutral communities. **Oikos**, v. 128, n. 12, p. 1704–1715, 2019.

LEIBOLD, M. A. et al. The metacommunity concept: a framework for multi-scale community ecology. **Ecology Letters**, v. 7, n. 7, p. 601–613, 2004.

LEVIN, S. A. The Problem of Pattern and Scale in Ecology: The Robert H. MacArthur Award Lecture. **Ecology**, v. 73, n. 6, p. 1943–1967, 1992.

MACARTHUR, R.; LEVINS, R. The Limiting Similarity, Convergence, and Divergence of Coexisting Species. **The American Naturalist**, v. 101, n. 921, p. 377–385, 1967.

MATTHEWS, T. J.; WHITTAKER, R. J. Neutral theory and the species abundance distribution: recent developments and prospects for unifying niche and neutral perspectives. **Ecology and Evolution**, v. 4, n. 11, p. 2263–2277, 2014.

MCCALLUM, M. L. Vertebrate biodiversity losses point to a sixth mass extinction. **Biodiversity and Conservation**, v. 24, n. 10, p. 2497–2519, 2015.

MCGILL, B. et al. Rebuilding community ecology from functional traits. **Trends in Ecology & Evolution**, v. 21, n. 4, p. 178–185, 2006.

MCGILL, B. J. A test of the unified neutral theory of biodiversity. **Nature**, v. 422, n. 6934, p. 881–885, 2003.

MELBOURNE, B. A.; HASTINGS, A. Extinction risk depends strongly on factors contributing to stochasticity. **Nature**, v. 454, n. 7200, p. 100–103, 2008.

MELBOURNE, B. A. Demographic stochasticity. In: Hastings, A. and Gross, L. J. **Encyclopedia of Theoretical Ecology**. Berkeley, University of California Press, p. 706–712, 2012.

MUNOZ, F.; HUNEMAN, P. From the Neutral Theory to a Comprehensive and Multiscale Theory of Ecological Equivalence. **The Quarterly Review of Biology**, 1, 2016.

ORROCK, J. L.; FLETCHER, R. J. Changes in community size affect the outcome of competition. **The American Naturalist**, v. 166, n. 1, p. 107–111, jul. 2005.

ORROCK, J. L.; WATLING, J. I. Local community size mediates ecological drift and competition in metacommunities. **Proceedings: Biological Sciences**, v. 277, n. 1691, p. 2185–2191, 2010.

PECL, G. T. et al. Biodiversity redistribution under climate change: Impacts on ecosystems and human well-being. **Science**, v. 355, eaai9214, 2017.

PIMM, S.; JONES, H.; DIAMOND, J. On the Risk of Extinction. **American Naturalist**, v. 132, p. 757–785, 1988.

REED, D. H.; HOBBS, G. R. The relationship between population size and temporal variability in population size. **Animal Conservation**, v. 7, n. 1, p. 1–8, 2004.

REID, A. J. *et al.* Emerging threats and persistent conservation challenges for freshwater biodiversity. **Biological Reviews**, v. 94, n. 3, p. 849–873, 2019.

ROHDE, K. Latitudinal Gradients in Species Diversity: The Search for the Primary Cause. **Oikos**, v. 65, n. 3, p. 514, 1992.

SCHUETTE, P. et al. Occupancy patterns and niche partitioning within a diverse carnivore community exposed to anthropogenic pressures. **Biological Conservation**, v. 158, p. 301–312, 2013.

SIQUEIRA, T. *et al.* Community size can affect the signals of ecological drift and niche selection on biodiversity. **Ecology**, v. 101, n. 6, 2020.

VELLEND, M. Conceptual synthesis in community ecology. **The Quarterly Review of Biology**, v. 85, n. 2, p. 183–206, 2010.

VELLEND, M. **The Theory of Ecological Communities** (MPB-57). Princeton University Press. 2016.

ZHANG, J. et al. Trophic interactions among vertebrate guilds and plants shape global patterns in species diversity. **Proceedings of the Royal Society B: Biological Sciences**, v. 285, n. 1883, p. 20180949, 2018.

2 CAPÍTULO I - High compositional dissimilarity among small communities is decoupled from environmental variation

Este manuscrito foi publicado na forma de artigo científico na revista Oikos e está formatado de acordo com as normas definidas pela revista.

Jacobi, C.M. and Siqueira, T. (2023), High compositional dissimilarity among small communities is decoupled from environmental variation. Oikos, 2023: e09802. <https://doi.org/10.1111/oik.09802>

High compositional dissimilarity among small communities is decoupled from environmental variation

Cristina Mariana Jacobi¹ and Tadeu Siqueira^{1,2}

1 Inst. of Biosciences, São Paulo State Univ. (UNESP), Rio Claro, SP, Brazil

2 School of Biological Sciences, Univ. of Canterbury, Christchurch, New Zealand

Abstract: Communities composed of small populations are predicted to be strongly influenced by stochastic demographic events and, thus, less affected by environmental selection than those composed of large populations. However, this prediction has only been tested with computer simulations, simplified controlled experiments, and limited observational data. Here, using multiple datasets on fish abundance in 541 streams we tested 1) if communities composed of small populations are more spatially variable and 2) less related to the environment variation than communities composed of large populations. We used process-based simulations to identify β -diversity metrics and community–environment measures that were appropriate to investigate the role of assembly processes along a gradient of community size. We show that variation in species composition among small communities is higher than among large communities and that the strength of community–environment relationships is weaker in small communities. Our results indicate that community size affects the strength of ecological drift and environmental selection in metacommunities. We thus suggest that further declines in the size of populations and ecosystems can make spatial variation in biodiversity more unpredictable.

Keywords: β -diversity, community size, drift, metacommunity, niche selection

2.1 Introduction

Small populations are relatively more prone to be affected by stochastic events than large populations (Otto and Whitlock 1997). While in larger populations stochastic demographic events tend to be balanced by the greater number of individuals, small populations have a higher probability to deviate from mean expectations (Melbourne 2012). For example, a small population can go locally extinct if all its individuals die before reproduction, an event that is less likely to occur in large populations. Thus, species relative abundance in communities composed of small populations tend to vary more in space and time independent of environmental variation – i.e. they tend to be more affected by ecological drift (Vellend 2016). Although much progress has been made recently towards the robustness of these predictions, this has been made mainly through simulations (Orrock and Watling 2010), simplified experiments (Gilbert and Levine 2017), and the analysis of limited observational data (Siqueira et al. 2020). Here we expand previous findings by not only focusing on spatial variation in species composition, but also by investigating the role of the environmental factors as drivers of that variation (i.e. community–environment relationship) along a gradient of community size in multiple independent data sets from different regions of the world.

In general, we can assume that while environmental selection deterministically assemble species through density-dependent and independent mechanisms, ecological drift causes variation in local species relative abundance through demographic stochasticity (Vellend 2016). Ecologists are increasingly recognizing that stochastic and deterministic processes interact with each other in many ways to form multispecies communities. For example, experimental research found that small plant communities showed high variation in species relative abundances and high levels of local extinction, resulting in lower local diversity and high variation in species composition (Gilbert and Levine 2017). Observational research found higher spatial variation in species composition among small communities of stream insects that resembled a stochastic model (Siqueira et al. 2020). These results are in line with model simulations that suggest that the effects of environmental selection can be overridden by the effects of ecological drift in small communities (Orrock and Watling 2010).

One prediction regarding the effects of ecological drift on communities that has received less attention concerns the decoupling of spatial variation in species composition and in environmental factors. If the role of ecological drift is strong in small communities, then community–environment relationships should be weak in metacommunities composed of small communities. That is, there should be a positive relationship between the strength of environmental selection and community size. This is not easy to test for various reasons, but mainly because of statistical issues, including lack of sufficient spatial replication (Siqueira et al. 2020) and zero-inflated data. The data required to test this prediction must include various replicated metacommunities from which individual measures of the strength of the community–environment relationship are estimated (Bini et al. 2014). Here, we used a global database of riverine fish abundances (Comte et al. 2020) to test the hypothesis that community size mediates the roles of ecological drift and environmental selection in driving community spatial variation (β -diversity). Specifically, we analyzed 32 metacommunities in four biogeographic realms (Australasia, Nearctic, Neotropics and Palearctic) that include 541 local communities.

First, we tested if spatial β -diversity was negatively related to community size. Our aim here was to test the generality of the findings by Siqueira et al. (2020) with more comprehensive data and considering recent research in β -diversity metrics. We expected that β -diversity would be higher among smaller communities and lower among larger communities, even after accounting for the potential effects of environmental heterogeneity and spatial extent. However, quantifying changes in β -diversity along a gradient of community size is challenging because almost all β -diversity metrics depend on both sampling effort and sample size. Thus, we must consider that community size itself can bias the estimates of β -diversity due to differences in the regional species pool size, species abundance distributions (SAD), and the number of individuals (Chase and Knight 2013, Engel et al. 2021). This is undesirable because a theoretically expected negative relationship between β -diversity and community size can be found simply because smaller communities are less likely to share species than larger communities due to sampling effects.

To select a β -diversity metric that was not affected by sampling effects related to community size, we used a process-based simulation model to simulate metacommunities with the same SAD, richness, and number of communities. The simulated metacommunities differed only in their overall size and were assembled

without any influence of environmental selection. A null relationship between β -diversity of simulated metacommunities and community size would indicate that the β -diversity metric does not change with community size without invoking environmental selection.

Second, we tested the importance of community size as a mediator of the relative role of environmental selection versus demographic stochasticity by regressing a measure of the strength of the community–environment relationship within each metacommunity against the median community size. Here, we used community incidence data to measure the strength of the community–environment relationship instead of abundance data to avoid that the relationship could increase with community size simply because statistical power increases with abundance (Gwinn et al. 2016). According to previous research (Magurran and Henderson 2003, Coyle et al. 2013), we expected that infrequent and occasional species should have a weak association with the environmental conditions in the community. Thus, we selected the core species of each community to measure the strength of the community–environment relationship, since they are abundant and widespread, and are more likely to be affected by niche-selection (Magurran 2007). We expected that the spatial variation in species composition of larger metacommunities would have a stronger relationship with the environment indicating the preponderance of deterministic environmental selection over demographic stochasticity in communities composed of larger populations. For this second prediction, we also used a process-based simulation model to investigate if the strength of the community–environment relationship would not increase with community size just because larger communities provide more statistical power to detect an association.

2.2 Material and Methods

Species data

We used publicly available fish abundance data sampled in riverine ecosystems in different continents, but mainly in Europe and North America (RivFishTIME – Comte et al. 2020). In the RivFishTIME database, sampling sites are organized within river basins. Each site we selected was assigned to a Strahler

stream order within the HydroRIVERS network (Lehner and Grill 2013). We also performed a visual inspection of orders attributed to sites located more than 1000 m away from the network, changing to the appropriate order, if needed. Then, we defined a community as the set of species occurring in a site and a metacommunity as the set of sites within basin delineations according to HydroBASINS (level 7 – Lehner and Grill 2013). We only considered sites that included direct fish counts and selected the most recent sampling event for each metacommunity that included at least 10 communities in first to third-order streams, excluding basins with less than five species. These procedures resulted in 32 metacommunities. Data selection was made in R ver. 4.2.1 (www.r-project.org) using the packages *mapview* (Appelhans et al. 2021) *raster* (Hijmans 2021a), *rgdal* (Bivand et al. 2021), *sf* (Pebesma 2018), *sp* (Pebesma and Bivand 2005, Bivand et al. 2013) and *dplyr* (Wickham et al. 2023).

Environmental data

For consistency and because local water variables (e.g. pH, temperature, dissolved oxygen) were not available, we obtained environmental variables relevant to freshwater ecosystems by snapping site geographic coordinates in two global remote-sensing datasets, HydroATLAS (Linke et al. 2019) and free-flowing rivers (FFR) (Grill et al 2019). Both datasets consider a reach as the smallest unit in the hydrological network, being characterized as the line segment between two confluences. Spatial data gathered by HydroATLAS include mean annual natural discharge in $\text{m}^3 \text{ s}^{-1}$ (Döll et al 2003), average elevation (m) (Robinson et al. 2014), average annual air temperature ($^{\circ}\text{C}$), average annual precipitation (mm) (Hijmans et al. 2005), and average human footprint index (Venter et al. 2016) within the local catchment that drains directly into the reach. From the FFR dataset we obtained the connectivity status index (CSI) that was calculated based on five pressure factors (river fragmentation, flow regulation, sediment trapping, water consumption and infrastructure development). These factors are associated with the four dimensions of freshwater connectivity (longitudinal, lateral, vertical and temporal). The CSI provides a characterization of connectivity to every individual river reach, ranging from 0 to 100% connected. We also used a categorical variable (CSI_FF2) based on CSI that indicates if a river reach belongs to a river with free-flowing, good connectivity or degraded status. Together, data from HydroATLAS and FFR were used as our environmental variables since they reflect hydrologic, physiographic, climatic and

anthropogenic features of the freshwater environment. These variables were also selected in R using packages (www.r-project.org) mapview, raster, rgdal, sf, sp and dplyr.

Community descriptors

We measured the size of each community as the number of individuals by site and called community size the median community size within each metacommunity. Thus, larger metacommunities were those formed by communities with more individuals and smaller metacommunities were those formed by communities with fewer individuals.

We used a process-based simulation model to select a β -diversity metric that was not dependent on sampling effects or directly affected by community size. To do this, we simulated spatially explicit metacommunities without the influence of environmental selection with the package mobsim (May et al. 2021). Simulated metacommunities had the same species abundance distributions (SAD, log-normal), richness, and the number of local communities, but differed in terms of total size (taken from a uniform distribution that varied according to the range of number of individuals sampled in the empirical metacommunities, Supporting information). We also represented the intraspecific aggregation caused by dispersal limitation considering that all communities were equally likely to be colonized by the initial member of each species, but once they were, subsequent members of that species were more likely to colonize surrounding locations. Metacommunities were simulated with different numbers of species and sites, according to values in the empirical data (Supporting information). For each of these combinations, we estimated different β -diversity metrics and regressed these β -diversity metrics against community size. Among the many β -diversity metrics we tested, rank-abundance difference was the only one that was not affected by community size in a neutral-like simulation scenario, and thus it was the metric selected to be used in this study. In the supporting information, we describe further metrics that failed the tests (Supporting information).

To describe variation in species rank-abundance patterns for the empirical metacommunities, we measured the median rank difference of each metacommunity as the difference in the rank of species between all pairs of communities within a metacommunity and divided it by the number of unique species in each pair. Rank

difference values were bound between 0 and 0.5 (maximum rank change). The R package *codyn* (Hallett et al. 2020) was used to measure rank-abundance differences. We fitted a multiple linear Gaussian model to describe the relationship between rank differences and community size, including metacommunity environmental heterogeneity and spatial extent as covariates (Supporting information for model fit assumptions). These covariates were added to take into account that larger regions could be more environmentally heterogeneous and have less connected communities that could confound the effects of community size on β -diversity. We quantified environmental heterogeneity within metacommunities using the standardized environmental variables of each community in a distance-based analysis of homogeneity of multivariate dispersions (Euclidean distance). This was done using the *betadisper* function from the *vegan* package in R (Oksanen et al. 2020). To measure the spatial extent of each metacommunity we calculated their centroids and measured the average distance of each of their communities to it. This analysis was also based on Euclidean distance calculated in meters with the *geosphere* package in R (Hijmans 2021b).

Community–environment relationships

To investigate the strength of the community–environment relationships within metacommunities and relate to the median community size we also first employed a process-based model to define which statistic would not be affected by community size itself. To do this, we simulated metacommunities that were equally affected by demographic stochasticity and environmental selection regardless of their size. The simulated metacommunities were colonized by species with similar niche breadth (0.2) and that followed a Gaussian response to one environmental variable. Species environmental optimum ranges were taken from a uniform distribution ranging from 0 to 5. We simulated ecological drift by drawing the abundance of every species in each local community from a Poisson distribution whose mean was given by the deterministic influence of environmental selection. We assigned different sizes to the metacommunities estimating species expected abundance in optimum environments from a random interval. Then we modeled the strength of the community–environment relationship within metacommunities. An unbiased measure of the strength of community–environment relationships should not be related to community size.

We tested a generalized linear model (GLM) to relate the strength of the community–environment relationship with community size, but we found that the relationship increased with community size (Supporting information), indicating that the GLM explanatory power increased with community size even for metacommunities assembled by equal contributions of drift and niche selection. Our second approach was to run a hierarchical modelling of species communities (HMSC) (Ovaskainen 2017), using the package HMSC (Tikhonov et al. 2021). HMSC is a powerful method as it can increase the predictive power identifying biotic interactions and the influence of missing environmental covariates (Tikhonov et al. 2020). The posterior distribution of the HMSC model was sampled with four Markov chain Monte Carlo (MCMC) chains. Each chain was run for 150 000 iterations, of which the first 50 000 were removed as burn-in. We obtained 1000 samples per chain that were recorded every 100th step (thin = 100) of the iterations. Then, we measured the explanatory power of the model obtaining a median coefficient of discrimination (Tjur R²) (Tjur 2009) for each simulated metacommunity that was regressed against their community size. We did not find a relationship between the HMSC Tjur R² and community size (Supporting information). Thus, we used the HMSC approach to answer our question using the same steps described above. We used a probit model to relate the response matrix, composed of the incidence of the core species of each community (10% most abundant), to the matrix of environmental variables. Because we did not include any random effects in the model, Tjur R² represented the total amount of variance explained only by the environmental variables (fixed effect). We also examined the MCMC convergence by examining the potential scale reduction factors (Gelman and Rubin 1992) of the model parameters.

2.3 Results

As expected, we found a negative relationship between rank-abundance difference and community size (Fig. 1a, Table 1). Neither environmental heterogeneity nor spatial extent were related to this β -diversity metric, indicating that 38.8% of the variation in rank-abundance difference was due uniquely to community size. These results indicate that fish composition in small communities was more

different among each other within a metacommunity compared to metacommunities composed of larger communities.

The MCMC convergence of the HMSC models was satisfactory, with the potential scale reduction factors for the β -parameters varying between 0.999 and 1.014. The coefficient of determination (Tjur R^2) varied between 0.10 and 0.67 (median = 0.20; standard deviation [SD] = 0.14). Considering only the total amount of variance in species abundance explained by environmental variables, average annual air temperature (median = 0.19; SD = 0.05), average elevation (median = 0.18; SD = 0.05), and connectivity status index (CSI; median = 0.16; SD = 0.07) explained the largest proportion of variance across all species (Supporting information). The highest values of explained variance, however, were found for mean annual natural discharge (0.47) and average human footprint index (0.42; Supporting information). As predicted, we found a positive relationship between Tjur R^2 and community size (Fig. 1b; standardized coefficient = 0.386, t-value = 2.29, R^2 = 0.149, p = 0.029), indicating that the strength of the community–environment relationship was lower in metacommunities composed of smaller communities.

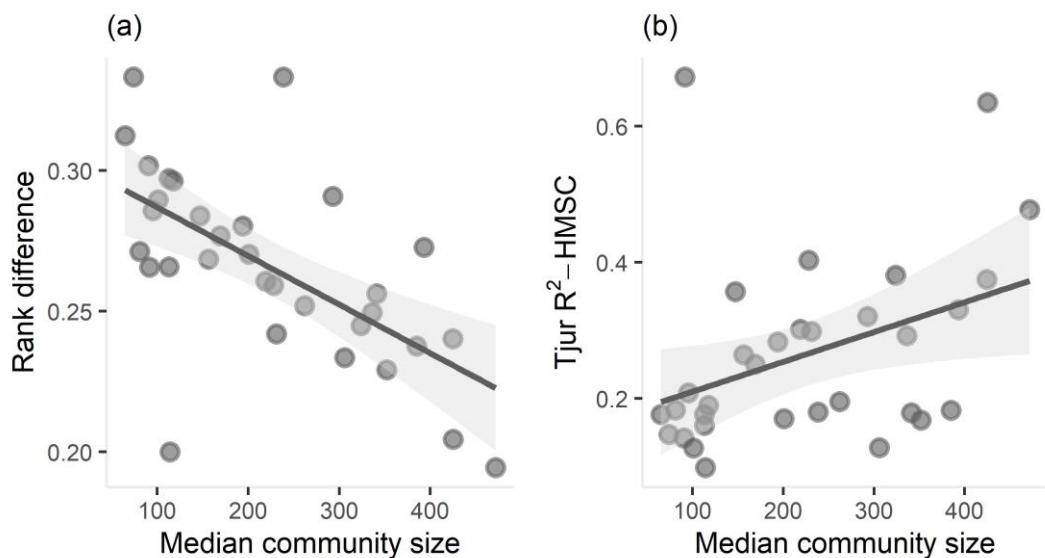


Figure 1. Relationships between (a) rank-abundance difference and median community size; and (b) between the strength of the community–environment relationship within metacommunities against median community size. Gray bands represent the 95% confidence interval around the regression line.

Table 1. Model statistics of the relationship of rank-abundance difference with community size (CS), environmental variation (EV) and spatial extent (SE) in riverine fish metacommunities. SD coeff. = standardized regression coefficient; R² and Adj. R² = coefficient of determination and adjusted coefficient of determination, respectively.

Predictor	SD coeff.	t-value	p-value	R ²	Adj R ²
CS	-0.657	-4.291	<0.001	0.388	0.368
EV	0.044	0.303	0.764		
SE	0.112	0.732	0.470		

2.4 Discussion

The persistence and stability of populations and metapopulations are positively related to their size (Otto and Whitlock 1997, Reed and Hobbs 2004, Terui et al. 2018). Expanding on this perspective and on previous findings (Siqueira et al. 2020) our analyses of intercontinental data support the prediction that spatial variation in species composition among small communities is higher than among large communities. We also found support for the prediction that the strength of the relationship between species composition and environmental variation is weak in metacommunities composed of small communities. Taken together, these results indicate that community size mediates the role of ecological drift and environmental selection as drivers of metacommunity spatial dynamics. More specifically, we suggest that ecological drift plays a substantial role in the spatial dynamics of small fish communities, decoupling species–environment relationships. Such decoupling indicates that further declines in the size of fish communities due to environmental change might add more stochasticity to their dynamics, making them less predictable in the face of future environmental scenarios.

Spatial variation in species composition occurs due to changes in various aspects of community structure such as abundance, evenness or richness (Avolio et al. 2019). We found a negative relationship between changes in species abundance distribution (rank-abundance differences) and median community size. This result

means that species switched positions in the rank-abundance curve (from rarer to more common and vice versa) more often in metacommunities composed of small communities than in those composed of large communities. Spatial variation in species abundance within metacommunities occurs due to a combination of processes, including dispersal, selection, and demographic stochasticity (Vellend 2016). Our results indicate that whereas changes in species abundance in large communities were mostly because of environmental variation (mainly temperature, elevation, and discharge), changes in species abundance in small communities were mostly caused by stochastic processes. In this sense, our work can be considered an extension of previous work that identified a positive relationship between population size and population stability (Reed and Hobbs 2004) and ecosystem size (and complexity) and metapopulation stability (Terui et al. 2018). Although there have been suggestions of a relationship between spatial and temporal variation in biodiversity properties (e.g. spatial–temporal beta diversity, Steiner and Leibold 2004), the degree to which such correlations are causal is still to be resolved (Stegen et al. 2013).

Random and unpredicted changes in species composition can happen entirely or partially due to demographic events that are independent of environmental selection (Vellend and Agrawal 2010, Shoemaker 2020). For example, smaller communities are more likely to be affected by births and deaths, changing species relative abundance independently of species fitness (Vellend and Agrawal 2010). Previous studies based on mechanistic simulation models suggested that when random demographic events play a major role in community dynamics, species competitive differences become less important to shape community composition and inferior competitors become increasingly represented in local communities (Orrock and Fletcher 2005, Orrock and Watling 2010). Siqueira et al. (2020) also found that small insect communities exhibit higher variation in species composition and that this variation was similar to a null expectation. As community size increases, the role of demographic events that are random regarding species identity should decrease relative to deterministic assembly. In this situation, metacommunity dynamics may be mainly governed by species sorting and patch dynamics (Orrock and Watling 2010) leading to deterministic exclusion of the inferior competitors and lowering β -diversity in large metacommunities.

The positive relationship between the strength of community–environment coupling and community size indicates that environmental selection plays a minor role in the dynamics of small communities. This relationship further suggests that environmental factors might have limited power to predict species identity or species abundances in small communities. Evidence of weak relationships between species composition and environmental factors is abundant in community ecology. For example, a comprehensive analysis of stream invertebrate communities across the continental United States showed that β -diversity was weakly related to environmental distances in many of the ecoregions studied (Bini et al. 2014). Heino et al. (2015) also assessed the role of environmental factors in driving β -diversity in stream insect communities in different regions of the world and found that environmental variables were poor predictors of species composition, explaining on average 13% of the spatial variation in community composition. Although we did not re-analyze data from previous studies reporting weak community–environment relationships, our results suggest that, besides the already known causes (e.g. ‘snapshot’ sampling, lack of statistical power, unaccounted legacy effects), ecological drift acting more strongly upon small communities might be an additional cause for such weak relationships.

As it is usually the case in observational, correlative studies, ours also have limitations. Stream communities can be structured by many environmental factors that vary at the local scale, such as flow, nutrient concentration, turbidity, pH (Allan et al. 2021). The set of variables we used is not necessarily the best one to represent community–environment relationships, as these variables were not measured at the same spatial scale used to sample communities. But, even with such data, we detected a positive relationship between the strength of the community–environment relationship and community size. We thus speculate that the inclusion of additional local environmental variables could affect the observed relationship in two ways: 1) the model explanatory power will increase similarly for all metacommunities as more variables are added to it. In this case, the relationship would remain positive, with a similar slope but higher intercept. 2) The role of stochasticity might be so strong in small communities that the strength of the community–environment relationship would not increase for them, but only for the large communities. In this case, the slope of the regression line would change to a more positive relationship, while the intercept would remain the same. Despite these

limitations, some of our models explained more than 30% (and in some cases up to 60%) of the variance in species composition. This result indicates that large scale variables such as average annual air temperature, mean annual natural discharge, and average human footprint index can be used to explain variation in fish occurrence (Chen and Olden 2020, Comte et al. 2021), especially for species in large communities. Another issue to bear in mind is that by relying solely on the median, we might have missed how other aspects of community size influence β -diversity and community–environment relationships. For example, other aspects of the distribution of population sizes within communities (e.g. range; SD) could represent different processes that interact with demographic stochasticity – an issue that deserves further investigation.

Relying on a robust data set that spans 541 communities, across 32 river basins in four biogeographic realms, and robust methods to quantify spatial variation in species composition, we provide evidence for a general pattern in nature – species composition varies more among small communities than among large communities. This pattern is likely driven by the stronger influence of demographic stochastic events on small than on larger communities, as the dynamics of small communities are weakly linked to spatial variation in environmental factors. On the other hand, metacommunities composed of large communities seem to be assembled more deterministically as their species composition were more related to environmental variation. As the effects of demographic stochasticity should manifest also, maybe even more strongly, on community temporal dynamics, future studies should analyze the relationship between temporal β -diversity and community size. We also suggest the prevalence of the relationships we describe here should be tested with data from terrestrial metacommunities, where community size might be more directly linked to ecosystem size. In a changing world where ecosystems have been fragmented and reduced, high variation in species composition among small communities that is decoupled from environmental variation points to further challenges to ecosystem management.

Acknowledgements – CMJ is grateful to Luciano B. Lima for the support in spatial analysis.

Funding – CMJ thanks the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior-Brasil (CAPES) – Finance Code 001 for funding. TS was supported by grant no. 21/00619-7, São Paulo Research Foundation (FAPESP) and by grant no. 309496/2021-7, Brazilian Natl Council for Scientific and Technological Development (CNPq).

2.5 Data availability statement

Data are available from the Zenodo Digital Repository: <https://doi.org/10.5281/zenodo.7703264> (Jacobi and Siqueira 2023).

2.6 References

- Allan, J. D., Castillo, M. M. and Capps, K. A. 2021. Stream ecology: structure and function of running waters. – Springer Nature.
- Appelhans, T., Detsch, F., Reudenbach, C. and Woullauer, S. 2021. mapview: interactive viewing of spatial data in R. – R package ver. 2.10.0. <https://CRAN.R-project.org/package=mapview>.
- Avolio, M. L., Carroll, I. T., Collins, S. L., Houseman, G. R., Hallett, L. M., Isbell, F., Koerner, S. E., Komatsu, K. J., Smith, M. D. and Wilcox, K. R. 2019. A comprehensive approach to analyzing community dynamics using rank abundance curves. – *Ecosphere* 10: e02881.
- Bini, L. M., Landeiro, V. L., Padial, A. A., Siqueira, T. and Heino, J. 2014. Nutrient enrichment is related to two facets of beta diversity for stream invertebrates across the United States. – *Ecology* 95: 1569–1578.
- Bivand, R., Pebesma, E. and Gomez-Rubio. 2013. Applied spatial data analysis with R, 2nd edn. – Springer.
- Bivand, R., Keitt, T. and Rowlingson, B. 2021. rgdal: bindings for the ‘Geospatial’ data abstraction library. – R package ver. 1.5-23. <https://CRAN.R-project.org/package=rgdal>.
- Chase, J. M. and Knight, T. M. 2013. Scale-dependent effect sizes of ecological drivers on biodiversity: why standardised sampling is not enough. – *Ecol. Lett.* 16: 17–26.
- Chen, K. and Olden, J. D. 2020. Threshold responses of riverine fish communities to land use conversion across regions of the world. – *Global Change Biol.* 26: 4952–4965.

- Comte, L. et al. 2020. RivFishTIME: a global database of fish time-series to study global change ecology in riverine systems. – *Global Ecol. Biogeogr.* 30: 38–50.
- Comte, L., Olden, J. D., Tedesco, P. A., Ruhi, A. and Giam, X. 2021. Climate and land-use changes interact to drive long-term reorganization of riverine fish communities globally. – *Proc. Natl Acad. Sci. USA* 118: e2011639118.
- Coyle, J. R., Hurlbert, A. H. and White, E. P. 2013. Opposing mechanisms drive richness patterns of core and transient bird species. – *Am Nat.* 181: E83–E90.
- Döll, P., Kaspar, F. and Lehner, B. 2003. A global hydrological model for deriving water availability indicators: model tuning and validation. – *J. Hydrol.* 270: 105–134.
- Engel, T., Blowes, S. A., McGlinn, D. J., May, F., Gotelli, N. J., McGill, B. J. and Chase, J. M. 2021. Using coverage-based rarefaction to infer non-random species distributions. – *Ecosphere* 12: e03745.
- Gelman, A. and Rubin, D. B. 1992. Inference from iterative simulation using multiple sequences. – *Stat. Sci.* 7: 457–472.
- Gilbert, N. and Levine, J. M. 2017. Ecological drift and the distribution of species diversity. – *Proc. R. Soc. B* 284: 20170507.
- Grill, G. et al. 2019. Mapping the world's free-flowing rivers. – *Nature* 569: 215–221.
- Gwinn, D. C., Allen, M. S., Bonvehio, K. I., Hoyer, M. V. and Beesley, L. S. 2016. Evaluating estimators of species richness: the importance of considering statistical error rates. – *Methods Ecol. Evol.* 7: 294–302.
- Hallett, L., Avolio, M. L., Carroll, I. T., Jones, S. K., MacDonald, A. A. M., Flynn, D. F. B., Slaughter, P., Ripplinger, J., Collins, S. L., Gries, C. and Jones, M. B. 2020. codyn: community dynamics metrics. – R package ver. 2.0.5. <https://github.com/NCEAS/codyn>
- Heino, J. et al. 2015. A comparative analysis reveals weak relationships between ecological factors and beta diversity of stream insect metacommunities at two spatial levels. – *Ecol. Evol.* 5: 1235–1248.
- Hijmans, R. J. 2021a. *_geosphere: spherical trigonometry_*. – R package ver. 1.5-14. <https://CRAN.R-project.org/package=geosphere>.
- Hijmans, R. J. 2021b. *raster: geographic data analysis and modeling*. – R package ver. 3.4-13, <https://CRAN.R-project.org/package=raster>.
- Hijmans, R. J., Cameron, S. E., Parra, J. L., Jones, P. G. and Jarvis, A. 2005. Very high resolution interpolated climate surfaces for global land areas. – *Int. J. Climatol.* 25: 1965–1978.

- Jacobi, C. M. and Siqueira, T. 2023. Data from: High compositional dissimilarity among small communities is decoupled from environmental variation. – Zenodo Digital Repository, <https://doi.org/10.5281/zenodo.7703264>.
- Lehner, B. and Grill, G. 2013. Global river hydrography and network routing: baseline data and new approaches to study the world's large river systems. – *Hydrol. Processes* 27: 2171–2186.
- Linke, S., Lehner, B., Dallaire, C. O., Ariwi, J., Gril, G., Anand, M., Beames, P., Burchard-Levine, V., Maxwell, S., Moidu, H., Tan, F. and Thieme, M. 2019. Global hydro-environmental sub-basin and river reach characteristics at high spatial resolution. – *Sci. Data* 6: 283.
- Magurran, A. and Henderson, P. 2003. Explaining the excess of rare species in natural species abundance distributions. – *Nature* 422: 714–716.
- Magurran, A. E. 2007. Species abundance distributions over time. – *Ecol.Lett.* 10: 347–354.
- May, F. 2021. mobsim: spatial simulation and scale-dependent analysis of biodiversity changes. – R package ver. 0.2.0. <https://github.com/MoBiodiv/mobsim>.
- Melbourne, B. A. 2012. – In: Hastings, A. and Berkeley, L. G. (eds), *Stochasticity, demographic. Encyclopedia of theoretical ecology*. Univ. of California Press, 706–712.
- Oksanen, J. et al. 2020. vegan: community ecology package. – R package ver. 2: 5-7. <https://CRAN.R-project.org/package=vegan>.
- Orrock, J. L. and Fletcher, R. J. Jr. 2005. Changes in community size affect the outcome of competition. – *Am Nat.* 166: 107–111.
- Orrock, J. L. and Watling, J. I. 2010. Local community size mediates ecological drift and competition in metacommunities. – *Proc. R. Soc. B* 277: 2185–2191.
- Otto, S. P. and Whitlock, M. C. 1997. Fixation of beneficial mutations in a population of changing size. – *Genetics* 146: 723–733.
- Ovaskainen, O., Tikhonov, G., Norberg, A., Blanchet, F. G., Duan, L., Dunson, D., Roslin, T. and Abrego, N. 2017 How to make more out of community data? A conceptual framework and its implementation as models and software. – *Ecol. Lett.* 20: 561–576.
- Pebesma, E. 2018. Simple features for R: standardized support for spatial vector data. – *R J.* 10: 439–446.
- Pebesma, E. J. and Bivand, R. S. 2005. Classes and methods for spatial data in R. – *R News* 5: 9–13.

- Reed, D. H. and Hobbs, G. R. 2004. The relationship between population size and temporal variability in population size. – *Anim. Conserv. Forum.* 7: 1–8.
- Robinson, N., Regetz, J. and Guralnick, R. P. 2014. EarthEnv-DEM90: a nearly-global, void-free, multi-scale smoothed, 90 m digital elevation model from fused ASTER and SRTM data. – *ISPRS J. Photogramm. Remote. Sens.* 87: 57–67.
- Shoemaker, L. G. et al. 2020. Integrating the underlying structure of stochasticity into community ecology. – *Ecology* 101: e02922.
- Siqueira, T., Saito, V. S., Bini, L. M., Melo, A. S., Petsch, D. K., Landeiro, V. L., Tolonen, K. T., Jyrkänkallio-Mikkola, J., Soininen and Heino, J. 2020. Community size can affect the signals of ecological drift and niche selection on biodiversity. – *Ecology* 101: e03014.
- Stegen, J. C., Freestone, A. L., Crist, T. O., Anderson, M. J., Chase, J. M., Comita, L. S., Cornell, H. V., Davies, K. F., Harrison, S. P., Hurlbert, A. H., Inouye, B. D., Kraft, N. J. B., Myers, J. A., Sanders, N. J., Swenson, N. G. and Vellend, M. 2013. Drivers of spatial and temporal turnover. – *Global Ecol. Biogeogr.* 22: 202–212.
- Steiner, C. F. and Leibold, M. A. 2004. Cyclic assembly trajectories and scale-dependent productivity–diversity relationships. – *Ecology* 85: 107–113.
- Terui, A., Ishiyama, N., Urabe, H., Ono, S., Finlay, J. O. and Nakamura, F. 2018. Metapopulation stability in branching river networks. – *Proc. Natl Acad. Sci. USA* 115: E5963–E5969.
- Tikhonov, G., Opedal, Ø. H., Abrego, L., Lehikoinen, A., de Jonge, M. M. J., Oksanen, J. and Ovaskainen, O. 2020. Joint species distribution modelling with the r-package Hmsc. – *Methods Ecol. Evol.* 11: 442–447.
- Tikhonov, G., Ovaskainen, O., Oksanen, J., de Jonge, M., Opedal, O. and Dallas, T. 2021. hmsc: hierarchical model of species communities. – R package ver. 3.0-11. <https://CRAN.R-project.org/package=Hmsc>.
- Tjur, T. 2009. Coefficients of determination in logistic regression models—a new proposal: the coefficient of discrimination. – *Am. Stat.* 63: 366–372.
- Vellend, M. 2016. The theory of ecological communities. Mpb series, edn 57. – Princeton Univ. Press.
- Vellend, M. and Agrawal, A. 2010. Conceptual synthesis in community ecology. – *Q. Rev. Biol.* 85: 183–206.
- Venter, O., Sanderson, E. W., Magrath, A., Allan, J. R., Beher, J., Jones, K. R., Possingham, H. P., Laurance, W. F., Wood, P., Fekete, B. M., Levy, M. A. and Watsom, J. E. M. 2016. Global terrestrial human footprint maps for 1993 and 2009. – *Sci. Data* 3: 160067.

Wickham, H. et al. 2023. dplyr: a grammar of data manipulation. –
<https://dplyr.tidyverse.org>, <https://github.com/tidyverse/dplyr>.

2.7 Supporting Information

2.7.1 Material and methods

Table S1 - General features of each dataset used in the empirical analyses at both regional and local scales.

Metacommunity ID	Regional			Local			
	# sites	# individuals	γ -diversity	α -diversity (mean)	α -diversity (median)	# individuals (mean)	# individuals (median)
2070021230	10	4538	12	7	6	454	471.5
2070021360	15	5559	15	7	7	371	393
2070469050	12	5398	19	5	2	450	238.5
2070470930	10	15956	21	8	8	1596	293
2070513970	29	4112	32	7	7	142	101
5070070110	29	4340	31	9	8	150	113
5070070390	12	3636	23	10	10	303	336
5070070440	21	4307	30	9	8	205	194
5070070450	13	3990	26	10	11	307	324
5070468810	10	2087	14	5	4	209	91.5
6070015200	22	2411	17	7	8	110	114
6070709820	16	4374	27	9	8	273	113
6070713620	22	3161	40	10	9	144	95.5
7070042610	12	2633	44	18	18	219	201
7070317870	11	3323	41	15	14	302	219
7070327640	12	756	6	2	2	63	65
7070360890	31	2460	7	2	2	79	74
7070415980	10	2971	33	10	8	297	117.5
7070450470	13	1220	21	7	7	94	81
7070451570	18	2882	27	9	9	160	147
7070453370	32	3967	31	7	6	124	90.5
7070620210	21	8395	47	20	20	400	306
7070630060	11	5356	34	16	16	487	425
7070642870	21	8429	36	14	15	401	352
7070643030	11	2704	34	15	15	246	228
7070647250	18	5814	46	16	17	323	262
7070647350	14	3651	42	16	15	261	231
7070648760	10	3928	44	18	18	393	424.5
7070651920	12	1978	41	18	19	165	156.5
7070656320	22	8021	43	16	16	365	341
7070663220	23	10421	29	15	15	453	385
7070684690	18	4501	48	17	16	250	169

2.7.2 Process-based simulation models

Spatial variation in species composition

As our first goal was to investigate how spatial variation in species composition within metacommunities varied along a gradient of median community size, we needed to make sure that our metrics representing spatial variation in species composition would not be affected by sampling effort and sample size. For example, larger communities have a higher probability to share more species and thus exhibit a lower variation in species composition than smaller communities because they have more individuals, but not necessarily because they are more similar.

So, using the package *mobsim* (May et al 2018), we simulated metacommunities with the same SAD (log-normal distribution), species richness and number of communities. They were assembled without any environmental filter and differed only in their overall size (taken from a uniform distribution that varied between 756 and 15956 individuals, according to the range of individuals sampled in the empirical metacommunities; Table S1). We represented the intraspecific aggregation caused by dispersal limitation considering that all communities were equally likely to be colonized by the initial member of each species, but once they were, subsequent members of that species were more likely to colonize surrounding locations. To investigate the potential effect of metacommunity characteristics in the tested metrics, we varied the number of species and sites in the simulations. The values of these parameters covered the range of variation of the empirical metacommunities (6-48 species and 10-32 sites, Table S1).

We then measured six β -diversity metrics within each simulated metacommunity and regressed β -diversity values against the median community size of the simulated metacommunities. Because we hypothesized that environmental selection should be stronger in larger communities and ecological drift should be stronger in smaller communities, we expected that an unbiased β -diversity metric would have no relationship with community size in the simulation, as environmental selection played no role in community assembly.

Our first candidate β -diversity metric was Baselga's (2010) multiple-site Total Beta Diversity based on presence-absence data. This metric accounts both for the spatial turnover and the nestedness components of β -diversity. It was carried out

in R with the function `beta.multi` (Baselga et al. 2022), using Jaccard dissimilarity. The second candidate β -diversity metric was C-score (Stone & Roberts 1990), which provides an average measure of co-occurrence for all species pairs. Higher C-score values indicate fewer co-occurrence between species in the matrix. The C-score was measured with the package `bipartite` (Dormann et al. 2008) in R. Next, we used two β -diversity metrics that could be corrected for under-sampling: the Jaccard-Chao and the Shannon. The correction applied to these metrics uses accumulation curves and should be more effective than the original metrics to remove γ -dependence (Cao et al 2021). The fifth candidate metric was a recently developed metric called β_c (Engel et al 2021), which was developed to represent differences in species composition even when there are differences in the size of the species pool and associated sampling effects.

Finally, as changes in species relative abundances are a good predictor of community change (Dornelas et al. 2014, Jones et al. 2017, Avolio et al. 2019), we measured how the rank-abundance of communities varied among each other within each metacommunity, and regressed their median against median community size.

Community-environment relationship

To investigate our second prediction, we had to be sure that a measure of the strength of the community-environment relationship would not itself be a statistical consequence of community size. For example, it is possible that a strong community-environment relationship emerges in large communities simply because it might be easier to model more individuals than a matrix full of zeros, irrespective of the underlying assembly process. Thus, we used a second process-based simulation in which metacommunities of different sizes were assembled by the same amount of demographic stochasticity and environmental selection. Now, we simulated metacommunities composed of species that followed a Gaussian response to a single environmental variable. They had the same niche breadth (0.2) but differed in their environmental optima (i.e., niche position; taken from a uniform distribution ranging from 0 to 5). We simulated ecological drift by drawing the abundance of every species in each local community from a Poisson distribution which mean was given by the deterministic influence of environmental selection. We assigned different sizes to the metacommunities estimating species expected abundance in optimum environments from a random interval.

We estimated the strength of the community-environment relationship, but only considering the incidence of the 10% most abundant species of each community. We tested two modeling approaches to estimate the strength of the community-environment relationship. First, we used a generalized linear model (GLM) and pseudo R^2 values as a measure of the strength of the community-environment relationship. We then used Hierarchical Modeling of Species Communities (HMSC) (Ovaskainen 2017) and Tjur R^2 values as a measure of the strength of the community-environment relationship. An appropriate measure of community-environment relationship would show no relationship with community size in simulated metacommunities under equal contribution of drift and selection.

2.7.3 Results

Spatial variation in species composition

We found that the first five (except rank-abundance difference) β -diversity metrics were related (varying from negative to positive) to median community size (Table S2). These results indicate that these five β -diversity metrics are affected by median community size independently of the underlying assembly process.

We observed no relationship between rank-abundance difference and community size for the range of parameter values in our simulations (Table S2). So, as the lack of relationship was consistent, we concluded that rank difference was the most appropriate β -diversity metric that we could use in our study to investigate the relationship between variation in species composition and community size.

Table S2 - Model statistics of the relationships between β -diversity metrics (Jaccard, C-score, Jaccard-Chao, Shannon, β_c and rank difference) and community size for 30 simulated metacommunities covering the range of variation of species and sites from the empirical metacommunities. SD coeff. = standardized regression coefficient; R^2 and Adj. R^2 = coefficient of determination and adjusted coefficient of determination, respectively. Some metrics were not calculated when the number of species and sites were small (empty lines).

β -diversity metric	SD coeff.	F-statistic	p-value	R^2
Metacommunities simulated with 6 species and sites				
Jaccard - Baselga				
C-score				
Jaccard-Chao				
Shannon	-0.316	3.114	0.088	0.100
β_c				
Rank difference	-0.152	0.664	0.422	0.023
Metacommunities simulated with 10 species and sites				
Jaccard - Baselga	-0.749	35.7	<0.001	0.560
C-score				
Jaccard-Chao	-0.617	17.17	<0.001	0.380
Shannon	-0.176	0.894	0.352	0.031
β_c	-0.583	11.85	0.002	0.34
Rank difference	0.073	0.150	0.702	0.005
Metacommunities simulated with 20 species and sites				
Jaccard - Baselga	-0.857	77.17	<0.001	0.734
C-score	0.350	3.903	0.058	0.122
Jaccard-Chao	-0.706	27.88	<0.001	0.499
Shannon	-0.346	3.82	0.061	0.120
β_c	-0.642	19.68	<0.001	0.413
Rank difference	0.221	1.434	0.241	0.049
Metacommunities simulated with 30 species and sites				
Jaccard - Baselga	-0.927	171.3	<0.001	0.859
C-score	-0.819	56.95	<0.001	0.670
Jaccard-Chao	-0.919	152.1	<0.001	0.844
Shannon	-0.737	33.21	<0.001	0.543
β_c	-0.226	1.513	0.229	0.052
Rank difference	0.252	1.894	0.180	0.063
Metacommunities simulated with 40 species and sites				
Jaccard - Baselga	-0.954	282.3	<0.001	0.910
C-score	-0.895	112.2	<0.001	0.800
Jaccard-Chao	-0.938	204.4	<0.001	0.879
Shannon	-0.674	23.26	<0.001	0.454

β -diversity metric	SD coeff.	F-statistic	p-value	R ²
Bc	0.513	9.996	0.004	0.263
Rank difference	0.144	0.595	0.447	0.021
Metacommunities simulated with 48 species and sites				
Jaccard - Baselga	-0.929	175.7	<0.001	0.863
C-score	-0.870	87.57	<0.001	0.758
Jaccard-Chao	-0.930	179.2	<0.001	0.865
Shannon	-0.783	44.38	<0.001	0.613
Bc	0.569	11.47	0.002	0.323
Rank difference	-0.069	0.136	0.715	0.005

Community-environment relationships

We found that the GLM pseudo-R² increased with community size (SD. coeff. = 0.678, R² = 0.459, p = <0.001, Fig. S1a), whereas there was no evidence of a relationship between HMSC Tjur R² and community size (SD. coeff. = 0.247, R² = 0.061, p = 0.188, Fig. S1b). So, we decided to use the HMSC to test our second prediction.

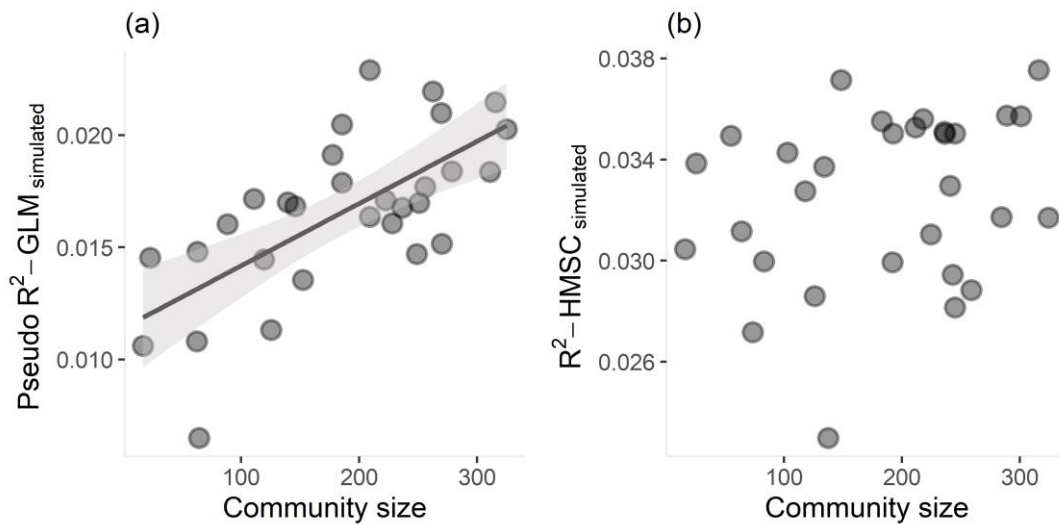


Figure S1. Relationship of the strength of the community-environment relationships within simulated metacommunities measured with a Generalized Linear Model (GLM) (a) and with Hierarchical Modelling of Species Communities (HMSC) (b) against median community size. Gray bands represent the 95% confidence interval around the regression line.

Model fit assumptions

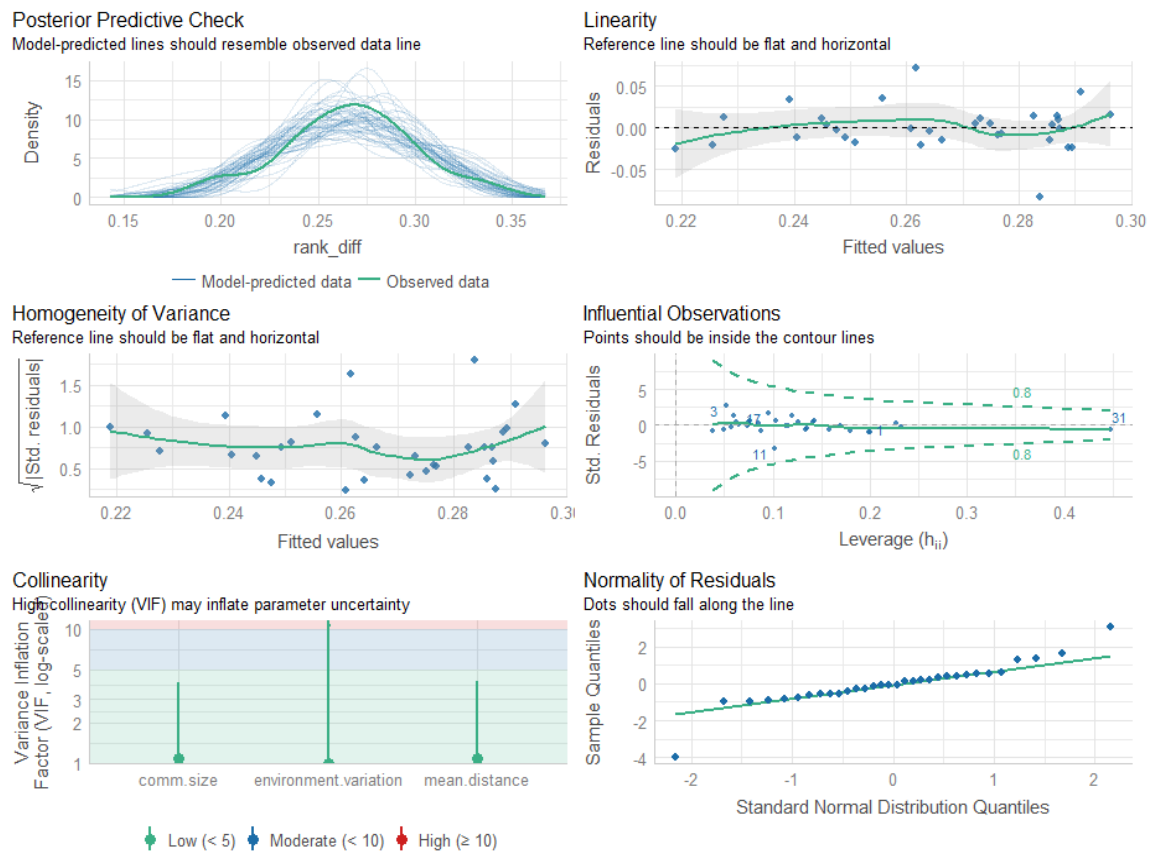


Figure S2. Model fit assumptions of the multiple linear model used to describe rank difference values as a function of community size, environmental heterogeneity and spatial extent of the metacommunities. All plots were done using the performance package (Lüdtke et al. 2021) in R.

Relative importance of each environmental variable to explain community composition

Table S3. Partitioning species' variance explained by the environmental variables across all metacommunities, considering the 10% most abundant species of each community. Environmental variables: discharge (Dis), elevation (Ele), temperature (Tmp), precipitation (Pre), human footprint (hft), connectivity status index (csi) and a measure that indicates if a river reach belongs to a river with free-flowing, good connectivity or degraded status (csiff2).

Species	Dis	Ele	Tmp	Pre	hft	csi	csiff2
Metacommunity ID 2070021230							
<i>Cottus gobio</i>	0.151	0.206	0.220	0.113	0.223	0.086	0.000
<i>Salmo trutta</i>	0.119	0.164	0.194	0.130	0.274	0.119	0.000
<i>Phoxinus phoxinus</i>	0.168	0.215	0.276	0.104	0.156	0.081	0.000
Metacommunity ID 2070021360							
<i>Salmo trutta</i>	0.335	0.106	0.162	0.069	0.056	0.235	0.037
<i>Cottus gobio</i>	0.209	0.134	0.159	0.061	0.134	0.239	0.065
<i>Phoxinus phoxinus</i>	0.183	0.195	0.185	0.090	0.111	0.178	0.058
<i>Barbatula barbatula</i>	0.146	0.140	0.307	0.144	0.072	0.141	0.049
<i>Perca fluviatilis</i>	0.155	0.142	0.191	0.146	0.103	0.203	0.061
Metacommunity ID 2070469050							
<i>Salmo trutta</i>	0.171	0.182	0.193	0.168	0.141	0.112	0.033
<i>Salmo salar</i>	0.119	0.194	0.193	0.142	0.130	0.188	0.035
<i>Cottus gobio</i>	0.171	0.204	0.165	0.173	0.138	0.109	0.040
<i>Barbatula barbatula</i>	0.110	0.197	0.206	0.152	0.112	0.183	0.038
<i>Squalius cephalus</i>	0.168	0.186	0.173	0.164	0.141	0.103	0.064
Metacommunity ID 2070470930							
<i>Cottus gobio</i>	0.291	0.145	0.183	0.123	0.132	0.081	0.044
<i>Phoxinus phoxinus</i>	0.327	0.118	0.168	0.151	0.122	0.077	0.036
<i>Barbatula barbatula</i>	0.243	0.148	0.154	0.176	0.148	0.090	0.041
<i>Perca fluviatilis</i>	0.200	0.164	0.164	0.156	0.136	0.129	0.051
<i>Squalius cephalus</i>	0.187	0.161	0.176	0.161	0.123	0.146	0.047
Metacommunity ID 2070513970							
<i>Barbatula barbatula</i>	0.154	0.277	0.286	0.134	0.075	0.074	0.000
<i>Gobio gobio</i>	0.105	0.239	0.245	0.127	0.118	0.166	0.000
<i>Phoxinus phoxinus</i>	0.137	0.281	0.281	0.129	0.080	0.092	0.000
<i>Rutilus rutilus</i>	0.054	0.247	0.215	0.157	0.239	0.087	0.000
<i>Squalius cephalus</i>	0.118	0.236	0.247	0.117	0.109	0.173	0.000
<i>Rhodeus sericeus</i>	0.085	0.210	0.317	0.111	0.195	0.081	0.000
<i>Alburnus alburnus</i>	0.166	0.200	0.200	0.155	0.109	0.170	0.000
<i>Abramis brama</i>	0.105	0.221	0.207	0.177	0.111	0.178	0.000

Species	Dis	Ele	Tmp	Pre	hft	csi	csiff2
<i>Rineloricaria aequalicuspis</i>	0.087	0.265	0.271	0.228	0.084	0.065	0.000
<i>Cyanocharax itaimbe</i>	0.096	0.244	0.226	0.245	0.089	0.100	0.000
<i>Pareiorhaphis hypselurus</i>	0.084	0.274	0.274	0.229	0.073	0.066	0.000
<i>Pareiorhaphis nudulus</i>	0.176	0.249	0.266	0.206	0.056	0.047	0.000
<i>Epactionotus bilineatus</i>	0.099	0.219	0.232	0.212	0.139	0.099	0.000
Metacommunity ID 6070709820							
<i>Poecilia reticulata</i>	0.467	0.163	0.072	0.103	0.118	0.078	0.000
<i>Knodus moenkhausii</i>	0.238	0.198	0.050	0.097	0.129	0.288	0.000
<i>Hypostomus ancistroides</i>	0.333	0.154	0.095	0.184	0.138	0.096	0.000
<i>Astyanax lacustris</i>	0.373	0.149	0.124	0.105	0.175	0.075	0.000
<i>Aspidoras fuscoguttatus</i>	0.275	0.156	0.094	0.154	0.137	0.185	0.000
<i>Gymnotus sylvius</i>	0.282	0.170	0.107	0.113	0.118	0.211	0.000
<i>Moenkhausia sanctaefilomenae</i>	0.295	0.183	0.101	0.104	0.122	0.196	0.000
<i>Serrapinnus notomelas</i>	0.260	0.256	0.096	0.116	0.156	0.116	0.000
Metacommunity ID 6070713620							
<i>Piabina argentea</i>	0.134	0.151	0.156	0.162	0.221	0.176	0.000
<i>Hypostomus ancistroides</i>	0.147	0.160	0.187	0.188	0.195	0.123	0.000
<i>Knodus moenkhausii</i>	0.265	0.164	0.233	0.192	0.072	0.074	0.000
<i>Corydoras aeneus</i>	0.162	0.153	0.222	0.154	0.131	0.177	0.000
<i>Astyanax lacustris</i>	0.099	0.163	0.249	0.202	0.192	0.095	0.000
<i>Poecilia reticulata</i>	0.092	0.142	0.129	0.150	0.288	0.199	0.000
<i>Oligosarcus pinto</i>	0.196	0.163	0.155	0.150	0.144	0.192	0.000
<i>Gymnotus sylvius</i>	0.130	0.156	0.238	0.158	0.130	0.187	0.000
<i>Serrapinnus notomelas</i>	0.125	0.255	0.141	0.188	0.128	0.163	0.000
<i>Hemigrammus marginatus</i>	0.158	0.159	0.162	0.185	0.126	0.210	0.000
<i>Imparfinis schubarti</i>	0.125	0.135	0.110	0.184	0.214	0.233	0.000
Metacommunity ID 7070042610							
<i>Luxilus albeolus</i>	0.093	0.134	0.208	0.342	0.089	0.134	0.000
<i>Etheostoma nigrum</i>	0.121	0.172	0.201	0.250	0.136	0.120	0.000
<i>Nocomis leptcephalus</i>	0.096	0.272	0.190	0.253	0.089	0.099	0.000
<i>Notropis altipinnis</i>	0.111	0.143	0.201	0.271	0.110	0.163	0.000
<i>Lepomis auritus</i>	0.087	0.173	0.192	0.223	0.135	0.190	0.000
<i>Etheostoma olmsted</i>	0.214	0.142	0.180	0.173	0.120	0.173	0.000
<i>Anguilla rostrata</i>	0.216	0.160	0.194	0.161	0.104	0.166	0.000
<i>Lythrurus matutinus</i>	0.165	0.202	0.178	0.220	0.089	0.145	0.000
<i>Noturus insignis</i>	0.124	0.155	0.206	0.236	0.123	0.156	0.000
<i>Notropis procne</i>	0.094	0.149	0.192	0.295	0.088	0.182	0.000
<i>Cyprinella analostana</i>	0.103	0.166	0.193	0.254	0.104	0.180	0.000
<i>Enneacanthus gloriosus</i>	0.114	0.153	0.248	0.155	0.125	0.204	0.000
Metacommunity ID 7070317870							
<i>Percina maculata</i>	0.204	0.065	0.080	0.147	0.338	0.113	0.052
<i>Notropis dorsalis</i>	0.191	0.093	0.130	0.176	0.178	0.118	0.115
<i>Etheostoma nigrum</i>	0.219	0.110	0.097	0.133	0.125	0.197	0.120
<i>Margariscus margarita</i>	0.205	0.068	0.143	0.150	0.144	0.209	0.082
<i>Luxilus cornutus</i>	0.219	0.077	0.071	0.197	0.156	0.215	0.065

Species	Dis	Ele	Tmp	Pre	hft	csi	csiff2
<i>Notropis volucellus</i>	0.204	0.092	0.115	0.171	0.130	0.168	0.120
<i>Notemigonus crysoleucas</i>	0.199	0.075	0.111	0.125	0.152	0.220	0.116
<i>Chrosomus eos</i>	0.205	0.104	0.075	0.145	0.173	0.215	0.083
Metacommunity ID 7070327640							
<i>Oncorhynchus clarkii</i>	0.120	0.274	0.232	0.092	0.204	0.079	0.000
<i>Salvelinus confluentus</i>	0.133	0.217	0.171	0.102	0.287	0.091	0.000
<i>Salmo trutta</i>	0.171	0.244	0.186	0.107	0.182	0.111	0.000
<i>Salvelinus fontinalis</i>	0.238	0.229	0.174	0.113	0.125	0.121	0.000
Metacommunity ID 7070360890							
<i>Prosopium williamsoni</i>	0.106	0.234	0.282	0.177	0.086	0.115	0.000
<i>Oncorhynchus clarkii</i>	0.109	0.175	0.194	0.076	0.229	0.217	0.000
<i>Salmo trutta</i>	0.058	0.188	0.201	0.156	0.240	0.157	0.000
<i>Salvelinus confluentus</i>	0.125	0.203	0.186	0.154	0.168	0.164	0.000
<i>Salvelinus fontinalis</i>	0.144	0.171	0.247	0.077	0.144	0.217	0.000
<i>Oncorhynchus mykiss</i>	0.098	0.255	0.314	0.135	0.101	0.098	0.000
Metacommunity ID 7070415980							
<i>Rhinichthys atratulus</i>	0.176	0.108	0.104	0.112	0.234	0.266	0.000
<i>Campostoma anomalum</i>	0.102	0.126	0.153	0.140	0.202	0.276	0.000
<i>Notropis stramineus</i>	0.105	0.127	0.148	0.141	0.207	0.271	0.000
<i>Pimephales promelas</i>	0.103	0.143	0.128	0.216	0.164	0.246	0.000
<i>Notropis dorsalis</i>	0.157	0.157	0.099	0.252	0.136	0.199	0.000
<i>Semotilus atromaculatus</i>	0.162	0.140	0.096	0.142	0.227	0.234	0.000
<i>Hybognathus hankinsoni</i>	0.213	0.140	0.102	0.171	0.152	0.222	0.000
<i>Catostomus commersonii</i>	0.209	0.126	0.114	0.179	0.154	0.218	0.000
<i>Culaea inconstans</i>	0.142	0.184	0.111	0.153	0.152	0.258	0.000
<i>Umbra limi</i>	0.172	0.140	0.109	0.173	0.171	0.235	0.000
<i>Lepomis cyanellus</i>	0.170	0.128	0.123	0.180	0.163	0.236	0.000
Metacommunity ID 7070450470							
<i>Etheostoma olmstedii</i>	0.084	0.198	0.226	0.208	0.145	0.138	0.000
<i>Etheostoma nigrum</i>	0.089	0.197	0.212	0.219	0.146	0.137	0.000
<i>Semotilus atromaculatus</i>	0.048	0.206	0.233	0.237	0.139	0.138	0.000
<i>Cottus bairdii</i>	0.105	0.183	0.204	0.210	0.159	0.138	0.000
<i>Rhinichthys cataractae</i>	0.151	0.186	0.167	0.163	0.198	0.135	0.000
<i>Rhinichthys obtusus</i>	0.067	0.175	0.178	0.179	0.228	0.173	0.000
Metacommunity ID 7070451570							
<i>Campostoma anomalum</i>	0.118	0.249	0.200	0.263	0.090	0.080	0.000
<i>Etheostoma caeruleum</i>	0.085	0.244	0.200	0.269	0.100	0.102	0.000
<i>Semotilus atromaculatus</i>	0.130	0.249	0.254	0.226	0.073	0.068	0.000
<i>Catostomus commersonii</i>	0.127	0.220	0.221	0.213	0.130	0.089	0.000
<i>Rhinichthys obtusus</i>	0.168	0.247	0.243	0.194	0.094	0.055	0.000
<i>Rhinichthys cataractae</i>	0.182	0.208	0.208	0.163	0.092	0.148	0.000
Metacommunity ID 7070453370							
<i>Catostomus commersonii</i>	0.092	0.171	0.190	0.227	0.179	0.104	0.038
<i>Rhinichthys cataractae</i>	0.137	0.237	0.243	0.162	0.067	0.132	0.022
<i>Semotilus atromaculatus</i>	0.192	0.179	0.221	0.267	0.064	0.059	0.018

Species	Dis	Ele	Tmp	Pre	hft	csi	csiff2
77070647250							
<i>Nocomis leptocephalus</i>	0.390	0.078	0.060	0.083	0.052	0.303	0.033
<i>Luxilus albeolus</i>	0.304	0.137	0.107	0.114	0.111	0.167	0.059
<i>Lepomis auritus</i>	0.176	0.099	0.225	0.164	0.137	0.148	0.050
<i>Luxilus cerasinus</i>	0.341	0.101	0.088	0.109	0.127	0.182	0.054
<i>Fundulus rathbuni</i>	0.205	0.169	0.119	0.119	0.077	0.253	0.058
<i>Semotilus atromaculatus</i>	0.300	0.171	0.109	0.123	0.099	0.135	0.062
<i>Gambusia holbrooki</i>	0.290	0.122	0.186	0.115	0.088	0.151	0.046
<i>Etheostoma olmstedii</i>	0.363	0.120	0.079	0.175	0.061	0.148	0.054
<i>Notropis procne</i>	0.218	0.118	0.145	0.132	0.179	0.159	0.050
<i>Lepomis macrochirus</i>	0.229	0.153	0.094	0.116	0.177	0.185	0.046
<i>Notropis hudsonius</i>	0.235	0.145	0.134	0.146	0.090	0.192	0.058
<i>Lythrurus ardens</i>	0.229	0.147	0.133	0.143	0.088	0.205	0.055
<i>Notropis altipinnis</i>	0.250	0.129	0.115	0.130	0.143	0.172	0.060
Metacommunity ID 7070647350							
<i>Etheostoma olmstedii</i>	0.071	0.177	0.115	0.170	0.128	0.339	0.000
<i>Erimyzon oblongus</i>	0.094	0.138	0.148	0.111	0.226	0.284	0.000
<i>Nocomis leptocephalus</i>	0.108	0.156	0.176	0.089	0.193	0.277	0.000
<i>Notropis altipinnis</i>	0.063	0.129	0.082	0.249	0.202	0.275	0.000
<i>Lepomis auritus</i>	0.100	0.184	0.136	0.108	0.172	0.300	0.000
<i>Notropis alborus</i>	0.078	0.143	0.220	0.143	0.166	0.251	0.000
<i>Lepomis macrochirus</i>	0.300	0.141	0.185	0.100	0.146	0.129	0.000
<i>Lepomis cyanellus</i>	0.095	0.171	0.132	0.178	0.134	0.290	0.000
<i>Gambusia holbrooki</i>	0.134	0.181	0.171	0.137	0.163	0.213	0.000
Metacommunity ID 7070648760							
<i>Lepomis macrochirus</i>	0.313	0.136	0.149	0.090	0.127	0.131	0.053
<i>Lepomis auritus</i>	0.271	0.107	0.146	0.133	0.166	0.131	0.046
<i>Nocomis leptocephalus</i>	0.278	0.171	0.256	0.067	0.074	0.118	0.036
<i>Noturus insignis</i>	0.273	0.130	0.225	0.077	0.087	0.143	0.066
<i>Luxilus albeolus</i>	0.319	0.153	0.139	0.072	0.117	0.140	0.060
<i>Etheostoma flabellare</i>	0.241	0.130	0.174	0.114	0.116	0.161	0.065
<i>Nocomis raneyi</i>	0.222	0.177	0.143	0.096	0.096	0.190	0.078
<i>Cyprinella analostana</i>	0.217	0.242	0.133	0.128	0.084	0.133	0.063
<i>Notropis procne</i>	0.209	0.250	0.138	0.127	0.081	0.134	0.059
Metacommunity ID 7070651920							
<i>Notropis cummingsae</i>	0.053	0.114	0.294	0.135	0.117	0.287	0.000
<i>Etheostoma olmstedii</i>	0.130	0.168	0.174	0.157	0.126	0.245	0.000
<i>Notemigonus crysoleucas</i>	0.099	0.148	0.202	0.157	0.128	0.266	0.000
<i>Lepomis auritus</i>	0.081	0.138	0.303	0.102	0.108	0.268	0.000
<i>Gambusia holbrooki</i>	0.123	0.138	0.222	0.160	0.248	0.109	0.000
<i>Lepomis macrochirus</i>	0.133	0.121	0.221	0.148	0.140	0.236	0.000
<i>Noturus insignis</i>	0.108	0.183	0.182	0.145	0.128	0.253	0.000
<i>Anguilla rostrata</i>	0.151	0.149	0.166	0.139	0.179	0.216	0.000
<i>Notropis altipinnis</i>	0.096	0.202	0.191	0.145	0.145	0.222	0.000
Metacommunity ID 7070656320							

Species	Dis	Ele	Tmp	Pre	hft	csi	csiff2
<i>Luxilus coccogenis</i>	0.106	0.203	0.279	0.156	0.155	0.101	0.000
<i>Lepomis auritus</i>	0.084	0.224	0.220	0.246	0.149	0.078	0.000
<i>Nocomis leptocephalus</i>	0.049	0.224	0.238	0.330	0.083	0.076	0.000
<i>Semotilus atromaculatus</i>	0.130	0.223	0.183	0.199	0.173	0.093	0.000
<i>Campostoma anomalum</i>	0.323	0.156	0.221	0.159	0.077	0.063	0.000
<i>Cyprinella chloristia</i>	0.117	0.311	0.240	0.165	0.084	0.083	0.000
<i>Fundulus rathbuni</i>	0.104	0.232	0.255	0.235	0.071	0.104	0.000
<i>Etheostoma olmstedii</i>	0.151	0.244	0.197	0.203	0.103	0.102	0.000
<i>Lepomis macrochirus</i>	0.189	0.228	0.211	0.174	0.112	0.086	0.000
<i>Notropis chlorocephalus</i>	0.062	0.183	0.196	0.383	0.119	0.056	0.000
<i>Clinostomus funduloides</i>	0.186	0.206	0.193	0.157	0.113	0.145	0.000
Metacommunity ID 7070663220							
<i>Nocomis leptocephalus</i>	0.236	0.226	0.248	0.085	0.079	0.127	0.000
<i>Etheostoma brevispinum</i>	0.085	0.320	0.241	0.162	0.112	0.079	0.000
<i>Cyprinella chloristia</i>	0.077	0.309	0.198	0.105	0.220	0.091	0.000
<i>Cyprinella pyrrhomelas</i>	0.094	0.197	0.236	0.145	0.265	0.063	0.000
<i>Hybopsis hypsinotus</i>	0.095	0.220	0.202	0.106	0.185	0.191	0.000
<i>Semotilus atromaculatus</i>	0.110	0.313	0.230	0.151	0.114	0.082	0.000
<i>Lepomis auritus</i>	0.112	0.255	0.208	0.239	0.120	0.066	0.000
<i>Notropis scepticus</i>	0.082	0.280	0.222	0.163	0.169	0.083	0.000
Metacommunity ID 7070684690							
<i>Semotilus atromaculatus</i>	0.128	0.196	0.237	0.086	0.102	0.214	0.038
<i>Nocomis leptocephalus</i>	0.067	0.276	0.127	0.042	0.081	0.393	0.014
<i>Notropis chiliticus</i>	0.125	0.323	0.127	0.041	0.034	0.336	0.014
<i>Notropis altipinnis</i>	0.163	0.211	0.157	0.090	0.088	0.253	0.039
<i>Etheostoma olmstedii</i>	0.073	0.235	0.261	0.114	0.059	0.233	0.024
<i>Lepomis auritus</i>	0.186	0.186	0.206	0.048	0.076	0.273	0.025
<i>Perca flavescens</i>	0.119	0.236	0.227	0.085	0.156	0.121	0.056
<i>Noturus insignis</i>	0.123	0.234	0.220	0.085	0.156	0.125	0.057
<i>Lepomis marginatus</i>	0.126	0.199	0.224	0.104	0.055	0.253	0.040
<i>Semotilus lumbee</i>	0.172	0.171	0.195	0.083	0.140	0.201	0.039
<i>Notropis cummingsae</i>	0.158	0.191	0.237	0.071	0.065	0.241	0.038
<i>Lepomis macrochirus</i>	0.105	0.158	0.208	0.110	0.160	0.221	0.036
<i>Aphredoderus sayanus</i>	0.112	0.149	0.201	0.121	0.173	0.207	0.037

2.7.4 Supporting information references

Avolio, M. L., Carroll, I. T., Collins, S. L., Houseman, G. R., Hallett, L. M., Isbell, F., Koerner, S. E., Komatsu, K. J., Smith, M. D. and Wilcox, K. R. 2019. A comprehensive approach to analyzing community dynamics using rank abundance curves. *Ecosphere*, 10: e02881.

- Baselga, A. 2010. Partitioning the turnover and nestedness components of beta diversity. *Glob. Ecol. Biogeogr.*, 19: 134-143.
- Baselga, A., Orme, D., Villeger, S., De Bortoli, J., Leprieur, F. and Logez, M. 2022. *_betapart*: Partitioning Beta Diversity into Turnover and Nestedness Components_. R package version 1.5.6. <<https://CRAN.R-project.org/package=betapart>>.
- Cao, K., Svenning, J. -C., Yan, C., Zhang, J., Mi, X. and Ma, K. 2021. Undersampling correction methods to control γ -dependence for comparing β -diversity between regions. 102: 9.
- Dornelas, M., Gotelli, N. J., McGill, B., Shimadzu, H., Moyes, F., Sievers, C. and Magurran, A. E. 2014. Assemblage time series reveal biodiversity change but not systematic loss. *Science*, 344: 296-9.
- Dormann, C. F., Gruber, B. and Fruend, J. 2008. Introducing the bipartite Package: Analysing Ecological Networks. *R news* Vol 8/2, 8-11.
- Engel, T., Blowes, S. A., McGlinn, D. J., May, F., Gotelli, N. J., McGill, B. J. and Chase, J. 2021. Using coverage-based rarefaction to infer non-random species distributions. *Ecosphere*, 12:e03745.
- Jones, S. K., Ripplinger, J. and Collins, S. L. 2017. Species reordering, not changes in richness, drives long-term dynamics in grassland communities. *Ecol. Lett.*, 20: 1556-1565.
- Lüdecke, D., Mattan, S., Ben-Shachar, Patil, I., Waggoner, P. and Makowski, D. 2021. performance: An R Package for Assessment, Comparison and Testing of Statistical Models. *J. Open Source Software*, 6(60): 3139.
- May, F., Gerstner, K., McGlinn, D. J., Xiao, X. and Chase, J. M. 2018. mobsim: An R package for the simulation and measurement of biodiversity across spatial scales. *Methods Ecol Evol.*, 9: 1401-1408.

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

Stone, L. and Roberts, A. 1990. The checkerboard score and species distributions. *Oecologia* 85: 74–79.

3 CAPÍTULO II - Unveiling the temporal signatures of demographic stochasticity from populations to metacommunities

Este manuscrito foi submetido na forma de artigo científico para a revista *Ecology* e está formatado de acordo com as normas definidas pela revista.

Unveiling the temporal signatures of demographic stochasticity from populations to metacommunities

Cristina Mariana Jacobi¹ and Tadeu Siqueira^{1,2}

¹ Institute of Biosciences, São Paulo State University (UNESP), Rio Claro, SP 13506-900, Brazil;

² School of Biological Sciences, University of Canterbury, Christchurch, New Zealand.

Abstract: The temporal stability of ecological properties tends to increase with spatial scale and levels of biological organization, which is mostly associated with deterministic processes. However, random fluctuations caused by demographic stochasticity in small populations might extend to communities and metacommunities, potentially affecting stability propagation across biological levels and spatial scales. Here, we tested this hypothesis by combining process-based simulations and statistical modeling of 468 sites distributed across 39 regions, sampled from 1981 to 2019, to investigate how fish communities and metacommunities changed over time at local and regional scales. We found that species-rich communities in highly seasonal environments were more variable over time. However, the major driver of compositional temporal variability was community size. Communities comprising smaller populations were more temporally variable than those comprising larger populations. This relationship was weaker at the regional scale, suggesting a dampening effect at the metacommunity level. Our results suggest that the potential effects of demographic stochasticity, which are undoubtedly stronger in small populations, might extend beyond populations, leaving different signals in the temporal variability of ecological properties. These effects appear to be stronger and consistent within small communities and weaker in metacommunities. Our study advances the knowledge of how populational-demographic stochasticity might affect biodiversity temporal dynamics across scales.

KEYWORDS: random demography; community size; compositional variability; environmental variability; spatial scale; temporal variability.

3.1 Introduction

Theoretical and observational research posit that the temporal variability of ecological properties (e.g., coefficient of variation of biomass) decreases with the increase of spatial scale, levels of biological organization, and trophic levels (Kéfi et al. 2019; Wang et al. 2019; Hammond et al. 2020; Siqueira et al. 2024). The explanations for these relationships are mainly based on deterministic processes, such as compensatory dynamics (Gonzalez and Loreau 2009, Brown et al. 2016), the role of mobile predators (McCann et al. 2005), and the Moran effect (Steiner et al. 2013). However, small populations experience high temporal variability due to random demographic events (Reed and Hobbs 2004; Melbourne and Hastings 2008). Thus, as a process that unfolds over time, could demographic stochasticity increase the temporal variability of communities and metacommunities? Or do its effects also weaken from populations to metacommunities? Understanding how random demographic events may affect biodiversity temporal variability at larger scales is important to better forecast how ecosystems will respond to global changes.

Demographic stochastic events such as birth, death, and emigration are inherent elements of population dynamics (Lande 1988; 1993; Otto and Whitlock 1997; Whitlock 2004; Willi, van Buskirk, and Hoffmann 2006; Melbourne and Hastings 2008). While various forces influence population demography, their impacts are consolidated into population size. Small populations are more susceptible to the effects of random events, making them more temporally variable than larger populations (Otto and Whitlock 1997; Reed and Hobbs 2004; Whitlock 2004). Expanding on this understanding, recent research suggests that the effects of local demographic stochastic events might extend to larger spatial scales and levels of biological organization. For example, there is mounting evidence indicating that demographic stochasticity increases spatial beta diversity among local communities comprised of small populations (small communities; Orrock and Watling 2010; Gilbert and Levine 2017; Siqueira et al. 2020) and weakens community-environment relationships (Jacobi and Siqueira 2023).

Although there is limited research about the effects of demographic stochasticity on the temporal variability of multispecies communities, most studies suggest that small communities tend to exhibit greater temporal variability (temporal beta diversity) compared to larger ones because small populations have few

individuals to buffer against random fluctuations in their demography (Orrock and Fletcher 2005; Pedruski, Fussmann, and Gonzalez 2015; Gilbert and Levine 2017; Legault, Fox, and Melbourne 2019). One possible explanation for such relationship is that when demographic stochasticity plays a strong role in community assembly, niche differences become less important, as the likelihood of any demographic event happening is similar for all individuals, regardless of their fitness differences (Orrock and Watling 2010, Legault et al. 2019). For this reason, it is reasonable to expect that small local communities should have higher temporal variability in their species composition (i.e., high temporal beta diversity) than large communities. However, stochastic variability in species composition might also be caused by an interaction between trait-diversity, dispersal and population size (Arim et al. 2023). Thus, further empirical research is still needed to understand the effects of demographic stochasticity in the temporal assembly of ecological communities.

Considering that small communities may experience high temporal variability in species composition, it is possible that metacommunities formed by small communities may also be highly variable in time. Simulations and experiments analyzing biodiversity variability have found that metacommunity size is a key factor shaping species composition temporal variability (LeCraw, Srivastava, and Romero 2014; Suzuki and Economo 2024). Thus, large temporal fluctuations in species composition within the metacommunity should happen if random extinctions (or abundance changes) of species occur independently across different sites within the metacommunity. Alternatively, if demographic stochasticity is not widespread within the metacommunity, its effects might be negligible compared to the influence of deterministic processes like compensatory dynamics (Gonzalez and Loreau 2009; Loreau 2010; Brown, Downing, and Leibold 2016).

In light of the ongoing declines in population sizes and the elevated rates of species extinction (McCallum 2015; He et al. 2019; Almond, Grooten, and Petersen 2020), we ask whether the effects of population-level demographic stochasticity propagate to higher levels of biological organization and influence biodiversity temporal dynamics at larger spatial scales. To do so, we analyzed how temporal variability in species composition was related to the size of communities at local and regional scales. If small communities are more variable than large ones simply due to their size, this would indicate a stronger influence of demographic stochasticity on these communities. However, analyzing this relationship is

challenging as, aside from the need for a significant amount of independent temporally and spatially replicated data across a gradient of community size, it is crucial that the metrics used to quantify temporal variation in species composition are not inherently affected by differences in community size. Thus, we first used a process-based simulation model to test if metrics commonly used to quantify variation in species composition could indicate the effects of demographic stochasticity via their relationship with community size. We then applied these metrics to time-series observational data and statistical models to understand the relationships of temporal variability in species composition at both the community and metacommunity levels with community size and other potential predictors, including temporal variation in environmental conditions, time series length, the number of samples in the time series, and species richness.

We expected to find a negative relationship between temporal variability in species composition within local communities and community size, suggesting that local species composition varies more over time within small communities. At the metacommunity level, we had two competitive expectations: (1) either a negative (2) or no relationship between temporal variability in spatial beta-diversity and regional community size. Evidence supporting the latter would indicate that the effects of local demographic stochasticity are weakened over time at the regional scale.

3.2 Material and Methods

Data

We ran our empirical analyses using data from two large databases. From RivFishTIME (Comte et al. 2021), we obtained time series count data of riverine fish. TerraClimate (Abatzoglou et al. 2018) provided high-resolution monthly data for environmental variables, from which we calculated annual averages for precipitation, maximum air temperature, and minimum air temperature (a good proxy for water temperature; Stefan and Preud'homme 1993).

We defined a metacommunity as the set of sites within basin delineations (HydroBASINS level 7 – Lehner and Grill 2013) and assigned a Strahler stream order for each sampled site using information from the HydroRIVERS network (Lehner and Grill 2013). After that, we selected metacommunities that met the following criteria:

(1) were composed of at least five communities in first to third-order streams, (2) were sampled at least four times in different years, and (3) had at least five species per metacommunity. When dealing with metacommunity data comprising multiple sampling events per year, we selected the sampling date with the highest number of sampled sites. These steps resulted in 468 communities distributed within 39 metacommunities, sampled from 1981 to 2019, located in the Australasia (12), Nearctic (12), and Palearctic (15) biogeographical realms (Appendix S1: Figure S1). These metacommunities are composed on average of 12 communities (standard deviation = 9) sampled, on average, 11 times in the time-series (standard deviation = 5) with an average temporal extent of 14 years (standard deviation = 5). All data selection and manipulation were made in R v. 4.2.1 (www.r-project.org) using the packages `ncdf4` (Pierce 2023), `mapview` (Appelhans et al. 2022) `raster` (Hijmans 2023), `sf` (Pebesma 2018; Pebesma and Bivand 2023), `sp` (Pebesma and Bivand 2005, Bivand et al. 2013), and `tidyverse` (Wickham et al. 2019).

Representing the signature of demographic stochasticity

Because the effects of demographic stochasticity are negatively correlated with population size, we used the median of community size over time as a proxy to represent the signature of demographic stochasticity. We measured the size of each community as the median number of individuals over time (local community size), considering that communities composed of fewer individuals are also composed, on average, of smaller populations than communities with more individuals. Finally, to represent the signature of demographic stochasticity at the regional scale, we measured the size of each metacommunity as the median community size within each metacommunity over time (regional community size).

Metrics of temporal variability

A major challenge involved in the use of community size as a proxy for the signature of demographic stochasticity is that many community metrics can be mathematically related to sample size regardless of the underlying assembly process (Beck, Holloway, and Schwanghart 2013; Chase and Knight 2013; Barwell, Isaac, and Kunin 2015; Cao et al. 2021). Thus, to select appropriate metrics of temporal variability that could be modelled against community size, we first developed a process-based simulation model according to the following steps.

We first simulated a purely neutral dynamics in 39 metacommunities of varying sizes over 1000-time steps with the *mcbrnet* package (Terui and Pomeranz 2023). These simulations were performed using different levels of dispersal within the metacommunities, which mirrored the empirical dataset in terms of community and species numbers. Given that the assembly of the simulated metacommunities was entirely neutral, there should be no relationship between temporal variability in species composition and community size. If any metric of temporal variability was related with size, we understood it as unsuitable for our objective. Secondly, we selected 10 time steps and measured temporal variability at both local (within each community) and regional (among communities) scales using different metrics: (i) temporal variability in species abundance measuring the coefficient of variation of abundance at each organizational level based on the partitioning framework developed by Wang et al. (2019); (ii) temporal variability in species composition at the local and regional scales employing the metric proposed by Lamy et al. (2021) and implemented with the *ltmc* package (Sokol & Lamy 2022); (iii) temporal beta diversity as the median of species rank changes (Avolio et al. 2019) within each community over time, using the *RAC_change* function from the *codyn* package (Hallett et al. 2020). Additionally, we used *codyn* to assess spatial beta diversity at the metacommunity level via the *RAC_difference* function. The median difference in species rank among communities was calculated at each time step. Finally, we regressed all these metrics of temporal variability described above against the median community size over time and compared the outcomes. The performance of each metric is informed in the Results section. Further details on the process-based simulation are available in the supplementary information.

Predictors of temporal variability in species composition

We measured several metrics that are expected to play a role in explaining temporal variability in species composition. This included assessing the number of samples collected over time (number of samples) and the temporal extent of each sampling (time series length) for each community and metacommunity. We also measured the coefficient of variation (CV) of maximum (CV *tmax*) and minimum (CV *tmin*) temperature and precipitation (CV *ppt*) as potential predictors of community variability over time. These metrics were measured by dividing the standard deviation of temperature and precipitation values at each site over time by the mean of these

values. We also estimated the median asymptotic richness of each community over time to be included as a predictor of community variability. Asymptotic richness was estimated with the iNEXT package (Hsieh et al. 2022), which combines extrapolation and interpolation techniques.

At the metacommunity level, we also measured environmental synchrony as a proxy for the Moran effect (Siqueira et al. 2024). We measured environmental synchrony within metacommunities by calculating the average Pearson correlation of each environmental variable (syn tmax, syn tmin, syn ppt) between communities over time. A high correlation or environmental synchrony would indicate that the environmental conditions being analyzed changed similarly across sites, while a low synchrony indicates that environmental conditions vary more independently across sites. Finally, we measured the spatial connectivity within each metacommunity by calculating the mean Euclidean distance, in meters, between the central point of each metacommunity and its respective communities. The smaller the mean distance between communities, the greater the proximity centrality between them (Erős et al. 2012). We used the geosphere package to perform this calculation (Hijmans 2022).

Statistical analysis

To investigate how temporal variability in the species composition of local communities was related to community size, we fitted a Generalized Linear Mixed Model (GLMM) with metacommunity identity as a random effect. We modeled temporal variability in the species composition as the response variable and included as fixed-effects the median of local community size over time, estimated species richness (median of estimated species richness over time), the coefficients of variation of local environmental variables over time (CV tmax, CV tmin, CV ppt), the number of samples and the time series length of each community. All variables were log transformed and we assumed a Gaussian distribution. We used the lme4 package (Bates et al. 2015) to fit this GLMM model and estimated the marginal (due to fixed effects) and the conditional (due to fixed and random effects) R^2 values with the MuMIn package (Bartoń 2023). We also estimated p-values associated with the fixed effects using Wald chi-square tests with the Anova function in the car package (Fox and Weisberg 2019).

To investigate whether temporal variability in species composition within metacommunities was related to regional community size, we used a Generalized

Linear Model (GLM). We included regional community size (median community size within each metacommunity over time), spatial connectivity, environmental synchrony (syn tmax, syn tmin, syn ppt), number of metacommunity samples over time, and the time series length of each metacommunity as predictors. This model was also described by a Gaussian distribution and the variables were not log transformed. Model assumptions were examined via a visual assessment, utilizing the `check_model` function from the `performance` package (Lüdtke et al. 2021), and the `simulateResiduals` function from the `DHARMA` package (Hartig 2022). All analyses were performed in R version 4.2.1 (www.r-project.org). Data and codes used in this research are available at Zenodo (Jacobi 2024).

3.3 Results

Metrics of temporal variability in simulated metacommunities

Our process-based simulation models indicated that while the temporal variability in abundance had a consistent negative relationship with community size both at the local and regional scales, temporal variability in species composition measured by the metric proposed by Lamy et al. (2021) exhibited a negative relationship with community size at the local scale (Appendix S1: Table S1). Thus, these metrics of temporal variability were not adequate to be used in our study. The species rank change metric was the only metric that was not related to community size at the local scale (Appendix S1: Table S1). The species rank difference was positively related to community size in most simulations at the regional scale; however, the explanatory power of the models (R^2) was consistently low across all simulation scenarios (mean $R^2 = 0.03$; Appendix S1: Table S1). Conversely, the LTMC metric had a negative relationship with regional community size in some simulations, while in others, no such relationship was observed (Appendix S1: Table S1). However, when a relationship was present, it exhibited a higher explanatory power compared to the metric based on species rank differences.

Thus, considering that the metrics based on change in species rank yielded the most robust results regarding community size variations in a neutral scenario, we used them to analyze the empirical data. More specifically, to represent temporal variability in species composition for each community, we used the species

rank change metric. To represent temporal variability in species composition in metacommunities we used the coefficient of variation (CV) of the rank difference of each year over time.

Relationships in the empirical dataset

Despite a small violation of model assumptions (residuals slightly deviated from the model predicted values; Appendix S1: Figure S2; S3), the model describing temporal variability in species composition at the local scale explained 34% of the variation in community turnover (fixed effects = 29%, random effects = 5%). Temporal variability in community composition was negatively related to local community size (Wald $X^2_{[1, 458]} = 112.4725$, p-value = $2.2e-16$, Figure 1a), indicating that species change more in rank position over time in smaller communities. Such result was not observed in the null model, which reinforces the significance of observing this relationship in the real dataset. Compositional variability also increased with the increase of species richness (Wald $X^2_{[1, 458]} = 93.2697$, p-value = $2.2e-16$, Figure 1b), CV of the minimum temperature (Wald $X^2_{[1, 458]} = 9.3190$, p-value = 0.002, Figure 1c), and of the CV of precipitation (Wald $X^2_{[1, 458]} = 7.8240$, p-value = 0.005, Figure 1d). These relationships indicate that more diverse communities in more variable environments were also more temporally variable in their species composition.

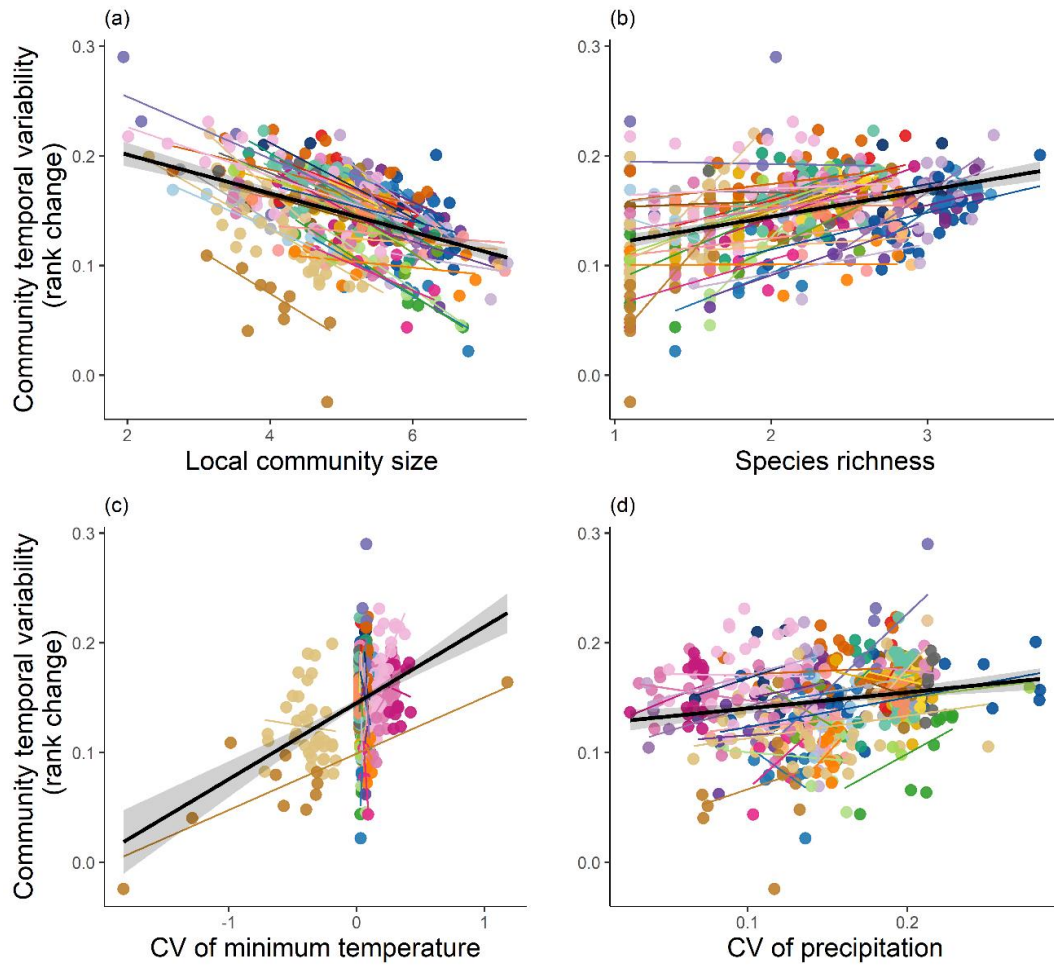


Figure 1. Relationship between temporal variability in species composition of local communities (rank change) and (a) local community size, (b) species richness, (c) coefficient of variation of minimum temperature (CV t_{min}), and (d) coefficient of variation of precipitation (CV ppt). Points represent local communities; colors represent the metacommunity to which the local communities belong to. All variables are in log scale. Grey bands represent the 95% confidence interval.

The model describing temporal variability in species composition at the regional scale exhibited high collinearity among the predictors of environmental synchrony. To address this issue, we refitted the model by including only one environmental synchrony variable at a time and then selected the model with the lowest Akaike information criterion value corrected for small sample size (AIC_c). The selected model was the one that included synchrony of precipitation as a predictor of environmental synchrony (Appendix S1: Table S2). Following this adjustment, the model showed no issues with assumptions (Appendix S1: Figure S4) and explained

16% of temporal variability in species composition. The only predictor variable associated with p -value < 0.05 in the model was syn ppt ($\beta = 0.0820$; std. error = 0.0367; t -value $_{[38, 33]} = -2.232$; $p = 0.0324$; Figure 2a; Appendix S1: Table S3), indicating that greater synchronicity in precipitation across communities reduces the temporal variability in regional species composition. While no relationship was found between temporal variability and regional community size, the data exhibited a slight negative trend (Figure 2b; Appendix S1: Table S3), which may suggest that a negative relationship could emerge with a larger sample. In summary, our results indicate that the relationship between temporal variability in species composition and size was statistically strong only for the local scale.

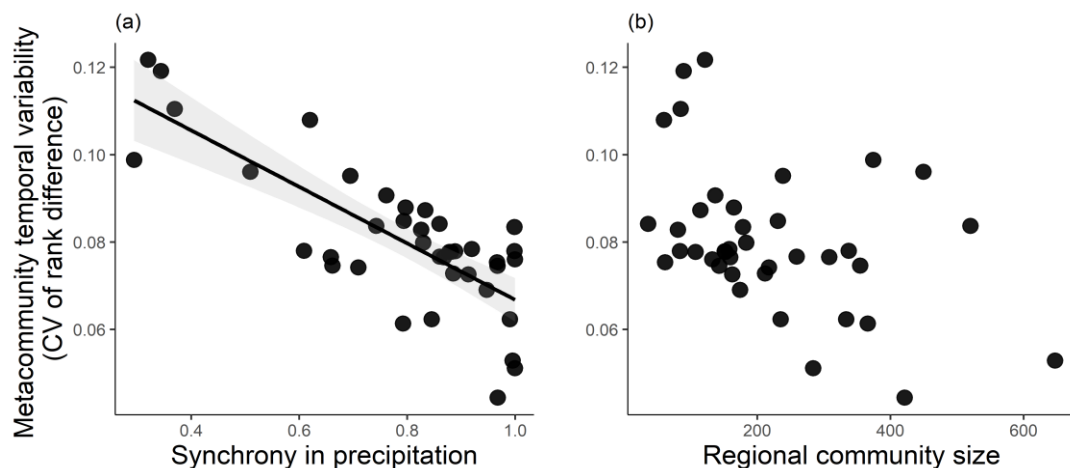


Figure 2. Relationship between temporal variability in species composition in metacommunities (CV of rank difference) and (a) synchrony in precipitation among localities, and (b) regional community size. Grey bands represent the 95% confidence interval.

3.4 Discussion

Our research investigated the potential for random fluctuations driven by demographic stochasticity to extend beyond the population level and influence the dynamics of higher organizational levels of biological organization, such as communities and metacommunities. Agreeing with our predictions, we found a

consistent negative relationship between temporal variability in species composition and community size, but mainly at the local scale. Although statistically weak, there was also a negative trend between temporal variability in species composition and community size at the regional scale. Together, our results suggest that the potential effects of demographic stochasticity, which are undoubtedly stronger in small populations, might leave different signals in the temporal variability of ecological properties, being stronger and consistent within small communities and weaker in metacommunities.

Populations comprising few individuals tend to fluctuate more over time because independent random demographic events can cause substantial changes in both their mean size and variance compared to populations with many individuals (Lande 1993; Legendre et al. 1999; Melbourne and Hastings 2008; Fauvergue et al. 2012). In line with this knowledge, our finding of a negative relationship between temporal variability in species composition and community size suggests that communities comprised of small populations are also more likely to vary in time. This means that species abundances within these communities change more over time than within larger communities, relatively to their size and number of species. This result provides empirical support to previous theoretical and experimental findings that indicated how random demographic events in small populations can lead to temporal changes in the compositional turnover of individual communities (Gilbert and Levine 2017; Legault, Fox, and Melbourne 2019). These studies suggest that communities comprising small populations undergo more pronounced demographic fluctuations, preventing any species from maintaining a constant competitive advantage over others (Orrock and Fletcher 2005). Consequently, the smaller the community, the more the effects of demographic stochasticity increase community variability over time.

Our results also indicated that the influence of demographic stochasticity on community composition weakens or even disappears over time as we move from the local to the regional scale. While at the local scale stochastic events are likely to cause major shifts in the species composition of small communities, at the regional scale they are likely to be counterbalanced by other factors such as the dispersal of individuals among patches and population spatial asynchrony (Wang and Loreau 2016; Thompson et al. 2020). To influence regional species composition, random demographic events that influence the species composition of local communities

would need to occur concurrently and similarly in different communities, resulting in the loss (or change in the abundance) of the same species across the entire metacommunity. However, due to the inherently random nature of demographic stochasticity (Hubbell 2001; Lande, Engen, and Sæther 2003), its impact is unlikely to be uniform across communities. Even if demographic stochasticity were widespread within the metacommunities we studied, its effects on local communities would have been asynchronous, meaning that the species affected by its effects would vary from one community to another, thus keeping regional species composition relatively stable.

Although our analyses did not indicate a statistical relationship between temporal variability in species composition at the regional scale with metacommunity size, a visual inspection of the scatter plot suggested a negative trend between these variables. As the first study of its kind, we caution that our analyses, derived from a limited sample size (39 metacommunities), should be approached with care before dismissing the potential propagation of the effects of random demography to the metacommunity level. Further studies comprising more metacommunities, longer time-series, and different response and predictor variables might clarify this potential relationship. For example, Suzuki and Economo (2024) found that increasing the number of patches in simulated metacommunities led to less temporal variability, implying that larger metacommunities could be more temporally stable.

Beyond community size, multiple abiotic and biotic factors can influence the temporal variability of multispecies communities (Dunson and Travis 1991; Araújo and Luoto 2007; Valencia et al. 2020). After adjusting for the effect of community size, we found that communities with more species were also the ones with more compositional variability over time. Although this finding contrasts with the existing research on the temporal stability of ecological properties, it can be explained by the type of variability being used here. The temporal variability of aggregate ecological properties, such as total biomass or abundance, is expected to decrease with the increase in the number of species due to the statistical averaging of fluctuations in species abundances (Doak et al. 1998; Xu et al. 2021). However, our analyses were based on the temporal variability of species composition (rank of species abundances). As found recently by Arim et al. (2023), high functional and taxonomic diversity allows for a wider array of responses to environmental fluctuations, potentially resulting in varying levels of success for different species over time or

under different conditions. Differential responses of species to stochastic fluctuations in the environment can lead to changes in species dominance over time and increase the variability in species rank abundance (Gonzalez and Descamps-Julien 2004). This is likely the reason why we found that the coefficient of variation (CV) of both minimum temperature and precipitation were positively related to the temporal variability in species composition of local communities.

Demographic stochastic events are a natural component of population dynamics (Otto and Whitlock 1997, Whitlock 2004, Willi et al. 2006). Our research indicates that such population-level events can have consequences for the temporal dynamics of local communities and potentially to metacommunities. So far, most previous research on this topic highlighted the effects of demographic stochasticity on either population temporal dynamics or community spatial dynamics (Lande 1993; Legendre et al. 1999; Vindenes and Engen 2017; Siqueira et al. 2020; Feng et al. 2022; Jacobi and Siqueira 2023). We extend previous findings here by showing with empirical data that demographic stochasticity can also generate greater temporal variability in species composition, but mainly at the local scale. We thus suggest that the effects of demographic stochasticity are scale-dependent, being strongest at the local scale and decreasing in intensity with the increase of spatial scale or level of biological organization.

Acknowledgements

We thank Holly Harris for commenting on a previous version of this manuscript. CMJ thanks the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior-Brasil (CAPES) – Finance Code 001 for funding. TS was supported by grant no. 21/00619-7, São Paulo Research Foundation (FAPESP) and by grant no. 309496/2021-7, Brazilian Natl Council for Scientific and Technological Development (CNPq).

Open Research statement: Data and code used in this research can be found in Jacobi (2024) at Zenodo: [10.5281/zenodo.11242948](https://doi.org/10.5281/zenodo.11242948)

3.5 References

- Abatzoglou, John T., Solomon Z. Dobrowski, Sean A. Parks, and Katherine C. Hegewisch. 2018. "TerraClimate, a High-Resolution Global Dataset of Monthly Climate and Climatic Water Balance from 1958–2015." *Scientific Data* 5 (1): 170191. <https://doi.org/10.1038/sdata.2017.191>.
- Almond, R. E. A., M. Grooten, and T. Petersen. 2020. *Bending the Curve of Biodiversity Loss*. Living Planet Report 2020. Gland: WWF.
- Appelhans, Tim, Detsch Florian, Reudenbach, Christoph, and Woellauer, Stefan. 2022. "_mapview: Interactive viewing of spatial data in R_." <https://CRAN.R-project.org/package=mapview>.
- Araújo, Miguel B., and Miska Luoto. 2007. "The Importance of Biotic Interactions for Modelling Species Distributions under Climate Change." *Global Ecology and Biogeography* 16 (6): 743–53. <https://doi.org/10.1111/j.1466-8238.2007.00359.x>.
- Arim, Matías, Verónica Pinelli, Lucía Rodríguez-Tricot, Esteban Ortiz, Mariana Illarze, César Fagúndez-Pachón, and Ana I. Borthagaray. 2023. "Chance and Necessity in the Assembly of Plant Communities: Stochasticity Increases with Size, Isolation and Diversity of Temporary Ponds." *Journal of Ecology* 111 (8): 1641–55. <https://doi.org/10.1111/1365-2745.14119>.
- Avolio, Meghan L., Ian T. Carroll, Scott L. Collins, Gregory R. Houseman, Lauren M. Hallett, Forest Isbell, Sally E. Koerner, Kimberly J. Komatsu, Melinda D. Smith, and Kevin R. Wilcox. 2019. "A Comprehensive Approach to Analyzing Community Dynamics Using Rank Abundance Curves." *Ecosphere* 10 (10): e02881. <https://doi.org/10.1002/ecs2.2881>.
- Bartoń, Kamil. 2023. "_MuMIn: Multi-model inference_." <https://CRAN.R-project.org/package=MuMIn>.
- Barwell, Louise J., Nick J. B. Isaac, and William E. Kunin. 2015. "Measuring B-diversity with Species Abundance Data." *The Journal of Animal Ecology* 84 (4): 1112–22. <https://doi.org/10.1111/1365-2656.12362>.
- Bates, Douglas, Maechler, Martin, Bolker, Ben, and Walker, Steve. 2015. "Fitting linear mixed-effects models using lme4." *Journal of Statistical Software* 67(1), 1-48. doi:10.18637/jss.v067.i01.
- Beck, Jan, Jeremy D. Holloway, and Wolfgang Schwanghart. 2013. "Undersampling and the Measurement of Beta Diversity." *Methods in Ecology and Evolution* 4 (4): 370–82. <https://doi.org/10.1111/2041-210x.12023>.
- Bivand, Roger S., Pebesma, Edzer, and Gomez-Rubio, V. 2013. "Applied spatial data analysis with R", Second edition. Springer, NY. <https://asdar-book.org/>

- Brown, Bryan L., Amy L. Downing, and Mathew A. Leibold. 2016. "Compensatory Dynamics Stabilize Aggregate Community Properties in Response to Multiple Types of Perturbations." *Ecology* 97 (8): 2021–33. <https://doi.org/10.1890/15-1951.1>.
- Cao, Ke, Jens-Christian Svenning, Chuan Yan, Jintun Zhang, Xiangcheng Mi, and Keping Ma. 2021. "Undersampling Correction Methods to Control γ -Dependence for Comparing β -Diversity between Regions." *Ecology* 102 (9): e03448. <https://doi.org/10.1002/ecy.3448>.
- Chase, Jonathan M., and Tiffany M. Knight. 2013. "Scale-Dependent Effect Sizes of Ecological Drivers on Biodiversity: Why Standardised Sampling Is Not Enough." *Ecology Letters* 16 (s1): 17–26. <https://doi.org/10.1111/ele.12112>.
- Comte, Lise, Juan Carvajal-Quintero, Pablo A. Tedesco, Xingli Giam, Ulrich Brose, Tibor Erős, Ana F. Filipe, et al. 2021. "RivFishTIME: A Global Database of Fish Time-series to Study Global Change Ecology in Riverine Systems." Edited by Volker Bahn. *Global Ecology and Biogeography* 30 (1): 38–50. <https://doi.org/10.1111/geb.13210>.
- Doak, D. F., D. Bigger, E. K. Harding, M. A. Marvier, R. E. O'Malley, and D. Thomson. 1998. "The Statistical Inevitability of Stability-Diversity Relationships in Community Ecology." *The American Naturalist* 151 (3): 264–76. <https://doi.org/10.1086/286117>.
- Dunson, William A., and Joseph Travis. 1991. "The Role of Abiotic Factors in Community Organization." *The American Naturalist*, November. <https://doi.org/10.1086/285270>.
- Erős, Tibor, Julian D. Olden, Robert S. Schick, Dénes Schmera, and Marie-Josée Fortin. 2012. "Characterizing Connectivity Relationships in Freshwaters Using Patch-Based Graphs." *Landscape Ecology* 27 (2): 303–17. <https://doi.org/10.1007/s10980-011-9659-2>.
- Fauvergue, Xavier, Elodie Vercken, Thibaut Malausa, and Ruth A Hufbauer. 2012. "The Biology of Small, Introduced Populations, with Special Reference to Biological Control." *Evolutionary Applications* 5 (5): 424–43. <https://doi.org/10.1111/j.1752-4571.2012.00272.x>.
- Feng, Tao, Hongjuan Zhou, Zhipeng Qiu, and Yun Kang. 2022. "Impacts of Demographic and Environmental Stochasticity on Population Dynamics with Cooperative Effects." *Mathematical Biosciences* 353 (November):108910. <https://doi.org/10.1016/j.mbs.2022.108910>.
- Fox, John, and Weisberg, Sanford. 2019. *An {R} Companion to Applied Regression*, Third Edition. Thousand Oaks CA: Sage. URL: <https://socialsciences.mcmaster.ca/jfox/Books/Companion/>
- Gilbert, Benjamin, and Jonathan M. Levine. 2017. "Ecological Drift and the Distribution of Species Diversity." *Proceedings of the Royal Society B*:

- Biological Sciences* 284 (1855): 20170507.
<https://doi.org/10.1098/rspb.2017.0507>.
- Gonzalez, Andrew, and Blandine, Descamps-Julien. 2004. "Population and Community Variability in Randomly Fluctuating Environments." *Oikos* 106 (1): 105–16. <https://doi.org/10.1111/j.0030-1299.2004.12925.x>.
- Gonzalez, Andrew, and Michel Loreau. 2009. "The Causes and Consequences of Compensatory Dynamics in Ecological Communities." *Annual Review of Ecology, Evolution, and Systematics* 40 (1): 393–414.
<https://doi.org/10.1146/annurev.ecolsys.39.110707.173349>.
- Hallett, Lauren, Avolio, Meghan, Carroll, Ian, Jones, Sydney, MacDonald, Andrew, Flynn, Dan, Slaughter, Peter, Ripplinger, Julie, Collins, Scott, Gries, Corinna, and Jones, Matthew. 2020. "_codyn: community dynamics metrics_".
<https://doi.org/10.5063/F1N877Z6>, <https://github.com/NCEAS/codyn>.
- Hammond, Matthew, Michel Loreau, Claire De Mazancourt, and Jurek Kolasa. 2020. "Disentangling Local, Metapopulation, and Cross-community Sources of Stabilization and Asynchrony in Metacommunities." *Ecosphere* 11 (4): e03078.
<https://doi.org/10.1002/ecs2.3078>.
- Hartig, Florian. 2022. "_DHARMA: Residual diagnostics for hierarchical (Multi-Level / Mixed) regression models_". <https://CRAN.R-project.org/package=DHARMA>.
- He, Fengzhi, Christiane Zarfl, Vanessa Bremerich, Jonathan N. W. David, Zeb Hogan, Gregor Kalinkat, Klement Tockner, and Sonja C. Jähnig. 2019. "The Global Decline of Freshwater Megafauna." *Global Change Biology* 25 (11): 3883–92. <https://doi.org/10.1111/gcb.14753>.
- Hijmans, Robert. 2022. "_geosphere: spherical trigonometry_". <https://CRAN.R-project.org/package=geosphere>.
- Hijmans, Robert. 2023. "_raster: Geographic data analysis and modeling.". <https://CRAN.R-project.org/package=raster>.
- Hsieh, T. C., Ma, K. H. and Chao, Anne. 2022. "iNEXT: iNterpolation and EXTrapolation for species diversity."
<http://chao.stat.nthu.edu.tw/wordpress/software-download/>.
- Hubbell, Stephen P. 2001. *The Unified Neutral Theory of Biodiversity and Biogeography*. Nachdr. Monographs in Population Biology 32. Princeton, NJ: Princeton Univ. Press.
- Jacobi, Cristina Mariana, and Tadeu Siqueira. 2023. "High Compositional Dissimilarity among Small Communities Is Decoupled from Environmental Variation." *Oikos* 2023 (8): e09802.

- Jacobi, Cristina Mariana. 2024. "Codes and data: unveiling temporal signatures of demographic stochasticity from populations to metacommunities [Data set]". Zenodo. <https://doi.org/10.5281/zenodo.11242949>
- Kéfi, Sonia, Virginia Domínguez-García, Ian Donohue, Colin Fontaine, Elisa Thébault, and Vasilis Dakos. 2019. "Advancing Our Understanding of Ecological Stability." *Ecology Letters* 22 (9): 1349–56. <https://doi.org/10.1111/ele.13340>.
- Lamy, Thomas, Nathan I. Wisnoski, Riley Andrade, Max C. N. Castorani, Aldo Compagnoni, Nina Lany, Luca Marazzi, et al. 2021. "The Dual Nature of Metacommunity Variability." *Oikos* 130 (12): 2078–92. <https://doi.org/10.1111/oik.08517>.
- Lande, Russell. 1988. "Genetics and Demography in Biological Conservation." *Science* 241 (4872): 1455–60. <https://doi.org/10.1126/science.3420403>.
- Lande, Russell. 1993. "Risks of Population Extinction from Demographic and Environmental Stochasticity and Random Catastrophes." *The American Naturalist* 142 (6): 911–27.
- Lande, Russell, Steinar Engen, and Bernt-Erik Sæther. 2003. *Stochastic Population Dynamics in Ecology and Conservation*. Oxford University Press.
- LeCraw, Robin M., Diane S. Srivastava, and Gustavo Q. Romero. 2014. "Metacommunity Size Influences Aquatic Community Composition in a Natural Mesocosm Landscape." *Oikos* 123 (8): 903–11. <https://doi.org/10.1111/oik.01253>.
- Legault, Geoffrey, Jeremy W. Fox, and Brett A. Melbourne. 2019. "Demographic Stochasticity Alters Expected Outcomes in Experimental and Simulated Non-neutral Communities." *Oikos* 128 (12): 1704–15. <https://doi.org/10.1111/oik.06028>.
- Legendre, Stéphane, Jean Clobert, Anders P. Møller, and Gabriele Sorci. 1999. "Demographic Stochasticity and Social Mating System in the Process of Extinction of Small Populations: The Case of Passerines Introduced to New Zealand." *The American Naturalist* 153 (5): 449–63. <https://doi.org/10.1086/303195>.
- Lehner, Bernhard, and Günther Grill. 2013. "Global River Hydrography and Network Routing: Baseline Data and New Approaches to Study the World's Large River Systems." *Hydrological Processes* 27 (15): 2171–86. <https://doi.org/10.1002/hyp.9740>.
- Loreau, Michel. 2010. "Linking Biodiversity and Ecosystems: Towards a Unifying Ecological Theory." *Philosophical Transactions of the Royal Society B: Biological Sciences* 365 (1537): 49–60. <https://doi.org/10.1098/rstb.2009.0155>.

- Lüdecke, Daniel, Ben-Shachar, Mattan S., Patil, Indrajeet, Waggoner, Philip, and Makowski, Dominique. 2021. "performance: An R package for assessment, comparison and testing of statistical models." *Journal of Open Source Software*, 6(60), 3139. <https://doi.org/10.21105/joss.03139>
- McCallum, Malcolm L. 2015. "Vertebrate Biodiversity Losses Point to a Sixth Mass Extinction." *Biodiversity and Conservation* 24 (10): 2497–2519. <https://doi.org/10.1007/s10531-015-0940-6>.
- McCann, Kevin S., Rasmussen, Joseph B, and Umbanhowar, James. 2005. "The dynamics of spatially coupled food webs." *Ecology Letters* 8: 513–523.
- Melbourne, Brett A., and Alan Hastings. 2008. "Extinction Risk Depends Strongly on Factors Contributing to Stochasticity." *Nature* 454 (7200): 100–103. <https://doi.org/10.1038/nature06922>.
- Orrock, John L., and Robert J. Fletcher. 2005. "Changes in Community Size Affect the Outcome of Competition." *The American Naturalist* 166 (1): 107–11. <https://doi.org/10.1086/430641>.
- Orrock, John L., and James I. Watling. 2010. "Local Community Size Mediates Ecological Drift and Competition in Metacommunities." *Proceedings: Biological Sciences* 277 (1691): 2185–91.
- Otto, Sarah P, and Michael C Whitlock. 1997. "The Probability of Fixation in Populations of Changing Size." *Genetics* 146 (2): 723–33. <https://doi.org/10.1093/genetics/146.2.723>.
- Pebesma, Edzer, and Bivand, Roger. 2005. "Classes and methods for spatial data in R." *R News* 5 (2), <https://cran.r-project.org/doc/Rnews/>.
- Pebesma, Edzer. 2018. "Simple Features for R: Standardized support for spatial vector data." *The R Journal* 10 (1), 439-446, <https://doi.org/10.32614/RJ-2018-009>
- Pebesma, Edzer, and Bivand, Roger. 2023. *Spatial data science: with applications in R (1st ed.)*. Chapman and Hall/CRC. <https://doi.org/10.1201/9780429459016>
- Pedruski, Michael T., Gregor F. Fussmann, and Andrew Gonzalez. 2015. "Predicting the Outcome of Competition When Fitness Inequality Is Variable." *Royal Society Open Science* 2 (8): 150274. <https://doi.org/10.1098/rsos.150274>.
- Pierce, David. 2023. "_ncdf4: Interface to Unidata netCDF (Version 4 or Earlier) Format Data Files_." <https://CRAN.R-project.org/package=ncdf4>.
- R Core Team 2022. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. <https://www.R-project.org/>.

- Reed, David H., and Gayla R. Hobbs. 2004. "The Relationship between Population Size and Temporal Variability in Population Size." *Animal Conservation* 7 (1): 1–8. <https://doi.org/10.1017/S1367943004003476>.
- Siqueira, Tadeu, Charles P. Hawkins, Julian D. Olden, Jonathan Tonkin, Lise Comte, Victor S. Saito, Thomas L. Anderson, et al. 2024. "Understanding Temporal Variability across Trophic Levels and Spatial Scales in Freshwater Ecosystems." *Ecology* 105 (2): e4219. <https://doi.org/10.1002/ecy.4219>.
- Siqueira, Tadeu, Victor S. Saito, Luis M. Bini, Adriano S. Melo, Danielle K. Petsch, Victor L. Landeiro, Kimmo T. Tolonen, Jenny Jyrkänkallio-Mikkola, Janne Soininen, and Jani Heino. 2020. "Community Size Can Affect the Signals of Ecological Drift and Niche Selection on Biodiversity." *Ecology* 101 (6). <https://doi.org/10.1002/ecy.3014>.
- Sokol, Eric, and Lamy, Thomas. 2022. "_ltmc: Long-term metacommunity analysis_."
- Stefan, Heinz G., and Eric B. Preud'homme. 1993. "Stream Temperature Estimation from Air Temperature1." *JAWRA Journal of the American Water Resources Association* 29 (1): 27–45. <https://doi.org/10.1111/j.1752-1688.1993.tb01502.x>.
- Steiner, Christopher, F., Stockwell, Richard, D., Kalaimani, Vidhya, and Agel, Zakaria. 2013. "Population Synchrony and Stability in Environmentally Forced Metacommunities". *Oikos* 122: 1195–1206. <https://doi.org/10.1111/j.1600-0706.2012.20936.x>
- Suzuki, Yuka, and Evan P. Economo. 2024. "The Stability of Competitive Metacommunities Is Insensitive to Dispersal Connectivity in a Fluctuating Environment." *The American Naturalist*, January, 729601. <https://doi.org/10.1086/729601>.
- Terui, A. and Pomeranz, J. F. 2023. "_mcbrnet: Metacommunity simulation in branching networks_." <https://github.com/aterui/mcbrnet>.
- Thompson, Patrick L., Laura Melissa Guzman, Luc De Meester, Zsófia Horváth, Robert Ptacnik, Bram Vanschoenwinkel, Duarte S. Viana, and Jonathan M. Chase. 2020. "A Process-Based Metacommunity Framework Linking Local and Regional Scale Community Ecology." *Ecology Letters* 23 (9): 1314–29. <https://doi.org/10.1111/ele.13568>.
- Valencia, Enrique, Francesco de Bello, Thomas Galland, Peter B. Adler, Jan Lepš, Anna E-Vojtkó, Roel van Klink, et al. 2020. "Synchrony Matters More than Species Richness in Plant Community Stability at a Global Scale." *Proceedings of the National Academy of Sciences* 117 (39): 24345–51. <https://doi.org/10.1073/pnas.1920405117>.
- Vindenes, Yngvild, and Steinar Engen. 2017. "Demographic Stochasticity and Temporal Autocorrelation in the Dynamics of Structured Populations." 2017. <https://onlinelibrary.wiley.com/doi/abs/10.1111/oik.03958>.

- Wang, Shaopeng, Thomas Lamy, Lauren M. Hallett, and Michel Loreau. 2019. "Stability and Synchrony across Ecological Hierarchies in Heterogeneous Metacommunities: Linking Theory to Data." *Ecography* 42 (6): 1200–1211. <https://doi.org/10.1111/ecog.04290>.
- Wang, Shaopeng, and Michel Loreau. 2016. "Biodiversity and Ecosystem Stability across Scales in Metacommunities." Edited by Ferenc Jordan. *Ecology Letters* 19 (5): 510–18. <https://doi.org/10.1111/ele.12582>.
- Whitlock, Michael C. 2004. "Selection and Drift in Metapopulations." In *Ecology, Genetics and Evolution of Metapopulations*, 153–73. Elsevier. <https://doi.org/10.1016/B978-012323448-3/50009-X>.
- Wickham, Hadley, Averick, Mara, Bryan, Jennifer, Chang, Winston, McGowan, Lucy D., François, Romain, Golemund, Garrett, Hayes, Alex, Henry, Lionel, Hester, Jim, Kuhn, Max, Pedersen, Thomas L., Miller, Evan, Bache, Stephan M., Müller, Kirill, Ooms, Jeroen, Robinson, David, Seidel, Dana P., Spinu, Vitalie, Takahashi, Kohske, Vaughan, Davis, Wilke, Claus, Woo, Kara, and Yutani, Hiroaki. 2019. "Welcome to the tidyverse." *Journal of Open Source Software*, *4*(43), 1686. <https://doi.org/10.21105/joss.01686>.
- Willi, Yvonne, Josh van Buskirk, and Ary A. Hoffmann. 2006. "Limits to the Adaptive Potential of Small Populations." *Annual Review of Ecology, Evolution, and Systematics* 37:433–58.
- Xu, Qianna, Xian Yang, Ying Yan, Shaopeng Wang, Michel Loreau, and Lin Jiang. 2021. "Consistently Positive Effect of Species Diversity on Ecosystem, but Not Population, Temporal Stability." Edited by Jonathan Chase. *Ecology Letters* 24 (10): 2256–66. <https://doi.org/10.1111/ele.13777>.

3.6 Supplementary Information: Appendix S1

Journal: Ecology

Unveiling the temporal signatures of demographic stochasticity from populations to metacommunities

Cristina Mariana Jacobi & Tadeu Siqueira

3.6.1 Supplementary Information Text

Process-based simulation model

To select temporal variability metrics that could be used to investigate the role of demographic stochasticity, we simulated neutral temporal dynamics in metacommunities of varying sizes and examined the relationship between temporal variability metrics and community size. By simulating neutral metacommunities, we eliminated niche selection effects on species composition. As communities were assembled under neutral dynamics, any relationship between temporal variability and community size would suggest that a given variability metric was not appropriate.

Initially, we simulated random branching networks using the function `brnet` from the `mcbrnet` package (Terui & Pomeranz, 2023). The branching networks were then employed in the `mcsim` function of the `mcbrnet` package to simulate our metacommunities. We generated 39 metacommunities comprising the same range of community (5-30) and species (5-48) numbers as observed in the datasets used for statistical modeling. Given that we were interested in simulating a neutral dynamic, we excluded environmental variation and niche differences from the simulation. We randomly varied the number of species among the metacommunities to generate a gradient of metacommunity sizes. We also incorporated a species carrying capacity value within communities to restrict growth and prevent excessive size increase over time. The carrying capacity of each population was consistent within communities of a given metacommunity but varied randomly among metacommunities (ranging from 50 to 150 individuals). We also introduced variability in species abundance within communities by simulating a disturbance. To achieve this, we set the disturbance

probability to 1 and the intensity to 0.43 for all metacommunities. We arrived at this disturbance value by adjusting it to match metacommunity sizes similar to those in real datasets. Additionally, we conducted simulations considering different dispersal values within metacommunities (0.1, 0.5, and 1) to observe if this could impact the relationships with size. Then, we ran each simulation for a total of 1000 time steps. We selected the time-steps 100, 200, 300, 400, 500, 600, 700, 800, 900, and 1000 as our temporal samples to calculate metrics related to the temporal variability in species composition.

We started measuring the temporal variability in species abundance at both the community and metacommunity levels to subsequently relate them to the median community size. We assessed temporal variability in species abundance by measuring the temporal variability of aggregate abundance at each biological level based on the framework proposed by Wang et al. (2019). However, we excluded this metric from our future analyses as it demonstrated a negative relationship with community size, both at the local and regional scales (Table. S1). Then, we decided to test some metrics for temporal variability in species composition. We measured both community and metacommunity variability using the approach proposed by Lamy et al. (2021) with the *ltmc* package. We also assessed temporal variability in the rank of species (Avolio et al. 2019), measuring the rank change within each community over time and the rank difference between communities at each time point. This was done using the *codyn* package (Hallett et al.) Considering multiple simulations conducted with various dispersal rates and different seeds, we observed that both at the local and regional scales, the metrics of species rank variability were less influenced by community size compared to the *ltmc* metrics (Table S1). When the metrics of species rank variability showed an association with size, the relationship exhibited low explanatory power. In contrast, the local *ltmc* metric consistently showed a negative relationship with community size, while the regional *ltmc* demonstrated higher explanatory power when related to community size.

Table S1 - Statistics obtained in the process-based simulation model by relating the metrics of temporal variability in species composition with the median community size. Simulations were conducted using different seeds and dispersal rates. The p-value represents the significance of each relationship, and the explanatory power was measured by R^2 .

Temporal variability metric	set.seed	Dispersal	Slope	p-value	R^2	Adjusted R^2
Local abundance (CV)	1234	0.1	-1.7E-04	2.5E-14	0.7958	0.7903
Regional abundance (CV)	1234	0.1	-4.1E-05	0.0002	0.3084	0.2897
LTMC local	1234	0.1	-1.2E-04	0.0055	0.1901	0.1683
LTMC regional	1234	0.1	-6.4E-06	0.0985	0.0721	0.0470
Rank change	1234	0.1	1.0E-06	0.8411	6.3E-05	-0.0015
Rank difference	1234	0.1	1.1E-05	0.0017	0.0252	0.0227
Local abundance (CV)	1234	0.5	-1.3E-04	5.2E-15	0.8123	0.8072
Regional abundance (CV)	1234	0.5	-4.4E-05	0.0006	0.2777	0.2581
LTMC local	1234	0.5	-3.8E-05	0.0132	0.1549	0.1321
LTMC regional	1234	0.5	-1.2E-06	0.6211	0.0067	-0.0202
Rank change	1234	0.5	8.6E-06	0.0364	0.0069	0.0053
Rank difference	1234	0.5	4.2E-06	0.301	0.0028	0.0002
Local abundance (CV)	1234	1	1.1E-01	2.7E-13	0.7682	0.7619
Regional abundance (CV)	1234	1	-5.1E-02	7.7E-05	0.3482	0.3306
LTMC local	1234	1	-3.7E-05	0.0108	0.1632	0.1406
LTMC regional	1234	1	-2.3E-06	0.405	0.0188	-0.0077
Rank change	1234	1	5.0E-07	0.9071	2.2E-05	-0.0016
Rank difference	1234	1	1.2E-05	0.0091	0.0174	0.0149
Local abundance (CV)	111	0.1	-1.9E-04	4.0E-11	0.6967	0.6885
Regional abundance (CV)	111	0.1	-6.0E-05	3.6E-05	0.3736	0.3567
LTMC local	111	0.1	-1.3E-04	0.0001	0.328	0.3098
LTMC regional	111	0.1	-9.0E-06	0.0021	0.2274	0.2065
Rank change	111	0.1	-2.3E-07	0.9565	4.8E-06	-0.0016
Rank difference	111	0.1	1.4E-05	1.5E-06	0.0580	0.0556
Local abundance (CV)	111	0.5	-1.4E-04	9.5E-12	0.7192	0.7116
Regional abundance (CV)	111	0.5	-4.8E-05	6.8E-06	0.4258	0.4103
LTMC local	111	0.5	-4.6E-05	0.0003	0.3016	0.2827
LTMC regional	111	0.5	-6.7E-06	0.0060	0.1868	0.1648
Rank change	111	0.5	6.3E-06	0.1187	0.0039	0.0023
Rank difference	111	0.5	2.1E-05	2.1E-08	0.0778	0.0754
Local abundance (CV)	111	1	-1.4E-04	2.7E-11	0.7028	0.6948
Regional abundance (CV)	111	1	-4.7E-05	1.3E-07	0.5329	0.5203
LTMC local	111	1	-4.3E-05	0.0003	0.3063	0.2875
LTMC regional	111	1	-5.5E-06	0.0255	0.1278	0.1042
Rank change	111	1	-4.8E-07	0.9029	2.4E-05	-0.0016
Rank difference	111	1	1.7E-05	1.4E-05	0.0476	0.0451

3.6.2 Supplementary Figures

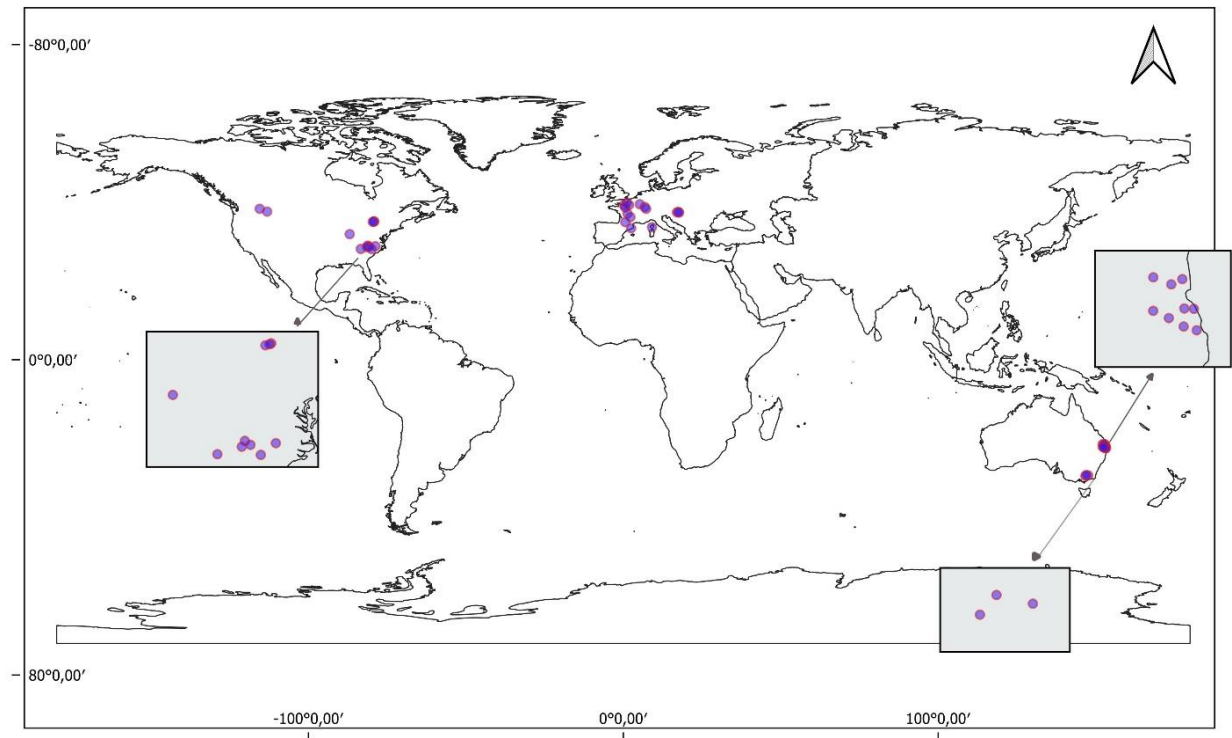


Figure S1 - Geographic distribution of the 39 metacommunities selected in our study, located in Australasia (12), Nearctic (12), and Palearctic (15) biogeographical realms.

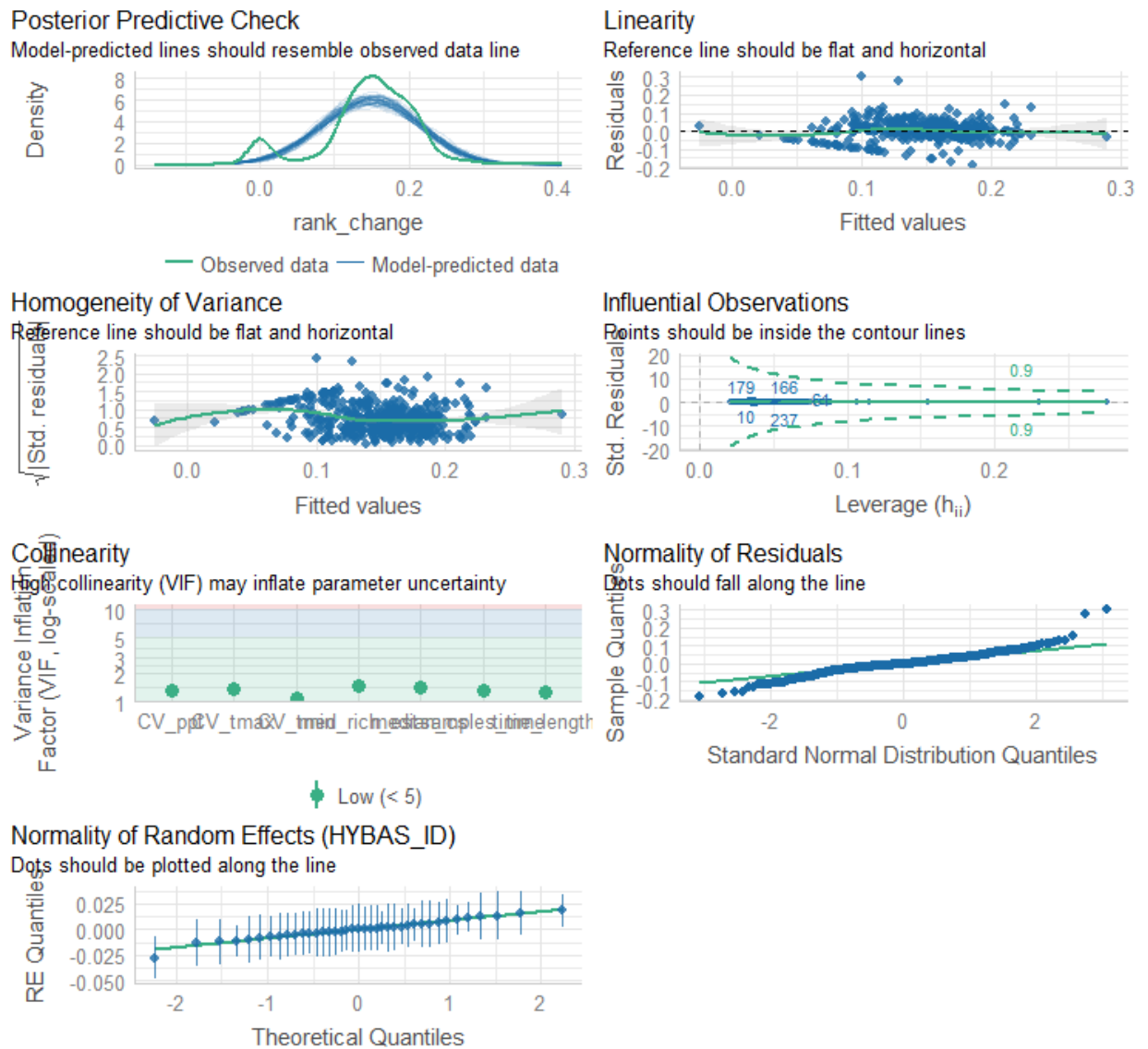


Figure S2 – Visual check of various assumptions of the model explaining temporal variability in species composition in local communities.

An alternative way to assess a GLMM is by examining the residuals from the fitted model, a process that can be done with `simulateResiduals` function from the DHARMA package (Hartig, 2022). This function generates a qq-plot (Fig. S3 - left panel) to identify overall deviations from the expected distribution, typically including tests for distribution correctness (KS test), dispersion, and outliers. Additionally, it produces the `plotResiduals` (Fig. S3 - right panel), which generates a plot of residuals against predicted values, aiding in visualizing deviations from uniformity. We observed that the residuals deviated slightly from the model's predictions (Fig. S3;

the closer the solid line to the dashed line, the better the fit). However, a residual pattern does not suggest the model is unsuitable (Hartig, 2022). With a considerable number of data points, residual diagnostics are likely to show significance, as achieving a perfectly fitting model is improbable (Hartig, 2022).

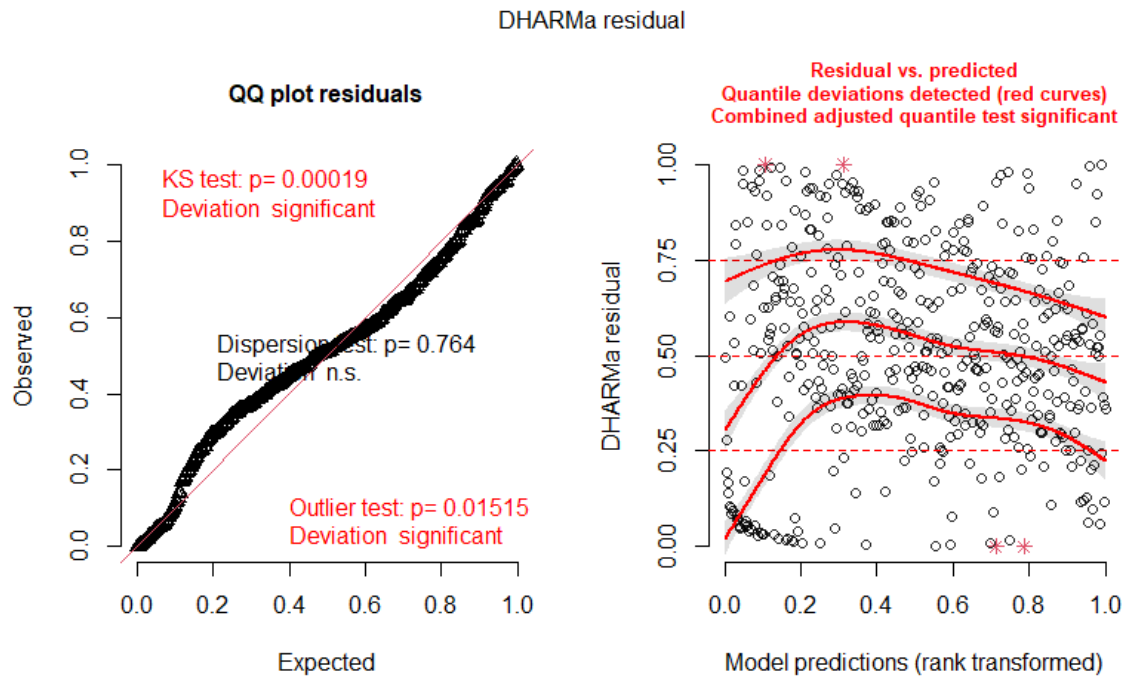


Figure S3 – Output from the simulateResiduals function of the DHARMA package, providing a visual aid in detecting deviations from uniformity for the model that explains the temporal variability in species composition in local communities (model one). Red stars represent simulation outliers.

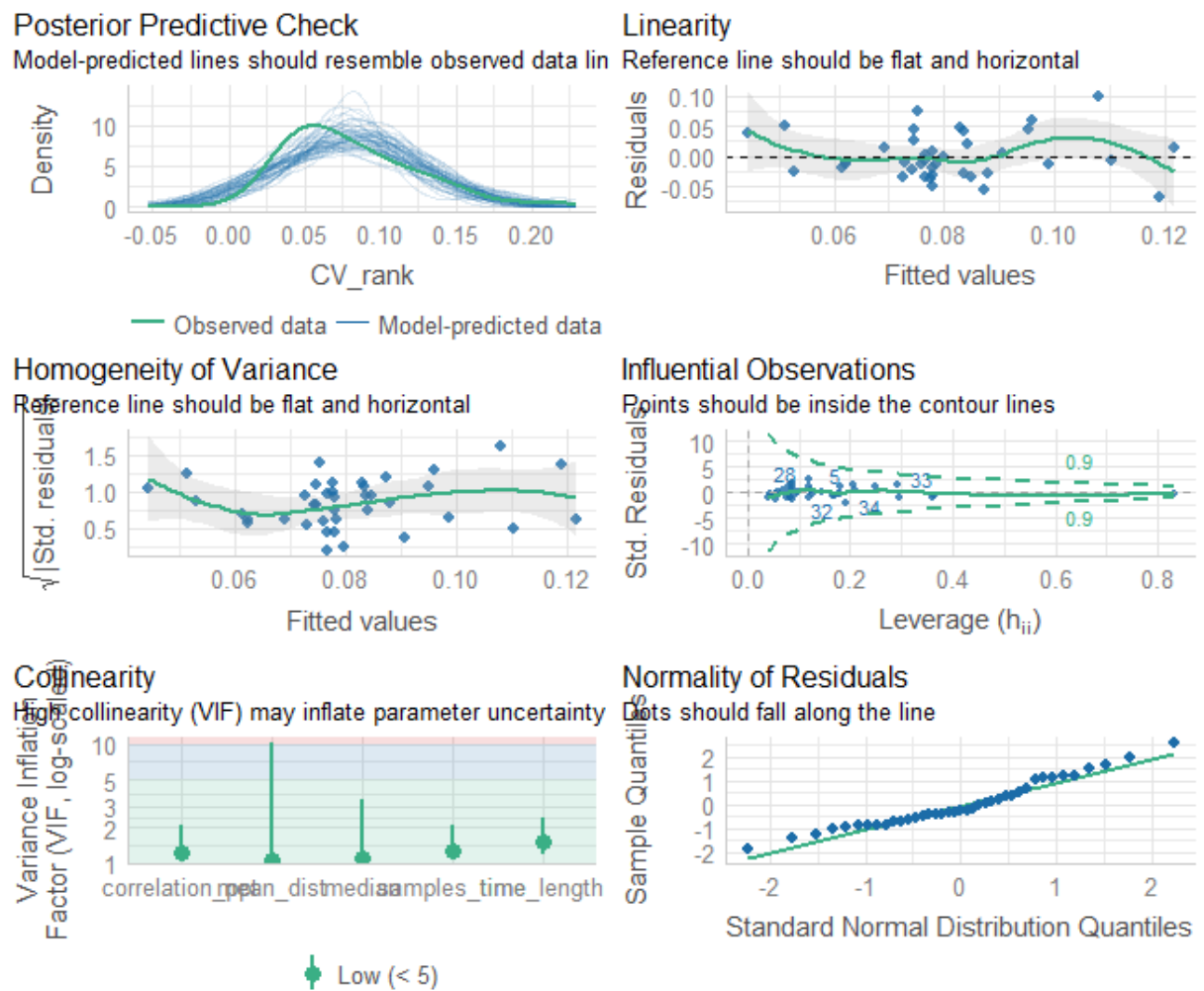


Figure S4 – Visual check of various assumptions from the model that explains the temporal variability in species composition in metacommunities.

3.6.3 Supplementary Table

Table S2. Model selection for generalized linear model of temporal variability in species composition at the regional scale (CV of rank difference) as a function of regional community size (Reg. size), spatial connectivity (Connectivity), number of metacommunity samples over time (Number of samples), time series length of each metacommunity (Time series length), synchrony in precipitation (Syn ppt), synchrony in maximum temperature (Syn tmax) and synchrony in minimum temperature (Syn tmin). AIC_c is Akaike's Information Criterion corrected for small sample sizes, ΔAIC_c indicates the difference between a model's AIC_c value and the AIC_c value for the best-fitting model.

Model	Reg. size	Connectivity	Number of samples	Time series length	Syn ppt	Syn tmax	Syn tmin	AIC _c	ΔAIC_c
M1	x	x	x	x	x			-128.6	0
M2	x	x	x	x			x	-128.28	0.32
M3	x	x	x	x		x		-128.02	0.58

Table S3 - Model statistics summary for the metacommunity GLM. Here, the response variable was the temporal variability in species composition at the regional scale (CV of rank difference). Model coefficients (and their standard errors), t-values, and p-values are reported for each predictor variable and the intercept.

	Estimate	Std. Error	t-value	p-value
Intercept	1.769e-01	4.538e-02	3.899	0.000447
Regional size	-4.074e-05	5.001e-05	-0.815	0.421058
Connectivity	-1.387e-09	1.438e-07	-0.010	0.992362
Syn ppt	-8.202e-02	3.674e-02	-2.232	0.032488
Number of samples	3.278e-04	1.521e-03	0.216	0.830645
Time series length	-1.865e-03	1.720e-03	-1.085	0.285962

3.6.4 Supporting information references

- Hartig, F. 2022. DHARMA: Residual Diagnostics for Hierarchical (Multi-Level / Mixed) Regression Models. <http://florianhartig.github.io/DHARMA/>
- Hallett L, Avolio M, Carroll I, Jones S, MacDonald A, Flynn D, Slaughter P, Ripplinger J, Collins S, Gries C, Jones M 2020. `_codyn`: Community Dynamics Metrics_. <https://doi.org/10.5063/F1N877Z6>, <https://github.com/NCEAS/codyn>.
- Lamy, T., N. I. Wisnoski, R. Andrade, M. C. N. Castorani, A. Compagnoni, N. Lany, L. Marazzi, et al. 2021. The dual nature of metacommunity variability. *Oikos* 130:2078–2092.
- Terui A, Pomeranz JF 2023. `_mcbrnet`: Metacommunity simulation in branching networks_. <https://github.com/aterui/mcbrnet>.
- Wang, S., Lamy, T., Hallett, L. M. and Loreau, M. 2019. Stability and synchrony across ecological hierarchies in heterogeneous metacommunities: linking theory to data. *Ecography* 42: 1200–1211.

CONSIDERAÇÕES FINAIS

Esta tese traz importantes contribuições para a Ecologia de Comunidades revelando como os efeitos da estocasticidade demográfica podem se estender a escalas espaciais maiores e a níveis mais elevados de organização biológica. Os resultados que encontrei são baseados na análise de extensas bases de dados espaciais e temporais de comunidades de peixes de riacho, aliadas a análise de variáveis ambientais e do uso de métricas robustas. No capítulo I, encontrei que a diminuição do tamanho das comunidades não só aumenta a variação espacial na composição de espécies, como sugerido em estudos anteriores, mas também que essa variação está menos relacionada às condições ambientais. Isso sugere que declínios na abundância das espécies devem aumentar a aleatoriedade na composição de espécies, de forma a tornar comunidades menores mais suscetíveis a extinções locais e menos previsíveis com base em variações ambientais. Essa observação contribui para a ideia de um continuum de forças entre os processos ecológicos, pois, à medida que a deriva ecológica ganha força, a seleção enfraquece. Além disso, a menor previsibilidade de comunidades menores deve dificultar a criação de modelos preditivos de distribuição e abundância de espécies, tornando o manejo e a conservação da biodiversidade ainda mais desafiadores frente a alterações ambientais. Por outro lado, a compreensão de que a composição de espécies em comunidades menores é mais suscetível a variações aleatórias pode ajudar a direcionar os esforços de conservação para essas regiões e assim garantir a proteção da diversidade desses locais.

No capítulo II, eu encontrei que as flutuações populacionais geradas pela estocasticidade demográfica podem deixar sinais diferentes na variabilidade temporal de propriedades ecológicas. Esses sinais são dependentes da escala espacial considerada, sendo mais fortes em escala local, resultando em maior variabilidade temporal na composição de espécies de comunidades pequenas. Com o aumento da escala, os efeitos da estocasticidade demográfica sobre a variabilidade temporal da composição de espécies parecem diminuir em intensidade. Provavelmente, em metacomunidades, os eventos estocásticos são contrabalançados por outros fatores, como a dispersão de indivíduos entre comunidades e a assincronia espacial das populações. Além disso, eu observei que

a variabilidade temporal na composição de espécies também aumenta em comunidades com maior riqueza de espécies e localizadas em ambientes mais variáveis, provavelmente porque nem todas as espécies reagem da mesma forma a mudanças no ambiente. Em metacomunidades, a variabilidade temporal na composição de espécies foi maior quando a sincronia ambiental foi menor entre as comunidades.

Esta tese também trouxe uma perspectiva metodológica importante ao demonstrar a necessidade de avaliar se as métricas usadas para medir a diversidade de espécies podem refletir a influência da estocasticidade demográfica sem que essa influência seja confundida com efeitos de diferenças de tamanho das amostras. Sugiro que estudos futuros avaliando estas relações utilizem simulações para controlar o efeito do tamanho da amostra, evitando interpretações errôneas.

Para continuarmos avançando o conhecimento na área, sugiro o desenvolvimento de estudos que avaliem o papel do acaso em ambientes diferentes (por exemplo o ambiente terrestre) e utilizando diferentes variáveis resposta e preditora para representar a força de processos determinísticos e estocásticos. Por exemplo, é possível investigar como o número de comunidades, o grau de conectividade e o tamanho dos ecossistemas afetam a força da estocasticidade demográfica. Por fim, acredito que compreender o papel do acaso é fundamental para modelar a dinâmica das comunidades de forma mais realista e tomar decisões de manejo mais assertivas. Ações que priorizem o aumento do tamanho das comunidades, por exemplo, podem evitar que estas variem de maneira imprevisível e que estejam mais sujeitas a extinções locais.

ANEXO I

➔ Lista de publicações que participei como autora ou coautora de forma paralela ao meu doutorado:

VILLAMARÍN, F.; JARDINE, T. D.; BUNN, S. E.; MALVASIO, A.; PIÑA, C. I.; **JACOBI, C. M.**; ARAÚJO, D. D.; DE BRITO, E. S.; CARVALHO, F. M.; DA COSTA, I. D.; VERDADE, L. M.; LARA, N.; DE CAMARGO, P. B.; MIORANDO, P. S.; PORTELINHA, T. C. G.; MARQUES, T. S.; MAGNUSSON, W. E. Body size predicts ontogenetic nitrogen stable-isotope ($\delta^{15}\text{N}$) variation, but has little relationship with trophic level in ectotherm vertebrate predators. **Scientific Reports**, 14, 14102. 2024.

NYHOLT, K.; JARDINE, T. D.; VILLAMARÍN, F.; **JACOBI, C. M.**; HAWES, J. E.; CAMPOS-SILVA, J. V.; SRAYKO, S.; MAGNUSSON, W. E. High rates of mercury biomagnification in fish from Amazonian floodplain-lake food webs. **Science of The Total Environment**, Volume 833, 155161, ISSN 0048-9697, 2022.

➔ Capítulo de livro aceito para publicação:

JACOBI, C. M.; GUIMARÃES, T. F. R.; BECKER, F. G. Representatividade de ecossistemas aquáticos continentais e da ictiofauna ameaçada em Unidades de Conservação do Rio Grande do Sul. *In: Biodiversidade e conservação de ecossistemas aquáticos do sul do Brasil*. Pelotas: USEB, 2024.