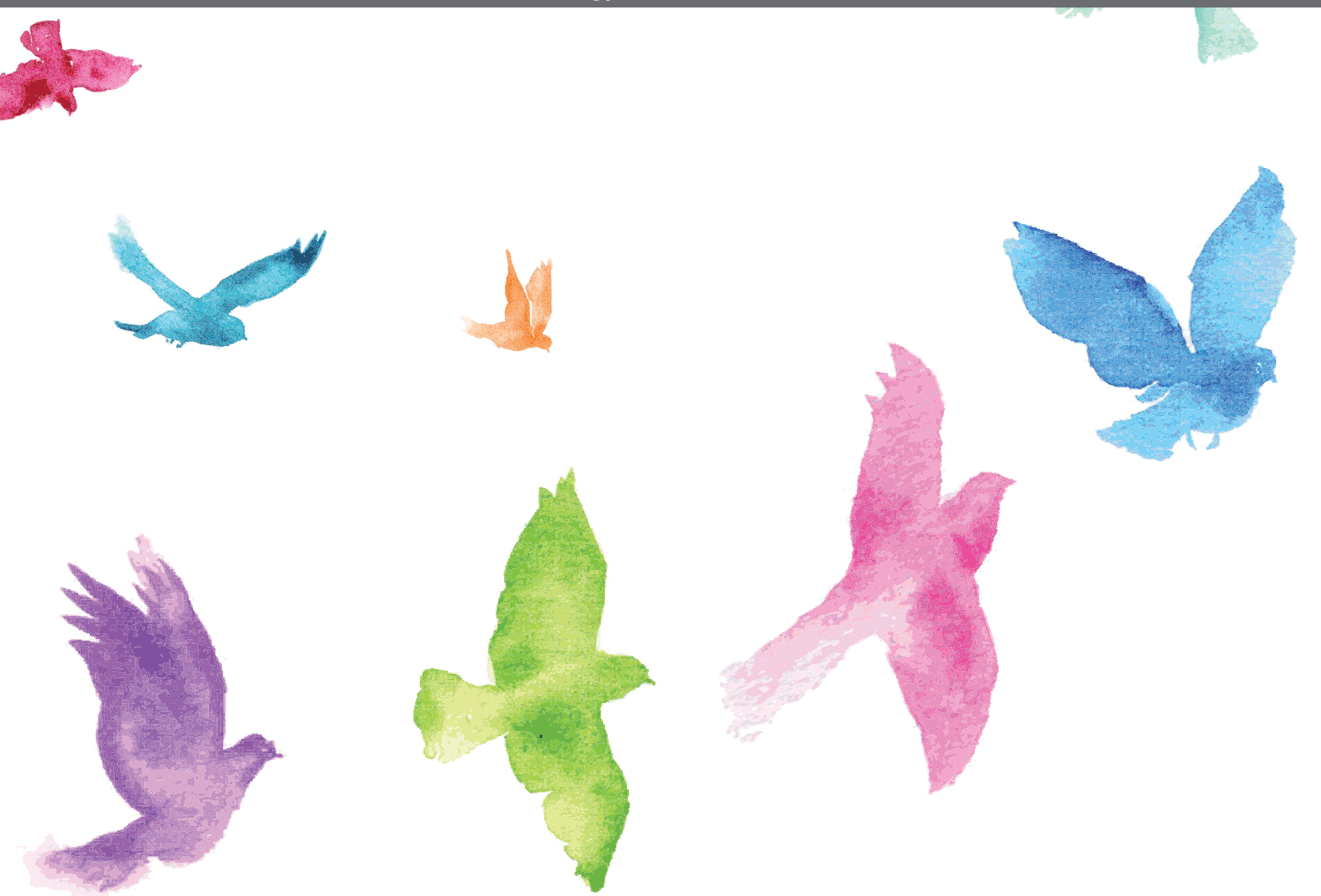
THEORETICAL APPROACHES TO COMMUNITY ECOLOGY

EDITED BY: Luís Borda-de-Água, Paulo A. V. Borges and John Maxwell Halley

PUBLISHED IN: Frontiers in Ecology and Evolution



frontiers Research Topics

frontiers

ISSN 1664-8714
ISBN 978-2-88974-424-4
DOI 10.3389/978-2-88974-424-4

About Frontiers

Frontiers is more than just an open-access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers Journal Series

The Frontiers Journal Series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the Frontiers Journal Series operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to Quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews.
Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the Frontiers Journals Series: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area! Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers Editorial Office: frontiersin.org/about/contact

THEORETICAL APPROACHES TO COMMUNITY ECOLOGY

Topic Editors:
Luís Borda-de-Água, Centro de Investigacao em Biodiversidade e Recursos Geneticos (CIBIO-InBIO), Portugal
Paulo A. V. Borges, University of the Azores, Portugal
John Maxwell Halley, University of Ioannina, Greece

Citation: Borda-de-Água, L., Borges, P. A. V., Halley, J. M., eds. (2022). Theoretical Approaches to Community Ecology. Lausanne: Frontiers Media SA. doi: 10.3389/978-2-88974-424-4

Table of Contents

frontiers in Ecology and Evolution

EDITORIAL
published: 11 January 2022
doi: 10.3389/fevo.2021.824432

Editorial: Theoretical Approaches to Community Ecology

Luís Borda-de-Água[1,2,3]*, *Paulo A. V. Borges*[4] *and John M. Halley*[5]

[1] *Centro de Investigação em Biodiversidade e Recursos Genéticos (CIBIO), InBIO Laboratório Associado, Universidade do Porto, Vairão, Portugal,* [2] *Centro de Investigação em Biodiversidade e Recursos Genéticos (CIBIO), InBIO Laboratório Associado, Instituto Superior de Agronomia, Universidade de Lisboa, Lisboa, Portugal,* [3] *BIOPOLIS Program in Genomics, Biodiversity and Land Planning, Centro de Investigação em Biodiversidade e Recursos Genéticos (CIBIO), Vairão, Portugal,* [4] *cE3c—Centre for Ecology, Evolution and Environmental Changes/Azorean Biodiversity Group, Faculty of Agriculture and Environment, Department of Environmental Sciences and Engineering, Universidade dos Açores, Angra Do Heroísmo, Portugal,* [5] *BET-EcoLab, Department of Biological Applications and Technology, University of Ioannina, Ioannina, Greece*

Keywords: analytical methods, biodiversity, community ecology, patterns and processes, theoretical ecology

Editorial on the Research Topic

Theoretical Approaches to Community Ecology

Our understanding of the processes determining community diversity and how to interpret community patterns is a work in progress. Since the landmark works of MacArthur and Wilson (1963, 1967) on Island Biogeography nearly 60 years ago, we have seen many exciting theoretical advances in community ecology. But is it enough? The last two decades have seen remarkable growth in the availability of data, especially biodiversity data, via satellite technology, next-generation sequencing and an explosion of open access online databases. At the same time, the expanding rate and complexity of key drivers of biodiversity loss, such as habitat fragmentation, invasive species and climatic change brings unprecedented threats to ecological communities (Borges et al., 2019). The conjunction of these changes is creating both new problems and new applications for Community Ecology. Due to this combination of recent challenges and opportunities, this is an especially appropriate moment to consider the role of theory in the development of new approaches to Community Ecology, of biodiversity patterns and the processes that influence them.

Theory is an important part of Community Ecology. Ecological communities are so complex that even a small part of any of them presents a world in itself, with diversity and complexity beyond the scope of any model to describe fully (Cardoso et al. in this special issue). So, every community theory is incomplete (Halley and Pimm in this special issue). Nevertheless, even crude theories provide essential guidance towards fuller understanding. Theory with a strong mathematical basis is a characteristic of any mature science. It describes in a mathematical way the first principles and mechanisms by which processes operate and interact, and this enables the science to renew itself when exposed to new types of data or when confronted with new types of problems. Such theory is thus necessary for ecology in general, and for Community Ecology in particular, because it enables Community Ecology to develop as a predictive science. In this way, Community Ecology can itself evolve and provide tools and perspectives that will assist our societies to respond to current and future environmental challenges.

This Research Topic explores the state-of-the-art in theoretical Community Ecology and reveals significant new research on the patterns and processes that drive communities. For this issue, we brought together researchers from ecology but also researchers with different backgrounds with common interests in Ecology, such as Physics, Mathematics and Computer Science. Thus, these articles present insights into Community Ecology from various perspectives, with different theoretical and methodological approaches, covering studies on communities at all scales, from

OPEN ACCESS

Edited and reviewed by:
György Barabás,
Linköping University, Sweden

***Correspondence:**
Luís Borda-de-Água
lbagua@gmail.com

Specialty section:
This article was submitted to Models in Ecology and Evolution, a section of the journal Frontiers in Ecology and Evolution

Received: *29 November 2021*
Accepted: *20 December 2021*
Published: *11 January 2022*

Citation:
Borda-de-Água L, Borges PAV and Halley JM (2022) Editorial: Theoretical Approaches to Community Ecology. Front. Ecol. Evol. 9:824432. doi: 10.3389/fevo.2021.824432

individuals to ecosystems, as well as multi-scale approaches. The papers were organised into five different categories: "community modelling," "patterns of species abundance and rarity," "land-use change and habitat fragmentation," "modelling multi-species dynamics," and "speciation and evolution."

The first two articles in the section Community modelling are clearly methodological. Cardoso et al. brings a novel approach using artificial intelligence and machine learning to detect patterns and predict outcomes in ways that often resemble human reasoning. They used a particular method called symbolic regression that allows the evolution of human-interpretable formulas to explain natural laws. For example, Cardoso et al. were able to find a new general model that explains why some islands have more species than others, being able to simulate the GDM model (Whittaker et al., 2008). This publication shows that evolving free-form equations purely from data, often without prior human inference or hypotheses, may represent a powerful tool in the arsenal of a discipline as complex as ecology. These authors also applied their approach successfully to other models in community ecology, testing the robustness of symbolic regression at extreme levels of noise when searching for the species-area relationship and finding new formulas for species richness estimators. The next methodological advance comes from the manuscript of Carvalho and Cardoso in which a new framework is proposed to decompose overall differentiation among hypervolumes into two distinct components: replacement of space (niche shifts) and net differences between space amplitudes (niche contraction/expansion processes). The authors demonstrate that it is possible to disentangle different processes underlying niche partitioning between coexisting species, offering novel insights to understand the drivers of niche partitioning and allowing the address of new niche- and trait-based questions in community ecology.

Eldijk et al. propose a simple eco-evolutionary model, the Community-Wide Rescue (CWR) model, in which when a community faces environmental deterioration, each species within the community is forced to undergo adaptation or become extinct. The authors show that when implementing their neutral individual based simulation model rare species become extinct very rapidly, which raises concerns with the ongoing global changes, as many ecosystems face irreversible human induced environmental change.

Finally in this topic, Halley and Pimm developed the niche-hypervolume concept of the community into a powerful model of community dynamics. The authors lay out a framework to answer similar questions to those addressed by Hubbell's neutral theory, but with a niche-based perspective and explore the behaviour using both mathematical analysis and computer simulations. This dynamic model reproduces key patterns of communities, such as lognormal species abundance distributions, lognormality of populations in the time domain, $1/f$-noise population fluctuations, extinction debt and logarithmic species-time curves. It also provides a powerful framework to explore significant ideas in ecology, such as community turnover and the drift of ecological communities into evolutionary time.

Several articles in the section Patterns of Species abundance and rarity used the species abundance distribution (SAD) as their main tool to describe and understand biodiversity patterns. For instance, the main objective of Tsafack et al. was to assess the influence of environmental conditions on the SADs of carabid beetles. They studied communities from three grassland ecosystems (desert, typical and meadow steppes) in China, and fitted the SADs using a variety of models: the meta-community zero sum multinomial, the lognormal, Fisher's logseries, and unimodal and multimodal gambin distributions. Importantly, they found that the parameters of the distributions were a function of the scale of analysis. Overall, the communities at different scales were equally well fitted by the meta-community zero-sum multinomial and Fisher's logseries.

Also realising the importance of looking at communities at different scales, Alirezazadeh et al. analysed the functional diversity of arthropod communities in the Azores. The main argument for using functional diversity was that some species play similar roles and therefore grouping species according to their characteristics may be more relevant to understanding community dynamics. However, the authors moved away from the typical approach of looking at only one scale and focused on characterising patterns of functional diversity across spatial scales, similar to what is routinely done with taxonomic diversity, such as the species-area relationship (SAR). In addition to studying the equivalent of SAR to functional groups, Alirezazadeh et al. applied, and further developed, methods previously introduced by some of the authors to study the scaling properties of SADs (e.g., Borda-de-Água et al., 2012). They found that scales of functional diversity are quantitatively different from taxonomic diversity and discuss the implications of this finding if the dynamics of communities are driven primarily by the characteristics of their functional diversity.

The importance of studying the shape of the SADs is also emphasised by Antão et al. These authors examined 11 large-scale datasets of different taxa and compared the fit provided by different models at different scales. They used the logseries, the lognormal and a multimodal distribution. A major result of this work is that multimodality is present for a wide range of scales, and concluded that two of the major theories, namely the Neutral Theory of Biodiversity and Biogeography (Hubbell, 2001) and the Maximum Entropy Theory (Harte, 2011) of ecology are unable to justify the variability of SAD shapes observed. The authors conclude that critical developments are still required in order to fully understand the shapes of the SADs.

Finally in this section, Brush and Harte emphasise the importance of relating observed patterns with the underlying mechanisms and processes. To this end, the authors developed a model that integrates pattern, by spatially distributing individuals of a species based on the rules of the Maximum Entropy Theory of Ecology (Harte, 2011), and mechanism, through the relationship between the probability of an individual to die and the abundance of its species. In order to compare the results of the model with real data, they used two data sets, one with information on the abundance of tropical tree and shrub species in a 50 ha plot in Barro Colorado Island and another with information on plant abundance on 64 m^2 serpentine grassland

plot. The authors conclude that their model predicts that an increase in intraspecific negative density dependence is related to more random spatial patterning of species spatial distributions.

One key driver of the erosion of biodiversity is habitat fragmentation and degradation (Borges et al., 2019). Three articles in this special issue, in the section Land-use change and habitat fragmentation, deal with this important theme. Staude et al. used a space-for-time substitution approach, surveying bird communities in native grassland sites in 31 regions with differing levels of habitat conversion. Their results support predictions of the theory of the ideal free distribution, and suggests that native habitat remnants are very important temporary refugia for specialist bird species. The results obtained can partly explain the segregation of habitat specialist and generalist birds observed in many agricultural landscapes, in which higher abundance of specialists are found in native grassland patches.

Changing from empirical to a modelling approach Chetcuti et al. used a multi-landcover, landscape-scale, individual-based model to show that the gamma diversity of species for whom the focal land cover is a suitable habitat, declined under fragmentation *per se*. Interestingly, specialists were unaffected by fragmentation *per se*. These authors also found that that beta-diversity and gamma-diversity increased overall even without differences among species in habitat specialisations. The findings of Chetcuti et al. will help to inform the fragmentation *per se* debate, showing that effects on biodiversity can be negative or positive, depending on species' competitive abilities and dependency on the fragmented land cover.

Finally, Wayman et al. used a generalised dissimilarity modelling approach to demonstrate that there is moderate and unique proportion of the variance explained by geographical distance *per se*, which could highlight the role of dispersal limitation in community dissimilarity. Moreover, they observed a key role for environmental filtering both for taxonomic nestedness and functional nestedness, with important conservation implications in the face of a warming climate and future land use change.

Communities of organisms are made up of populations of individual species. Thus, all community dynamics ultimately rest on the dynamics of populations. For example, the rise and fall of species richness in a community may be seen as the number of species in a community having numbers greater than some threshold. This special issue has three such papers, in the section Modelling multi-species dynamics.

Seto and Iwasa explored how thermodynamics illuminates population interactions in microbial communities. Many traditional model populations, especially for microbial systems, can be divided into producers and consumers. However, most investigations are focused on systems based around solar energy, while many communities are based around other forms of energy, such as hot springs, hydrothermal vent systems or even gut microbes. Does it really matter, which energy is the basis of a community? Seto and Iwasa argue that it does and they introduce theoretical approaches based on differential equations that incorporate thermodynamics to highlight characteristic interactions between species in the microbial community. In contrast to light-dependent producers, which compete with one another for light, producers using Gibbs free energy (chemical energy) can have cooperative interactions with each other through the effects of the relative quantities of products and reactants on the available chemical energy. This is an exciting insight into how communities work at the microbial level with non-solar inputs.

Akimenko also used a resource-consumer model for interacting populations. However, in this case, the system of coupled differential equations was generalised in the direction of seeing what are the effects of age-structure and delays in the basic equations. The model of consumer population dynamics was described by a delayed transport equation, and the dynamics of resource patches are described by ODEs with saturated intake rate. The delay described the digestion period of generalist consumers and is included in the calorie intake rate, which impacts the consumer's fertility and mortality and affects what happens if the behaviour of the food resource species changes when consumer population grows. Usually, the incorporation of delays into systems of consumer-resource equations, causes increasing instability. However, here the delay caused by the digestion period of a generalist consumer did not cause local asymptotical instabilities of consumer population at the semi-trivial and non-trivial equilibria. These theoretical results (supported by simulations) also have applications to metapopulation dynamics as well as to multispecies models.

He et al. again focus on a microbial system but this time it is one based on parasitism. Their study of frequency-dependent competition between strains of *Plasmodium falciparum* also belongs in the fields of evolution and the epidemiology of malaria. The focus here is on how populations of *P. falciparum* exhibit a vast diversity of the *var* gene encoding its major surface antigen, with each parasite comprising multiple copies from this diverse gene pool. The authors used a combination of stochastic agent-based models, principal component analysis and network analyses to test various theoretical results, where assembly of local parasite diversity occurs under frequency-dependent selection.

Both papers in the section Speciation and evolution focus on speciation. Interestingly, both concluded that dispersal ability plays a major role in determining the characteristics of the observed phylogenetic trees, their regions of clustering, and the associated speciation rates.

Tao et al. address the question of why some lineages have higher speciation rates than others when colonising a given region. In order to study the radiation of different lineages the authors developed a spatially explicit neutral model with protracted speciation; importantly, this model was free from the complexity of explicit niches. The main parameters controlling the radiation sizes were the geographic isolation and dispersal ability. The authors found an intermediate dispersal value where lineages may differ considerable in radiation size: if for stochastic reasons dispersal is very low, this leads to rapid speciation events that were named a "radiation cascade." Tao et al. suggest that such radiation events may be common, and not necessarily only a feature of neutral models, and they may occur without the need to invoke niche differentiation.

Also using a theoretical model, Kallimanis et al. studied whether metrics of phylogenetic clustering are able to detect

a taxon speciation history. In this study, the feature that led to important outcomes was dispersal. When dispersal was not present, then only when the incipient species was present in the speciation region could some metrics identify regions of high speciation. On the other hand, when dispersal was present and speciation occurred in a spatially continuous way, then there was a connexion between the speciation rate and the clustering of the phylogenetic tree; if speciation happened randomly in space, then such a relationship was not observed. The authors conclude that phylogenetic clustering can only be an indicator of speciation under a limited set of conditions.

Altogether, the articles in this special issue show that theory community ecology is a vibrant area of research, employing a wide variety of approaches. It is important that continuing theoretical efforts in these directions take advantage of concurrent and increasingly rapid progress in data collection of data. We believe that the papers of this special issue will foster such progress and stimulate further advances in theoretical approaches in community ecology and collaboration among researchers using different methodological approaches.

AUTHOR CONTRIBUTIONS

All authors have made substantial contributions to the Editorial and in organising and editing the Research Topic.

FUNDING

LB-d-A and PAVB were financed through Portuguese national funds through FCT—Fundação para a Ciência e a Tecnologia, I.P. (LB-d-A: under the Norma Transitória-DL57/2016/CP1440/CT0022; PAVB: UIDB/00329/2020-2024).

REFERENCES

Borda-de-Água, L., Borges, P. A. V., Hubbell, S. P., and Pereira, H. M. (2012). Spatial scaling of species abundance distributions. *Ecography* 35, 549–556. doi: 10.1111/j.1600-0587.2011.07128.x

Borges, P. A. V., Gabriel, R., and Fattorini, S. (2019). "Biodiversity erosion: causes and consequences," in *Life on Land. Encyclopedia of the UN Sustainable Development Goals*, eds W. Leal Filho, A. Azul, L. Brandli, P. Özuyar, and T. Wall (Cham: The Springer Nature), 1–10.

Harte, J. (2011). *Maximum Entropy and Ecology: A Theory of Abundance, Distribution, and Energetics*. Oxford: Oxford University Press.

Hubbell, S. P. (2001). *The Unified Neutral Theory of Biodiversity and Biogeography (MPB-32)*. Princeton, NJ: Princeton University Press.

MacArthur, R. H., and Wilson, E. O. (1963). An equilibrium theory of insular zoogeography. *Evolution* 17, 373–387. doi: 10.1111/j.1558-5646.1963.tb03295.x

MacArthur, R. H., and Wilson, E. O. (1967). *The Theory of Island Biogeography*. Princeton, NJ: Princeton University Press.

Whittaker, R. J., Triantis, K. A., and Ladle, R. J. (2008). A general dynamic theory of oceanic island biogeography. *J. Biogeogr.* 35, 977–994. doi: 10.1111/j.1365-2699.2008.01892.x

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Ecology and Evolution

METHODS
published: 30 July 2020
doi: 10.3389/fevo.2020.00243

Decomposing the Causes for Niche Differentiation Between Species Using Hypervolumes

José Carlos Carvalho[1,2,3†] *and Pedro Cardoso*[2,3*†]

[1] Department of Biology, CBMA – Centre of Molecular and Environmental Biology, University of Minho, Braga, Portugal, [2] Azorean Biodiversity Group, CE3C – Centre for Ecology, Evolution and Environmental Changes, Universidade dos Açores, Angra do Heroísmo, Portugal, [3] LIBRe – Laboratory for Integrative Biodiversity Research, Finnish Museum of Natural History, University of Helsinki, Helsinki, Finland

Hutchinson's n-dimensional hypervolume concept holds a central role across different fields of ecology and evolution. The question of the amount of hypervolume overlap and differentiation between species is of great interest to understand the processes that drive niche dynamics, competitive interactions and, ultimately, community assembly. A framework is proposed to decompose overall differentiation among hypervolumes into two distinct components: niche shifts and niche contraction/expansion processes. Niche shift corresponds to the replacement of space between the hypervolumes occupied by two species, whereas niche contraction/expansion processes correspond to net differences between the amount of space enclosed by each hypervolume. A procedure to implement non-continuous trait data in the estimation of *n*-dimensional hypervolumes is proposed. Hypervolumes were constructed for three Darwin' finches, *Geospiza conirostris*, *Geospiza magnirostris,* and *Geospiza difficilis* using intraspecific trait data. Results showed that significant niche shifts, not niche contraction, occurred between *G. conirostris* and *G. magnirostris* in Genovesa island, where they live in sympatry. This means that *G. conirostris* occupied a different niche space and not a reduced space on Genovesa. *G. difficilis* was well differentiated from the other two species. The proposed framework allows disentangling different processes underlying niche partitioning between coexisting species. This framework offers novel insights to understand the drivers of niche partitioning strategies among coexisting species.

Keywords: fundamental niche, hypervolume, morphospace, niche contraction, niche shift, realized niche

OPEN ACCESS

Edited by:
John Maxwell Halley, University of Ioannina, Greece

Reviewed by:
Stefano Mammola, National Research Council (CNR), Italy
Harald Bugmann, ETH Zürich, Switzerland
Giorgos Kokkoris, School of Environment, University of the Aegean, Greece

****Correspondence:***
Pedro Cardoso
pedro.cardoso@helsinki.fi

† These authors have contributed equally to this work

Specialty section:
This article was submitted to Models in Ecology and Evolution, a section of the journal Frontiers in Ecology and Evolution

Received: *22 January 2020*
Accepted: *06 July 2020*
Published: *30 July 2020*

Citation:
Carvalho JC and Cardoso P (2020) Decomposing the Causes for Niche Differentiation Between Species Using Hypervolumes. Front. Ecol. Evol. 8:243. doi: 10.3389/fevo.2020.00243

INTRODUCTION

The concept of species fundamental niche proposed by Hutchinson (1957) is at the roots of many ecological and evolutionary theories (e.g., Holt, 2009; Blonder, 2018). Hutchinson formalized this concept as a multidimensional hypervolume, defined by a set of *n* independent variables that represent biologically relevant axes. According to his perspective, each point within this *n*-dimensional space corresponds to a possible state of the environment permitting the species to survive.

Hutchinsonian niches can be estimated by quantifying the functional trait hypervolume occupied by the individuals of a given species. In essence, a set of functional traits, representing major axes of ecological strategies, are selected. Then, the hypervolume is constructed to estimate the morphospace occupied by the species. The hypervolume, representing the intraspecific trait variability, is commonly interpreted as the realized niche of the species (e.g., Pigot et al., 2016).

The question of the amount of morphospace overlap and differentiation between species is of great interest to understand the processes that drive niche dynamics, competitive interactions and, ultimately, community assembly (e.g., Ricklefs and Cox, 1977; Stubs and Wilson, 2004; Kraft et al., 2008). The hypervolume overlap corresponds to the shared space between two species, whilst hypervolume differentiation corresponds to the sum of the unique fractions of space belonging to each species. Niche differentiation between species may be caused by two distinct processes: niche shifts and niche contraction (or expansion). Niche shifts are determined by the replacement of niche space between species. In this case, species change their niche position in order to occupy a different niche space and reduce overlap. For example, changes of morphological traits related to food acquisition have been interpreted as a strategy to reduce niche overlap and avoid competition between species (Huey et al., 1974). On the other hand, niche contraction (or expansion) processes are determined by differences in niche breadth (e.g., Pulliam, 1986). Niche contraction occurs when a species reduces its niche breadth. For example, when faced with more intense competition from another species, many organisms restrict their utilization of shared microhabitats and/or other resources (Pianka, 2000). This might give them advantage exploring these resources over other, generalist, species, as they adapt to explore them to the fullest. Niche expansion occurs when a species augments its niche breadth when presented with ecological opportunity (McCormack and Smith, 2008). Compelling examples of niche expansion are given by studies in islands where species were released from competition (e.g., Lister, 1976). Thus, to understand the factors that drive niche differentiation between species it is important to disentangling niche shifts from niche expansion (or contraction) processes.

In this paper, we propose an integrated framework for partitioning hypervolumes via pairwise comparisons into niche shifts and niche contraction (or expansion). We also provide guidelines to incorporate non-continuous trait data into hypervolume estimation. We present a case study to illustrate how these concepts can be applied to decompose the causes for niche differentiation between species.

MATERIALS AND METHODS

Hypervolume Partitioning

To decompose hypervolume differentiation into different fractions, we use the set theory notation, in the manner of Hutchinson. Given a pairwise hypervolume system composed by hypervolumes hi and hj, the total space occupied by them is given by their union ($hi \cup hj$). Total pairwise hypervolume corresponds to the sum of hypervolume overlap given by the interception ($hi \cap hj$) plus hypervolume differentiation, denoted by $hi \Delta hj$.

Thus, a pairwise hypervolume system can be partitioned according to the equation:

$$hi \cup hj = hi \Delta hj + hi \cap hj \quad (1)$$

The differentiation between the hypervolumes hi and hj ($hi \Delta hj$) corresponds to the sum of their unique fractions:

$$hi \Delta hj = hi \setminus hj + hj \setminus hi \quad (2)$$

This can also be obtained by:

$$hi \Delta hj = hi \cup hj - hi \cap hj. \quad (3)$$

Total differentiation ($hi \Delta hj$) can occur as a consequence of two distinct processes: (i) the differentiation that results from the replacement of space between hypervolumes; and, (ii) differentiation that results from the net difference between the amounts of space enclosed by each hypervolume.

The net difference between both hypervolumes is given by $|hi \setminus hj - hj \setminus hi|$. The replacement component can be obtained by subtracting the difference fraction from total differentiation, using Eq. 2:

$$hi \setminus hj + hj \setminus hi - |hi \setminus hj - hj \setminus hi| \quad (4)$$

This expression can be written in the form:

$$\begin{aligned} &\max(hi \setminus hj,\ hj \setminus hi) + \min(hi \setminus hj,\ hj \setminus hi) - \\ &[\max(hi \setminus hj,\ hj \setminus hi) - \min(hi \setminus hj,\ hj \setminus hi)]. \end{aligned} \quad (5)$$

By simplifying the expression, we obtain the replacement component: $2\min(hi \setminus hj, hj \setminus hi)$.

Thus, the total differentiation component can be additively partitioned in two components:

$$hi \Delta hj = 2\min(hi \setminus hj,\ hj \setminus hi) + |hi \setminus hj - hj \setminus hi|. \quad (6)$$

A total partitioning of $hi \cup hj$ can be obtained by the equation (see **Figure 1**):

$$hi \cup hj = hi \cap hj + 2\min(hi \setminus hj,\ hj \setminus hi) + |hi \setminus hj - hj \setminus hi|. \quad (7)$$

The terms in Eqs 4 and 5 can be scaled in relation to the total space occupied by the hypervolume pairwise system, thus obtaining the following equivalency:

$$\begin{aligned} 1 = (hi \cap hj + hi \Delta hj)/hi \cup hj = [hi \cap hj + 2\min \\ (hi \setminus hj,\ hj \setminus hi) + |hi \setminus hj - hj \setminus hi|]/hi \cup hj. \end{aligned} \quad (8)$$

This equivalency can be summarized in the equation:

$$1 = H_{overlap} + H\delta_{total} = H_{overlap} + H\delta_{repl} + H\delta_{diff} \quad (9)$$

where, $H_{overlap}$ refers to the proportion of overlap between hypervolumes (similarity), $H\delta_{total}$ refers to total differentiation (δ) between both hypervolumes, $H\delta_{repl}$ corresponds to the fraction of differentiation that results from the replacement of space between hypervolumes and $H\delta_{diff}$ is the fraction of

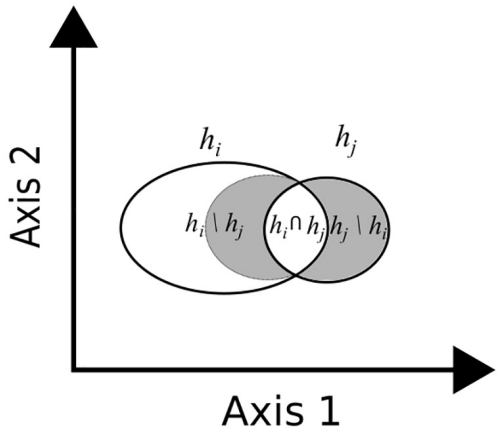


FIGURE 1 | Diagram showing hypervolume partitioning in two dimensions. The overlap between hypervolumes *hi* and *hj* is given by the interception (*hi* ∩ *hj*) and the differentiation (*hi* Δ *hj*) is given by the sum of their unique fractions, *hi* *hj* and *hj* *hi*. Note that: *hi* Δ *hj* = 2min(*hi* *hj*, *hj* *hi*) + |*hi* *hj* – *hj* *hi*|, where 2min(*hi* *hj*, *hj* *hi*) corresponds to the differentiation due to the replacement of space (gray shaded area) and |*hi* *hj* – *hj* *hi*| refers to the net difference between both hypervolumes (white area in the *hi* *hj* space).

differentiation that results from the net differences between the spaces occupied by both hypervolumes. The relative importance of each component could be expressed in terms of percentage of total differentiation by calculating the fractions $H\delta_{repl}/H\delta_{total}$ and $H\delta_{diff}/H\delta_{total}$.

Assessing Intraspecific Trait Variability and Niche Partitioning

Here, we advocate that intraspecific trait variability can be used to estimate niche parameters. The underlying assumption of a trait-based approach is that traits reflect species adaptations to the environment (Diaz and Cabido, 2001) and, hence, should be a useful tool to quantify species' niches (Violle and Jiang, 2009). In other words, we are assuming that the morphospace occupied by each species is a surrogate for its realized niche.

Consider **N** individuals belonging to a given species **s** and measure a set of **T** traits for each individual. The matrix **N** × **T** can be used to construct the hypervolume, which is assumed to describe the realized niche of the species. Therefore, the niche breadth of species **s** is given by the extent of the hypervolume.

Obviously, hypervolumes can be calculated for a set of species (**S**). The set operations can be applied to hypervolumes using Eqs 1–7 for each species pair. This allows obtaining a pairwise partitioning of the niche species pairs. Therefore, the terms in Eq. 7 mean: $H_{overlap}$ = niche overlap between two species; $H\delta_{total}$ = total niche differentiation between two species; $H\delta_{repl}$ = differentiation due to the replacement of niche space; $H\delta_{diff}$ = differentiation due to differences of niche breadth between species. Thus, higher values of $H\delta_{repl}$ are indicative of niche shifts between species, whereas higher values of $H\delta_{diff}$ are indicative of niche contraction/expansion of one of the species in relation to the other.

Incorporating Different Data Types in Hypervolume Estimation

A limitation to the current hypervolume algorithms, is that they require datasets with continuous variables (Blonder et al., 2014; Blonder, 2018). We suggest that this limitation may be overcome by applying already available statistical methods to the matrix of species × traits (**S** × **T**), prior to hypervolume estimation.

Multivariate techniques, specially developed to deal with mixed data, may be used to ordinate the **S** × **T** matrix (e.g., Hill and Smith, 1976; Kiers, 1994). These ordination techniques are similar to a principal component analysis, allowing to extract several orthogonal axes (components) representing the variation of the **S** × **T** matrix (e.g., PCAMIX, Chavent et al., 2017). These axes can be interpreted in relation to the original variables (traits), by examining the matrix of eigenvectors in a similar way to a PCA. Then, the resulting axes are used as the new variables to compute hypervolumes (**Figure 2**). If the aim is to include all the variation of the **S** × **T** matrix in the analyses, one should retain all the axes. However, most probably some of these components represent little variation and are difficult to interpret. Therefore, a more parsimonious strategy would be to select only those axes that provide a clear interpretable result.

The distance-based framework proposed by Laliberté and Legendre (2010) is an alternative to consider. Briefly, this approach is based on three steps: (i) an appropriate distance measure (Gower dissimilarity measure; Gower, 1971; Podani, 1999; Legendre and Legendre, 2012) is computed from **S** × **T** matrix; (ii) this distance matrix is analyzed through principal coordinate analysis (PCoA); and, (3) the resulting PCoA axes are used as the new variables to compute hypervolumes (**Figure 2**). A comprehensive description of PCoA is given in Legendre and Legendre (2012). Nevertheless, PCoA may be difficult to implement in the context of hypervolume estimation, since it tends to produce n-1 eigenvectors (*n* = the number of observations). This means that for some ecological datasets, the calculation of hypervolumes based on all PCoA axes may be impractical or even impossible. Therefore, one may choose the first few axes of the PCoA, provided that they represent a large propotion of variation of the original data set. Another possibility is to use non-metric multidimensional scaling (NMDS), instead of PCoA, to analyze the Gower's distance matrix. In NMDS one could choose a small number of axes to represent the underlying dissimilarity between observations. The NMDS tries to find, interactively, the optimal configuration in the reduced space that maximizes the observation distances in the configuration and the original distances (for a comprehensive review see Legendre and Legendre, 2012). Frequently, a small number of axes is sufficient to represent the original distances.

All techniques allow to extract orthogonal axes, which eliminates the potential correlation among trait variables, a pre-requisite of hypervolume methods. Independently of which technique is used to transform the **S** × **T** matrix, an additional point to consider is giving weights to trait variables, because a single categorical trait can be coded by several dummy variables. Also, prior weights can be given to trait variables, allowing to differentiate traits that contribute more to the

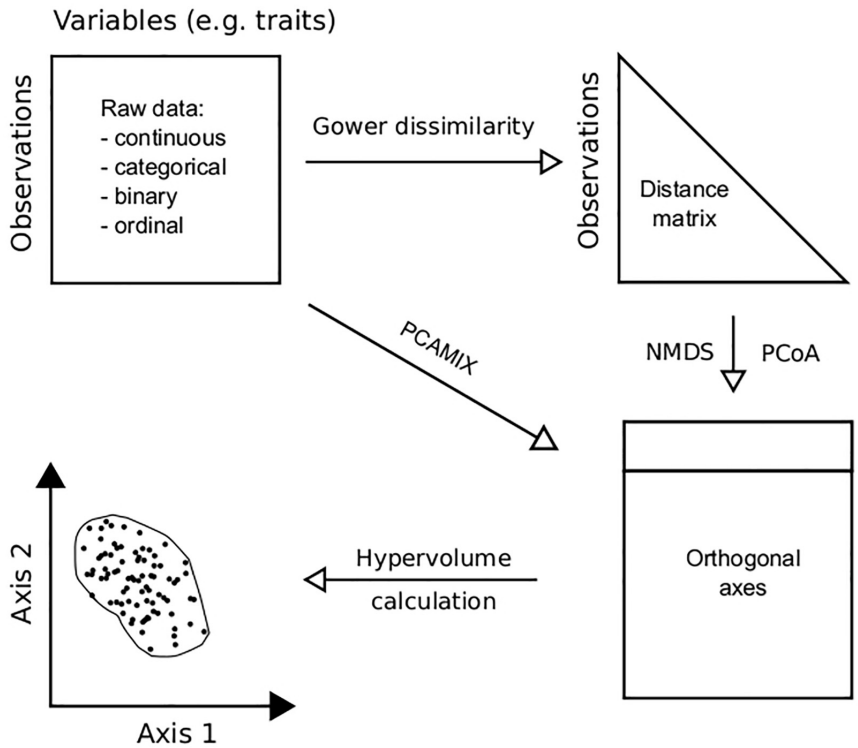


FIGURE 2 | Diagram showing the construction of axes to be used in hypervolume estimation. The axes are obtained by ordination for mixed data types (PCAMIX; see Chavent et al., 2017) or by performing a PCoA or NMDS carried on a Gower dissimilarity matrix of individuals X traits.

biological performance of the species. Contrary to PCAMIX, the information of each trait is lost during the PCoA and NMDS, because the $\mathbf{S} \times \mathbf{T}$ matrix is transformed into a dissimilarity matrix of individuals, prior to the calculation of hypervolumes. Thus, the interpretation of these shapes is relative to the axes of the PCoA and NMDS and not in relation to the traits. Therefore, questions about which traits contribute more or less to a given hypervolume parameter are difficult to address.

Case Study

We present a demonstration analysis of hypervolume partitioning for three species of Darwin's finches: *Geospiza conirostris*, *Geospiza magnirostris* and *G. difficilis*. *G. conirostris* and *G. magnirostris* species inhabit Island Genovesa, but only the former is present in Island Española. A well-known hypothesis for the two species is that changes of morphological traits induced by interspecific competition occurred on Genovesa and, therefore, they should have evolved to occupy dissimilar niches (Lack, 1945, 1947; Grant and Grant, 1982). Therefore, we hypothesize that niche differentiation between these species was determined by the replacement of space (niche shift) and not the contraction/expansion of niches. *G. difficilis* inhabits Island Genovesa, however it is much shorter than *G. conirostris* and *G. magnirostris*. Hence, we expect that *G. difficilis* occupies a well differentiated niche space from the other two species. We tested these hypotheses on Genovesa and Española islands with morphometric data collected by Lack (1945, 1947): WingL, wing length; BeakH, Beak height; UbeakL, upper beak length and N-UBkL, nostril upper beak length. Data is available from Dryad Digital Repository: < https://doi.org/10.5061/dryad.150 >. The original measurements (in mm) were log10-transformed prior to the construction of hypervolumes. Hypervolumes for each species were calculated using a gaussian kernel density estimator with the hypervolume R package with default parameters (see Blonder et al., 2018 for details). Then, hypervolume decomposition (Eqs 1–7) was carried with the function "kernel.beta" of the BAT package (Cardoso et al., 2015). Hypervolumes are reported in units of SDs to the power of the number of trait dimensions used.

To demonstrate how to estimate hypervolumes with a dataset with mixed type variables (continuous and non-continuous), a second analysis was performed. For the case, the WingL was transformed into a ranked ordered variable with 10 levels. Then, a Gower's dissimilarity matrix was calculated and a NMDS was performed. The NMDS axes were used to estimate the hypervolumes for each species and to carry out hypervolume decomposition.

RESULTS

As hypothesized, hypervolumes calculated using morphometric data for *Geospiza conirostris* and *Geospiza magnirostris* were well differentiated on Genovesa Island (**Figure 3**), the niche overlap

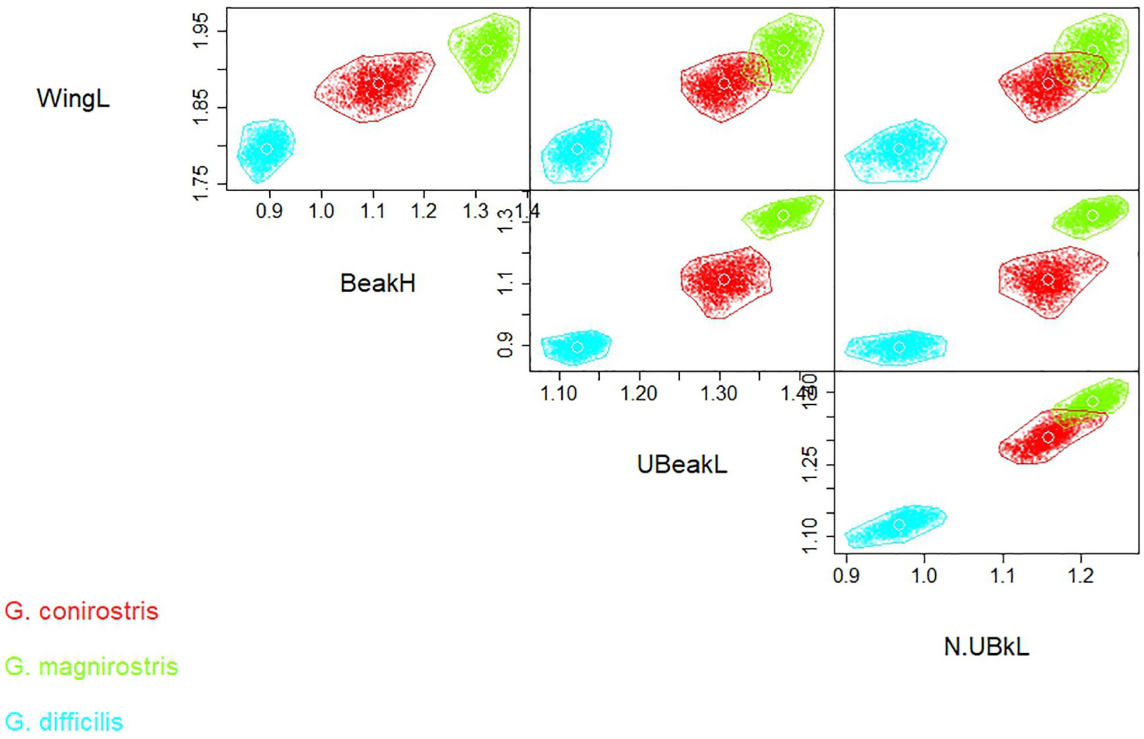


FIGURE 3 | Niche partitioning for three species of Darwin's finches *Geospiza conirostris*, *Geospiza magnirostris,* and *Geospiza difficilis* for Genovesa Island. Hypervolumes shown as 2D projections for all combinations of trait axis (WingL, wing length; BeakH, Beak height; UbeakL, upper beak length and N-UBkL, nostril upper beak length) after log-transformation of the original measurements.

between these species being null ($H_{overlap}$ = 0). By decomposing total differentiation ($H\delta_{total}$ = 1) into replacement and differences between niche breadths, we found that such differentiation was mostly caused by niche replacement ($H\delta_{repl}$ = 0.609) and not contraction/expansion, although the latter was also considerable ($H\delta_{diff}$ = 0.391). By comparing the niche of *G. conirostris* on Genovesa and Española, we found that they overlap slightly ($H_{overlap}$ = 0.246) and the differentiation between both islands was mainly due to the replacement of niche space ($H\delta_{repl}$ = 0.709), not contraction/expansion ($H\delta_{diff}$ = 0.045). Therefore, the replacement component contributed 94% to the total differentiation, whilst differences in niche space contributed only 6% to total differentiation. This result reinforces the interpretation that the presence of *G. magnirostris* on Genovesa Island led to a significant shift (not contraction) of *G. conirostris* niche.

In relation to *Geospiza difficilis*, we found that its niche was highly differentiated from both *G. conirostris* ($H\delta_{total}$ = 1) and *G. magnirostris* ($H\delta_{total}$ = 1). The differentiation between *G. difficilis* and *G. conirostris* was caused, almost in equal proportions, by the replacement of niche space ($H\delta_{repl}$ = 0.511) and differences of niche breadth ($H\delta_{diff}$ = 0.489). However, the replacement of niche space was clearly the dominant component ($H\delta_{repl}$ = 0.879; $H\delta_{diff}$ = 0.121) of the differentiation between *G. difficilis* and *G. magnirostris*.

The analyses carried out with mixed type variables (continuous and non-continuous) are presented in **Supplementary Material**. Results were similar to that obtained using only continuous data.

DISCUSSION

Hutchinson's fundamental niche concept holds a central role across different fields of ecology and evolution (e.g., Holt, 2009; Blonder, 2018). Conceptually, the species niche corresponds to an *n*-dimensional hypervolume enclosing the range of conditions under which the species can survive and reproduce. Niche overlap occurs when species use the same resources or environmental conditions. In this regard, a crucial question is which processes determine the differentiation between species niches and, ultimately, determine species coexistence (MacArthur and Levins, 1967). In this paper, we propose a novel framework to partitioning overall differentiation between hypervolumes into two distinct fractions: niche shifts corresponding to the differentiation that results from the replacement of space between hypervolumes and niche contraction/expansion, corresponding to the differentiation that results from the net differences between the space enclosed by hypervolumes.

We illustrate our framework with a classic dataset (Lack, 1945, 1947). Our results revealed that differentiation between *Geospiza conirostris* and *Geospiza magnirostris* occurred mostly by niche shifts processes and not by the niche contraction of *G. conirostris*

on Island Genovesa, where both species live in sympatry. This means that *G. conirostris* occupied a different niche space and not a reduced space on Genovesa. These results are consistent with the hypothesis proposed by Lack (1947) that morphology of beaks was causally influenced by interspecific competition for food resources. This hypothesis received considerable support by analyzing the food habits of these species (Grant and Grant, 1982). These authors showed that *G. magnirostris* feeds almost entirely on large-hard seeds, whilst *G. conirostris* exploits Opuntia and arthropods on Genovesa. Moreover, on Española Island, where *G. magnirostris* is absent, *G. conirostris* consume large hard seeds. Therefore, it seems that the presence of *G. magnirostris* induces a shift in the diet of *G. conirostris* and, consequently, on its beak morphology.

As expected, our results showed that *Geospiza difficilis* was clearly differentiated from the other two species. However, distinct processes determined the differentiation of *G. difficilis*, from both *G. conirostris* and *G. magnirostris*. In relation to the former, the replacement of space (niche shifts) and differences of niche breadth (contraction or expansion) were both important, whilst in relation to the latter, the replacement component dominated. This result emphasizes the importance of considering the different components of differentiation to understand the processes that determine niche partitioning among species.

The functional roles and the sensitivity to environmental changes of the vast majority of taxa are usually unknown (the so-called Hutchinsonian shortfall; Cardoso et al., 2011). Therefore, a trait-based quantification of species niche can provide a practical way to assess niche parameters for a potentially large number of species, making it possible to understand and predict species niche responses to environmental changes (Violle and Jiang, 2009). Trait-based approaches have been used successfully to integrate functional ecology with community assembly and coexistence theories (e.g., McGill et al., 2006; Ackerly and Cornwell, 2007; Kraft et al., 2008). A major assumption of this approach is that trait dissimilarity is related to decreasing niche overlap among co-existent species (e.g., Stubs and Wilson, 2004), a rationale related to the limiting similarity principle (MacArthur and Levins, 1967). This premise assumes, implicitly, that the trait space is a good surrogate for the niche space. However, linking trait patterns to niche differentiation remains a challenge and, ultimately, depends on which traits were included in the analysis (D'Andrea and Ostling, 2016). Thus, an informed choice of traits should be made, based on their presumed adaptation to the species' ecological niche.

Although, the framework proposed has been tested in the context of functional and community ecology, we advocate that it could be also a useful tool in other areas of ecological niche modeling (e.g., Blonder, 2018; Mammola, 2019). For example, bioclimatic hypervolumes can be constructed and compared between species or within the same species across different geographic areas or time periods. This may allow to identify which processes determine niche changes (shifts or contractions/expansion of niche space) along climatic gradients or in response to climatic changes between time periods.

In conclusion, an integrated framework was proposed to partition overall differentiation between hypervolumes into distinct fractions, replacement of space (niche shifts) and net differences between space amplitudes (niche contraction/expansion processes). This framework allows quantifying the relative importance of each process in determining the differentiation of species niches. In that sense, the method offers novel insights to understand the drivers of niche partitioning strategies among coexistence species. We expect that the methods here presented will allow to address a wider range of niche- and trait-based questions than was possible to date.

DATA AVAILABILITY STATEMENT

The datasets generated for this study are available on request to the corresponding author.

AUTHOR CONTRIBUTIONS

All authors listed have made a substantial, direct and intellectual contribution to the work, and approved it for publication.

FUNDING

JC was supported by the Grant No. MSCA-IF-EF-ST 706482.

ACKNOWLEDGMENTS

We thank the three reviewers for their valuable comments that improved this manuscript considerably.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fevo.2020.00243/full#supplementary-material

REFERENCES

Ackerly, D. D., and Cornwell, W. K. (2007). A trait-based approach to community assembly: partitioning of species trait values into within- and among-community components. *Ecol. Lett.* 10, 135–145. doi: 10.1111/j.1461-0248.2006.01006.x

Blonder, B. (2018). Hypervolume concepts in niche and trait-based ecology. *Ecography* 41, 1441–1455. doi: 10.1111/ecog.03187

Blonder, B., Lamanna, C., Violle, C., and Enquist, B. J. (2014). The n-dimensional hypervolume. *Global Ecol. Biogeogr.* 23, 595–609. doi: 10.1111/geb.12146

Blonder, B., Morrow, C. B., Maitner, B., Harris, D. J., Lamanna, C., Violle, C., et al. (2018). New approaches for delineating n-dimensional hypervolumes. *Methods Ecol. Evol.* 9, 305–319. doi: 10.1111/2041-210x.12865

Cardoso, P., Pekár, S., Jocqué, R., and Coddington, J. A. (2011). Global patterns of guild composition and functional diversity of spiders. *PLoS One* 6:e21710. doi: 10.1371/journal.pone.0021710

Cardoso, P., Rigal, F., and Carvalho, J. C. (2015). BAT—biodiversity assessment Tools, an R package for the measurement and estimation of alpha and beta taxon, phylogenetic and functional diversity. *Methods Ecol. Evol.* 6, 232–236. doi: 10.1111/2041-210x.12310

Chavent, M., Kuentz-Simonet, V., Labenne, A., and Saracco, J. (2017). Multivariate analysis of mixed data: the PCAmixdata R package. arXiv:1411.4911.

D'Andrea, R., and Ostling, A. (2016). Challenges in linking trait patterns to niche differentiation. *Oikos* 125, 1369–1385. doi: 10.1111/oik.02979

Diaz, S., and Cabido, M. (2001). Vive la difference: plant functional diversity matters to ecosystem processes. *Trends Ecol. Evol.* 16, 646–655. doi: 10.1016/s0169-5347(01)02283-2

Gower, J. C. (1971). A general coefficient of similarity and some of its properties. *Biometrics* 27, 857–871.

Grant, B. R., and Grant, P. R. (1982). Niche shifts and competition in darwin's finches: geospiza conirostris and congeners. *Evology* 36, 637–657. doi: 10.1111/j.1558-5646.1982.tb05432.x

Hill, M. O., and Smith, A. J. E. (1976). Principal component analysis of taxonomic data with multi-state discrete characters. *Taxon* 25, 249–255. doi: 10.2307/1219449

Holt, R. D. (2009). Bringing the Hutchinsonian niche into the 21st century: ecological and evolutionary perspectives. *Proc. Natl. Acad. Sci. U.S.A.* 106, 19659–19665. doi: 10.1073/pnas.0905137106

Huey, R. B., Pianka, E., Egan, M., and Coons, L. (1974). Ecological shifts in sympatry: kalahari fossorial lizards (*Typhlosaurus*). *Ecology* 55, 304–316. doi: 10.2307/1935218

Hutchinson, G. E. (1957). Concluding remarks. *Cold Spring Harbor. Symposia Q. Biol.* 22, 415–427.

Kiers, H. A. L. (1994). Simple structure in component analysis techniques for mixtures of qualitative and quantitative variables. *Psychometrika* 56, 197–212. doi: 10.1007/bf02294458

Kraft, N. J. B., Valencia, R., and Ackerly, D. D. (2008). Functional traits and niche-based tree community assembly in an amazonian forest. *Science* 322, 580–582. doi: 10.1126/science.1160662

Lack, D. L. (1945). The Galapagos finches (Geospizinae). *Occ. Pap. Calif. Acad. Sci.* 21, 1–159.

Lack, D. L. (1947). *Darwin's Finches: An Essay on the General Biological Theory of Evolution*. Cambridge, MA: Cambridge University Press.

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

Legendre, P., and Legendre, L. (2012). *Numerical Ecology*. Amsterdam: Elsevier.

Lister, B. C. (1976). The nature of niche expansion in West Indian Anolis lizards I: ecological consequences of reduced competition. *Evolution* 30, 659–676. doi: 10.1111/j.1558-5646.1976.tb00947.x

MacArthur, R. H., and Levins, R. (1967). The limiting similarity, convergence and divergence of coexisting species. *Am. Nat.* 101, 377–385. doi: 10.1086/282505

Mammola, S. (2019). Assessing similarity of n- dimensional hypervolumes: which metric to use? *J. Biogeogr.* 46, 2012–2023. doi: 10.1111/jbi.13618

McCormack, J. E., and Smith, T. B. (2008). Niche expansion leads to small-scale adaptive divergence along an elevation gradient in a medium-sized passerine bird. *Proc. R. Soc. B* 275, 2155–2164. doi: 10.1098/rspb.2008.0470

McGill, B. J., Enquist, B. J., Weiher, E., and Westoby, M. (2006). Rebuilding community ecology from functional traits. *Trends Ecol. Evol.* 21, 178–185. doi: 10.1016/j.tree.2006.02.002

Pianka, E. R. (2000). *Evolutionary Ecology*, <edition>6th Edn. San Francisco, CA: Benjamin-Cummings.

Pigot, A. L., Trisos, C. H., and Tobias, J. A. (2016). Functional traits reveal the expansion and packing of ecological niche space underlying an elevational diversity gradient in passerine birds. *Proc. R. Soc. B* 283:20152013. doi: 10.1098/rspb.2015.2013

Podani, J. (1999). Extending Gower's general coefficient of similarity to ordinal characters. *Taxon* 48, 331–340. doi: 10.2307/1224438

Pulliam, H. R. (1986). Niche expansion and contraction in a variable environment. *Amer. Zool.* 26, 71–79. doi: 10.1093/icb/26.1.71

Ricklefs, R. E., and Cox, G. W. (1977). Morphological similarity and ecological overlap among passerine birds on St. Kitts, British West Indies. *Oikos* 29, 60–66.

Stubs, W. J., and Wilson, J. B. (2004). Evidence for limiting similarity in a sand dune community. *J. Ecol.* 92, 557–567. doi: 10.1111/j.0022-0477.2004.00898.x

Violle, C., and Jiang, L. (2009). Towards a trait-based quantification of species niche. *J. Plant Ecol.* 2, 87–93. doi: 10.1093/jpe/rtp007

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Ecology and Evolution

ORIGINAL RESEARCH
published: 29 October 2020
doi: 10.3389/fevo.2020.552268

Uniting Community Ecology and Evolutionary Rescue Theory: Community-Wide Rescue Leads to a Rapid Loss of Rare Species

Timo J. B. van Eldijk†, Karen Bisschop† and Rampal S. Etienne†*

Groningen Institute for Evolutionary Life Sciences, Faculty of Science and Engineering, University of Groningen, Groningen, Netherlands

OPEN ACCESS

Edited by:
Luís Borda-de-Água, Universidade do Porto, Portugal

Reviewed by:
Masato Yamamichi, The University of Queensland, Australia
Robert Dan Holt, University of Florida, United States

****Correspondence:***
Timo J. B. van Eldijk
t.j.b.van.eldijk@rug.nl

†ORCID:
Timo J. B. van Eldijk
orcid.org/0000-0002-2164-1443
Karen Bisschop
orcid.org/0000-0001-7083-2636
Rampal S. Etienne
orcid.org/0000-0003-2142-7612

Specialty section:
This article was submitted to Models in Ecology and Evolution, a section of the journal Frontiers in Ecology and Evolution

Received: *15 April 2020*
Accepted: *01 October 2020*
Published: *29 October 2020*

Citation:
van Eldijk TJB, Bisschop K and Etienne RS (2020) Uniting Community Ecology and Evolutionary Rescue Theory: Community-Wide Rescue Leads to a Rapid Loss of Rare Species. Front. Ecol. Evol. 8:552268. doi: 10.3389/fevo.2020.552268

Most ecological communities are facing changing environments, particularly due to global change. When migration is impossible, adaptation to these altered environments is necessary to survive. Yet, we have little theoretical understanding how ecological communities respond both ecologically and evolutionarily to such environmental change. Here we introduce a simple eco-evolutionary model, the Community-Wide Rescue (CWR) model, in which a community faces environmental deterioration and each species within the community is forced to undergo adaptation or become extinct. We assume that all species in the community are equivalent except for their initial abundance. This individual based simulation model thus combines community ecology and evolutionary rescue theory. We show that under Community-Wide Rescue a rapid loss of rare species occurs. This loss occurs due to competition and a limited supply of beneficial mutations. The rapid loss of rare species provides a testable prediction regarding the impact of Community-Wide Rescue on species abundance distributions in ecological communities.

Keywords: neutral theory of biodiversity, community rescue, evolutionary rescue, adaptation to environmental change, species abundance distributions, antibiotic resistance, microbial community evolution, extinction of rare species

INTRODUCTION

Many ecosystems face abrupt human-induced environmental change and evolutionary adaptation might be the only way to avoid extinction when migration is difficult (Vitousek et al., 1997; IPCC, 2014). Understanding precisely how ecological communities respond to abruptly changing environments is therefore paramount. This calls for models that predict how an ecological community composed of many different species adapts to such a deteriorated environment (Hoffmann and Sgrò, 2011). Such models of community-wide adaptation are not only relevant from the perspective of global change, but they are also important to understand the response of any community to environmental change, such as the microbiome of a medical patient undergoing a prolonged treatment with antibiotics. In this case, not just a single pathogenic bacterium faces a changed environment, but a complex community consisting of many thousands of species (Arumugam et al., 2011; Cho and Blaser, 2012), must adapt to avoid extinction. Whilst many models exist that study how a population of a single species, or a community composed of two

species, adapts to environmental change (Hoffmann and Sgrò, 2011; Martin et al., 2013; Northfield and Ives, 2013; Osmond and De Mazancourt, 2013; Cortez and Yamamichi, 2019), fewer models exist that describe the response of an entire community composed of multiple species to an altered environment, although there are some examples (De Mazancourt et al., 2008; Bell, 2017; Lasky, 2019). Furthermore, empirical results, describing community wide adaptation, such as those presented by Bell and Gonzalez (2011), Low-Décarie et al. (2015), Bell et al. (2019), and Roodgar et al. (2019), are clearly calling for such models.

Evolutionary rescue theory models situations in which a population can only escape extinction if it adapts. In a classical evolutionary rescue scenario, where the environment in which a population resides deteriorates, the population starts declining as a result. Extinction can then only be averted if a mutant establishes that has a positive growth rate in the new environment; i.e., the population is rescued. This process results in the well-known U-shaped curve of population size over time (Gomulkiewicz and Holt, 1995; Gonzalez et al., 2012; Orr and Unckless, 2014). Most models of evolutionary rescue focus on deriving the probability of the occurrence of such a rescue event given a certain initial population size, a rate of population decline, and a mutation rate. Evolutionary rescue theory could even be a useful tool to predict the emergence of antibiotic resistance (Martin et al., 2013; Alexander et al., 2014).

Here, we explore a new scenario in which not a single population, but a whole community composed of many different species faces a deteriorated environment, causing the populations of each species to decline. Only those species in which a rescue mutant with a positive growth rate establishes, remain in the community. In other words, evolutionary rescue occurs on a community-wide basis. We examine the effect of this process on species abundance distributions.

We present a parsimonious model of this Community-Wide Rescue (CWR) process. It describes the change in species abundances, during and after community-wide evolutionary rescue. We assume that all species are equivalent; they all start with the same negative growth rate and all have the same fixed mutation rate toward a phenotype with a positive growth rate. These assumptions are inspired by those made in the neutral model of biodiversity (Hubbell, 1997). The neutral model has been shown to be able to explain various patterns of species abundances, and has become a baseline model for community diversity patterns when species differences or species asymmetries are ignored (Alonso et al., 2006; Rosindell et al., 2011; Wennekes et al., 2012; Scheffer et al., 2018). However, because we include an explicit mutational process that introduces a different growth rate, our Community-Wide Rescue model is not strictly neutral. We compare our results with those of two null models: neutral models in which the community dynamics are solely governed by ecological drift. The first null model has a constant community size, whilst the second null model mimics the decrease in community size that occurs during Community-Wide Rescue.

The aim of this paper is to construct and explore a simple model for the CWR process, and to examine how under this model CWR affects the patterns of species abundances within a community. We quantify these patterns using Rank Abundance Curves (RAC, also known as rank abundance diagrams or distributions, RAD, McGill et al., 2007). It is well known that many different mechanisms can generate similar RACs, and hence RACs should be interpreted with caution (Chave et al., 2002). We aim to see if this general pattern also holds for our CWR model, or if perhaps RACs are informative about the (past) occurrence of CWR. We show that CWR causes a loss of rare species from the community, due to a limited supply of beneficial mutations and competition. In rare species, their low abundance limits their supply of beneficial mutations that can rescue them from extinction, whilst they face increased competition with more common species that have already undergone such beneficial mutations. However, RACs produced by the CWR process could equally well have been produced by a neutral model. In addition, as RACs proved uninformative, we also examined the rate at which CWR changes the relative species abundances (i.e., alters the RAC) and compare this to the rate at which ecological drift alters species abundance patterns. We show that CWR causes an extremely rapid loss of rare species. Such insights are crucial to understand the effects of environmental change on ecological communities.

MATERIALS AND METHODS

Our model of the CWR process is a continuous-time individual-based stochastic model, where birth, death, and mutation events are simulated using the Doob-Gillespie algorithm (Gillespie, 1976). We assume that all species are equivalent except for their initial densities. This assumption is unlikely to hold in a natural community, but its simplicity allows us to focus on the key ingredients of the CWR process. Furthermore, we consider a single closed community, i.e., there is no immigration. It is worth noting that this implies that the observed dynamics are transient in nature, when time goes to infinity all species will eventually go extinct due ecological drift. This assumption of no migration allows us to more clearly see the effect of CWR in a single (local) community. In the CWR model, the community consists of several species, each with an initial abundance that is drawn using the sampling formula for standard neutral communities (Etienne, 2005). Initially, all individuals of each species have the same negative growth rate. We call an individual with this negative growth rate a "resident." The initial community thus represents a community immediately after a drastic environmental change, in which the populations of all species are declining and unless adaptation occurs extinction is inevitable for all species. However, each resident individual can undergo a mutation to become a mutant individual, this occurs with a rate μ (note that this process implicitly assumes asexual inheritance). Again, the value for μ is the same, regardless of the species to which an individual belongs. All mutants have the same positive growth rate. Hence, we assume the simplest possible model of evolutionary rescue, as posited by Orr and Unckless (2008) and Martin et al. (2013): only a single mutational step is required to achieve a positive growth rate and this

mutation has a constant fitness effect. μ could for example represent the mutation rate toward antibiotic resistance, see also Martin et al. (2013).

The growth rates of the residents and mutants are implemented as follows. The death rates for the residents and the mutants are equal and given by d. We assume that the birth rate for both mutants and residents depends on total community size (i.e., total number of individuals in the community of all species combined),

$$b = b_0(1 - \frac{N_{tot}}{K}) \quad (1)$$

where b_0 is the rate of birth in a pristine community (no other individuals present). This parameter b_0 is different between residents and mutants (hence, we have $b_{0,res}$ and $b_{0,mut}$). We assume that $b_{0,res} < d$ so that the resident always has a negative growth rate and $b_{0,mut} > d$, so that the mutants always have a positive growth rate. Parameter K is the number of individuals at which the birth rate is equal to 0, and N_{tot} is the total number of individuals (summed across species, including both residents and mutants) in the community. It is important to note that K is not the sole parameter controlling the carrying capacity of the community; this is determined by the interplay of b_0, d, and K and is given by K (1- d/b_0). Our model deviates from standard neutral models in that we do not impose a zero-sum constraint (otherwise the community cannot decline), and that instead we have community-wide density-dependent birth. Haegeman and Etienne (2008) showed that community-level density-dependence in immigration and birth does not affect the predictions on the species abundance distributions, so we do not strongly deviate from a standard neutral model in this sense. The default parameter set for simulating the CWR model was $b_{0,res}$ = 0.05, $b_{0,mut}$ = 0.6, d = 0.1, K = 16000, and μ = 0.0005. The initial species abundances for all simulations were generated with the sampling formula for standard neutral communities as derived by Etienne (2005) using a community size of 16000, a fundamental biodiversity number, θ, of 200 and a migration parameter, I, of 40. In this neutral model, the fundamental biodiversity number controls the species abundance distribution in the regional species pool, whilst the migration parameter governs the frequency of migration from the regional species pool to the local species pool. For a more complete description the reader is referred to Etienne and Olff (2004) and Etienne (2005). Here this model is simply used to generate a reasonable initial species abundance distribution. The exact same initial species abundance distribution was used for all simulations, unless stated otherwise. All simulations, plots and analysis were performed using R version 3.5.1 (R Core Team, 2014). All new simulation code is provided in the CWERNI R-package that is available at: https://github.com/DeadParrot69/CWERNI.

To answer the question whether an endpoint RAC from a CWR community can be distinguished from a RAC generated by a neutral community, we used a simulation, fitting, and re-simulation approach. First, we simulated a community using CWR, with the default parameters. Subsequently, we fitted a neutral model to the RAC using the SADISA-package (Haegeman and Etienne, 2017). From this fit we obtain a log-likelihood, which in essence is a measure of the goodness of fit of the neutral model on the RAC generated using CWR. To generate a distribution of log-likelihoods with which to compare the log-likelihood of the neutral model fit on the CWR RAC, the parameters obtained from the neutral model fit were used to perform 500 neutral model simulations (Etienne, 2005). Then, the SADISA-package (Haegeman and Etienne, 2017) was used on each of these neutral simulations to fit a neutral model. This created a distribution of log-likelihoods for these neutral model simulations. Subsequently we determined whether the log-likelihood obtained from the neutral model fit on the CWR RAC falls outside or inside the distribution of the log-likelihoods obtained through neutral model fits on neutral model simulations. Instead of the log-likelihood we also looked at the distribution of two different diversity indices, the Shannon entropy (Rényi entropy, α = 1) and the collision entropy (Rényi entropy, α = 2) of the simulated communities. This process was repeated ten times each time with a newly drawn neutral starting community. We note that the model underlying the SADISA estimates is subtly different from that used to perform the re-simulations. The SADISA estimator makes an independent species assumption, whilst the code used for the simulations instead assumes a zero-sum assumption, but it has been shown that the RACs that these model produce are indistinguishable (Haegeman and Etienne, 2008, 2017).

In order to place the rate of rare species loss due to CWR into context, we compared it to the rate of rare species loss in a local community due to ecological drift in two truly neutral models. The first is a simple neutral model (SN) without a CWR process, where the birth and death rates are equal to those of the mutant in the CWR model. This model is thus a neutral model of the local community without immigration or speciation; it only describes the loss of species through ecological drift.

We expected the CWR model to show a decrease in total community size before recovery (due to the negative growth rate of the residents). Such a decrease in local community size, can accelerate the rate of rare species loss through ecological drift. Therefore, we also constructed a neutral model similar to SN, but where the basic birth rate for all species is set to a value less than the death rate for a predetermined time interval. This induces a steady decrease in total community size from the start of the simulation until the end of the interval. We chose the length of the interval, such that the community size decrease is similar to that observed during CWR. We call this model the variable-birth neutral model (VBN).

We simulated the three models (CWR, SN, and VBN) for 100 units of time. This was a sufficient number for all residents to go extinct in the CWR model, see also **Supplementary Figure 6**. When there are no more residents in the community the evolutionary rescue process is considered complete as all species have either undergone adaptation or gone extinct; hence we chose to simulate for 100 units of time. Each model was simulated 500 times. The SN model was simulated using the parameters b_0 = 0.6, d = 0.1, and K = 16000 (i.e., the same parameters as the mutants in the CWR model). For the VBN model we set the basic birth rate of all the species in the community (b_0) equal to $b_{0,\ res}$, during the first twenty units of time. After this time interval,

which was tuned so as to create a decrease in total community size similar or perhaps even slightly larger in nature than that in the CWR community, we set the basic birth rate equal to $b_{0,mut}$. The other parameters were the same as in the SN model. To study the RAC of a community at different stages of CWR, we plotted the resulting RACs at different points in time: $t = 15$, $t = 30$, $t = 50$, $t = 75$, and $t = 100$.

In models examining evolutionary rescue, the mutation rate and the establishment probability of the mutant are known to determine the probability of evolutionary rescue (Martin et al., 2013). Therefore, to gain more insight into our CWR model, we wanted to examine the effect of the mutation rate (μ) and mutant birth rate ($b_{0,mut}$), on the CWR process. by respectively varying the mutation rates (μ = 0.00005, 0.0005, 0.005, 0.05) and the mutant birth rate ($b_{0,mut}$ = 0.2, 0.4, 0.6, 0.8), and leaving all other parameters the same as in the default parameter set. Again, we ran 500 independent simulations for each set of parameters.

In our CWR model we assumed that $b_{0,res} < d$ so that the resident always has a negative growth rate. If this condition is not satisfied, one is no longer modeling evolutionary rescue. However, one can imagine a scenario in which $b_{0,res} > d$, for example when a bacterial community is confronted with sub-inhibitory concentrations of antibiotics. In such a community the species are not doomed to go extinct, but residents are simply replaced by fitter mutants, in essence a community-wide selective sweep. Such situations might be much more common than strict evolutionary rescue scenarios, so examining this situation could extend the applicability of our model. Therefore, we also studied a selective sweep model, derived from our CWR model, in which the only difference is that $b_{0,res} > d$, resulting in both a resident and a mutant with a positive net growth rate, whilst the mutant still has a higher net growth rate than the resident. We performed 500 simulations of this model using the parameter set $b_{0,res}$ = 0.3 and all other parameters the same as in the default CWR model parameter set.

RESULTS

The loss of rare species in the CWR community (**Figures 1A,B**) is much faster than in the neutral (SN) community (**Figures 1C,D**). In other words, the CWR process causes a very rapid loss of rare species, when compared to the rate of rare species loss from a local community due to ecological drift. Furthermore, the rate of rare species loss in the CWR model is also much larger than in the VBN model (**Figures 1E,F**). Because the VBN model has a variable carrying capacity tuned to create a decrease in total community size similar to the one observed in the CWR model, we can conclude that the rapid loss of rare species in the CWR model is not just due to ecological drift being accelerated by a decrease in total community size. In addition, the observed rapid loss of rare species occurs consistently in a relatively wide range of community sizes (between K = 1000 and K = 16000, see **Supplementary Figures 12–14**).

The same pattern is evident if one examines the figures showing the RACs at different time points for each of the three models (**Figure 2** and **Supplementary Figures 7**, **8**). Furthermore, by closely examining **Figure 2** one can see exactly at which point during the CWR process the loss of rare species occurs. During the first stage of CWR a community-wide decline occurs that does not greatly alter the shape of the RAC (**Figure 2A**). It is only as the first mutants begin to invade and the total community size starts to rebound (**Figure 2F**) that the shape of the RAC begins to change and that the loss of rare species starts to occur (**Figure 2B**). The loss of rare species continues after the community size has stabilized (**Figures 2C,D**). Once the residents have disappeared from the population, the shape of the RAC is fairly stable (**Figures 2D–F**).

The mutation rate has a strong influence on the results (**Figure 3**). If the mutation rate is very high, rescue becomes so likely that all species undergo rescue and there is no loss of rare species beyond the effects of normal ecological drift in a neutral community without speciation/immigration (**Figures 3A,B**). By contrast, if the mutation rate is very low, almost none of the species in the community undergo rescue (**Figure 3E**) and in some cases not a single rescue mutant manages to establish itself in the community (**Figure 3F**). Therefore, intermediate mutation rates seem to be required for CWR to impact the RAC and create a loss of rare species greater than that produced by ecological drift alone. In other words, the rate of rare species loss during CWR depends on the mutation rate.

Increasing $b_{0,mut}$ i.e., increasing the fitness advantage of the mutant, does not seem to influence the loss of rare species, as the RAC's obtained after the CWR process, for different values of $b_{0,mut}$ are indistinguishable (**Figure 4**). However, increasing $b_{0,mut}$ does seem to increase the speed of the rescue process. In particular, if $b_{0,mut}$ is higher, the recovery phase of the rescue process proceeds much faster, due to the higher maximal growth rate of the mutant. It should be noted that increasing $b_{0,mut}$ also increases the net carrying capacity of the rescued population, because despite a constant K, the net carrying capacity, given by K (1- d/b_0), is the density where the net birth rate is equal to the death rate. Despite this increased carrying capacity, the recovery phase is still much faster in the simulations with a high $b_{0,mut}$.

In the selective sweep model, the residents have a positive net growth rate, i.e., instead of CWR, the resident with a positive growth rate is replaced by a mutant with an even higher growth rate. As can be seen in **Figure 5B**, there is only a very minor decrease in the total community size during this replacement process. However as can be seen in **Figure 5A**, the rate of rare species loss in the selective sweep model is much higher than in the neutral SN and VBN models. In other words, when compared to ecological drift, community wide adaptation can cause a very rapid loss of rare species, just like CWR.

The fitting and re-simulation approach using the log-likelihoods of neutral model fits showed that the log-likelihood of a neutral model fit on the CWR model results consistently fell within the distribution of log-likelihoods obtained from neutral model simulations (**Supplementary Figure 9A**). A similar result was obtained when instead of log-likelihoods, the values of the Shannon entropy and the Rényi entropy of the RACs were used; the values of the Shannon entropy and the Rényi entropy estimated from the CWR RAC consistently fell inside those estimated on neutral model simulations simulated using

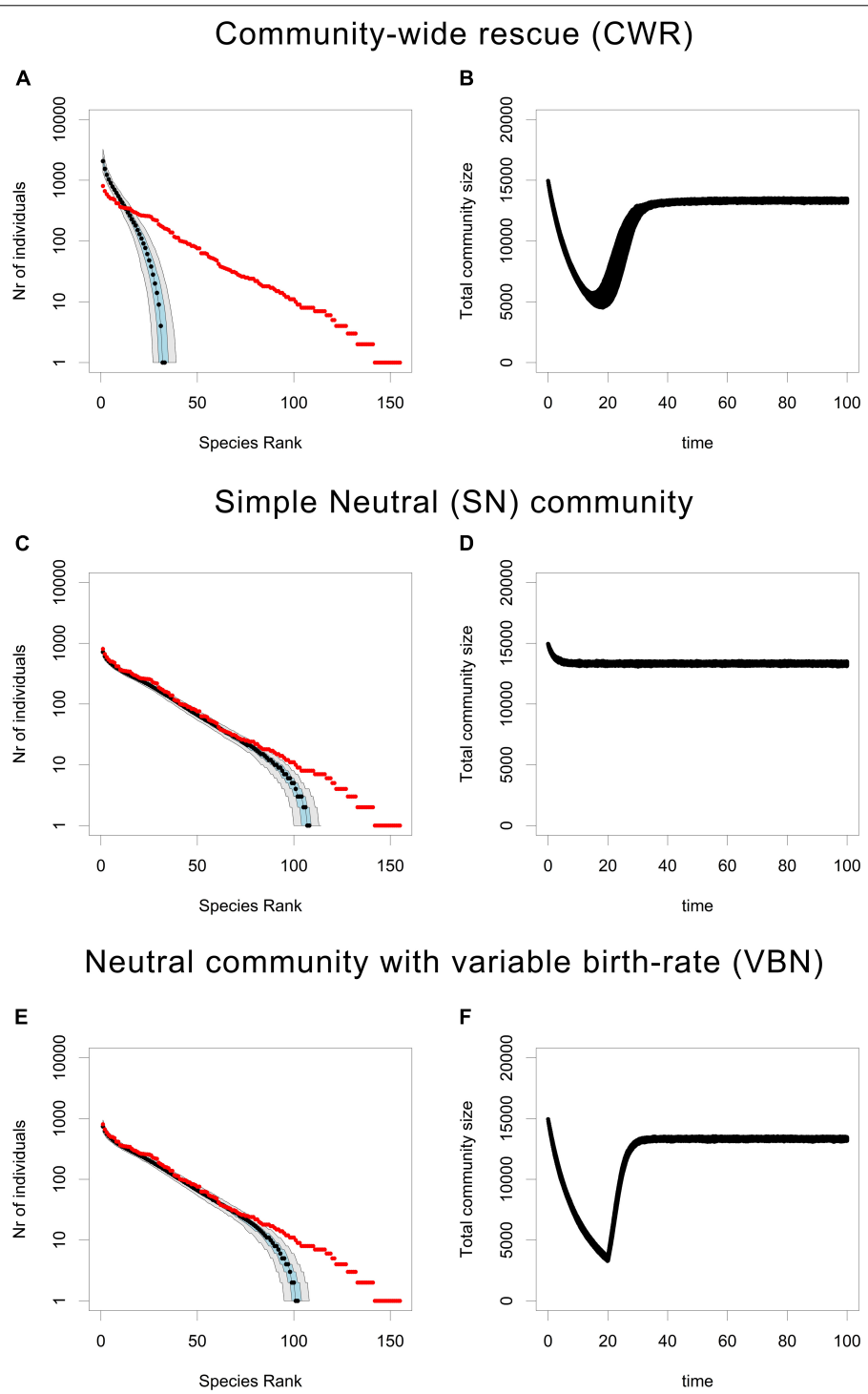


FIGURE 1 | Rank Abundance Curves and total community size under the CWR, SN, and VBN models. **(A,C,E)** Show the RACs produced after 100 units of time by the CWR model, the SN model and the VBN model, where the median is shown in black, the 25th and the 75th percentile are shown in blue and the 5th and the 95th percentile are shown in gray, the initial community is plotted in red. **(B,D,F)** Show the accompanying trajectories of total community size for each simulation over time. All plots are based on 500 simulations. Parameters for **(A,B)**: $b_{0,res} = 0.05$, $b_{0,mut} = 0.6$, $d = 0.1$, $K = 16000$, and $\mu = 0.0005$, for **(C,D)**: $b_0 = 0.6$, $d = 0.1$ and $K = 16000$, for **(E,F)**: $b_0 = 0.05$ for t between 0 and 20, $b_0 = 0.6$ for t between 20 and 100, $d = 0.1$, and $K = 16000$.

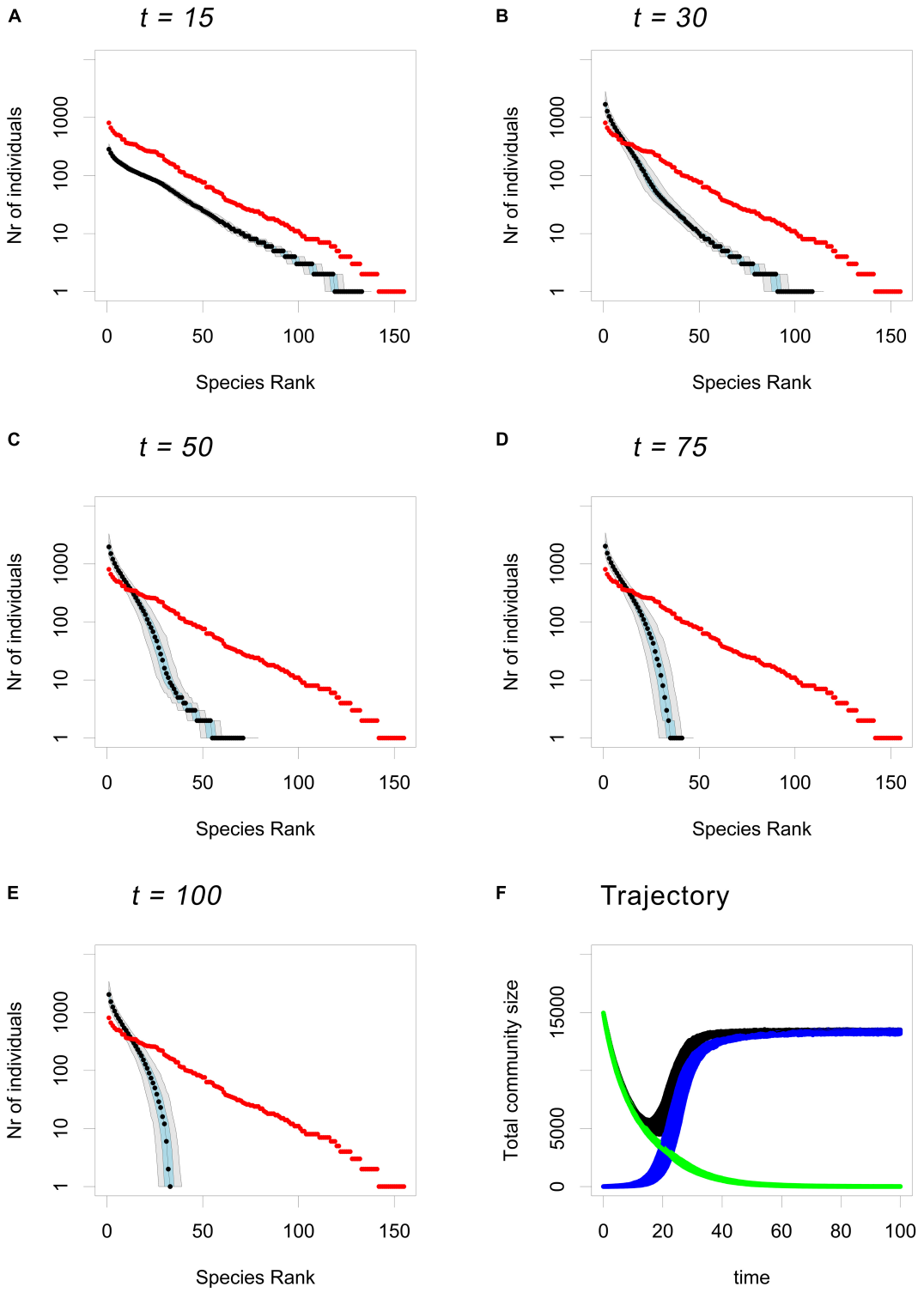


FIGURE 2 | Time trajectory of the RAC under the CWR model. Plots based on 500 CWR simulations using the (default) parameters $b_{0,res}$ = 0.05, $b_{0,mut}$ = 0.6, d = 0.1, K = 16000 and μ = 0.0005. **(A–E)** Show the RAC of the community at t = 15, t = 30, t = 50, t = 75, and t = 100 respectively, were the median is shown in black, the 25th and the 75th percentile are shown in blue and the 5th and the 95th percentile are shown in gray, and the input community is plotted in red. **(F)** Shows the trajectories of the total community size (black), the total number of residents in the communities (green) and the total number of mutants in the community (blue).

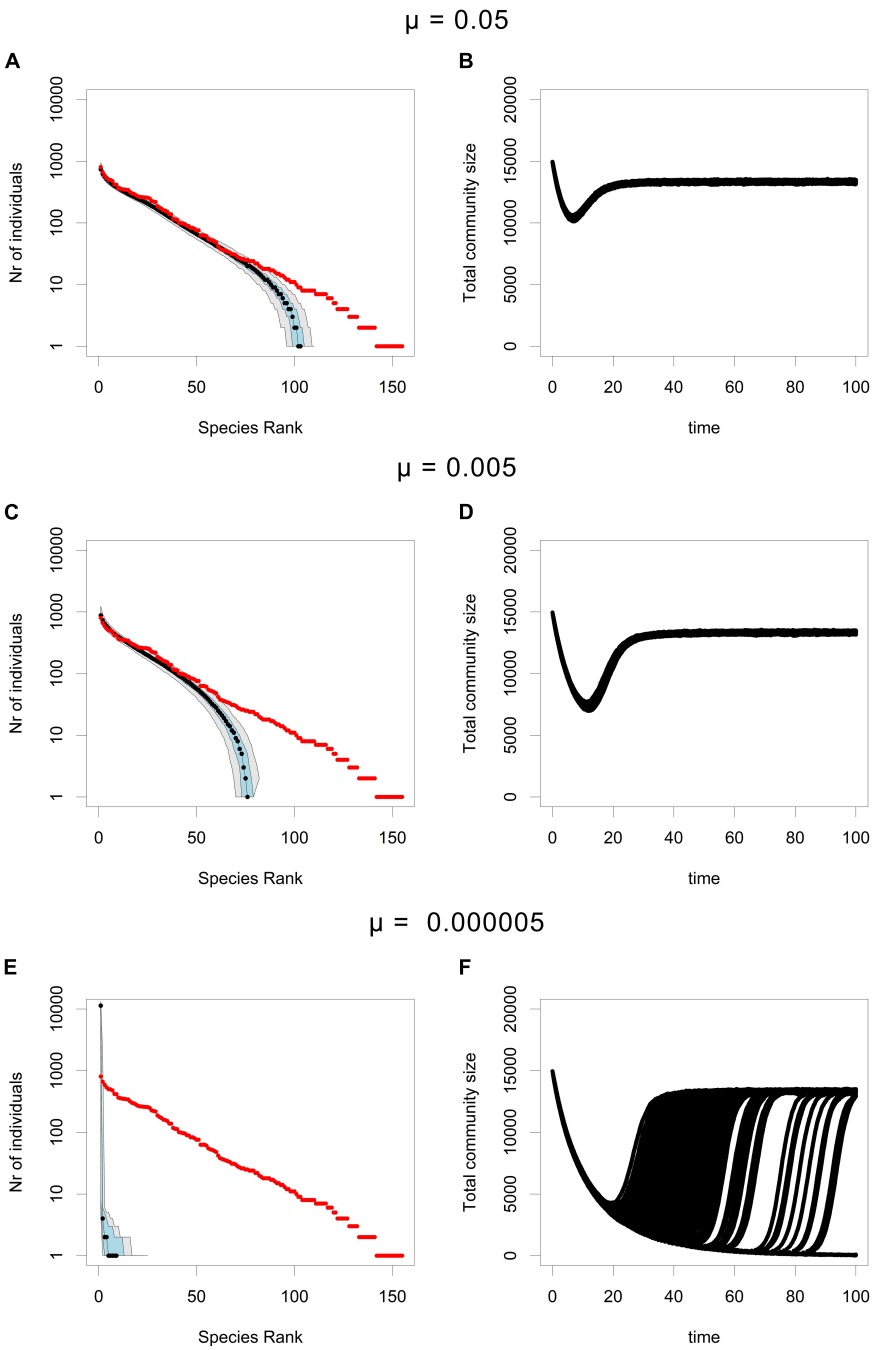


FIGURE 3 | RACs **(A,C,E)** and total community size trajectories **(B,D,F)** under the CWR process, for different mutation probabilities (μ), for each different mutation rate 500 simulations were performed. In the RAC plots the median is shown in black, the 25th and the 75th percentile are shown in blue, the 5th and the 95th percentile are shown in gray and the initial community is plotted in red. Simulations were performed using the parameters $b_{0,res}$ = 0.05, $b_{0,mut}$ = 0.6, d = 0.1, and K = 16000. For **(A,B)** μ = 0.05, in **(C,D)** μ = 0.005, and in **(E,F)** μ = 0.000005. The total community size trajectories were plotted for of each of the 500 simulations, hence the separation of these trajectories at low mutation rates **(F)**.

neutral model parameters estimated from the CWR RACs (**Supplementary Figures 9B,C**). Both of these results imply that there is no information in an endpoint RAC alone that would allow one to determine whether that RAC had been created by a neutral process or a CWR process.

DISCUSSION

We have shown that a single endpoint RAC does not allow one to determine whether that RAC had been created by a neutral process or a CWR process. This conclusion is in accordance with

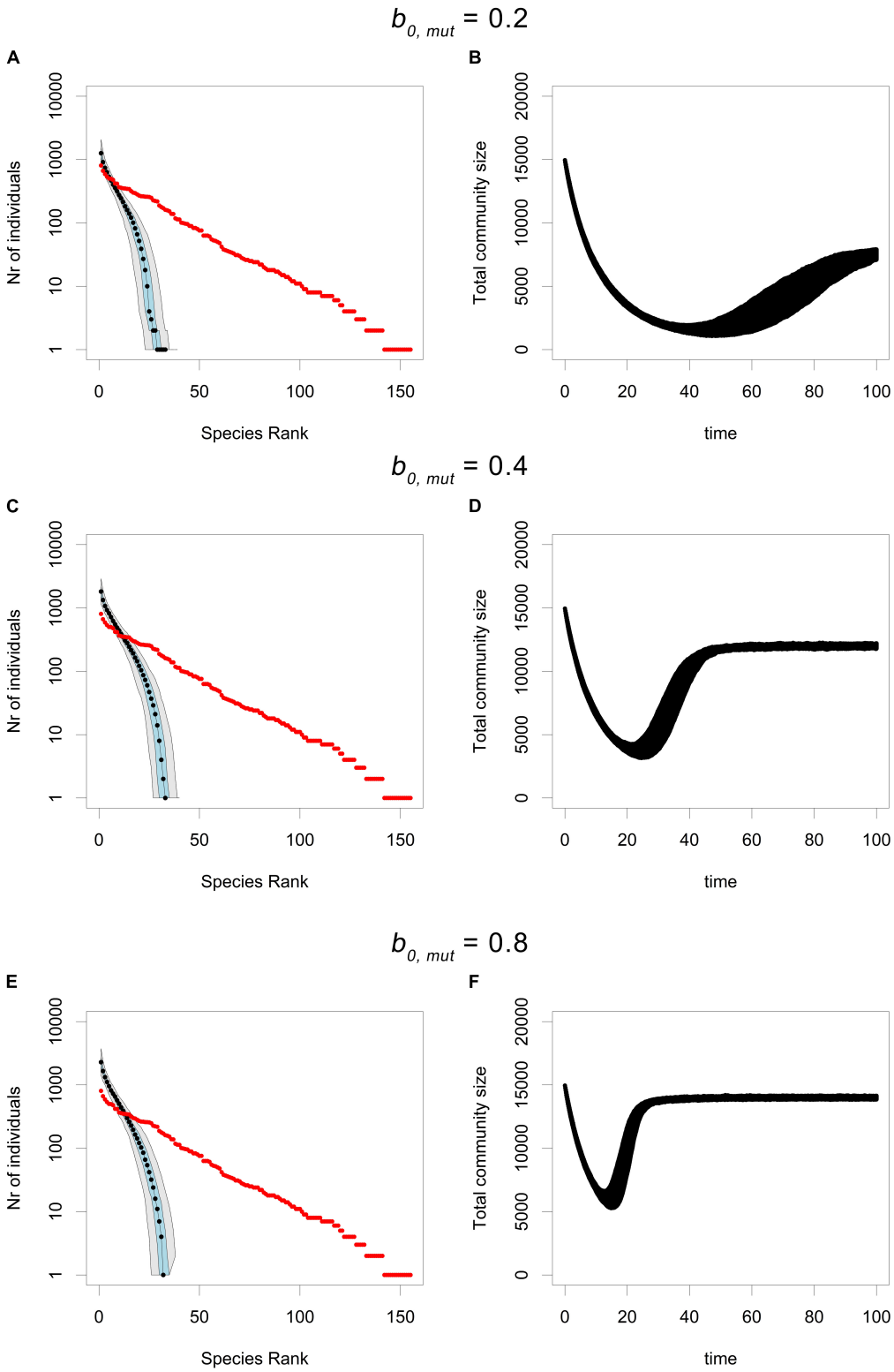


FIGURE 4 | The RAC under CWR processes with different mutant birth probabilities ($b_{0,mut}$). All simulations were performed using the parameters $b_{0,res}$ = 0.05, d = 0.1, K = 16000, and μ = 0.0005. In **(A,B)** $b_{0,mut}$ = 0.2, for **(C,D)** $b_{0,mut}$ = 0.4 and in **(E,F)** $b_{0,mut}$ = 0.8. Panels **(A,C,E)** show the RAC's after 100 units of time, where the median is shown in black, the 25th and the 75th percentile are shown in blue and the 5th and the 95th percentile are shown in gray, with the input community plotted in red. **(B,D,F)** Display the trajectories of total community size over time. All plots are based on 500 simulations.

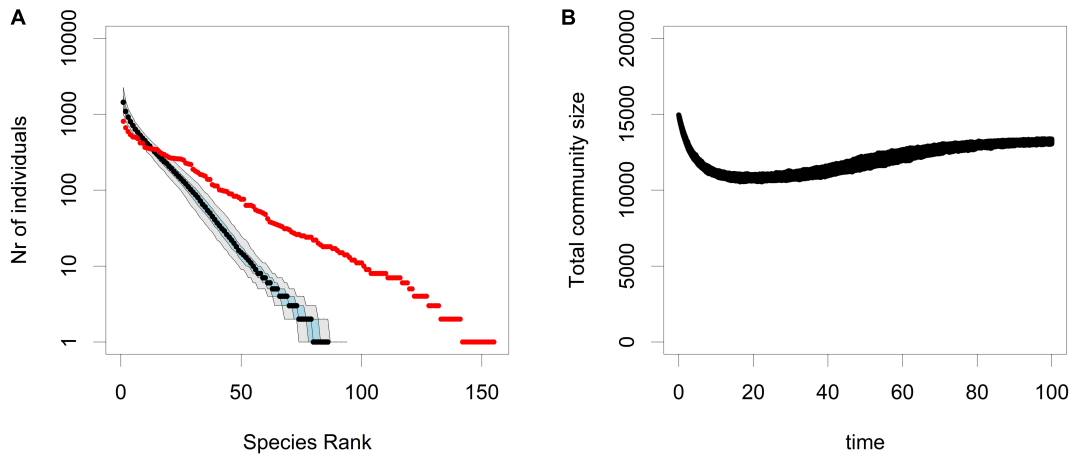


FIGURE 5 | The RAC **(A)** and the total community size **(B)** for simulations of the CWR model where the strict conditions of CWR are relaxed and reflect a scenario where the net growth rate of the residents is positive, yet still lower than that of the mutants. This causes a selective sweep during which the residents are replaced by the mutants, because the mutants have a higher fitness. In the RAC plot the median is shown in black, the 25th and the 75th percentile are shown in blue and the 5th and the 95th percentile are shown in gray, and the input community is plotted in red. These plots are based on 500 simulations. The parameters used were $b_{0,res} = 0.3$, $b_{0,mut} = 0.6$, $d = 0.1$, $K = 16000$ and $\mu = 0.0005$.

the general pattern in the literature; whilst some non-neutral processes, such as trait based environmental filtering (Jabot, 2010), can be detected by examining species abundances, many different non-neutral processes can generate surprisingly similar RACs (Chave et al., 2002).

The most striking outcome of our modeling effort is that CWR (**Figures 1A,B**) causes a very rapid loss of rare species, when compared to ecological drift (**Figures 1C,D**). This holds even if one accounts for the increase in ecological drift due to a decrease in total community size as in the VBN model (**Figures 1E,F**). Furthermore, this result is shown for a wide range of community sizes (between $K = 1000$ and $K = 16000$, see **Supplementary Figures 12–14**). In a neutral model governed by ecological drift, rare species are more likely to go extinct simply due to their lower abundance. However, in the CWR model rare species have a higher probability of going extinct, because their low abundance also means that they will have a lower probability of producing a beneficial mutant before going extinct. In other words, for rare species the supply of beneficial mutations is limited by their low abundance. This dependence of the probability of rescue on the initial population density is well characterized in standard models of evolutionary rescue and has also been demonstrated empirically (Holt, 1990; Bell and Gonzalez, 2009; Martin et al., 2013). Low abundance causes a low probability of a beneficial mutant occurring, because mutation occurs on a per-capita basis, i.e., the probability of a beneficial mutation arising during a certain time interval depends on the product of μ and the population size of the species, so during the same time interval a mutation is less likely to occur in a species with a small population size. Furthermore, the time to extinction for rare species is lower, so there is less time for a mutant to arise before the rare species goes extinct.

However, there is another effect, hypothesized by Bell (2017), which contributes to the loss of rare species: competition. This represents a crucial difference between our model and standard models of evolutionary rescue (Martin et al., 2013). In our model the birth rate of all species is governed by the total number of individuals in the community (regardless of their species), all species compete with each other (community-level density dependence). So, a species that has undergone rescue will increase the total number of individuals in the community. This causes the birth-rate of the remaining species to decrease. For the species that have not yet undergone rescue, this accelerates their decay, decreasing the time available to find a mutant before going extinct. In other words, the evolutionary rescue of one species, promotes the extinction of its competitors (Bell, 2017). Rare species that do manage to produce a mutant will tend to do so relatively late in the simulation, because their low abundance gives them a low probability of producing a mutant per unit of time. On the other hand, species with a high abundance that manage to produce a mutant will tend to do so relatively early on in the simulation, thereby promoting the extinction of the rare species through competition. It is interesting to contrast these results with those of De Mazancourt et al. (2008), who showed that on a community level biodiversity can inhibit adaptation, due to competitive interactions. In our model, the fact that rare species fail to adapt is also partly driven by competitive interactions, in that sense reaffirming the general result that competition can inhibit adaptation. However, the crucial difference is that in the model of De Mazancourt et al. (2008) these competitive interactions are driven by explicit assumptions about the ecology of each species, whilst in our model species are ecologically equivalent except for their initial abundance.

In our model, the limited supply of beneficial mutations at low abundance and competition between the species, together disproportionally promote the extinction of rare species during CWR. These two effects are also crucial to understand how

changing the mutation rate impacts the CWR process (**Figure 3**). From standard models of evolutionary rescue it follows that a high mutation rate results in a high probability of rescue (Martin et al., 2013). Furthermore, in our model a high mutation rate implies that mutations occur at very similar times for different species, limiting the competitive advantage of common species that rescue early. Therefore, if the mutation rate is too high, almost all species undergo rescue and very little rare species loss occurs (**Figures 3A,B**). For very low mutation rates the opposite holds true and very few species undergo rescue (**Figures 3E,F**). It should be noted that at very low mutation rates, in some cases not even a single species undergoes rescue. So, in other words, a lower mutation rate causes a greater loss of rare species, yet if the mutation rate is too low no rescue occurs and the entire community goes extinct.

The influence of the mutant birth rate ($b_{0,mut}$) on the CWR process (**Figure 4**) is not as straightforward. Based on standard models of evolutionary rescue, increasing the mutant birth rate should increase the fixation probability of the mutant and thereby increase the probability of rescue. Furthermore, increasing the mutant birth rate should also increase the competitive advantage of those species that rescue early. However, contrary to our expectations, we observed that an increase in the mutant birth rate does not cause an increase in the loss of rare species. Instead, an increase in the mutant birth rate only seems to affect the speed of the CWR process. This is in part due to the fact that an increase in the mutant birth rate also increases the overall carrying capacity of the community. This increase could offset the competitive advantage of the species that rescue early. Because they grow faster, the equilibrium community size is also larger. However, this increase in the community size does not influence the fixation probability of the mutant as derived in classical models of evolutionary rescue. Therefore, the fact that increasing the mutant birth rate does not increase the loss of rare species indicates that competition between early and late rescuing species is the more dominant mechanism responsible for the loss of rare species. This emphasizes the added value of our current modeling approach for understanding evolutionary rescue in a multi-species context.

We also created a different model based on the CWR model where we allowed the residents to have a positive growth rate ($b_{0,res} > d$). Relaxing this assumption implies that this model does not reflect a strict evolutionary rescue scenario, as this requires a decaying resident population. This model represents a community-wide selective sweep, during which residents with a positive growth rate are replaced by mutants with an even higher growth rate. When comparing this selective sweep model (**Figures 5A,B**) to the neutral SN and VBN model it is evident that the community-wide selective sweep causes a rapid loss of rare species when compared to ecological drift. However, the rate of rare species loss is lower than in the CWR model. Hence, one might conclude that community-wide adaptation in general leads to a loss of rare species, implying that our findings from the CWR model are more generally applicable. Furthermore, as there is no evolutionary rescue process in our selective sweep model, the only mechanism responsible is the competition between species that have found the high fitness mutant and those that have not. Hence the fact that competition alone is enough to cause the rapid rare species loss in the selective sweep model also indicates that competition is a more dominant mechanism of rare species loss in the CWR model.

As emphasized before, our CWR model assumes a simple model of evolutionary rescue. Most notably, rescue requires only a single mutation step, with a fixed positive fitness effect. For some situations these assumptions should provide a reasonable approximation. For example, the evolution of resistance to certain antibiotics requires only a single or very few mutations. In addition, the mechanisms underlying resistance can be quite similar across different species (Hooper and Jacoby, 2015). However, obviously these simple assumptions do not hold under all biological circumstances. So how would a more complex assumptions regarding mutation affect the outcome of our CWR model? Allowing multiple mutational steps of varying fitness effects would serve to make the competition during the rescue process more asymmetrical. Therefore, this would be expected to cause an even greater loss of rare species compared to our current CWR model.

The CWR model presented here assumes that all species are (initially) equal, differing only in their initial abundances, an assumption inspired by the neutral theory of biodiversity. Evidently this assumption is unlikely to strictly hold in natural communities, yet it allows us to create a relatively simple model. Furthermore, our model does not consider immigration and speciation. Future CWR models could include mutation probabilities, birth probabilities, and death probabilities that differ across species, and include migration and speciation. It will be interesting to see whether demographic rescue, by immigration, will counteract or aid evolutionary rescue by mutation.

It is striking that the change in the shape of the RAC produced by the CWR process i.e., one devoid of rare species is a pattern commonly observed by ecologists in "stressed" or disturbed communities (Bazzaz, 1975; Halloy and Barratt, 2007; Webb and Leighton, 2011). Additionally, antibiotic treatment also seems to cause a similar loss of rare species in the microbiome of patients, which persists long after the treatment (Sommer and Dantas, 2011). Interestingly, a study of benthic foraminifera during the Paleocene–Eocene thermal maximum by Webb et al. (2009) showed a decrease in richness, an increase in kurtosis, and a decrease in evenness during the Paleocene–Eocene thermal maximum, i.e., a change in the shape of the RAC that would also be consistent with a CWR scenario.

However, it is important to realize that there are countless other ecological explanations that may account for the loss of rare species in stressed environments. For example, rare species tend to be more specialized and are hence more sensitive to disturbance (Davies et al., 2004). Or the loss of a single keystone species can in turn lead to the loss of many rare species that may depend on it (Rapport et al., 1985). Thus, if rapid loss of rare species is observed, that is much faster than would be expected due to ecological drift, this does not *per se* imply an underlying CWR process.

However, this does not mean that CWR is a hypothetical process with little relevance, to real ecological communities.

Experimentalists are examining evolutionary rescue in a community context. The examples include microbiomes adapting to antibiotic treatment (Roodgar et al., 2019); soil microbial communities adapting to herbicides (Low-Décarie et al., 2015) and lacustrine plankton communities adapting to acidification (Bell et al., 2019). There are many more situations in which CWR could be considered as a potential mechanism for rare species loss, as many ecosystems face irreversible human induced environmental change on a community-wide level (Vitousek et al., 1997).

All in all, the current CWR model represents an initial exploration of CWR and could be considered as a baseline model regarding the effect of community-wide evolutionary rescue on species abundances. Yet, this simple model provides a clear testable prediction regarding the effect of CWR on species abundances: Community-Wide Rescue causes a very rapid loss of rare species.

DATA AVAILABILITY STATEMENT

The simulation code presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: https://github.com/DeadParrot69/CWERNI.

AUTHOR CONTRIBUTIONS

TE was primarily responsible for the conception of the idea, programming, simulation, analysis, and writing of this manuscript. KB contributed to the development of the ideas and provided extensive input regarding the writing of the manuscript. RE contributed to the development of the ideas, proposed modeling strategies, suggested analyses and provided extensive input regarding the writing of the manuscript.

FUNDING

TE wishes to thank the Erasmus Mundus Programme in Evolutonary biology (MEME) for the opportunities and funding provided. RE thanks the Netherlands Organisation for Scientfic Research (NWO) for funding through a VICI grant. This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (Grant agreement No. 789240).

ACKNOWLEDGMENTS

We are greatly indebted to Pedro Neves, Inès Daras, and Richel Bilderbeek for their valuable comments on the model and the manuscript. Furthermore, Rixt Heerschop is thanked for a proofreading and her help in preparing the figures. F. J. Weissing is thanked for his continous support. Guillaume Martin and Yoann Anciaux are thanked for their valuable lessons and discussions regarding evolutionary rescue theory. In addition, we thank S. Panish-Inq. for her unexpected comments. We would also like to thank Mika van Eldijk for providing computational resources when they were most needed. Furthermore, we would like to thank François Massol and five reviewers for their valuable comments which greatly improved the manuscript.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fevo.2020.552268/full#supplementary-material

REFERENCES

Alexander, H. K., Martin, G., Martin, O. Y., and Bonhoeffer, S. (2014). Evolutionary rescue: linking theory for conservation and medicine. *Evol. Appl.* 7, 1161–1179. doi: 10.1111/eva.12221

Alonso, D., Etienne, R. S., and McKane, A. J. (2006). The merits of neutral theory. *Trends Ecol. Evol.* 21, 451–457. doi: 10.1016/j.tree.2006.03.019

Arumugam, M., Raes, J., Pelletier, E., Le Paslier, D., Yamada, T., Mende, D. R., et al. (2011). Enterotypes of the human gut microbiome. *Nature* 473, 174–180.

Bazzaz, F. A. (1975). Plant species diversity in old-field successional ecosystems in Southern Illinois. *Ecology* 56, 485–488. doi: 10.2307/1934981

Bell, G. (2017). Evolutionary rescue. *Annu. Rev. Ecol. Evol. Syst.* 48, 605–627.

Bell, G., Fugère, V., Barrett, R., Beisner, B., Cristescu, M., Fussmann, G., et al. (2019). Trophic structure modulates community rescue following acidification. *Proc. R. Soc. B* 286:20190856. doi: 10.1098/rspb.2019.0856

Bell, G., and Gonzalez, A. (2009). Evolutionary rescue can prevent extinction following environmental change. *Ecol. Lett.* 12, 942–948. doi: 10.1111/j.1461-0248.2009.01350.x

Bell, G., and Gonzalez, A. (2011). Adaptation and evolutionary Rescue in metapopulations experiencing environmental deterioration. *Science* 1327, 1327–1331. doi: 10.1126/science.1203105

Chave, J., Muller-Landau, H. C., and Levin, S. A. (2002). Comparing classical community models: theoretical consequences for patterns of diversity. *Am. Nat.* 159, 1–23. doi: 10.2307/3079311

Cho, I., and Blaser, M. J. (2012). The human microbiome: at the interface of health and disease. *Nat. Rev. Genet.* 13, 260–270. doi: 10.1038/nrg3182

Cortez, M. H., and Yamamichi, M. (2019). How (co) evolution alters predator responses to increased mortality: extinction thresholds and hydra effects. *Ecology* 100:e02789.

Davies, K. F., Margules, C. R., and Lawrence, J. F. (2004). A synergistic effect puts rare, specialized species at greater risk of extinction. *Ecology* 85, 265–271. doi: 10.1890/03-0110

De Mazancourt, C., Johnson, E., and Barraclough, T. G. (2008). Biodiversity inhibits species' evolutionary responses to changing environments. *Ecol. Lett.* 11, 380–388. doi: 10.1111/j.1461-0248.2008.01152.x

Etienne, R. S. (2005). A new sampling formula for neutral biodiversity. *Ecol. Lett.* 8, 253–260. doi: 10.1111/j.1461-0248.2004.00717.x

Etienne, R. S., and Olff, H. (2004). A novel genealogical approach to neutral biodiversity theory. *Ecol. Lett.* 7, 170–175. doi: 10.1111/j.1461-0248.2004.00572.x

Gillespie, D. T. (1976). A general method for numerically simulating the stochastic time evolution of coupled chemical reactions. *J. Comput. Phys.* 22, 403–434. doi: 10.1016/0021-9991(76)90041-3

Gomulkiewicz, R., and Holt, R. D. (1995). When does evolution by natural selection prevent extinction? *Evolution* 49, 201–207. doi: 10.2307/2410305

Gonzalez, A., Ronce, O., Ferriere, R., and Hochberg, M. E. (2012). Evolutionary rescue: an emerging focus at the intersection between ecology and evolution. *Philos. Trans. R. Soc. B* 368:20120404. doi: 10.1098/rstb.2012.0404

Haegeman, B., and Etienne, R. S. (2008). Relaxing the zero-sum assumption in neutral biodiversity theory. *J. Theor. Biol.* 252, 288–294. doi: 10.1016/j.jtbi.2008.01.023

Haegeman, B., and Etienne, R. S. (2017). A general sampling formula for community structure data. *Methods Ecol. Evol.* 8, 1506–1519. doi: 10.1111/2041-210x.12807

Halloy, S. R. P., and Barratt, B. I. P. (2007). Patterns of abundance and morphology as indicators of ecosystem status: a meta-analysis. *Ecol. Complexity* 4, 128–147. doi: 10.1016/j.ecocom.2007.04.002

Hoffmann, A. A., and Sgrò, C. M. (2011). Climate change and evolutionary adaptation. *Nature* 470, 479–485.

Holt, R. D. (1990). The microevolutionary consequences of climate change. *Trends Ecol. Evol.* 5, 311–315. doi: 10.1016/0169-5347(90)90088-u

Hooper, D. C., and Jacoby, G. A. (2015). Mechanisms of drug resistance: quinolone resistance. *Ann. N.Y. Acad. Sci.* 1354, 12–31. doi: 10.1111/nyas.12830

Hubbell, S. P. (1997). A unified theory of biogeography and relative species abundance and its application to tropical rain forests and coral reefs. *Coral Reefs* 16, S9–S21.

IPCC (2014). *Climate Change (2014): Synthesis Report*. Geneva: IPCC.

Jabot, F. (2010). A stochastic dispersal-limited trait-based model of community dynamics. *J. Theor. Biol.* 262, 650–661. doi: 10.1016/j.jtbi.2009.11.004

Lasky, J. R. (2019). Eco-evolutionary community turnover following environmental change. *Evol. Appl.* 12, 1434–1448. doi: 10.1111/eva.12776

Low-Décarie, E., Kolber, M., Homme, P., Lofano, A., Dumbrell, A., Gonzalez, A., et al. (2015). Community rescue in experimental metacommunities. *Proc. Natl. Acad. Sci. U.S.A.* 112, 14307–14312. doi: 10.1073/pnas.1513125112

Martin, G., Aguilée, R., Ramsayer, J., Kaltz, O., and Ronce, O. (2013). The probability of evolutionary rescue: towards a quantitative comparison between theory and evolution experiments. *Philos. Trans. R. Soc. B* 368:20120088. doi: 10.1098/rstb.2012.0088

McGill, B. J., Etienne, R. S., Gray, J. S., Alonso, D., Anderson, M. J., Benecha, H. K., et al. (2007). Species abundance distributions: moving beyond single prediction theories to integration within an ecological framework. *Ecol. Lett.* 10, 995–1015. doi: 10.1111/j.1461-0248.2007.01094.x

Northfield, T. D., and Ives, A. R. (2013). Coevolution and the effects of climate change on interacting species. *PLoS Biol.* 11:e1001685. doi: 10.1371/journal.pbio.1001685

Orr, H. A., and Unckless, R. L. (2008). Population extinction and the genetics of adaptation. *Am. Nat.* 172, 160–169. doi: 10.1086/589460

Orr, H. A., and Unckless, R. L. (2014). The population genetics of evolutionary rescue. *PLoS Genet.* 10:e1004551. doi: 10.1371/journal.pgen.1004551

Osmond, M. M., and De Mazancourt, C. (2013). How competition affects evolutionary rescue. *Philos. Trans. R. Soc. B* 368:20120085. doi: 10.1098/rstb.2012.0085

R Core Team (2014). *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing.

Rapport, D. J., Regier, H. A., and Hutchinson, T. C. (1985). Ecosystem behavior under stress. *Am. Nat.* 125, 617–640. doi: 10.1086/284368

Roodgar, M., Good, B. H., Garud, N. R., Martis, S., Avula, M., Zhou, W., et al. (2019). Longitudinal linked read sequencing reveals ecological and evolutionary responses of a human gut microbiome during antibiotic treatment. *bioRxiv* [Preprint]. doi: 10.1101/2019.12.21.886093

Rosindell, J., Hubbell, S. P., and Etienne, R. S. (2011). The unified neutral theory of biodiversity and biogeography at age ten. *Trends Ecol. Evol.* 26, 340–348. doi: 10.1016/j.tree.2011.03.024

Scheffer, M., van Nes, E. H., and Vergnon, R. (2018). Toward a unifying theory of biodiversity. *Proc. Natl. Acad. Sci. U.S.A.* 115, 639–641. doi: 10.1073/pnas.1721114115

Sommer, M. O. A., and Dantas, G. (2011). Antibiotics and the resistant microbiome. *Curr. Opin. Microbiol.* 14, 556–563. doi: 10.1016/j.mib.2011.07.005

Vitousek, P. M., Mooney, H. A., Lubchenco, J., and Melillo, J. M. (1997). Human domination of earth's ecosystems. *Science* 277, 494–499. doi: 10.1126/science.277.5325.494

Webb, A. E., Leighton, L., Schellenberg, S., Landau, E., and Thomas, E. (2009). Impact of the Paleocene-Eocene thermal maximum on deep-ocean microbenthic community structure: using rank-abundance curves to quantify paleoecological response. *Geology* 37, 783–786. doi: 10.1130/g30074a.1

Webb, A. E., and Leighton, L. R. (2011). "Exploring the ecological dynamics of extinciton," in *Quantifying the Evolution of Early Life*, eds M. Laflamme and S. Q. Dornbos (Cham: Springer), 185–220.

Wennekes, P. L., Rosindell, J., and Etienne, R. S. (2012). The neutral—niche debate: a philosophical perspective. *Acta Biotheor.* 60, 257–271. doi: 10.1007/s10441-012-9144-6

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Ecology and Evolution

REVIEW
published: 30 November 2020
doi: 10.3389/fevo.2020.602809

How Thermodynamics Illuminates Population Interactions in Microbial Communities

Mayumi Seto[1]* *and Yoh Iwasa*[1,2]

[1] Department of Chemistry, Biology, and Environmental Sciences, Nara Women's University, Nara, Japan, [2] Department of Bioscience, Graduate School of Science and Technology, Kwansei Gakuin University, Sanda-shi, Japan

In traditional population models of microbial ecology, there are two central players: producers and consumers (including decomposers that depend on organic carbon). Producers support surface ecosystems by generating adenosine triphosphate (ATP) from sunlight, part of which is used to build new biomass from carbon dioxide. In contrast, the productivity of subsurface ecosystems with a limited supply of sunlight must rely on bacteria and archaea that are able generate ATP solely from chemical or electric energy to fix inorganic carbon. These "light-independent producers" are frequently not included in traditional food webs, even though they are ubiquitous in nature and interact with one another through the utilization of the by-products of others. In this review, we introduce theoretical approaches based on population dynamics that incorporate thermodynamics to highlight characteristic interactions in the microbial community of subsurface ecosystems, which may link community structures and ecosystem expansion under conditions of a limited supply of sunlight. In comparison with light-dependent producers, which compete with one another for light, the use of Gibbs free energy (chemical energy) can lead cooperative interactions among light-independent producers through the effects of the relative quantities of products and reactants on the available chemical energy, which is termed abundant resource premium. The development of a population theory that incorporates thermodynamics offers fundamental ecological insights into subsurface microbial ecosystems, which may be applied to fields of study such as environmental science/engineering, astrobiology, or the microbial ecosystems of the early earth.

Keywords: microbial ecology, mutualism, mathematical models, abundant resource premiums, syntrophy, chemolithotrophy

OPEN ACCESS

Edited by:
John Maxwell Halley, University of Ioannina, Greece

Reviewed by:
Nikolaos Monokrousos, International Hellenic University, Greece
Stavros D. Veresoglou, Freie Universität Berlin, Germany

***Correspondence:**
Mayumi Seto
seto@ics.nara-wu.ac.jp

Specialty section:
This article was submitted to Models in Ecology and Evolution, a section of the journal Frontiers in Ecology and Evolution

Received: *04 September 2020*
Accepted: *10 November 2020*
Published: *30 November 2020*

Citation:
Seto M and Iwasa Y (2020) How Thermodynamics Illuminates Population Interactions in Microbial Communities. Front. Ecol. Evol. 8:602809. doi: 10.3389/fevo.2020.602809

INTRODUCTION

By the look of terrestrial ecosystems, life on Earth first appears to consist of abundant plant and animal communities. However, upon closer observation, there exists rich and diverse communities of fungi, bacteria, and archaea, many of which are not readily visible. According to Bar-On et al. (2018), the total biomass of the earth's ecosystems is approximately 550 gigatons of carbon (Gt C) with an estimated 81.8% of plants, 12.7% of bacteria, 2.2% of fungi, and 1.3% of archaea, whereas animals only comprise 0.4%. They also reported that approximately 70 Gt C of biomass is present in the marine sub-seafloor sediments and the oceanic crust, as well as terrestrial substratums deeper

than 8 m, which is dominated by bacteria and archaea. Total bacterial and archaeal cell numbers are estimated to be upwards of 1×10^{30} cells with nearly 3/4 residing deep in the subsurface (Flemming and Wuertz, 2019). Explorations of ocean drilling research projects have revealed that bacterial and archaeal communities are present at depths of 2.5 km below the ocean floor (Parkes et al., 1994; Ciobanu et al., 2014; Inagaki et al., 2015; Glombitza et al., 2016; Trembath-Reichert et al., 2017; Tanikawa et al., 2018). These findings suggest that there are widespread microbial communities in the subsurface realm that are not normally perceived.

Life cannot function without a constant supply of energy. As producers, plants harness sunlight as their primary energy source. However, in subsurface ecosystems where the availability of light is limited, the primary energy source would be derived from geochemical origins rather than solar energy. For example, hydrothermal vent systems are the natural plumbing systems that supply chemical substances as energy sources (e.g., hydrogen gas, methane, and dissolved iron) from the interior of the Earth. Unlike animals that depend solely on aerobic respiration for ATP generation, bacteria and archaea have evolved diverse metabolic processes that enable them to harness a variety of chemical reactions that use energy sources from geochemical origins (**Table 1**). These organisms synthesize ATP without relying on sunlight or organic substrates to produce their own biomass from inorganic carbon. By extending the ecological concept of producers to these microbes, they can be called "light-independent producers," which would include denitrifiers, nitrifiers, iron oxidizers/reducers, sulfate reducers, methanotrophs, methanogens, etc. These functional (or trait-based) groups may include species that depends on organic substrate as energy source, but microbes harnessing the reactions listed in **Table 1** are independent of organic substrate. Some light-independent producers may use molecular oxygen, although this may also be a photosynthetic by-product.

The presence of light-independent producers is not limited to subsurface ecosystems and also occurs in general and extreme environments. Due to their access to chemical substances, they are often observed in the interfaces between oxic (oxygen plentiful) and anoxic (oxygen-free) environments, and can move energy-source reactions forward. These environments are temporarily or spatially created, even in surface ecosystems such as sediment-water interfaces, soil aggregates, and biofilms (e.g., see review by Lau et al., 2018). Some light-independent producers can tolerate extreme pH or temperature conditions and are widely observed in mine drainage, soda lakes, hot springs, polar regions, and hydrothermal vents (Sorokin and Kuenen, 2005; Sattley and Madigan, 2006; Martin et al., 2008; Kimura et al., 2011). Although the contribution of these microbial communities to ecosystem productivity in surface ecosystems may be less important than the primary productivity of plants, light-independent producers drive important biogeochemical cycles by catalyzing energy-source reactions (Falkowski et al., 2008; Burgin et al., 2011).

The potential for subsurface primary production that is independent of solar energy has been suggested by several reviews (Gold, 1992; Stevens, 1997; Colman et al., 2017), but there are relatively few studies available that deal with the population dynamics and interspecific interactions among light-independent producers. The interspecific interactions among these producers are often associated with the use of chemical substances involved in energy-source reactions. Understanding these interactions requires not only knowledge of the biological interactions, but also the chemical interactions and energy that depends on the presence of the chemical substances in the system. Therefore, in this review, we introduce theoretical approaches based on population dynamics, which incorporates thermodynamics to highlight characteristic interactions in microbial communities in subsurface ecosystems on an energetic basis.

ENERGY AND NUTRITIONAL GROUPS

To expand our understanding of energy in ecosystems and how organisms use this energy for growth, we first summarize the concepts of energy and nutritional groups. Organisms, especially microorganisms, can be classified into primary nutritional groups based on the direct source of energy and carbon (**Figure 1**). As the definition of nutritional groups is somewhat ambiguous, we clarify them in light of recent research regarding bacterial and archaeal communities.

In ecology, energy flow has often been treated as a process that occurs in parallel with carbon flow. This is true for organisms that use carbon substrates as energy sources, however, the energy sources of microorganisms are not necessarily sunlight or carbon substrates. Energy is converted to ATP and then used for biomass synthesis. Producers use a part of ATP in the process of converting inorganic carbon to organic carbon (biomass). In addition to energy and carbon sources, we introduce the processes of ATP synthesis and carbon assimilation, which may highlight the differences between the carbon and energy needs of living organisms.

Classifications Based on Energy Source

In traditional textbooks, the direct energy source for life is described as either being sunlight or chemical reactions. Organisms that use sunlight are called phototrophs and those that harness energy from chemical reactions are called chemotrophs. According to these definitions, animals and fungi are classified as chemotrophs. Some chemotrophs use organic carbon for energy-source reactions (e.g., aerobic respirators and fermentation bacteria), and others only use inorganic substances (e.g., the reactions listed in **Table 1**). The former are termed chemoorganotrophs and the latter as chemolithotrophs. In addition to light and chemical energy, some microbes have been found to gain energy directly from electrons in the form of electric energy (Lovley, 2006, 2008; Thrash and Coates, 2008; Torres et al., 2010). As these electrons are supplied from the progression of the oxidation-reduction reaction (the redox reaction, a type of chemical reaction that involves the transfer of electrons), the distinction between chemical and electric energy gain is somewhat vague. Electron flow is also generated during photosynthesis and aerobic respiration, and is essential for one of the two ATP synthesizing processes described below.

TABLE 1 | Possible examples of energy-source reactions for the light-independent producers.

	Process	Reaction	Electron donor	Electron acceptor	Negative of standard Gibbs energy change of reaction, $-\Delta rG°$ (kJ per mol of electron donor)
I	Aerobic methane oxidation	$CH_4 + 2O_2 \rightarrow HCO_3^- + H_2O + H^+$	CH_4	O_2	790
II	Denitrification using hydrogen sulfide ion with SO_4^{2-} production	$HS^- + 1.6NO_3^- \rightarrow SO_4^{2-} + 0.8N_2 + 0.2H_2O + 0.6OH^-$	HS^-	SO_4^{2-}	721
III	Methane oxidation using manganese oxide	$CH_4 + 4MnO_2 + 7H^+ \rightarrow HCO_3^- + 4Mn^{2+} + 5H_2O$	CH_4	MnO_2	635
IV	Methane oxidation using iron oxide	$CH_4 + 8Fe(OH)_3 + 15H^+ \rightarrow HCO_3^- + 8Fe^{2+} + 21H_2O$	CH_4	$Fe(OH)_3$	598
V	Anaerobic ammonia oxidation (ANAMMOX)	$NH_4^+ + NO_2^- \rightarrow N_2 + 2H_2O$	NH_4^+	NO_2^-	362
VI	Complete ammonia oxidation (COMAMMOX)	$NH_4^+ + 2O_2 \rightarrow NO_2^- + H_2O + 2H^+$	NH_4^+	O_2	269
VII	Denitrification using hydrogen sulfide ion with S^0	$HS^- + 0.4NO_3^- + 1.4H^+ \rightarrow S^0 + 0.2N_2 + 1.2H_2O$	HS^-	NO_3^-	252
VIII	Ammonia oxidation	$NH_4^+ + 1.5O_2 \rightarrow NO_2^- + H_2O + 2H^+$	NH_4^+	O_2	190
IX	Nitrite reduction using hydrogen	$3H_2 + NO_2^- + 2H^+ \rightarrow NH_4^+ + H_2O$	H_2	NO_2^-	174
X	Nitrite oxidation	$NO_2^- + 0.5O_2 \rightarrow NO_3^-$	NO_2^-	O_2	79
XI	Aerobic Fe oxidation	$Fe^{2+} + 0.25O_2 + H^+ \rightarrow Fe^{3+} + 0.5H_2O$	Fe^{2+}	O_2	44
XII	Methanogenesis	$4H_2 + CO_2 \rightarrow CH_4 + 2H_2O$	H_2	CO_2	30
XIII	Fe oxidation using nitrate	$Fe^{2+} + 0.2NO_3^- + 2.4H_2O \rightarrow Fe(OH)_3 + 0.1N_2 + 1.8H^+$	Fe^{2+}	NO_3^-	26
XIV	Sulfate reduction using hydrogen	$H_2 + 0.25SO_4^{2-} \rightarrow 0.25H_2S + 0.5OH^- + 0.5H_2O$	H_2	SO_4^{2-}	18
XV	Methane oxidation using sulfate	$CH_4 + SO_4^{2-} \rightarrow HCO_3^- + HS^- + H_2O$	CH_4	SO_4^{2-}	17

For the calculation of $-\Delta_r G°$, the values of $\Delta_f G°$ at 25°C and 1 bar are cited from CHNOSZ package for R (http://chnosz.net). $-\Delta_r G°$ for ferrihydrite was taken from Majzlan et al. (2004).

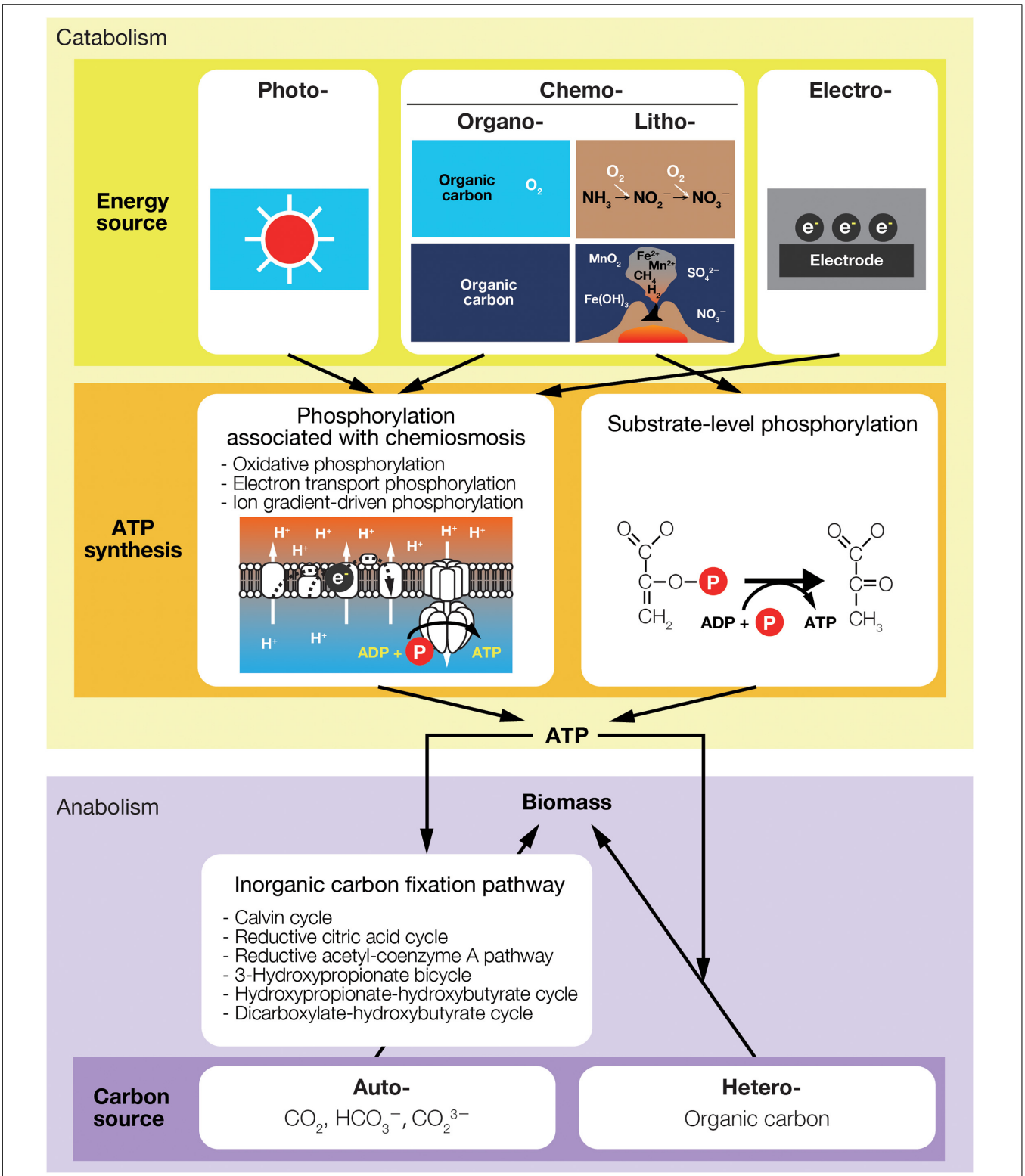


FIGURE 1 | Primary nutritional groups and types of energy sources, ATP synthesis, and carbon sources for life. The names of primary nutritional groups are built by combining the Greek terms meaning the direct source of energy [photo (light), chemo (chemical), or electro (electric)] and carbon [auto (using inorganic carbon) or hetero (using organic carbon)] and "troph(s)" meaning nourishment. Chemo-(auto or hetero)-trophs that use organic carbon for energy-source reactions (e.g., aerobic respirators and fermentation bacteria) are especially termed chemo-organo-(auto or hetero)-trophs and others that only use inorganic substances for energy-source reactions (e.g., the reactions listed in **Table 1**) are termed chemo-litho-(auto or hetero)-trophs.

ATP Synthesis

No matter which energy form is used, life eventually converts energy into ATP by adding one phosphate group to adenosine diphosphate (ADP), in a process called phosphorylation. Phosphorylation (i.e., ATP synthesis) in living organisms is achieved by two different biochemical pathways. Phosphorylation can occur via the direct transfer of a phosphate group from a phosphorylated intermediate metabolic compound to ADP. This is known as substrate-level phosphorylation and is exemplified by glycolysis (the first pathway of aerobic respiration) and fermentation (**Figure 1**). ADP-ribose-derived ATP synthesis in the cell nucleus might be considered as this type of phosphorylation (Wright et al., 2016). Alternatively, phosphorylation can be achieved by the electron flow generated from an energy-source reaction. The electron flow eventually creates ion gradients across membranes, or chemiosmosis, which enables ATP synthase to synthesize ATP (Mitchell, 1961; Dimroth, 1987; Russell, 2007; especially see Müller and Hess, 2017 for a discussion on the terms describing this type of phosphorylation). In this paper, we have termed the latter type as "phosphorylation associated with chemiosmosis." In terms of energy production, substrate-level phosphorylation is generally less efficient than phosphorylation associated with chemiosmosis. However, the faster substrate-consumption rate of substrate-level phosphorylation can surpass phosphorylation associated with chemiosmosis under some conditions (Vazquez et al., 2010; Zheng, 2012).

Photosynthesis and aerobic respiration partly share similar phosphorylation processes, which are associated with chemiosmosis of hydrogen ions when electrons move down the electron transport chain. For light reactions occurring in plants, electrons are released after the light absorption by photosynthetic pigments and drive chemiosmosis across the thylakoid membrane. Similarly, most energy-source reactions, including aerobic reactions and the reactions listed in **Table 1**, are involved in electron transfers and, therefore, categorized as redox reactions that can be spontaneously carried out in the absence of sunlight. For example, during aerobic respiration, 24 moles of electrons are translocated from 1 mole of glucose to oxygen:

$$C_6H_{12}O_6(\text{electron donor}) + 6H_2O \rightarrow 6CO_2 + 24H^+ + 24e^- \quad (1a)$$

$$6O_2(\text{electron acceptor}) + 24H^+ + 24e^- \rightarrow 12H_2O \quad (1b)$$

$$6O_2 + C_6H_{12}O_6 \rightarrow 6CO_2 + 6H_2O, \quad (1c)$$

where e^- indicates an electron. The overall reaction of Eq. (1c) is the well-known equation for aerobic respiration, which is the sum of the oxidation reaction (Eq. (1a)) that involves the loss of electrons from an electron donor ($C_6H_{12}O_6$) and the reduction reaction (Eq. (1b)) that is involved with the gain of electrons of an electron acceptor (O_2). Although the electron flow involved in Eq. 1 is eventually eliminated from Eq. 1c, this flow is used to generate ATP throughout chemiosmosis in mitocondria.

In general, chemical substances that act as electron acceptors are plentiful in oxic environments, while those that act as electron donors are abundant under anoxic conditions. Surface ecosystems have unique oxic environments where plants constantly supply the abundant organic substrate, which acts as an electron donor. In contrast, both electron acceptors and donors (except for organic substrates) come into play in oxic-anoxic interfaces and can supply energy for light-independent producers.

Classification Based on the Carbon Source

Some fraction of the energy stored as ATP is used to maintain the metabolism of an individual, while some is allocated to the growth of the population. Population growth is often quantified as the amount of carbon that is newly added to the overall population. For nutritional classifications, the source of carbon groups organisms into autotrophs that fix inorganic carbon to build biomass and heterotrophs that use organic carbon to synthesize biomass. Note that the classification of organisms according to carbon source is made based only on the form of carbon used as building blocks. It is independent of the direct energy source (light, chemical, or electric) to generate ATP, even though the carbon may be used in the energy-source reaction. The Calvin-Benson-Bassham (CBB) cycle is the only carbon-fixation pathway in eukaryotes (algae and plants). Some autotrophic bacteria also fix carbon dioxide in the CBB cycle, while others have evolved unique, inorganic carbon-fixation pathways (Berg et al., 2010; Berg, 2011; Hügler and Sievert, 2011).

According to these two classifications, plants and phytoplankton are photo-autotrophs, while animals and most fungi are chemo-heterotrophs. In subsurface ecosystems where there is no light, microbial communities are composed of chemo-autotrophs and chemo-heterotrophs. Here, we exclude discussions of photo-heterotrophic bacteria, but this nutritional group may be abundant and comprise approximately 10% of the microbial community in the euphotic zone in the upper ocean (Kolber et al., 2001). Some bacteria and archaea are known for their versatile metabolic strategies, which enable a cell to use both inorganic and organic carbon as building blocks, or the different types of energy source reactions (Eiler, 2006; Hügler and Sievert, 2011). These organisms that can switch metabolic strategies in response to environmental conditions are called mixotrophs.

For surface ecosystems, the primary and basic energy source is light. Plants and phytoplankton as producers harness light energy to synthesize ATP, part of which is used to fix carbon dioxide (inorganic carbon) into biomass. The direct energy source of consumers and decomposers is chemical as they generate ATP from aerobic respiration, a reaction using oxygen and organic substrates. Since both of these substrates are the by-products of photosynthesis, the original energy source for consumers and decomposers is light energy. Although consumers and decomposers also allocate organic carbon to the building blocks of biomass, carbon flows are a good indicator of energy flow

within feeding interactions because it is closely linked with the energy transfer among organisms that use organic substrates as their energy source.

In contrast, the ultimate energy source of subsurface ecosystems should be chemical energy (including electrical energy), which may not rely on the supply of organic substrates as the photosynthetic by-product from surface ecosystems. Direct or indirect carbon transfers from light-independent producers to chemoorganoheterotrophs that belong to the same nutritional group as animals can also be linked to energy transfers through the microbial community. However, the energy consumption of light-independent producers and the consequential interspecific interactions among them highlights the unique relationship between population growth and energy, which can be understood as the use of energy as opposed to organic carbon.

CHARACTERISTIC INTERSPECIFIC INTERACTIONS AMONG LIGHT-INDEPENDENT PRODUCERS

Of the interspecific interactions of light-dependent producers, resource competition is an important factor that determines growth, fitness, and productivity. Although the ultimate limiting resource for light-dependent producers can be the availability of essential nutrients (primarily nitrogen and phosphorus as the building blocks of biomass), water, and/or space, only those interactions associated with the availability of light energy as a limiting resource is discussed here. However, it should be noted that the availability of nutrients and water affect the ability of light-dependent producers to capture light energy. For example, a phosphorus deficiency in chloroplast stroma inhibits ATP synthase activity (Takizawa et al., 2008; Karlsson et al., 2015; Carstensen et al., 2018).

There is experimental, observational, and theoretical evidence indicating that plants and phytoplankton may specialize in separate light niches in response to different light environments (Litchman, 2000; Litchman and Klausmeier, 2001; Silvertown, 2004; Six et al., 2007; Sterck et al., 2011). Although some shade-tolerant plants may acquire new niches due to the presence of others, plants reside in the same canopy are often competitive and they never increase the light intensity, or light energy (**Figure 2A**).

The availability of light energy in a terrestrial ecosystem is mainly determined by the amount of sunlight received at the Earth's surface. An example of a model that describes the availability of light, L, under a canopy is,

$$L = L_0 f\left(\sum_i c_i B_i\right) \tag{2}$$

where L_0 is the light intensity at the top of the canopy (the light available prior to any consumption), c_i is the light adsorption coefficient of species i, and B_i is the biomass of species i. f describes the total light consumption by the existing species, which is a decreasing function with a maximum value of 1 that holds when no plants exist (Grovar, 1997). Any increase in B_i is associated with the consumption of light and, thus, decreases the availability of light and inhibits the light-dependent growth rate of other species.

Interspecific interactions via the use of chemical energy also significantly affect the growth of light-independent producers, although not always in a negative manner. The Gibbs energy, G, in thermodynamics is a measure of the availability of chemical energy in a system. The amount of Gibbs energy depends on all of the chemical substances present in a system and is described as:

$$G = \sum n_j \mu_j \tag{3}$$

where n_j is the number of moles of substance j, and $\mu_j = \left(\frac{\partial G}{\partial n_j}\right)_{T,P,n_{k\neq j}}$ denotes the partial molar Gibbs energy, or chemical potential, when temperature, pressure, and the amount of all other substances are held constant. μ_j depends not only on the substance j, but on other substances in the system, as the interaction of substances affects the availability of chemical potential μ_j. Unlike the change of L in Eq. (2), which is described by direct consumption by plant species, the change in G can only be indirectly described by changes in the molar concentrations of reactants and products associated with the progression of the energy-source reaction (**Figure 2B**).

The Gibbs energy change of a reaction, $\Delta_r G$, is an important and useful thermodynamic property that enables us to estimate the availability of chemical energy for microbes under a given environmental condition. $\Delta_r G$ denotes the amount of chemical energy dissipated from a reaction on a kJ per mole basis. Hence, a negative $\Delta_r G$ (i.e., $-\Delta_r G$) denotes the maximum available energy for species when it catalyzes one mole of reaction to dissipate energy. If species 1 generates energy from the following overall reaction, then:

$$A_1 + e^- \rightarrow A_2 \qquad \text{(A reduction)}, \tag{4a}$$

$$B_2 \rightarrow B_1 + e^- \qquad \text{(B oxidation)}, \tag{4b}$$

$$A_1 + B_2 \rightarrow A_2 + B_1 \qquad \text{(reaction 1 : the overall reaction of Eqs. (4a) and (4b))}, \tag{4c}$$

where X_1 and X_2 are two different compounds (or two different redox states) composed of the same element. X_2 has more

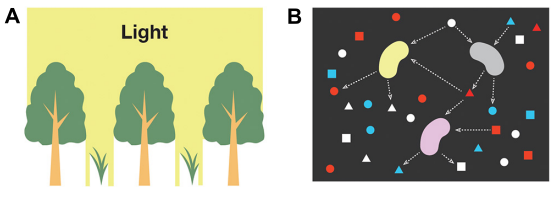


FIGURE 2 | Interspecific interactions associated with the utilization of **(A)** light and **(B)** chemical energy.

electrons than X_1. In Eq. (4c), B_2 serves as an electron donor and A_1 serves as an electron acceptor. The negative Gibbs energy change of this reaction is

$$-\Delta_r G = -\Delta_r G^\circ + RT \ln \frac{\{A_1\}\{B_2\}}{\{A_2\}\{B_1\}} \quad (4d)$$

where $-\Delta_r G^\circ$ is the standard Gibbs energy change of reaction 1 (kJ mol^{-1}), R is the gas constant (8.31×10^{-3} kJ K^{-1} mol^{-1}), T is the absolute temperature (K), and $\{X_j\}$ is the activity of X_j. Activity can be approximated as the molar concentration when the interaction forces of the substances can be ignored: $\{X_j\} \approx n_j/v$, where v is the volume. $^\circ$ indicates the standard-state condition where all reactants and products (chemical substances utilized for and released from the energy-source reaction) have activity = 1. $-\Delta_r G^\circ$ is given by,

$$-\Delta_r G^\circ = \Delta_f G_{A_1}{}^\circ + \Delta_f G_{B_2}{}^\circ - \left(\Delta_f G_{A_2}{}^\circ + \Delta_f G_{B_1}{}^\circ\right), \quad (4e)$$

where $\Delta_f G_x{}^\circ$ denotes the Gibbs energy change accompanied by the x formation (here, $x = A_1, A_2, B_1$, or B_2). $\Delta_f G_x{}^\circ$ is the relative level of Gibbs energy of x from all of the reference states of the elements composing x [e.g., the relative level of Gibbs energy of CH_4 in comparison with C(solid) as graphite and H_2 (gas)]. $\Delta_f G_x{}^\circ$ is a constant that is intrinsic to substance x under standard conditions. As $\Delta_r G^\circ$ is the difference between the sum of $\Delta_f G^\circ$ for the by-products and the sum of $\Delta_f G^\circ$ for the reactants, it is a reaction-specific quantity that is determined by the combination of reactants and by-products. $\Delta_f G_x{}^\circ$ values are available from the CHNOSZ package for R[1]. The second term of $-\Delta_r G$ in the right-hand side of Eq. (4d) indicates the effect of the relative quantities of products and reactants on the available chemical energy. This is termed "abundant resource premium" (ARP) (Seto and Iwasa, 2019a) and is especially important for the growth of light-independent producers, which results in characteristic mutualistic interactions, or obligately mutualistic metabolism.

The growth of light-independent producers always decreases the total amount of Gibbs energy of a system; similarly, the growth of light-dependent producers reduces the availability of light. However, unlike light consumption by plant species, which often results in competition, the progression of reaction 1 does not always decrease the availability of chemical energy for other light-independent producers. When species 2 harnesses a reaction that utilizes A_2 we get,

$$A_2 + C_2 \rightarrow A_3 + C_1 \text{ (reaction 2)}, \quad (5a)$$

where the negative Gibbs energy change of reaction 2 is

$$-\Delta_r G = -\Delta_r G^\circ + RT \ln \frac{\{A_2\}\{C_2\}}{\{A_3\}\{C_1\}}. \quad (5b)$$

To distinguish $\Delta_r G$ in Eqs. (4d) and (5b), from here we will use the symbol $\Delta_r G_i$, where subscript i denotes the reaction number. Because the supply of A_2 increases $-\Delta_r G_2$ due to the presence of the ARP term, species 2 may generate more energy when species 1 exists (**Figure 2B**). The consumption of A_2 also increases $-\Delta_r G_1$, which consequentially increases the energy gain of species 1. Hence, the consumption of the by-product of the energy-harvesting reaction can lead to mutualism between species 1 and 2 owing to the presence of ARP. An example of a mutualistic interaction which is similar to this type occurs between ammonia-oxidizers and nitrite-oxidizers in **Table 1** [the combination of reactions (VIII) and (X)], which are microbes interacting via the supply and consumption of nitrite ion ($NO_2{}^-$). Although these microbes compete for oxygen, one can predict that the presence of a nitrite-oxidizer may increase the energy acquisition per reaction of an ammonia-oxidizer under oxic conditions where the availability of oxygen may not limit the growth of nitrite-oxidizer and ammonia-oxidizer. These microbes have been confirmed to occur in clusters and be in frequent contact with each other within sludge flocs (Mobarry et al., 1996).

A more intuitive example of the mutualism between light-independent producers that completes material recycling within their energy-source reaction is a combination of Eq. (4c) and

$$A_2 + C_1 \rightarrow A_1 + C_2 \text{ (reaction 3)}, \quad (6a)$$

where the negative Gibbs energy change of reaction 3 is

$$-\Delta_r G_3 = -\Delta_r G_3{}^\circ + RT \ln \frac{\{A_2\}\{C_1\}}{\{A_1\}\{C_2\}}. \quad (6b)$$

An example can be found in the iron cycle between Fe^{2+} and Fe oxides (e.g., the combination of reaction (IV) and (XI) or (XIII) in **Table 1**) such as Fe $(OH)_3$ that is readily formed in the presence of Fe^{3+} at a circumneutral pH. Experimental studies have reported the possibility of micro-scale bacterial iron cycling at oxic-anoxic interfaces (Emerson and Revsbech, 1994; Sobolev and Roden, 2002; Roden et al., 2004). Another example is the sulfate cycle, which is closely associated with the methane and nitrogen cycles [e.g., the combination of reactions (II) and (XV) in **Table 1**] (Lau et al., 2016).

POPULATION DYNAMICS UNDER AVAILABLE CHEMICAL-ENERGY LIMITATIONS

The availability of chemical energy changes in response to changes in the amount of chemical substances in the surrounding environment, however, this effect has been neglected in population dynamics models in traditional microbial ecology. The population growth of light-dependent producers (plants and phytoplankton) is often limited by the availability of energy. Similarly, the population dynamics of light-independent producers are often limited by the acquisition rate of chemical energy. Interestingly, modeling this provides a framework for understanding mutualistic interactions among light-independent

[1]http://chnosz.net

producers, which has not been previously recognized in the ecological sciences. In a previous paper, we introduced a microbial population dynamics model in relation to energy production per reaction that included the ARP term (Seto and Iwasa, 2019a). For a microbial population that harvests energy from reaction i,

$$\frac{dx_i}{dt} = q_i \{c_i F_i (-\Delta_r G_i) - m_i\} x_i \tag{7a}$$

where F_i denotes the biomass-specific catalytic rate of the energy-source reaction, q_i is the biomass yield of species i for a given energy gain from ATP that is generated from an energy-source reaction, c_i is the fraction of energy that can be used for ATP synthesis ($0 < c_i < 1$) and excludes energy expenditure such as loss by heat transfer, and m_i is the rate of energy loss that occurs to maintain cellular activity. For simplicity, we call x_i the population density of the ith species, but it is more appropriate to call it the ith functional group, as multiple species typically share the same energy-source reaction. The term in curly brackets in Eq. (7a) describes the biomass-specific rate of the net energy gain. A reaction requiring two substrates (an electron acceptor and an electron donor) may be simultaneously limited by the low availability of both substrates. There are several proposed models that encompass dual-limitations from the abundance of electron donors and acceptors (Bae and Rittmann, 1996; Jin and Bethke, 2003, 2007; LaRowe et al., 2012). Here we have adopted the simplest model—the double-Monod model. For Eq. (4c) (reaction 1) it is described as,

$$F_1 = r_1 \frac{A_1}{K_{A_1} + A_1} \frac{B_2}{K_{B_2} + B_2} \tag{7b}$$

where r_1 is the maximum microbial catalytic rate of reaction 1, and K_X is the half-saturation constant for X.

The growth of chemoheterotrophs (e.g., aerobic bacteria and fermentation bacteria) is often described by the Monod equation, where the growth rate of microbial population x appears to be limited solely by the availability of the concentration of the organic substrate R:

$$\frac{dx}{dt} = Fx \quad \text{where} \quad F = \frac{rR}{K + R}. \tag{8}$$

Equation (7a) can be approximated as Eq. (8) if the following three conditions are satisfied: (i) the decreasing rate – mx is added to Eq. (8), (ii) the concentration of one of the two substrates is significantly higher than its half-saturation constant such that the corresponding factor is replaced by 1, and (iii) the effect of ARP on $-\Delta_r G$ can be ignored because of the high value of $-\Delta_r G^\circ$. Assumptions (ii) and (iii) are plausible for aerobic bacteria and fermentation bacteria. The incubation of aerobic bacteria is usually conducted under aerobic conditions (i.e., high oxygen concentration), whereas fermentation bacteria only require a single organic substrate, which is used to generate energy from substrate-level phosphorylation. Additionally, the energy-source reaction that uses the organic substrate generally has a high $-\Delta_r G^\circ$. The relative abundance of reactants and by-products can affect $-\Delta_r G$ only when the contribution of ARP on $-\Delta_r G$ is so large that $-\Delta_r G$ cannot be approximated as $-\Delta_r G^\circ$. When the availability of reactants is significantly higher than products, for example $(\{A1\}\{B2\})/(\{A2\}\{B1\}) = 10^6$, ARP is approximately 33.5 kJ mol^{-1} at 1 bar and 298.15 K. The ARP increases with increasing temperature if all other variables and parameters are fixed. However, the activities of chemical substances and $-\Delta_r G^\circ$ often depend on temperature sensitivity, which may occasionally decrease $-\Delta_r G$ with increasing temperature. As shown in Eq. 4e, the value of $-\Delta_r G^\circ$ is determined by the combination of reactants and products under constant temperature and pressure conditions. The $-\Delta_r G^\circ$ of light-independent producers is typically lower than 1000 kJ mol^{-1}, whereas $-\Delta_r G^\circ$ can exceed 2000 kJ mol^{-1} for aerobic respiration when glucose is used as an electron donor. This suggests that the ARP term does not affect the growth dynamics of microbes that harness aerobic respiration, but in the case of light-independent producers, it may have a significant effect on microbial growth and steady-state behavior.

Due to the ARP effect, the model's behaviors are qualitatively different from those observed in traditional microbial population growth models that follow the Monod-type equation (Seto and Iwasa, 2019a). The steady-state population density of the Monod-type equation increases with increasing r_i and decreasing K_X, as these changes increase the resource-utilizing ability or the resource consumption rate of the species. Meanwhile, the steady population density of the population dynamics that follow Eq. (7) is maximized at an intermediate r_i and K_X. This is because the increase in the resource-utilizing ability positively affects population growth when the abundance of the reactants is relatively larger than that of the by-products, but it negatively affects the energy acquisition per reaction as by-products accumulate in the surroundings, or $-\Delta_r G$ becomes low.

The unique steady-state response of light-independent producers to r_i and K_X may provide them with a distinctive resource-utilizing strategy. If several species compete for the same resource, the steady-state analysis of the Monod-type population dynamics model predicts that the species that can deplete the abundance of resources to the lowest level will exclude others at the steady-state (R^*-rule; Tilman, 1982), which may act as a major selective force for many microbial species. Following the R^*-rule, when all of the species have the same parameter values except for r, only the species with the highest r will eventually dominate the system and maintain the highest population-density level. In contrast, in our model, if several species harness the same energy-source reaction with the same parameter values except for r coexist within a system, the species with the highest r will also exhibit the fastest growth rate and dominate the system, but will maintain a relatively lower population density. Modeling studies have predicted that a small population is more strongly threatened by stochastic processes than a large population if the population is isolated, which would subsequently cause local extinction (Matthies et al., 2004). This implies that subsurface ecosystems may shape a

unique evolution of specialization in the resource utilization of light-independent producers.

ARP-DRIVEN MUTUALISTIC INTERACTIONS

Understanding two-species population dynamics is the first step toward identifying the complex interspecific interactions among light-independent producers. We previously undertook mathematical analyses of two models, the structures of which are illustrated in **Figure 3**. These models deal with mutualistic interspecific interactions that occur through the use of different chemical forms of an element labeled as A_1, A_2, and A_3 (Seto and Iwasa, 2019b, 2020), where a larger subscript number denotes having more electrons. Although light-independent producers can compete for the same resource(s) used in energy-source reactions, we focused on two cases of mutualism where, (i) species 2 uses the by-product of species 1 (the combination of Eqs. (4c) and (5a), i.e., reactions 1 and 2), and (ii) each species receives its own resources from the other species and provides them with the by-products as a resource (the combination of Eqs. (4c) and (6a), i.e., reactions 1 and 3). We call the former type of interspecific interaction a one-way interaction, and the latter a recycling interaction.

For both models, the presence of species 2 increases the steady-state biomass of species 1 and expands the realized niche of species 1 from its fundamental niche. For the one-way interaction model, the interspecific interaction can only be commensalism in the absence of ARP (i.e., the second term on the right-hand side of $-\Delta_r G$ is eliminated) because the presence of species 1, which supplies the resources for species 2, can expand the conditions for the survival of species 2 and its realized niche. This is a standard effect known as "niche construction" or "niche changing" (Erwin, 2008; Laland et al., 2016). The ARP term allows species 1 to increase the fitness of species 1 in the presence of species 2 because species 2 reduces the abundance of the by-product of species 1, resulting in the expansion of the realized niches of species 1. Hence, both species are able to live in conditions in which neither of them could survive alone. Meanwhile, for the recycling interaction model, the same tendency is true for a model with and without the ARP term. As confirmed by the analysis of the one-species model, an increase in the resource-utilizing ability of species in the two models eventually leads to a decline in steady-state population density. However, the presence of a mutualistic partner allows a species to keep its steady-state population density at a higher level than when it would if alone. This implies that a species with a higher resource-utilizing ability may be able to robustly survive in a system if a species that works in a mutually beneficial manner is present and in close proximity.

The mutualistic relationship observed between species 1 and 2 occurs when the population growth of both species depends on the amount of chemical energy including the ARP term. If the growth rate is limited by not the energy acquisition process (catabolism) but the biosynthetic process (anabolism), these two species may compete for essential nutrients to build biomass. This suggests that interspecific interactions among light-independent producers may change depending on which process ultimately limits the microbial growth rate.

The presence of a mutually beneficial partner increases the rate of the energy-source reaction, which may drive the geochemical process that proceeds very slowly in the absence of light-independent producers. Species 1 in **Figure 3B** is able to survive only in the presence of species 2 that sufficiently accelerates the reaction from A_1 to A_2 when it is unlikely to carry forward abiotically. A reaction spontaneously proceeds with releasing energy when $\Delta_r G < 0$ (i.e., $-\Delta_r G > 0$). However, the rate of reaction is not determined by the value of $\Delta_r G$, but, rather, it depends on activation energy and environmental factors (e.g., temperature, coexisting chemical substances, water availability, and pH), which affect the intensity of the competition between abiotic and microbial reactions (Kirby et al., 1999; Melton et al., 2014; Liu et al., 2017). Microbes, like enzymes in biochemical systems, can often accelerate the reaction rate by lowering the energetic barrier required for the reaction to commence. For instance, in the presence of the iron-oxidizing bacteria *Thiobacillus ferrooxidans*, the iron oxidation rate increases up to several orders of magnitude in comparison with the rate under abiotic controls (Lacey and Lawson, 1970; Nordstrom, 1985).

In contrast, rapid abiotic reactions may inhibit the growth of microbes that compete for the same resources. Under neutral pH conditions, it is hypothesized that the growth of iron-oxidizing bacteria is constrained under low oxygen conditions, where microbial iron oxidation can outcompete rapid abiotic iron oxidation (Druschel et al., 2008). Micro-scale iron recycling between iron-oxidizing and -reducing bacteria may enable iron-oxidizing bacteria to overcome abiotic iron oxidation.

DISCUSSION

The Productivity of Subsurface Microbial Ecosystems

The presence of abundant microbial biomass in subsurface ecosystems implies there is considerable productivity of light-independent producers. However, research on the productivity of subsurface microbial ecosystems has been advanced only recently due to limited access to the subsurface biosphere.

Estimations of the productivity of hydrothermal vent ecosystems based on calculations of $\Delta_r G$ have suggested that the productivity of light-independent producers contributes only a small fraction of the total photosynthetic biomass produced in the oceans (McCollom and Shock, 1997; McCollom, 2007). The vast biomass stock, despite low productivity in the subsurface biosphere, may have three explanations. First, the energy sources of subsurface ecosystems may be heavily dependent on organic substrates supplied by surface ecosystems. The dominance of heterotrophic microbes that use organic carbon to build biomass has been confirmed for some sedimentary subsurface ecosystems (Biddle et al., 2006; Morono et al., 2011). Second, microorganisms might adapt or acclimate to energetic stress by lowering the maintenance energy (m in our model). To thrive under limited access to energy sources, microbes in the

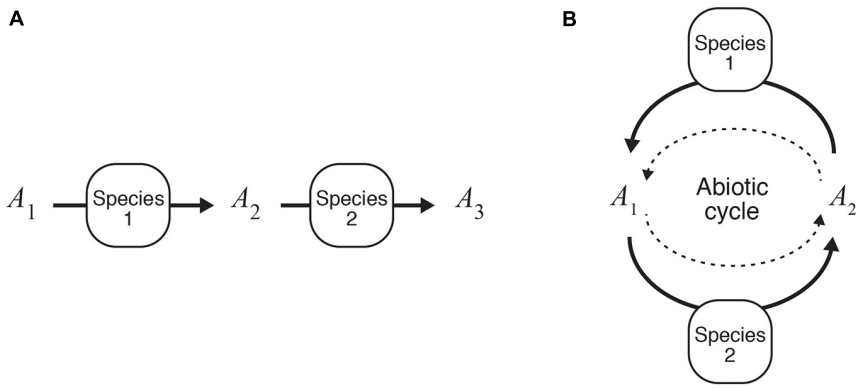


FIGURE 3 | Mutualistic catabolic interactions. **(A)** One-way interactions (Seto and Iwasa, 2019b). **(B)** Recycling interactions (Seto and Iwasa, 2020).

subsurface have often been observed to reduce maintenance energy by switching their metabolic modes; and, as a result, the mean cell-doubling time may then range from a few years to thousands of years (Jørgensen and Boetius, 2007; Valentine, 2007; Hoehler and Jørgensen, 2013). Finally, microbes under low oxygen environments require less energy to synthesize biomass than those under oxygen-rich conditions, which may result in relatively high biomass conversion efficiency (q in our model) in the subsurface biosphere. Nitrogen and sulfur sources in oxygen-rich conditions often exist as NO_3^- and SO_4^{2-}. To generate amino acids as building blocks of biomass, NO_3^- and SO_4^{2-} should be converted into ammonia and hydrogen sulfide ions, and, in the presence of oxygen, more energy is required to carry out these reactions (McCollom and Amend, 2005).

An interesting question is whether the biodiversity-ecosystem functioning (BEF) relationship in subsurface microbial ecosystems is similar across surface and subsurface ecosystems. Controlled experiments in terrestrial ecosystems have shown increased plant biomass and primary productivity with increasing plant diversity (Hector et al., 1999; Tilman et al., 2001), whereas a mechanistic model that incorporates a food-web structure has demonstrated that plant biomass does not always increase with plant diversity (Thébault and Loreau, 2003). Studies examining BEF relationships for soil microorganisms reported somewhat contradictory results; some studies emphasized the link between the physiological or functional diversity and ecosystem functions, especially process rates or activities (e.g., nitrification rate and denitrification activity), in soil microorganisms (Balser and Firestone, 2005; Bell et al., 2005; Webster et al., 2005; Philippot et al., 2013), and the other demonstrated the irrelevance of microbial diversity to ecosystem functions (especially carbon mineralization) (Balser and Firestone, 2005; Wertz et al., 2006). Ecosystem functions of microbes are often evaluated by the process rates of carbon and nitrogen, but as we introduced so far, the energy-source reactions of light-independent producers drive the flows of not only carbon and nitrogen but also other elements such as iron and sulfur. This may result in the underestimation of the effect of microbial diversity on their ecosystem functioning. The extension of our model would help to provide a theoretical framework for understanding the BEF relationship for light-independent producers and overall subsurface microbial ecosystems, including heterotrophic microbes. It would be interesting to address whether or how much the supply of organic substrates from surface ecosystems controls the productivity and biodiversity of these ecosystems, as suggested previously (Hubalek et al., 2016; Magnabosco et al., 2018).

Ecology of Light-Independent Producers as an Interdisciplinary Field

The application of a population dynamics model to the community of light-independent producers will develop our understanding of not only subsurface ecology, but environmental science/engineering, astrobiology, and the evolution of early life and ecosystems.

Microbial fuel cells are a technology that converts chemical energy to electricity by using electron transfer metabolisms (Du et al., 2007). While research has concentrated on the conversion of organic substrates to electricity, the application of light-independent producers is also expected to produce carbon compounds from CO_2 in a renewable manner. However, because of low product yields and productivity, the practical application of light-independent producers requires improved efficiency under industrial conditions (Claassens et al., 2016). For some microbial fuel cells that use organic substrates, higher power generation in co-cultures, as opposed to individual strains, has been reported (Logan et al., 2019). While more realistic modeling is needed, our approach will be useful in understanding productivity in co-cultures by linking population dynamics with the availability of chemical energy. This application will also be useful in understanding the different efficiencies of wastewater treatment by light-independent producers in pure and co-cultures (e.g., Parshina et al., 2005).

Theoretical, experimental, and field investigations have established the physical and chemical limits of microbial life under extreme conditions. In the field of astrobiology, the potential survival and existence of life on other planets has been debated based on these limitations. The activity of microbes, including light-independent producers, helps in the search for

life on extraterrestrial planets and in planetary protection against microbial contamination from and to Earth in the course of space missions. Although microbial life can be limited by multiple factors (e.g., the availability of water, gravity, temperature, pressure, and the intensity of UV/ionizing radiation) (Rummel et al., 2014; Moissl-Eichinger et al., 2016), the presence of the potential energy-source reaction is one important factor that restricts microbial activity. The limits on microbial growth are often measured based on single-culture experiments, but, as our results suggest, the presence of a mutualistic counterpart may expand limits on microbial growth. This suggests that the limits on microbial growth require an understanding of the effects of biological interactions, in addition to abiotic factors.

Although the energy source for the earliest life is still under debate, both phylogenetic and biochemical evidence suggest that the earliest life forms were chemoautotrophs (Pace, 1997; Lane et al., 2010; Judson, 2017), which may have used hydrogen and carbon dioxide as geochemical sources (Weiss et al., 2016). The early Earth was deficient in oxygen and organic matter previous to the evolution of oxygenic photosynthetic organisms, which limited the energy-source reactions for the light-independent producers listed in **Table 1**. One interesting question deals with how microbial metabolic diversity evolved and was constructed in the earliest ecosystem under extreme energy limitations. If the earliest life forms depended on chemical energy, the most likely place of origin would be the hydrothermal vent system, which creates an interface between thermal fluids with electron-abundant substances and the relatively electron-deficient substances in the bottom water. Another interesting question is how life could have abandoned this attractive energy source to expand the ecosystems. Our results suggest that the evolution of a mutualistic metabolic counterpart may have enabled life under extreme energy limitations to be freed from energetic constraints and invade other systems (Seto and Iwasa, 2019b, 2020). Especially for microbes harnessing reactions with low $-\Delta_r G^\circ$ and those that do not use either organic substrates or oxygen, the ARP may play an important role in potential growth, survival and community structure.

AUTHOR CONTRIBUTIONS

Both authors contributed to the writing of the manuscript.

FUNDING

This work was supported by JSPS Grant-in-Aid for Scientific Research (C) [Grant Number 19K06853] awarded to MS.

ACKNOWLEDGMENTS

We thank the following people for their helpful comments: K. Kadowaki, K. Koba, M. Kondoh, Y. Tachiki, and K. Uriu. We also thank NM and SDV for their comments.

REFERENCES

Bae, W., and Rittmann, B. E. (1996). A structured model of dual-limitation kinetics. *Biotechnol. Bioeng.* 49, 683–689. doi: 10.1002/(SICI)1097-0290(19960320)49:6<683::AID-BIT10<3.0.CO;2-7

Balser, T. C., and Firestone, M. K. (2005). Linking microbial community composition and soil processes in a California annual grassland and mixed-conifer forest. *Biogeochemistry* 73, 395–415. doi: 10.1007/s10533-004-0372-y

Bar-On, Y. M., Phillips, R., and Milo, R. (2018). The biomass distribution on Earth. *Proc. Natl. Acad. Sci. U.S.A.* 115, 6506–6511. doi: 10.1073/pnas.1711842115

Bell, T., Newman, J. A., Silverman, B. W., Turner, S. L., and Lilley, A. K. (2005). The contribution of species richness and composition to bacterial services. *Nature* 436, 1157–1160. doi: 10.1038/nature03891

Berg, I. A. (2011). Ecological aspects of the distribution of different autotrophic CO2 fixation pathways. *Appl. Environ. Microbiol.* 77, 1925–1936. doi: 10.1128/AEM.02473-10

Berg, I. A., Kockelkorn, D., Ramos-Vera, W. H., Say, R. F., Zarzycki, J., Hügler, M., et al. (2010). Autotrophic carbon fixation in archaea. *Nat. Rev. Microbiol.* 8, 447–460. doi: 10.1038/nrmicro2365

Biddle, J. F., Lipp, J. S., Lever, M. A., Lloyd, K. G., Sørensen, K. B., Anderson, R., et al. (2006). Heterotrophic Archaea dominate sedimentary subsurface ecosystems off Peru. *Proc. Natl. Acad. Sci. U.S.A.* 103, 3846–3851. doi: 10.1073/pnas.0600035103

Burgin, A. J., Yang, W. H., Hamilton, S. K., and Silver, W. L. (2011). Beyond carbon and nitrogen: how the microbial energy economy couples elemental cycles in diverse ecosystems. *Front. Ecol. Environ.* 9, 44–52. doi: 10.1890/090227

Carstensen, A., Herdean, A., Schmidt, S. B., Sharma, A., Spetea, C., Pribil, M., et al. (2018). The impacts of phosphorus deficiency on the photosynthetic electron transport chain. *Plant Physiol.* 177, 271–284. doi: 10.1104/pp.17.01624

Ciobanu, M. C., Burgaud, G., Dufresne, A., Breuker, A., Rédou, V., Ben Maamar, S., et al. (2014). Microorganisms persist at record depths in the subseafloor of the Canterbury Basin. *ISME J.* 8, 1370–1380. doi: 10.1038/ismej.2013.250

Claassens, N. J., Sousa, D. Z., Dos Santos, V. A. P. M., De Vos, W. M., and Van Der Oost, J. (2016). Harnessing the power of microbial autotrophy. *Nat. Rev. Microbiol.* 14, 692–706. doi: 10.1038/nrmicro.2016.130

Colman, D. R., Poudel, S., Stamps, B. W., Boyd, E. S., and Spear, J. R. (2017). The deep, hot biosphere: twenty-five years of retrospection. *Proc. Natl. Acad. Sci. U.S.A.* 114, 6895–6903. doi: 10.1073/pnas.1701266114

Dimroth, P. (1987). Sodium ion transport decarboxylases and other aspects of sodium ion cycling in bacteria. *Microbiol. Rev.* 51, 320–340. doi: 10.1128/mmbr.51.3.320-340.1987

Druschel, G. K., Emerson, D., Sutka, R., Suchecki, P., and Luther, G. W. (2008). Low-oxygen and chemical kinetic constraints on the geochemical niche of neutrophilic iron(II) oxidizing microorganisms. *Geochim. Cosmochim. Acta* 72, 3358–3370. doi: 10.1016/j.gca.2008.04.035

Du, Z., Li, H., and Gu, T. (2007). A state of the art review on microbial fuel cells: a promising technology for wastewater treatment and bioenergy. *Biotechnol. Adv.* 25, 464–482. doi: 10.1016/j.biotechadv.2007.05.004

Eiler, A. (2006). Evidence for the ubiquity of mixotrophic bacteria in the upper ocean: implications and consequences. *Appl. Environ. Microbiol.* 72, 7431–7437. doi: 10.1128/AEM.01559-06

Emerson, D., and Revsbech, N. P. (1994). Investigation of an iron-oxidizing microbial mat community located near Aarhus, Denmark: field studies. *Appl. Environ. Microbiol.* 60, 4022–4031.

Erwin, D. H. (2008). Macroevolution of ecosystem engineering, niche construction and diversity. *Trends Ecol. Evol.* 23, 304–310. doi: 10.1016/j.tree.2008.01.013

Falkowski, P. G., Fenchel, T., and Delong, E. F. (2008). The microbial engines that drive Earth's biogeochemical cycles. *Science* 320, 1034–1039. doi: 10.1126/science.1153213

Flemming, H. C., and Wuertz, S. (2019). Bacteria and archaea on Earth and their abundance in biofilms. *Nat. Rev. Microbiol.* 17, 247–260. doi: 10.1038/s41579-019-0158-9

Glombitza, C., Adhikari, R. R., Riedinger, N., Gilhooly, W. P., Hinrichs, K. U., and Inagaki, F. (2016). Microbial sulfate reduction potential in coal-bearing sediments down to ˜2.5 km below the seafloor off Shimokita Peninsula, Japan. *Front. Microbiol.* 7:1576. doi: 10.3389/fmicb.2016.01576

Gold, T. (1992). The deep, hot biosphere. *Proc. Natl. Acad. Sci. U.S.A.* 89, 6045–6049. doi: 10.1073/pnas.89.13.6045

Grovar, J. P. (1997). *Resource Competition*. London: Chapman and Hall.

Hector, A., Schmid, B., Beierkuhnlein, C., Caldeira, M. C., Diemer, M., Dimitrakopoulos, P. G., et al. (1999). Plant diversity and productivity experiments in European grasslands. *Science* 286, 1123–1127. doi: 10.1126/science.286.5442.1123

Hoehler, T. M., and Jørgensen, B. B. (2013). Microbial life under extreme energy limitation. *Nat. Rev. Microbiol.* 11, 83–94.

Hubalek, V., Wu, X., Eiler, A., Buck, M., Heim, C., Dopson, M., et al. (2016). Connectivity to the surface determines diversity patterns in subsurface aquifers of the Fennoscandian shield. *ISME J.* 10, 2447–2458. doi: 10.1038/ismej.2016.36

Hügler, M., and Sievert, S. M. (2011). Beyond the Calvin cycle: autotrophic carbon fixation in the ocean. *Annu. Rev. Mar. Sci.* 3, 261–289. doi: 10.1146/annurev-marine-120709-142712

Inagaki, F., Hinrichs, K. U., Kubo, Y., Bowles, M. W., Heuer, V. B., Hong, W. L., et al. (2015). Exploring deep microbial life in coal-bearing sediment down to ∼2.5 km below the ocean floor. *Science* 349, 420–424. doi: 10.1126/science.aaa6882

Jin, Q., and Bethke, C. M. (2003). A new rate law describing microbial respiration. *Appl. Environ. Microbial.* 69, 2340–2348. doi: 10.1128/AEM.69.4.2340

Jin, Q., and Bethke, C. M. (2007). The thermodynamics and kinetics of microbial metabolism. *Am. J. Soc.* 307, 643–677.

Jørgensen, B. B., and Boetius, A. (2007). Feast and famine – microbial life in the deep-sea bed. *Nat. Rev. Microbiol.* 5, 770–781. doi: 10.1038/nrmicro1745

Judson, O. P. (2017). The energy expansions of evolution. *Nat. Ecol. Evol.* 1:0138. doi: 10.1038/s41559-017-0138

Karlsson, P. M., Herdean, A., Adolfsson, L., Beebo, A., Nziengui, H., Irigoyen, S., et al. (2015). The *Arabidopsis* thylakoid transporter PHT4;1 influences phosphate availability for ATP synthesis and plant growth. *Plant J.* 84, 99–110. doi: 10.1111/tpj.12962

Kimura, S., Bryan, C. G., Hallberg, K. B., and Johnson, D. B. (2011). Biodiversity and geochemistry of an extremely acidic, low-temperature subterranean environment sustained by chemolithotrophy. *Environ. Microbiol.* 13, 2092–2104. doi: 10.1111/j.1462-2920.2011.02434.x

Kirby, C. S., Thomas, H. M., Southam, G., and Donald, R. (1999). Relative contributions of abiotic and biological factors in Fe(II) oxidation in mine drainage. *Appl. Geochem.* 14, 511–530. doi: 10.1016/S0883-2927(98)00071-7

Kolber, Z. S., Plumley, F. G., Lang, A. S., Beatty, J. T., Blankenship, R. E., VanDover, C. L., et al. (2001). Contribution of aerobic photoheterotrophic bacteria to the carbon cycle in the ocean. *Science* 292, 2492–2495. doi: 10.1126/science.1059707

Lacey, D. T., and Lawson, F. (1970). Kinetics of the liquid−phase oxidation of acid ferrous sulfate by the bacterium *Thiobacillus ferrooxidens*. *Biotechnol. Bioeng.* 12, 29–50. doi: 10.1002/bit.260120104

Laland, K., Matthews, B., and Feldman, M. W. (2016). An introduction to niche construction theory. *Evol. Ecol.* 30, 191–202. doi: 10.1007/s10682-016-9821-z

Lane, N., Allen, J. F., and Martin, W. (2010). How did LUCA make a living? Chemiosmosis in the origin of life. *Bioessays* 32, 271–280. doi: 10.1002/bies.200900131

LaRowe, D. E., Dale, A. W., Amend, J. P., and Van Cappellen, P. (2012). Thermodynamic limitations on microbially catalyzed reaction rates. *Geochim. Cosmochim. Acta* 90, 96–109.

Lau, M. C. Y., Kieft, T. L., Kuloyo, O., Linage-Alvarez, B., Van Heerden, E., Lindsay, M. R., et al. (2016). An oligotrophic deep-subsurface community dependent on syntrophy is dominated by sulfur-driven autotrophic denitrifiers. *Proc. Natl. Acad. Sci. U.S.A.* 113, E7929–E7936. doi: 10.1073/pnas.1612244113

Lau, M. P., Niederdorfer, R., Sepulveda-Jauregui, A., and Hupfer, M. (2018). Synthesizing redox biogeochemistry at aquatic interfaces. *Limnologica* 68, 59–70. doi: 10.1016/j.limno.2017.08.001

Litchman, E. (2000). Growth rates of phytoplankton under fluctuating light. *Freshw. Biol.* 44, 223–235. doi: 10.1046/j.1365-2427.2000.00559.x

Litchman, E., and Klausmeier, C. A. (2001). Competition of phytoplankton under fluctuating light. *Am. Nat.* 157, 170–187.

Liu, D., Zhu, W., Wang, X., Pan, Y., Wang, C., Xi, D., et al. (2017). Abiotic versus biotic controls on soil nitrogen cycling in drylands along a 3200 km transect. *Biogeosciences* 14, 989–1001. doi: 10.5194/bg-14-989-2017

Logan, B. E., Rossi, R., Ragab, A., and Saikaly, P. E. (2019). Electroactive microorganisms in bioelectrochemical systems. *Nat. Rev. Microbiol.* 17, 307–319. doi: 10.1038/s41579-019-0173-x

Lovley, D. R. (2006). Microbial fuel cells: novel microbial physiologies and engineering approaches. *Curr. Opin. Biotechnol.* 17, 327–332. doi: 10.1016/j.copbio.2006.04.006

Lovley, D. R. (2008). The microbe electric: conversion of organic matter to electricity. *Curr. Opin. Biotechnol.* 19, 564–571. doi: 10.1016/j.copbio.2008.10.005

Magnabosco, C., Lin, L. H., Dong, H., Bomberg, M., Ghiorse, W., Stan-Lotter, H., et al. (2018). The biomass and biodiversity of the continental subsurface. *Nat. Geosci.* 11, 707–717. doi: 10.1038/s41561-018-0221-6

Majzlan, J., Navrotsky, A., and Schwertmann, U. (2004). Thermodynamics of iron oxides: Part III. Enthalpies of formation and stability of ferrihydrite (∼$Fe(OH)_3$), schwertmannite (∼$FeO(OH)_{3/4}(SO_4)_{1/8}$), and ε-Fe_2O_3. *Geochim. Cosmochim. Acta* 68, 1049–1059. doi: 10.1016/S0016-7037(03)00371-5

Martin, W., Baross, J., Kelley, D., and Russell, M. J. (2008). Hydrothermal vents and the origin of life. *Nat. Rev. Microbiol.* 6, 805–814. doi: 10.1038/nrmicro1991

Matthies, D., Bräuer, I., Maibom, W., and Tscharntke, T. (2004). Population size and the risk of local extinction: empirical evidence from rare plants. *Oikos* 105, 481–488. doi: 10.1111/j.0030-1299.2004.12800.x

McCollom, T. M. (2007). Geochemical constraints on sources of metabolic energy for chemolithoautotrophy in ultramafic-hosted deep-sea hydrothermal systems. *Astrobiology* 7, 933–950.

McCollom, T. M., and Amend, J. P. (2005). A thermodynamic assessment of energy requirements for biomass synthesis by chemolithoautotrophic microorganisms in oxic and anoxic environments. *Geobiology* 3, 135–144. doi: 10.1111/j.1472-4669.2005.00045.x

McCollom, T. M., and Shock, E. L. (1997). Geochemical constraints on chemolithoautotrophic metabolism by microorganisms in seafloor hydrothermal systems. *Geochim. Cosmochim. Acta* 61, 4375–4391.

Melton, E. D., Swanner, E. D., Behrens, S., Schmidt, C., and Kappler, A. (2014). The interplay of microbially mediated and abiotic reactions in the biogeochemical Fe cycle. *Nat. Rev. Microbiol.* 12, 797–808. doi: 10.1038/nrmicro3347

Mitchell, P. (1961). Coupling of phosphorylation to electron and hydrogen transfer by a chemi-osmotic type of mechanism. *Nature* 191, 144–148. doi: 10.1038/191144a0

Mobarry, B. K., Wagner, M., Urbain, V., Rittmann, B. E., and Stahl, D. A. (1996). Phylogenetic probes for analyzing abundance and spatial organization of nitrifying bacteria. *Appl. Environ. Microbiol.* 62, 2156–2162.

Moissl-Eichinger, C., Cockell, C., and Rettberg, P. (2016). Venturing into new realms? Microorganisms in space. *FEMS Microbiol. Rev.* 40, 722–737. doi: 10.1093/femsre/fuw015

Morono, Y., Terada, T., Nishizawa, M., Ito, M., Hillion, F., Takahata, N., et al. (2011). Carbon and nitrogen assimilation in deep subseafloor microbial cells. *Proc. Natl. Acad. Sci. U.S.A.* 108, 18295–18300. doi: 10.1073/pnas.1107763108

Müller, V., and Hess, V. (2017). The minimum biological energy quantum. *Front. Microbiol.* 8:2019. doi: 10.3389/fmicb.2017.02019

Nordstrom, D. K. (1985). "The rate of ferrous iron oxidation in a stream receiving acid mine effluent," in *Selected Papers in the Hydrologic Sciences, US Geological Survey Water-Supply Paper* 2270, ed. S. Subitzky (CA, United States: University of California, San Diego), 113–119.

Pace, N. R. (1997). A molecular view of microbial diversity and the biosphere. *Science* 276, 734–740. doi: 10.1126/science.276.5313.734

Parkes, R. J., Cragg, B. A., Bale, S. J., Getlifff, J. M., Goodman, K., Rochelle, P. A., et al. (1994). Deep bacterial biosphere in Pacific Ocean sediments. *Nature* 371, 410–413. doi: 10.1038/371410a0

Parshina, S. N., Kijlstra, S., Henstra, A. M., Sipma, J., Plugge, C. M., and Stams, A. J. M. (2005). Carbon monoxide conversion by thermophilic sulfate-reducing bacteria in pure culture and in co-culture with *Carboxydothermus*

hydrogenoformans. *Appl. Microbiol. Biotechnol.* 68, 390–396. doi: 10.1007/s00253-004-1878-x

Philippot, L., Spor, A., Hénault, C., Bru, D., Bizouard, F., Jones, C. M., et al. (2013). Loss in microbial diversity affects nitrogen cycling in soil. *ISME J.* 7, 1609–1619. doi: 10.1038/ismej.2013.34

Roden, E. E., Sobolev, D., Glazer, B., and Luther, G. W. (2004). Potential for microscale bacterial Fe redox cycling at the aerobic-anaerobic interface. *Geomicrobiol. J.* 21, 379–391. doi: 10.1080/01490450490485872

Rummel, J. D., Beaty, D. W., Jones, M. a, Bakermans, C., Barlow, N. G., Boston, P. J., et al. (2014). A new analysis of Mars "Special Regions": findings of the second MEPAG special regions science analysis group (SR-SAG2). *Astrobiology* 14, 887–968. doi: 10.1089/ast.2014.1227

Russell, J. B. (2007). The energy spilling reactions of bacteria and other organisms. *J. Mol. Microbiol. Biotechnol.* 13, 1–11. doi: 10.1159/000103591

Sattley, W. M., and Madigan, M. T. (2006). Isolation, characterization, and ecology of cold-active, chemolithotrophic, sulfur-oxidizing bacteria from perennially ice-covered Lake Fryxell, Antarctica. *Appl. Environ. Microbiol.* 72, 5562–5568. doi: 10.1128/AEM.00702-06

Seto, M., and Iwasa, Y. (2019a). Population dynamics of chemotrophs in anaerobic conditions where the metabolic energy acquisition per redox reaction is limited. *J. Theor. Biol.* 467, 164–173. doi: 10.1016/j.jtbi.2019.01.037

Seto, M., and Iwasa, Y. (2019b). The fitness of chemotrophs increases when their catabolic by-products are consumed by other species. *Ecol. Lett.* 22, 1994–2005. doi: 10.1111/ele.13397

Seto, M., and Iwasa, Y. (2020). Microbial material cycling, energetic constraints and ecosystem expansion in subsurface ecosystems. *Proc. R. Soc. B Biol. Sci.* 287:20200610. doi: 10.1098/rspb.2020.0610

Silvertown, J. (2004). Plant coexistence and the niche. *Trends Ecol. Evol.* 19, 605–611. doi: 10.1016/j.tree.2004.09.003

Six, C., Finkel, Z. V., Irwin, A. J., and Campbell, D. A. (2007). Light variability illuminates niche-partitioning among marine picocyanobacteria. *PLoS One* 12:e1341. doi: 10.1371/journal.pone.0001341

Sobolev, D., and Roden, E. E. (2002). Evidence for rapid microscale bacterial redox cycling of iron in circumneutral environments. *Antonie Van Leeuwenhoek* 81, 587–597. doi: 10.1023/A:1020569908536

Sorokin, D. Y., and Kuenen, J. G. (2005). Chemolithotrophic haloalkaliphiles from soda lakes. *FEMS Microbiol. Ecol.* 52, 287–295. doi: 10.1016/j.femsec.2005.02.012

Sterck, F., Markesteijn, L., Schieving, F., and Poorter, L. (2011). Functional traits determine trade-offs and niches in a tropical forest community. *Proc. Natl. Acad. Sci. U.S.A.* 108, 20627–20632. doi: 10.1073/pnas.1106950108

Stevens, T. (1997). Lithoautotrophy in the subsurface. *FEMS Microbiol. Rev.* 20, 327–337. doi: 10.1016/S0168-6445(97)00015-6

Takizawa, K., Kanazawa, A., and Kramer, D. M. (2008). Depletion of stromal Pi induces high "energy-dependent" antenna exciton quenching (qE) by decreasing proton conductivity at CF O-CF1 ATP synthase. *Plant Cell Environ.* 31, 235–243. doi: 10.1111/j.1365-3040.2007.01753.x

Tanikawa, W., Tadai, O., Morono, Y., Hinrichs, K. U., and Inagaki, F. (2018). Geophysical constraints on microbial biomass in subseafloor sediments and coal seams down to 2.5 km off Shimokita Peninsula, Japan. *Prog. Earth Planet. Sci.* 5:58. doi: 10.1186/s40645-018-0217-2

Thébault, E., and Loreau, M. (2003). Food-web constraints on biodiversity-ecosystem functioning relationships. *Proc. Natl. Acad. Sci. U.S.A.* 100, 14949–14954. doi: 10.1073/pnas.2434847100

Thrash, J. C., and Coates, J. D. (2008). Review: direct and indirect electrical stimulation of microbial metabolism. *Environ. Sci. Technol.* 42, 3921–3931. doi: 10.1021/es702668w

Tilman, D. (1982). *Resource Competition and Community Structure*. Princeton, NJ: Princeton University Press.

Tilman, D., Reich, P. B., Knops, J., Wedin, D., Mielke, T., and Lehman, C. (2001). Diversity and productivity in a long-term grassland experiment. *Science* 294, 843–845. doi: 10.1126/science.1060391

Torres, C. I., Marcus, A. K., Lee, H. S., Parameswaran, P., Krajmalnik-Brown, R., and Rittmann, B. E. (2010). A kinetic perspective on extracellular electron transfer by anode-respiring bacteria. *FEMS Microbiol. Rev.* 34, 3–17. doi: 10.1111/j.1574-6976.2009.00191.x

Trembath-Reichert, E., Morono, Y., Ijiri, A., Hoshino, T., Dawson, K. S., Inagaki, F., et al. (2017). Methyl-compound use and slow growth characterize microbial life in 2-km-deep subseafloor coal and shale beds. *Proc. Natl. Acad. Sci. U.S.A.* 114, E9206–E9215. doi: 10.1073/pnas.1707525114

Valentine, D. L. (2007). Adaptations to energy stress dictate the ecology and evolution of the Archaea. *Nat. Rev. Microbiol.* 5, 316–323. doi: 10.1038/nrmicro1619

Vazquez, A., Liu, J., Zhou, Y., and Oltvai, Z. N. (2010). Catabolic efficiency of aerobic glycolysis: the Warburg effect revisited. *BMC Syst. Biol.* 4:58. doi: 10.1186/1752-0509-4-58

Webster, G., Embley, T. M., Freitag, T. E., Smith, Z., and Prosser, J. I. (2005). Links between ammonia oxidizer species composition, functional diversity and nitrification kinetics in grassland soils. *Environ. Microbiol.* 7, 676–684. doi: 10.1111/j.1462-2920.2005.00740.x

Weiss, M. C., Sousa, F. L., Mrnjavac, N., Neukirchen, S., Roettger, M., Nelson-Sathi, S., et al. (2016). The physiology and habitat of the last universal common ancestor. *Nat. Microbiol.* 1:16116. doi: 10.1038/nmicrobiol.2016.116

Wertz, S., Degrange, V., Prosser, J. I., Poly, F., Commeaux, C., Freitag, T., et al. (2006). Maintenance of soil functioning following erosion of microbial diversity. *Environ. Microbiol.* 8, 2162–2169. doi: 10.1111/j.1462-2920.2006.01098.x

Wright, R. H. G., Lioutas, A., Le Dily, F., Soronellas, D., Pohl, A., Bonet, J., et al. (2016). ADP-ribose-derived nuclear ATP synthesis by NUDIX5 is required for chromatin remodeling. *Science* 352, 1221–1225. doi: 10.1126/science.aad9335

Zheng, J. (2012). Energy metabolism of cancer: glycolysis versus oxidative phosphorylation. *Oncol. Lett.* 4, 1151–1157.

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Ecology and Evolution

METHODS
published: 11 December 2020
doi: 10.3389/fevo.2020.530135

Automated Discovery of Relationships, Models, and Principles in Ecology

Pedro Cardoso[1,2]*, *Vasco V. Branco*[1,3], *Paulo A. V. Borges*[2], *José C. Carvalho*[1,2,4], *François Rigal*[2,5], *Rosalina Gabriel*[2], *Stefano Mammola*[1,6], *José Cascalho*[7] and *Luís Correia*[3]

[1] Laboratory for Integrative Biodiversity Research (LIBRe), Finnish Museum of Natural History Luomus, University of Helsinki, Helsinki, Finland, [2] cE3c – Centre for Ecology, Evolution and Environmental Changes/Azorean Biodiversity Group, Departamento de Ciências Agrárias, Universidade dos Açores, Angra do Heroísmo, Portugal, [3] Laboratório de Sistemas Informáticos de Grande Escala, Faculdade de Ciências, Universidade de Lisboa, Lisbon, Portugal, [4] Molecular and Environmental Centre - Centre of Molecular and Environmental Biology, Department of Biology, University of Minho, Braga, Portugal, [5] Institut Des Sciences Analytiques et de Physico Chimie pour L'environnement et les Materiaux UMR5254, Comité National de la Recherche Scientifique - University de Pau et des Pays de l'Adour - E2S UPPA, Pau, France, [6] Molecular Ecology Group (MEG), Water Research Institute, National Research Council, Verbania Pallanza, Italy, [7] Departamento de Ciências Agrárias, Núcleo de Investigação e Desenvolvimento em e-Saúde (NIDes), Angra do Heroísmo, Portugal

OPEN ACCESS

Edited by:
Miguel A. Fortuna, Estación Biológica de Doñana (EBD), Spain

Reviewed by:
Carlos J. Melián, Swiss Federal Institute of Aquatic Science and Technology, Switzerland
Shai Pilosof, Ben-Gurion University of the Negev, Israel

****Correspondence:***
Pedro Cardoso
pedro.cardoso@helsinki.fi

Specialty section:
This article was submitted to Models in Ecology and Evolution, a section of the journal Frontiers in Ecology and Evolution

Received: *28 January 2020*
Accepted: *23 November 2020*
Published: *11 December 2020*

Citation:
Cardoso P, Branco VV, Borges PAV, Carvalho JC, Rigal F, Gabriel R, Mammola S, Cascalho J and Correia L (2020) Automated Discovery of Relationships, Models, and Principles in Ecology. Front. Ecol. Evol. 8:530135. doi: 10.3389/fevo.2020.530135

Ecological systems are the quintessential complex systems, involving numerous high-order interactions and non-linear relationships. The most used statistical modeling techniques can hardly accommodate the complexity of ecological patterns and processes. Finding hidden relationships in complex data is now possible using massive computational power, particularly by means of artificial intelligence and machine learning methods. Here we explored the potential of symbolic regression (SR), commonly used in other areas, in the field of ecology. Symbolic regression searches for both the formal structure of equations and the fitting parameters simultaneously, hence providing the required flexibility to characterize complex ecological systems. Although the method here presented is automated, it is part of a collaborative human–machine effort and we demonstrate ways to do it. First, we test the robustness of SR to extreme levels of noise when searching for the species-area relationship. Second, we demonstrate how SR can model species richness and spatial distributions. Third, we illustrate how SR can be used to find general models in ecology, namely new formulas for species richness estimators and the general dynamic model of oceanic island biogeography. We propose that evolving free-form equations purely from data, often without prior human inference or hypotheses, may represent a very powerful tool for ecologists and biogeographers to become aware of hidden relationships and suggest general theoretical models and principles.

Keywords: artificial intelligence, ecological complexity, evolutionary computation, genetic programming, species richness estimation, species-area relationship, species distribution modeling, symbolic regression

INTRODUCTION

Complexity is a term often used to characterize systems with numerous components interacting in ways such that their collective behavior is difficult to predict, but where emergent properties give rise to patterns, more or less simple but seldom linear (**Table 1**) (Holland, 1995; Mitchell, 2009). Complex systems science is therefore an effort to understand non-linear systems with multiple

TABLE 1 | Glossary of terms.

Artificial Intelligence (AI)–A scientific field concerned with the automation of activities we associate with human thinking (Russell et al., 2010).
Big data–Very large amount of structured or unstructured data, hard to model with general statistical techniques but with the potential to be mined for information.
Complex system–A system in which a large network of components organize, without any central controller and simple although non-linear rules of operation, into a complex collective behavior that creates patterns, uses information, and, in some cases, evolves, and learns (Mitchell, 2009).
General model–An equation that is found to be useful for multiple datasets, often but not necessarily, derived from a general principle. In most cases the formal structure of equations is kept fixed, while some parameters must be fitted for each individual dataset.
General principle–Refers to concepts or phenomenological descriptions of processes and interactions (Evans et al., 2013). May not have direct translation to any general model, but be a purely conceptual abstraction.
Genetic programming (GP)–A biologically-inspired method for getting computers to automatically create a computer program to solve a given problem (Koza, 1992). It is a type of evolutionary algorithm, where each solution to be tested (individual in a population of possible solutions) is a computer program.
Pareto front–A curve connecting a set of best solutions in a multi-objective optimization problem. If several conflicting objectives are sought (e.g., minimize both error and complexity of formulas), the Pareto front allows visualizing the set of best solutions (Smits and Kotanchek, 2005).
Symbolic regression (SR)–A function discovery approach for modeling of multivariate data. It is a special case of genetic programming, one where possible solutions are equations instead of computer programs.

connected components and how "the whole is more than the sum of the parts" (Holland, 1998). Biological systems probably are among the most complex (Solé and Goodwin, 2000), and among them, ecological systems are the quintessential complex systems (Anand et al., 2010). These are composed of individuals from different species, interacting and exchanging energy in multiple ways, furthermore, relating with the physical environment at different spatial and temporal scales in non-linear relationships. Consequently, ecology is dominated by idiosyncratic results, with most ecological processes being contingent on the spatial and temporal scales in which they operate. This makes it difficult to identify recurrent patterns, knowing also that pattern does not necessarily identify process (Lawton, 1996; Dodds, 2009; Passy, 2012). The most used exploratory (e.g., principal component analysis) and statistical modeling techniques (e.g., linear and non-linear regressions) can hardly reflect the complexity of ecological patterns and processes, often failing to find meaningful relationships in data. More flexible techniques, such as generalized additive models (GAMs), usually do not allow an easy interpretation of results and particularly of putative causal relationships (e.g., Sugihara et al., 2012). For ecological data, we require more flexible and robust, yet amenable to full interpretation, analytical methods, which can eventually lead to the discovery of general principles and models.

The aim of any ecological principle is to provide a robust model for exploring, describing, and predicting ecological processes regardless of taxon identity and geographic region (Lawton, 1996; Dodds, 2009). Finding a recurrently high goodness-of-fit for a model to an ecological pattern for most taxa and ecosystems is usually a compelling evidence of a mechanistic process controlling that pattern. When general principles are translated into robust models, general statistical methods are mostly abandoned in favor of these, of which only few examples exist in ecology (**Data Sheet 1**). Such general, widely applicable equations are mostly found by intellectual *tour de force*. Yet, they are only the tip of the iceberg, usually incorporating few of the variables increasingly available to ecologists and that could potentially explain such patterns.

The automation of techniques for collecting and storing ecological and related data, with increasing spatial and temporal resolutions, has become one of the central themes in ecology and bioinformatics. Yet, automated and flexible ways to synthesize such complex and big data were mostly lacking until recently (Martin et al., 2018; Chen et al., 2019; Desjardins-Proulx et al., 2019). Finding hidden relations within such data is now possible using massive computational power. New computer-intensive methods have been developed or are now available or possible (Reshef et al., 2011), including the broad field of artificial intelligence (AI) or machine learning (ML) which have produced a variety of approaches (Lu, 2019). Artificial intelligence includes a series of evolution-inspired techniques, brought together in the sub-field of evolutionary computation, of which the most studied and well-known probably are genetic algorithms (Holland, 1975). Genetic programming, namely in the form of symbolic regression (SR) (Koza, 1992), is a derivation of genetic algorithms that searches the space of mathematical equations without any constraints on their form. Hence, it provides the required flexibility to represent complex systems as presented by many ecological systems (**Figure 1**). Contrarily to traditional statistical techniques, symbolic regression searches for both the formal structure of equations and the fitting parameters simultaneously (Schmidt and Lipson, 2009). Finding the structure of equations is especially useful to discover general models, providing insights into the processes and eventually leading to the discovery of new and yet undiscovered principles. Fitting the parameters provides insight into the raw data and allow for specific predictions. Successful examples on the use of SR in ecology include modeling of land-use change (Manson, 2005; Manson and Evans, 2007), effects of climate change on populations (Tung et al., 2009; Larsen et al., 2014), community distribution (Larsen et al., 2012; Yao et al., 2014), predicting micro-organismal blooms (Muttil and Lee, 2005; Muttil and Chau, 2006; Jagupilla et al., 2015; Tromas et al., 2017), deriving vegetation indices (Almeida et al., 2015), forecasting the trophic evolution of lakes (Bertoni et al., 2016), using parasites as biological tags (Barrett et al., 2005), and even to revisit classical ecological models such as the Lotka–Volterra predator–prey equation (Martin et al., 2018; Chen et al., 2019).

The goal of this work is to explain, test, and show the usefulness of SR in uncovering hidden relationships within typical ecological datasets. To illustrate this, we used five case studies reflecting typical analytical problems faced by ecologists. In the first example, (i) we test the robustness of SR when finding the power law applied to the species-area relationship (SAR) with

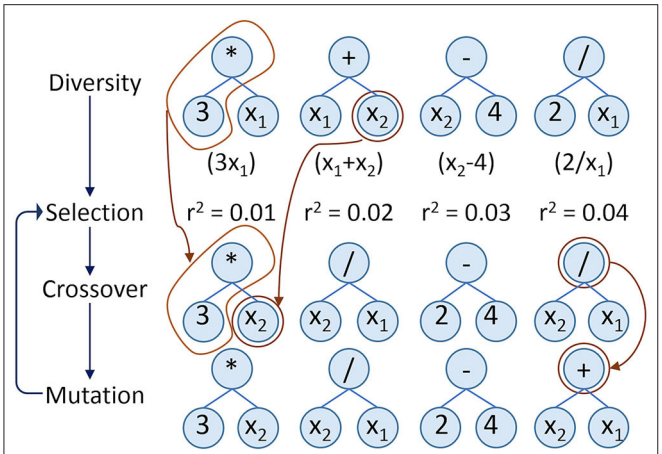


FIGURE 1 | Schematic representation of the symbolic regression workflow. The basic representation is a parse-tree where building blocks such as variables (in this case: x_1, x_2), parameters (integers or real numbers), and operators (e.g., +, –, ×, ÷) are connected forming functions (in parenthesis under the first line of trees). Initial equations are generated by randomly linking different building blocks. Equations are combined through crossover, giving rise to new equations with characteristics from both parents (arrows linking the first and second rows of trees). Equations with better fitness (e.g., R^2) have higher probabilities of recombining. To avoid loss of variability, a mutation step is added after crossover (arrows linking the second and third rows of trees). After multiple generations, evolution stops and a set of free-form equations best reflecting the input data is found.

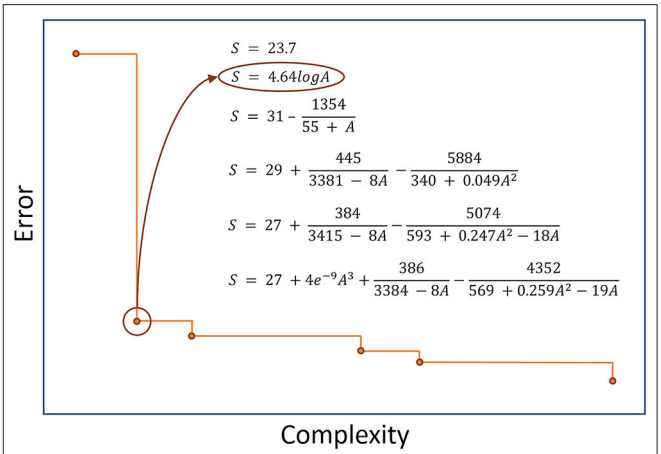


FIGURE 2 | Example of a Pareto front depicting error vs. complexity. This example reflects a symbolic regression search of the best species–area relationship for native spiders in the Azores (Portugal). The second formula is clearly the most promising, with both high accuracy (low error) and low complexity. In many occasions a single formula is not clearly best, in which case weights can be given to each of them through indices that simultaneously positively weight accuracy and negatively weight complexity (such as AIC or BIC) and/or multiple formulas presented as possible outcomes.

extreme levels of noise, even beyond the high levels typical of ecological datasets. In the next two examples, we demonstrate how SR can deal with complex datasets, namely to model (ii) species richness; and (iii) species spatial distributions. Finally, we illustrate how SR can be used to find general models in ecology, by using it to develop new formulas for (iv) species richness estimation; and v) the general dynamic model of oceanic island biogeography (GDM).

GENERAL METHODOLOGY

Symbolic regression works as a computational parallel to the evolution of species. A population of initial equations is generated randomly by combining different building blocks, such as the variables of interest (independent explanatory variables), algebraic operators (e.g., +, –, ÷, ×), analytic function types (exponential, log, power, etc.), constants and other ways to combine the data (e.g., Boolean or decision operators) (**Figure 1**). Being random, these initial equations almost invariably fail in describing the patterns or phenomena of interest, but some equations are slightly better than others. All are then combined through crossover, giving rise to new equations with characteristics from both parents. Equations with better fitness—as estimated using a chosen statistical measure such as R^2 or Akaike's Information Criterion (AIC; Akaike, 1974)—have a higher probability of recombining. To avoid new equations being bounded by initially selected building blocks or quickly losing variability along the evolutionary process, a mutation step (acting on any building block) is added to the process after crossover. After multiple generations, an acceptable level of accuracy by some of the equations is often attained and the researcher stops the process.

For this work we used the software Eureqa (Nutonian, Inc) (Schmidt, 2015), which provides an intuitive interface suitable also for non-expert SR users (Dubčáková, 2011). Although a commercial version is available, we used the freely available test version for all analyses. For each run, the software outputs a list of equations along an error/complexity Pareto-front (Smits and Kotanchek, 2005), with the most accurate equation for each level of complexity being shown (**Figure 2**). For the SR search we used only algebraic and analytic operators (+, –, ÷, ×, log, power) in all examples below, so that outputs could be most easily interpreted. The goodness-of-fit was evaluated using R^2 or AICc, depending on the question (see below). The Pareto-front often presents an "elbow," where near-minimum error meets near-minimum complexity. The equation in this inflection is closer to the origin of both axes and is a good starting point for further investigation—if both axes are in comparable qualitative scales. Often, however, this inflection point is not obvious, and a single formula is not clearly best. In such cases, weights can be given to each of them through indices that positively weight accuracy and negatively weight complexity, such as AICc (Akaike, 1974). However, in all cases it is important to check all formulas along the Pareto-front. Often equations or models that make immediate sense to the specific question may not be detected by these automated methods.

Each of the five case studies was analyzed independently and using different approaches to test the performance of SR against other methods. We often opted to use fully independent datasets for three reasons. First, we were looking for general formulas,

which should be tested against fully independent data. Second, it was a much more powerful and convincing way to demonstrate the method than using subsets of the same data, which necessarily have some common ground that facilitates the job of any method. Third, our datasets, as typical for many in ecology, often had very few observations, making it hard to do sub-setting in few cases.

CASE-STUDIES

Finding the Species-Area Relationship (SAR) With Extreme Levels of Noise

Typical ecological datasets not only have few data points, but these are prone to varying levels of noise. Noise can be due to natural phenomena, such as spillover from neighboring regions, unpredictable weather events, etc., lack of the most appropriate data to model the phenomena, or simply errors in measuring or sampling. Testing any novel method to its robustness to different sources and levels of noise is therefore imperative.

One of the most studied examples of SARs is their application to island biogeography (ISAR). The shape of ISARs has been modeled by many functions, but three of the simplest seem to be preferred in most cases, the power, exponential, and linear models (Triantis et al., 2012) (**Data Sheet 1**). The power model in particular includes island area (A) and two fitting parameters, c (the intercept) and z (the scaling of richness with area) (**Data Sheet 1**). Here we created 30 fictional islands each one corresponding to one of the 30 possible combinations resulting from five different areas (10, 100, 1000, 10,000, and 100,000 km^2), two typical values for c (1 and 10), and three typical values for z (0.2, 0.3, and 0.4). We then simulated sampling from these 30 islands, each with a sampled richness equal to the multiplication of the true richness value by five different levels of noise as given by the standard deviation of a sampling from a normal distribution with mean = 1 and sd = 0, 0.1, 0.2, 0.4, and 0.8. The theoretical richness of each island was then multiplied by 10 simulations of each noise level using this approach, providing a total of 50 search trials (**Data Sheet 2**). We must emphasize that with sd = 0.8 the noise was extreme and unreasonable, with for example islands predicted to have 100 species presenting anything between 0 and 199 species after noise was added. We evaluated the ability of SR to develop the power law by counting at each noise level how many times the usual formulation and a derivation without the intercept c were found among the 10 searches per level.

Our simulations using SR were able to find the power-law of the SAR even with the most extreme scenarios (**Data Sheet 2**). From 100% success rate with sd = 0 or 0.1, to 70% with sd = 0.2, 50% with sd = 0.4, and 40% with sd = 0.8. If we include the simpler formulation with no intercept, success rate was 100, 100, 100, 90, and 80%, respectively for sd = 0, 0.1, 0.2, 0.4, and 0.8.

Modeling Species Richness

Modeling and mapping the species richness of high diversity taxa at regional to large scales is often impossible without extrapolation from sampled to non-sampled sites. Here, we used an endemic arthropod dataset collected in Terceira Island, Azores. Fifty-two sites were sampled using pitfall traps for epigean arthropods (Cardoso et al., 2009), 13 in each of four land-use types: natural forest, exotic forest, semi-natural pasture, and intensively managed pasture. In this problem, given the size of the dataset, we used a 5-fold cross validation. We explained and predicted species richness per site using elevation, slope, annual average temperature, annual precipitation, and an index of disturbance with values ranging from 0 (absence of human presence) to 100 (dense urban environment) (Cardoso et al., 2013). For SR we ran each fold five times to minimize the risk that the formulas found represented local optima. We then reported the average and range of R^2 and AICc of the five partitions for both the training and test data.

As the response variable was count data, Generalized Linear Models (GLM) and Generalized Additive Models (GAM) with a Poisson error structure and a log link were used. We used the package MuMIn (Barton, 2015) and the *R* environment (R Core Team, 2015) for multi-model inference based on AICc (Hurvich and Tsai, 1989) values, using all variables plus all possible interactions for GLM. For fitting GAM, we used package gam (Hastie, 2015). The R^2 goodness of fit was used as the fitness measure. For each run of the SR (25 in total) we picked the formula at the inflection point of the Pareto-front (**Data Sheet 2**). Both R^2 and AICc were used to compare GLM and GAM with SR on the test datasets.

The model selected by GLM in all five k-folds was:

$$S = e^{(\mathrm{a}+\mathrm{b}H\,-\mathrm{c}P\,-\,\mathrm{d}D)}$$

where H = altitude, P = precipitation, D = disturbance; a, b, c, and d are fitting parameters with mean a = 1.894 (range: 1.116–2.577), mean b = 0.00419 (range: 0.00360–0.00574), mean c = 0.000972 (range: 0.000726–0.001212), and mean d = 0.0251 (range: 0.0118–0.0331). The mean training R^2 = 0.529 (range: 0.469–0.573) and mean training AICc = 104.151 (range: 103.055–105.634). The mean testing R^2 = 0.528 (range: 0.313–0.770) and mean testing AICc = 45.340 (range: 41.913–48.513).

The results of GAM were similar to the GLM, the algorithm selecting the most parsimonious formulation equivalent to a GLM. The SR results performed considerably better than GLM or GAM with a much simpler formula using a single variable (Disturbance) and much better training and testing results, with 23 out of 25 formulas chosen being in the form:

$$S = (a/D) - b$$

where a and b were fitting parameters with mean a = 140.787 (range: 134.700–145.775) and mean b = 1.325 (range: 1.078–1.483). The mean training R^2 = 0.603 (range: 0.576–0.644) and mean training AICc = 52.631 (range: 47.982–56.985). The mean testing R^2 = 0.601 (range: 0.449–0.737) and mean testing AICc = 19.088 (range: 13.042–23.432). We should emphasize the simplicity of interpretation of this formula, indicating that species richness essentially was inversely related with disturbance.

Modeling Species Distributions

Species distribution modeling (SDM) is one of the most widely used correlative statistical approaches to biodiversity assessment,

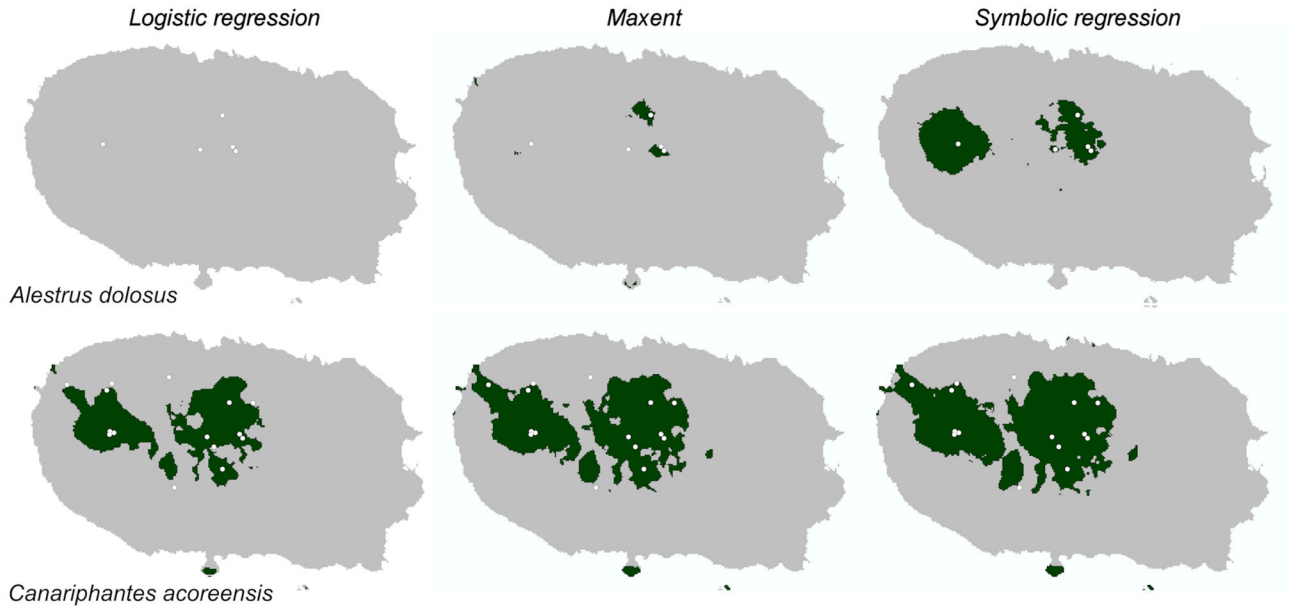


FIGURE 3 | Predicted distribution of two Azorean arthropods using three modeling methods. Observed locations (white dots) and predicted distribution (dark green areas) of *Alestrus dolosus* (Coleoptera, Elateridae) and *Canariphantes acoreensis* (Araneae, Linyphiidae) in the island of Terceira (Azores, Portugal) using logistic regression, maximum entropy, and symbolic regression.

for example to fill gaps in our knowledge on individual species distributions, predict species response to climate change, and the spread of invasive species (Araújo et al., 2019). As a case study, we modeled the potential distribution of two endemic Azorean species in Terceira Island: the rare forest click-beetle *Alestrus dolosus* (Crotch) (Coleoptera, Elateridae) and the abundant but mostly forest-restricted spider *Canariphantes acoreensis* (Wunderlich) (Araneae, Linyphiidae). We compared the performance of logistic regressions (GLM with a binomial error structure) and maximum entropy models (Maxent; Phillips et al., 2006) in predicting the distribution of the two species, to that of SR. GLM and MaxEnt are two of the most widely used approaches for SDM (Elith et al., 2006). Given the intrinsic differences between methods, we had to use different background datasets (**Data Sheet 3**). Maxent used the environmental maps of the islands with a resolution of 100 m, from where it extracted pseudo-absences. We then converted the probabilistic potential distribution maps to presence/absence using the maximum value of training sensitivity plus specificity as the threshold as recommended by Liu et al. (2005). Logistic regression and SR used presence/absence data from the 52 sampled sites. We conducted multi-model inference of logistic regression based on AICc values. In the SR run we included a step function, so that positive and negative values were converted to presence and absence (binary output), respectively. Absolute error, reflecting the number of incorrect classifications, was used as the fitness measure. As inflection points of the Pareto fronts were clear, the best SR formula for each species was chosen based on them (**Data Sheet 2**). Given the scarcity of occurrences of species in the dataset (from 10 to 35% of the data points) we opted for a balanced split of 50% for training and test sets. In all cases only the training data (26 sites) were used to obtain the models. Logistic GLM, Maxent, and SR were compared in their performance for predicting presence and absence of species on the 26 test sites using the Sensitivity, Specificity, and True Skill Statistic (TSS) (Allouche et al., 2006).

The potential distribution models were relatively similar for *C. acoreensis* but show marked differences for *A. dolosus* (**Figure 3**). Symbolic regression outperformed both other models for *A. dolosus* and was as good as Maxent for *C. acoreensis*, with both outperforming logistic regression (**Table 2**). The SR models were not only the best, presenting maximum values for TSS, but were also the easiest to interpret. *A. dolosus* was predicted to have adequate environmental conditions in all areas above 614 m elevation, being restricted to pristine native forest. *Canariphantes acoreensis* could potentially be present in all areas with disturbance values below 41.3, occurring not only in native forest but also in adjacent semi-natural grassland and humid exotic forest. The logistic regression and Maxent models used a large number of explanatory variables for *A. dolosus* yet performed worse on the test data than did SR (**Table 2**).

Developing Species Richness Estimators

Several asymptotic functions have been used to estimate species richness (Soberón and Llorente, 1993), including the Clench function (Clench, 1979), the negative exponential function, and the rational function (Ratkowsky, 1990) (**Data Sheet 1**). We used SR to rediscover or eventually find novel asymptotic models that would outperform them. Two independent datasets were used resulting from exhaustive and standardized sampling for

TABLE 2 | Species distribution models for two endemic arthropod species on the island of Terceira (Azores, Portugal).

Model	Formula	Sensitivity	Specificity	TSS
Alestrus dolosus				
Logistic regression	$1/(1+e^{-8469+0.432P+540.77T})$	0	**1**	0
Maxent	Uses all variables but *SI*, main is *D* (contribution = 74.1%)	0.5	**1**	0.5
Symbolic regression	*step* (H−614)	**1**	0.75	**0.75**
Canariphantes acoreensis				
Logistic regression	$1/(1+e^{-3.617+0.103D})$	0.667	**0.7**	0.367
Maxent	Uses only *D* (contribution = 100%)	**0.833**	0.65	**0.483**
Symbolic regression	*step* (41.3–D)	**0.833**	0.65	**0.483**

Accuracy statistics on an independent test dataset are given by the True Skill Statistic (TSS). H = altitude, SI = slope, T = average annual temperature, P = annual precipitation, and D = disturbance index. The step function in symbolic regression converts positive values inside parentheses to presence and negative values to absence. Best values in bold.

spiders in 1 ha plots, performed by 8 collectors during 320 h of sampling in a single hectare using five different methods (so-called sampling protocol "COBRA" —Conservation Oriented Biodiversity Rapid Assessment; Cardoso, 2009). The training dataset was from a mixed forest in Gerês (northern Portugal) and the test dataset was from a *Quercus* forest in Arrábida (southern Portugal) (Cardoso et al., 2008a,b).

Randomized accumulation curves for both sites were produced using the R package BAT (Cardoso et al., 2015). The true diversity of each site was calculated as the average between different non-parametric estimators (Chao 1 and 2, Jackknife 1 and 2). Because the sampled diversity in the training dataset reached a very high completeness but we wanted to simulate typically very incomplete sampling, datasets with 10, 20, 40, 80, and 160 randomly chosen samples were extracted and used, in addition to the complete 320 samples dataset, as independent runs in SR. Squared error was used as the fitness measure. Additionally, we imposed a strong penalty to non-asymptotic functions, although these were still allowed in the search process (see **Data Sheet 2** for details). The weighted and non-weighted scaled mean squared errors implemented in BAT (Cardoso et al., 2015) were used as accuracy measures.

For the training dataset, one asymptotic model was found by SR (**Data Sheet 2**):

$$S = \frac{aQ}{b+Q}$$

where a and b were fitting parameters. This model was in fact the Clench model with a different formulation (**Data Sheet 1**), where the asymptote was a. A second, slightly more complex but better fitting, model was found for partial datasets with 40 or more samples:

$$S = \frac{c+aQ}{b+Q}$$

where c is a third fitting parameter. The asymptote was again given by the value of a (**Figure 4**). This model was similar to the rational function (**Data Sheet 1**). It was found to outperform the Clench and negative exponential for both the training and testing datasets (**Table 3**).

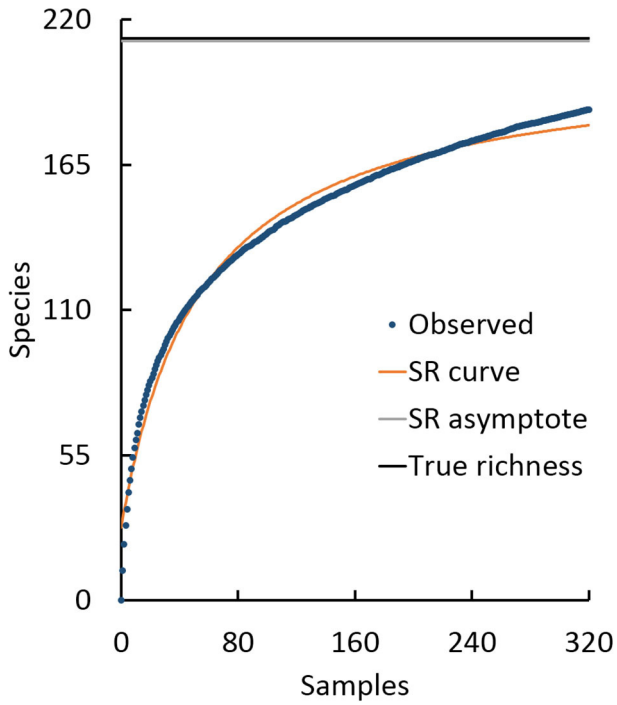


FIGURE 4 | Accumulation curve for spider sampling in Gerês (Portugal). The result of searching for the best fitting asymptotic formula using symbolic regression is also shown.

Developing the General Dynamic Model of Oceanic Island Biogeography (GDM)

The general dynamic model of oceanic island biogeography (GDM) was proposed to predict the responses of the key processes of immigration, speciation and extinction in volcanic islands, recognizing the role of geological processes in driving diversity on oceanic islands (Whittaker et al., 2008). Traditionally, the GDM is tested using a model where island species diversity is regressed as a function of area and age. Several different equations have been found to describe the GDM, extending the different SAR models with the addition

TABLE 3 | Comparison of three asymptotic equations used to estimate spider species richness in two forest sites.

Model	Raw accuracy	Weighted accuracy
Gerês (training)		
Observed	0.113	0.037
Clench	0.055	0.018
Negative exponential	0.115	0.049
Rational function	**0.045**	**0.012**
Arrábida (testing)		
Observed	0.103	0.031
Clench	0.038	0.010
Negative exponential	0.092	0.037
Rational function	**0.032**	**0.008**

*See **Data Sheet 1** for formulas. Raw accuracy is the scaled mean squared error considering the entire observed accumulation curve (each formula was fitted to the curves using 4–320 samples) and weighted accuracy is this value weighted by the sampling effort at each point in the curve (where effort is the ratio between number of individuals and observed species richness). Note that lower values (in bold) are better as they reflect the deviation from a perfect estimator.*

of a polynomial term using island age and its square (TT2), depicting the island's ontogeny. The first to be proposed was an extension of the exponential model (Whittaker et al., 2008), the power model extensions following shortly after (Fattorini, 2009; Steinbauer et al., 2013).

Our objective was to test if we could re-discover or eventually refine existing models for the ISAR and GDM from data alone. We used the Azores and Canary Islands spiders (Cardoso et al., 2010) as training data. To independently test the generality of models arising from spider data, we used bryophyte data from the same archipelagos (Aranda et al., 2014). The area and maximum time since emergence of each island were used as explanatory variables and the native species richness per island as the response variables (**Data Sheet 4**). The R^2 value was used as the fitness measure. The best SAR and GDM equations found by SR were chosen based on the inspection of the Pareto front (**Data Sheet 2**) but looking also for interpretability of the models. These were then compared with the existing models using AICc and the *R* package BAT (Cardoso et al., 2015).

For the Azorean spiders, the best fitting previous model (both highest R^2 and lowest AICc) for the ISAR was the exponential model (**Table 4**). The SR run discovered roughly the same model, indicating, however, that the intercept (*c* term) was adding unnecessary complexity. A similar ranking of models was verified for bryophytes in the same region, revealing the robustness of the new model.

For the Canary Islands, the best model for spiders was a linear function of area:

$$S = 75 + 0.047A$$

($R^2 = 0.364$, AICc = 65.631). Although it was easy to interpret, the explained variance was relatively low. The SR run reached a much higher explanatory power:

$$S = 112 - 1.002^A$$

($R^2 = 0.806$, AICc = 57.320). In this case though, the equation was over-fitting to the few available data (7 data points), as this function was erratic creating a biologically indefensible model. The reason the ISAR was hard to model for the Canary Islands spiders was because we were missing the major component Time (Cardoso et al., 2010). This was depicted by the GDM, of which the best of the current equations was found to be the power model described by Fattorini (2009) (**Table 4**). Nevertheless, using SR we were able to find an improved, yet undescribed, model (**Table 4**). This represented a general model expanding the linear SAR:

$$S = c + zA + XT - yT^2$$

When tested with Canarian bryophytes, this new formulation wss almost as good as the power model (**Table 4**).

DISCUSSION

Symbolic regression has the advantage over most standard regression methods (e.g., GLM) of being more flexible, allowing a good fitting to data with better interpretability, since results are in the form of mathematical formulas. GLMs and other similar techniques assume linear relations between variables or require a priori knowledge on the form of the relation (e.g., quadratic, cubic, interactions between variables, etc.).

SR also has one or more advantages over other, commonly used, highly flexible regression (e.g., GAMs) or machine learning techniques (e.g., neural-networks): (1) numerical, ordinal, and categorical variables are easily combined; (2) redundant variables are usually eliminated in the search process and only the most important are retained if anti-bloat measures (intended to reduce the complexity of equations) are used. Incidentally, this also releases the user from the problem of dealing with collinearity (Dormann et al., 2013); (3) the evolved equations are human-readable and interpretable; and (4) solutions are easily applied to new data.

Using SR, we were able to "distill" free-form equations and models that not only consistently outperform but are more intelligible than the ones resulting from rigid methods, such as GLM, or "black-boxes," such as Maxent. This was the case for both species richness and distribution models.

We were also able to re-discover and refine equations for estimating species richness based on sampling curves and the ISAR and GDM from data alone. All the examples presented in this work suggest that evolving free-form equations purely from data, often without prior human inference or hypotheses, may represent an under-explored but very powerful tool for ecologists and biogeographers, allowing the finding of hidden relationships in data and suggesting new ideas to formulate general theoretical principles.

The idea that SR is a powerful tool for reverse-engineering ecological theory from data is not new. Many examples reviewed in the introduction suggest that different authors across disparate disciplines understand this date back to the early 2000s. Recently, the potential of SR in ecology was discussed in two essays that showed how SR can be used to develop classic demographic time

TABLE 4 | Species area relationship (SAR) models for Azorean taxa and General Dynamic Models (GDM) of oceanic island biogeography for Canarian taxa.

Model	Formula	R^2	AICc
SAR Azorean Spiders (training)			
Power	$S = 13.379A^{0.438}$	0.642	32.505
Exponential	$S = 0.549 + 4.538 \log A$	**0.780**	28.102
Linear	$S = 19.357 + 0.017A$	0.435	36.604
SR	$S = 4.641 \log A$	**0.780**	**23.319**
SAR Azorean Bryophytes (testing)			
Power	$S = 181.625A^{0.803}$	0.666	78.085
Exponential	$S = -27.824 + 57.114 \log A$	**0.728**	76.208
Linear	$S = 196.215 + 0.259A$	0.617	79.295
SR	$S = 51.889 \log A$	0.722	**71.617**
GDM Canarian Spiders (training)			
Whittaker	$S = -185.589 + 41.732 \log A + 17.776T - 1.022T^2$	0.873	110.350
Fattorini	$\log S = 2.585 + 0.281 \log A + 0.157T - 0.009T^2$	0.941	105.025
Steinbauer	$\log S = 3.367 + 0.098 \log A + 1.502 \log T - 0.454 \log T^2$	0.814	113.007
SR	$S = 42.283 + 0.051A + 17.379T - T^2$	**0.952**	**61.505**
GDM Canarian Bryophytes (testing)			
Whittaker	$S = -176.599 + 66.602 \log A + 21.361T - 1.620T^2$	**0.773**	**125.214**
Fattorini	$\log S = 4.544 + 0.137 \log A + 0.126T - 0.009T^2$	**0.803**	**124.217**
Steinbauer	$\log S = 5.136 + 0.017 \log A + 1.063 \log T - 0.382 \log T^2$	0.612	128.963
SR	$S = 192.660 + 0.075A + 20.702T - 1.576T^2$	**0.785**	**124.841**

S = native species richness; A = area of the island; T = maximum time of emergence. Best models are indicated in bold.

series from data (Martin et al., 2018; Chen et al., 2019). Yet, our study is the first to bring together all these disparate applications, demonstrating the versatility of this tool by applying it to a range of diverse ecological and evolutionary problems, both theoretical and practical. These results suggest that the true potential for the use of SR in ecology and evolution is yet to be fulfilled.

From Particular to General Principles

Scientific fields such as physics rarely rely on general statistical inference methods such as linear regression for hypothesis testing. The complexity of ecology made such methods an imperative in most cases. Symbolic regression not only allows the discovery of relationships specific to particular datasets, but also the finding of general models, globally applicable to multiple systems of particular nature, as we tried to exemplify. SR has the significant advantage of generating a variety of expressions from the given function set. For example, other methods may be configured to fit a polynomial to the data, but the user has to specify the degree. In SR, the power function in addition to the four basic arithmetic operations, allows the generation of unlimited degree polynomials, therefore providing a wider exploration of the solution space.

As mentioned, SR is designed to optimize both the form of the equations and the fitting parameters simultaneously. The fitting parameters usually are specific to each dataset, but the form may give clues toward general principles. For example, all archipelagos will follow an ISAR, even if each archipelago will have its own c and z values. Although this aspect has not been explored in this study, we suggest two ways of finding general principles. First, as was hinted by our estimators' example, one may independently analyse multiple datasets from the same type of systems. From each dataset, one or multiple equations may arise. Many of these will be similar in form even if the fitting parameters are different. Terms repeated in several equations along the Pareto front or with different datasets tend to be meaningful (Schmidt and Lipson, 2009). We may then try to fit the most promising forms to all datasets optimizing the fitting parameters to each dataset and look for which forms seem to have general value over all data. Second, one may simultaneously analyse multiple datasets from the same type of systems but with a change to the general SR implementation. Instead of optimizing both form and fitting parameters, the algorithm may focus on finding the best form, with fitting parameters being optimized during the evaluation step of the evolution for each dataset independently. This parameter optimization could be done with standard methods such as quasi-newton or simplex (Nocedal and Wright, 1999). To our knowledge, this approach has yet to be implemented, but it would allow finding general models and possibly principles, independently of the idiosyncrasies of each dataset.

The Need for Human Inference

Many data mining techniques are regarded, and rightly so, as "black boxes." SR is transparent in this regard, as variables are related through human-interpretable formulas. This is particularly important if the goal is to find equations with both predictive and explanatory power, building the bridge between finding the pattern and explaining the driving process, or if a general principle is to be suggested.

Our results show that an automated discovery system can identify meaningful relationships in ecological data. Yet, as shown by our Canary Island spider SAR model, some equations might be very accurate but overfit the data. As with any relationship finding, either automated or human, correlation does not imply causation, and spurious relationships are not only possible but probable given complex enough data.

Although the method here presented is automated, it is part of a collaborative human–machine effort. The possibility of exploiting artificial intelligence working together with human expertise can be traced back to Engelbart (1962), where the term "augmented intelligence" was coined to designate such collaboration. It has been subsequently developed and extended to teamwork involving one or more artificial intelligence agents together with one or (many) more humans, in diverse domains such as robotic teams (Yanco et al., 2004) or collective intelligence for evolutionary multi-objective optimization (Cinalli et al., 2015). In ecological problems, human knowledge may play a fundamental role: (1) in the beginning of the process, when selecting input variables, building blocks and SR parameters; and (2) in the interpretation and validation of equations. The choice of equations along a machine-generated Pareto front should also take advantage of human expert knowledge to identify the most interesting models to explain the data. The researcher might then decide to disregard, accept, or check equation validity using other methods.

A *priori* Knowledge

To some extent, it is possible to select a priori the type of models the algorithm will search for by selecting the functions to include in the function set. The choice of the function set is very important. A function set lacking a relevant function for the model may delay evolution or prevent it from finding a proper solution altogether. A way to take advantage of human expert knowledge is to seed the initial population of expressions with some we know are related with the problem. For example, when searching for the GDM we could have given the algorithm multiple forms of the ISAR to seed the search process. This is a directive approach and must be done in a parsimonious way, to avoid directing evolution too strongly, possibly trapping it in a local minimum. Such an approach should be complemented with random expressions in the initial population to create the necessary pool material for evolution to well explore the solution space. Therefore, a priori knowledge in SR has a stronger influence than in other inference methods, such as Bayesian, where a less adequate prior may be overcome by enough data.

Fine-Tuning the Process

The number of options in SR is immense. Population size is positively correlated with variability of models and how well the search space is explored, but might considerably slow the search. Mutation rates are also positively correlated with variability, but rates that are too high might prevent the algorithm converging on the best models. The fitness measure depends on the specific problem and on the type of noise in the data.

The number of generations to let the search run is entirely dependent on the problem complexity and time available. Often the algorithm reaches some equation that makes immediate sense to the researcher and the process can be immediately stopped for further analysis of results. Sometimes several competing equations seem to make sense but are not entirely convincing, in which case some indicators can be used as a stop rule, such as high values of stability and maturity of the evolution process (Schmidt, 2015).

The speed with which evolution occurs is extremely variable, depending on factors including the complexity of the relationships, having the appropriate variables and building blocks and the level of noise in the data. Fortunately, the process is easily adaptable to parallel computing, as many candidate functions can be evaluated simultaneously, allowing the use of multiple cores and even computer clusters to speed the search of equations.

Caveats

The SR approach is fully data-driven. This means it requires high-quality data if meaningful relationships are to be found. Also, it makes no a priori assumptions, so the final result might make no (obvious) sense, leading to spurious inferences, particularly if data are scarce or poor-quality, or if the right building blocks are not provided. Additionally, SR suffers from the same limitations of evolutionary algorithms in general. In many cases the algorithm may get stuck in local minima of the search space, requiring time (or even a restart with different parameters) to find the global minimum. Finally, SR suffers from the problem of bloat, which consists on an excessive growth of the expressions. There are mitigating approaches, like introducing a penalty for long expressions in the fitness function, or doing a posteriori symbolic analysis and simplification. However, bloat is still a problem under research.

Nevertheless, the fact that SR produces human legible expressions turns out to be useful even in the case of very large expressions. The expert eye can usually distinguish relevant fragments from a variety of unmeaningful segments of long expressions. And those relevant fragments often spur new thoughts and experiments.

The Automation of Science?

The methods here presented can be powerful additions to theoretical and experimental ecology, even if new conceptual hypotheses have to be created to accommodate the new equations. Such models could even be the only available means of investigating complex ecological systems when experiments are not feasible or datasets get too big/complex to model, using traditional statistical techniques (e.g., Tromas et al., 2017).

This family of techniques has led several authors to suggest the "automation of science" (King et al., 2009), where computers are able to advance hypotheses, test them, and reach conclusions in largely unassisted processes. This falls into the realm of exploiting knowledge (or symbolic)

driven AI together with data driven AI, or also automated machine learning an approach that recently began to gain momentum (e.g., Zhuang et al., 2017). SR potential is high in this automated science avenue since it bridges well from data to symbolic representations. What is clear already is its capability of producing formulas that help researchers to focus on initially imperceptible but interesting relationships within datasets and therefore SR may guide the process of hypothesis creation.

DATA AVAILABILITY STATEMENT

Eureqa scripts and Pareto front are available in **Data Sheet 2**. Data for Case study Modeling species richness are in Cardoso et al. (2013). Train and test datasets used for species distribution modeling are in **Data Sheet 3**. Dataset used in the case study Developing species richness estimators are available in the R package BAT (Cardoso et al., 2015). Dataset used in case study Developing the general dynamic model of oceanic island biogeography (GDM) are derived from Aranda et al. (2014) and Cardoso et al. (2010) see **Data Sheet 4**.

AUTHOR CONTRIBUTIONS

PC conceived the original idea and led the writing of the manuscript with LC. PC and VB performed the analyses with contributions from LC, JCC, and FR. PB and RG contributed with data. All authors contributed to the writing of the manuscript.

FUNDING

PC and VB were supported by Kone Foundation. PB and FR were partly funded by the project FCT-PTDC/BIA-BIC/119255/2010 - Biodiversity on oceanic islands: toward a unified theory. LC was supported by FCT through LASIGE Research Unit, ref. UIDB, UIDP/00408/2020. SM acknowledges support from the European Commission through Horizon 2020 Marie Skłodowska-Curie Actions (MSCA) individual fellowships (Grant no. 882221).

ACKNOWLEDGMENTS

We thanked Robert Whittaker, Stano Pekár, Michael Lavine, and Otso Ovaskainen for comments on earlier versions of the manuscript; Carla Gomes and Ronan Le Bras for fruitful discussions around AI and ecology.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fevo.2020.530135/full#supplementary-material

Data Sheet 1 | Examples of general principles in ecology and of some of the respective statistical models.

Data Sheet 2 | Data and settings used for all Symbolic Regression analyses in the paper (Eureqa file: http://www.nutonian.com/products/eureqa/).

Data Sheet 3 | Train and test databases used for species distribution modeling.

Data Sheet 4 | Species, area, and age for each Canarian Island.

REFERENCES

Akaike, H. (1974). New look at statistical-model identification. *Ieee T. Automat. Contr.* 19, 716–723. doi: 10.1109/TAC.1974.1100705

Allouche, O., Tsoar, A., and Kadmon, R. (2006). Assessing the accuracy of species distribution models: prevalence, kappa and the true skill statistic (TSS). *J. Appl. Ecol.* 43, 1223–1232. doi: 10.1111/j.1365-2664.2006.01214.x

Almeida, J., dos Santos, J. A., Miranda, W. O., Alberton, B., Morellato, L. P. C., and Torres, R. D. (2015). Deriving vegetation indices for phenology analysis using genetic programming. *Ecol. Inform.* 26, 61–69. doi: 10.1016/j.ecoinf.2015.01.003

Anand, M., Gonzalez, A., Guichard, F., Kolasa, J., and Parrott, L. (2010). Ecological systems as complex systems: challenges for an emerging science. *Diversity* 2, 395–410. doi: 10.3390/d2030395

Aranda, S. C., Gabriel, R., Borges, P. A. V., Santos, A. M. C., de Azevedo, E. B., Patino, J., et al. (2014). Geographical, temporal and environmental determinants of bryophyte species richness in the Macaronesian islands. *PLoS ONE* 9:e101786. doi: 10.1371/journal.pone.0101786

Araújo, M. B., Anderson, R. P., Barbosa, A. M., Beale, C. M., Dormann, C. F., Early, R., et al. (2019). Standards for distribution models in biodiversity assessments. *Sci. Adv.* 5:eaat4858. doi: 10.1126/sciadv.aat4858

Barrett, J., Kostadinova, A., and Raga, J. A. (2005). Mining parasite data using genetic programming. *Trends Parasitol.* 21, 207–209. doi: 10.1016/j.pt.2005.03.007

Barton, K. (2015). *MuMIn: Multi-Model Inference*. Available online at: https://CRAN.R-project.org/package=MuMIn (accessed September, 2020).

Bertoni, R., Bertoni, M., Morabito, G., Rogora, M., and Callieri, C. (2016). A non-deterministic approach to forecasting the trophic evolution of lakes. *J. Limnol.* 75, 242–252. doi: 10.4081/jlimnol.2016.1374

Cardoso, P. (2009). Standardization and optimization of arthropod inventories—the case of Iberian spiders. *Biodivers. Conserv.* 18, 3949–3962. doi: 10.1007/s10531-009-9690-7

Cardoso, P., Aranda, S. C., Lobo, J. M., Dinis, F., Gaspar, C., and Borges, P. A. V. (2009). A spatial scale assessment of habitat effects on arthropod communities of an oceanic island. *Acta Oecol.* 35, 590–597. doi: 10.1016/j.actao.2009.05.005

Cardoso, P., Arnedo, M. A., Triantis, K. A., and Borges, P. A. V. (2010). Drivers of diversity in Macaronesian spiders and the role of species extinctions. *J. Biogeogr.* 37, 1034–1046. doi: 10.1111/j.1365-2699.2009.02264.x

Cardoso, P., Gaspar, C., Pereira, L. C., Silva, I., Henriques, S. S., da Silva, R. R., et al. (2008a). Assessing spider species richness and composition in Mediterranean cork oak forests. *Acta Oecol.* 33, 114–127. doi: 10.1016/j.actao.2007.10.003

Cardoso, P., Rigal, F., and Carvalho, J. C. (2015). BAT - Biodiversity Assessment Tools, an R package for the measurement and estimation of alpha and beta taxon, phylogenetic and functional diversity. *Methods Ecol. Evol.* 6, 232–236. doi: 10.1111/2041-210X.12310

Cardoso, P., Rigal, F., Fattorini, S., Terzopoulou, S., and Borges, P. A. V. (2013). Integrating Landscape disturbance and indicator species in conservation studies. *PLoS ONE* 8:e63294. doi: 10.1371/journal.pone.0063294

Cardoso, P., Scharff, N., Gaspar, C., Henriques, S. S., Carvalho, R., Castro, P. H., et al. (2008b). Rapid biodiversity assessment of spiders (Araneae) using semi-quantitative sampling: a case study in a Mediterranean forest. *Insect Conserv. Diver.* 1, 71–84. doi: 10.1111/j.1752-4598.2007.00008.x

Chen, Y., Angulo, M. T., and Liu, Y. Y. (2019). Revealing complex ecological dynamics via symbolic regression. *Bioessays* 41:1900069. doi: 10.1002/bies.201900069

Cinalli, D., Martí, L., Sanchez-Pi, N., and Garcia, A. C. B. (2015). "Collective preferences in evolutionary multi-objective optimization: techniques and potential contributions of collective intelligence," in *30th Annual*

ACM Symposium on Applied Computing (New York, NY), 133–138. doi: 10.1145/2695664.2695926

Clench, H. (1979). How to make regional lists of butterflies: some thoughts. *J. Lepid. Soc*. 33, 216–231.

Desjardins-Proulx, P., Poisot, T., and Gravel, D. (2019). Artificial intelligence and synthesis in ecology and evolution. *Front. Ecol. Evol*. 7:402. doi: 10.3389/fevo.2019.00402

Dodds, W. K. (2009). *Laws, Theories and Patterns in Ecology*. Berkeley, CA: University of California Press. doi: 10.1525/california/9780520260405.001.0001

Dormann, C. F., Elith, J., Bacher, S., Buchmann, C., Carl, G., Carré, G., et al. (2013). Collinearity: a review of methods to deal with it and a simulation study evaluating their performance. *Ecography* 3, 27–46. doi: 10.1111/j.1600-0587.2012.07348.x

Dubčáková, R. (2011). Eureqa: software review. *Genet. Program. Evol. M*. 12, 173–178. doi: 10.1007/s10710-010-9124-z

Elith, J., Graham, C. H., Anderson, R. P., Dudik, M., Ferrier, S., Guisan, A., et al. (2006). Novel methods improve prediction of species' distributions from occurrence data. *Ecography* 29, 129–151. doi: 10.1111/j.2006.0906-7590.04596.x

Engelbart, D. (1962). *Augmenting Human Intellect: A Conceptual Framework*. Summary Report AFOSR-3233. Stanford Research Institute, Menlo Park, CA, United States. doi: 10.21236/AD0289565

Evans, M. R., Grimm, V., Johst, K., Knuuttila, T., de Langhe, R., Lessells, C. M., et al. (2013). Do simple models lead to generality in ecology? *Trends Ecol. Evol*. 28, 578–583. doi: 10.1016/j.tree.2013.05.022

Fattorini, S. (2009). On the general dynamic model of oceanic island biogeography. *J. Biogeogr*. 36, 1100–1110. doi: 10.1111/j.1365-2699.2009.02083.x

Hastie, T. (2015). *gam: Generalized Additive Models*. Available online at: https://CRAN.R-project.org/package=gam (accessed September, 2020).

Holland, J. H. (1975). *Adaptation in Natural and Artificial Systems: An Introductory Analysis With Applications to Biology, Control, and Artificial Intelligence*. Ann Arbor, MI: University of Michigan Press.

Holland, J. H. (1995). *Hidden Order: How Adaptation Builds Complexity*. Reading, MA: Addison-Wesley.

Holland, J. H. (1998). *Emergence: From Chaos to Order*. Reading, MA: Addison-Wesley.

Hurvich, C. M., and Tsai, C. L. (1989). Regression and time series model selection in small samples. *Biometrika* 76, 297–307. doi: 10.1093/biomet/76.2.297

Jagupilla, S. C. K., Vaccari, D. A., Miskewitz, R., Su, T. L., and Hires, R. I. (2015). Symbolic regression of upstream, stormwater, and tributary *E. coli* concentrations using river flows. *Water Environ. Res*. 87, 26–34. doi: 10.1002/j.1554-7531.2015.tb00138.x

King, R. D., Rowland, J., Oliver, S. G., Young, M., Aubrey, W., Byrne, E., et al. (2009). The automation of science. *Science* 324, 85–89. doi: 10.1126/science.1165620

Koza, J. R. (1992). *Genetic Programming : On the Programming of Computers by Means of Natural Selection*. Cambridge, MA: MIT Press.

Larsen, P. E., Cseke, L. J., Miller, R. M., and Collart, F. R. (2014). Modeling forest ecosystem responses to elevated carbon dioxide and ozone using artificial neural networks. *J. Theor. Biol*. 359, 61–71. doi: 10.1016/j.jtbi.2014.05.047

Larsen, P. E., Field, D., and Gilbert, J. A. (2012). Predicting bacterial community assemblages using an artificial neural network approach. *Nat. Methods* 9, 621–625. doi: 10.1038/nmeth.1975

Lawton, J. H. (1996). Patterns in ecology. *Oikos* 75, 145–147. doi: 10.2307/3546237

Liu, C. R., Berry, P. M., Dawson, T. P., and Pearson, R. G. (2005). Selecting thresholds of occurrence in the prediction of species distributions. *Ecography* 28, 385–393. doi: 10.1111/j.0906-7590.2005.03957.x

Lu, Y. (2019). Artificial intelligence: a survey on evolution, models, applications and future trends. *J. Manag. Anal*. 6, 1–29. doi: 10.1080/23270012.2019.1570365

Manson, S. M. (2005). Agent-based modeling and genetic programming for modeling land change in the Southern Yucatan Peninsular Region of Mexico. *Agr. Ecosyst. Environ*. 111, 47–62. doi: 10.1016/j.agee.2005.04.024

Manson, S. M., and Evans, T. (2007). Agent-based modeling of deforestation in southern Yucatan, Mexico, and reforestation in the Midwest United States. *Proc. Natl. Acad. Sci. U.S.A*. 104, 20678–20683. doi: 10.1073/pnas.0705802104

Martin, B. T., Munch, S. B., and Hein, A. M. (2018). Reverse-engineering ecological theory from data. *P Roy. Soc. B-Biol. Sci*. 285:20180422. doi: 10.1098/rspb.2018.0422

Mitchell, M. (2009). *Complexity: A Guided Tour*. Oxford England, New York, NY: Oxford University Press.

Muttil, N., and Chau, K. (2006). Neural network and genetic programming for modelling coastal algal blooms. *Int. J. Environ. Pollut*. 28, 223–238. doi: 10.1504/IJEP.2006.011208

Muttil, N., and Lee, J. H. W. (2005). Genetic programming for analysis and real-time prediction of coastal algal blooms. *Ecol. Model*. 189, 363–376. doi: 10.1016/j.ecolmodel.2005.03.018

Nocedal, J., and Wright, S. J. (1999). *Numerical Optimization*. New York, NY: Springer. doi: 10.1007/b98874

Passy, S. I. (2012). A hierarchical theory of macroecology. *Ecol. Lett*. 15, 923–934. doi: 10.1111/j.1461-0248.2012.01809.x

Phillips, S. J., Anderson, R. P., and Schapire, R. E. (2006). Maximum entropy modeling of species geographic distributions. *Ecol. Model*. 190, 231–259. doi: 10.1016/j.ecolmodel.2005.03.026

R Core Team (2015). *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing.

Ratkowsky, D. A. (1990). *Handbook of Nonlinear Regression Models*. New York, NY: M. Dekker.

Reshef, D. N., Reshef, Y. A., Finucane, H. K., Grossman, S. R., McVean, G., Turnbaugh, P. J., et al. (2011). Detecting novel associations in large data sets. *Science* 334, 1518–1524. doi: 10.1126/science.1205438

Russell, S. J., Norvig, P., and Davis, E. (2010). *Artificial Intelligence: A Modern Approach, 3rd Edn*. Upper Saddle River, NJ: Prentice Hall.

Schmidt, M., and Lipson, H. (2009). Distilling free-form natural laws from experimental data. *Science* 324, 81–85. doi: 10.1126/science.1165893

Schmidt, M. L. H. (2015). *Eureqa*. Available online at: https://www.nutonian.com/ (accessed September, 2020).

Smits, G. F., and Kotanchek, M. (2005). "Pareto-front exploitation in symbolic regression," in *Genetic Programming Theory and Practice II*, eds U.-M. O'Reilly, T. Yu, R. Riolo, and B. Worzel (Boston, MA: Springer), 283–299. doi: 10.1007/0-387-23254-0_17

Soberón, J., and Llorente, J. (1993). The use of species accumulation functions for the prediction of species richness. *Conserv. Biol*. 7, 480–488. doi: 10.1046/j.1523-1739.1993.07030480.x

Solé, R. V., and Goodwin, B. C. (2000). *Signs of Life: How Complexity Pervades Biology*. New York, NY: Basic Books.

Steinbauer, M. J., Klara, D., Field, R., Reineking, B., and Beierkuhnlein, C. (2013). Re-evaluating the general dynamic theory of oceanic island biogeography. *Front. Biogeogr*. 5, 185–194. doi: 10.21425/F5FBG19669

Sugihara, G., May, R., Ye, H., Hsieh, C. H., Deyle, E., Fogarty, M., et al. (2012). Detecting causality in complex ecosystems. *Science* 338, 496–500. doi: 10.1126/science.1227079

Triantis, K. A., Guilhaumon, F., and Whittaker, R. J. (2012). The island species-area relationship: biology and statistics. *J. Biogeogr*. 39, 215–231. doi: 10.1111/j.1365-2699.2011.02652.x

Tromas, N., Fortin, N., Bedrani, L., Terrat, Y., Cardoso, P., Bird, D., et al. (2017). Characterizing and predicting cyanobacterial blooms in an 8-year amplicon sequencing time course. *ISME J*. 11, 1746–1763. doi: 10.1038/ismej.2017.58

Tung, C. P., Lee, T. Y., Yang, Y. C. E., and Chen, Y. J. (2009). Application of genetic programming to project climate change impacts on the population of Formosan Landlocked Salmon. *Environ. Modell. Softw*. 24, 1062–1072. doi: 10.1016/j.envsoft.2009.02.012

Whittaker, R. J., Triantis, K. A., and Ladle, R. J. (2008). A general dynamic theory of oceanic island biogeography. *J. Biogeogr*. 35, 977–994. doi: 10.1111/j.1365-2699.2008.01892.x

Yanco, H. A., Drury, J. L., and Scholtz, J. (2004). Beyond usability evaluation: analysis of human-robot interaction at a major robotics competition. *Hum-Comput. Interact*. 19, 117–149. Available online at: https://www.nist.gov/publications/beyond-usability-evaluation-analysis-human-robot-interaction-major-robotics-competition

Yao, M. J., Rui, J. P., Li, J. B., Dai, Y. M., Bai, Y. F., Hedenec, P., et al. (2014). Rate-specific responses of prokaryotic diversity and structure to nitrogen deposition in the *Leymus chinensis* steppe. *Soil Biol. Biochem*. 79, 81–90. doi: 10.1016/j.soilbio.2014.09.009

Zhuang, Y., Wu, F., Chen, C., and Pan, Y. (2017). Challenges and opportunities: from big data to knowledge in AI 2.0. *Front. Inform. Tech. El.* 18, 3–14. doi: 10.1631/FITEE.1601883

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Ecology and Evolution

ORIGINAL RESEARCH
published: 17 December 2020
doi: 10.3389/fevo.2020.607619

Habitat Fragmentation Increases Overall Richness, but Not of Habitat-Dependent Species

Jordan Chetcuti[1], William E. Kunin[2] and James M. Bullock[3]*

[1] Department of Botany, School of Natural Sciences, Trinity College Dublin, Dublin, Ireland, [2] School of Biology, Faculty of Biological Sciences, University of Leeds, Leeds, United Kingdom, [3] UK Centre for Ecology & Hydrology, Wallingford, United Kingdom

Debate rages as to whether habitat fragmentation leads to the decline of biodiversity once habitat loss is accounted for. Previous studies have defined fragmentation variously, but research needs to address "fragmentation *per se,*" which excludes confounding effects of habitat loss. Our study controls for habitat area and employs a mechanistic multi-species simulation to explore processes that may lead some species groups to be more or less sensitive to fragmentation *per se*. Our multi-land-cover, landscape-scale, individual-based model incorporates the movement of generic species, each with different land cover preferences. We investigate how fragmentation *per se* changes diversity patterns; within (alpha), between (beta) and across (gamma) patches of a focal-land-cover, and if this differs among species groups according to their specialism and dependency on this focal-land-cover. We defined specialism as the increased competitive ability of specialists in suitable habitat and decreased ability in less suitable land covers compared to generalist species. We found fragmentation *per se* caused an increase in gamma diversity in the focal-land-cover if we considered all species regardless of focal-land-cover preference. However, critically for conservation, the gamma diversity of species for whom the focal land cover is suitable habitat declined under fragmentation *per se*. An exception to this finding occurred when these species were specialists, who were unaffected by fragmentation *per se*. In general, focal-land-cover species were under pressure from the influx of other species, with fragmentation *per se* leading to a loss of alpha diversity not compensated for by increases in beta diversity and, therefore, gamma diversity fell. The specialist species, which were more competitive, were less affected by the influx of species and therefore alpha diversity decreased less with fragmentation *per se* and beta diversity compensated for this loss, meaning gamma diversity did not decrease. Our findings help to inform the fragmentation *per se* debate, showing that effects on biodiversity can be negative or positive, depending on species' competitive abilities and dependency on the fragmented land cover. Such differences in the effect of fragmentation *per se* would have important consequences for conservation. Focusing conservation efforts on reducing or preventing fragmentation in areas with species vulnerable to fragmentation.

Keywords: species diversity, fragmentation *per se*, individual-base model, specialisation, habitat association, habitat preference, movement ecology, simulation

OPEN ACCESS

Edited by:
John Maxwell Halley, University of Ioannina, Greece

Reviewed by:
Alessandro Chiarucci, University of Bologna, Italy
Athanasios Kallimanis, Aristotle University of Thessaloniki, Greece

****Correspondence:***
Jordan Chetcuti
chetcutj@tcd.ie

Specialty section:
This article was submitted to Models in Ecology and Evolution, a section of the journal Frontiers in Ecology and Evolution

Received: *17 September 2020*
Accepted: *30 November 2020*
Published: *17 December 2020*

Citation:
Chetcuti J, Kunin WE and Bullock JM (2020) Habitat Fragmentation Increases Overall Richness, but Not of Habitat-Dependent Species. Front. Ecol. Evol. 8:607619. doi: 10.3389/fevo.2020.607619

INTRODUCTION

Humans have modified over 75% of the global land area, and the resulting habitat loss and degradation are recognized as the principal drivers of biodiversity declines (IPBES, 2018). A major consequence of landscape modification is that, aside from declines in habitat area, previously large blocks of natural habitats have become fragmented into smaller patches within a matrix of human-modified land-use such as farms and cities (Haddad et al., 2015). It is clear that habitat loss reduces species diversity, simply by shrinking the areas in which species using that habitat can live (MacArthur and Wilson, 1967; Hodgson et al., 2011; Keil et al., 2015). However, the effect of fragmenting habitats is less clear. "Fragmentation *per se*" (FPS) refers to the effects of fragmentation after taking account of, or in the absence of, habitat loss (Fahrig, 2003). Put another way, if there is no habitat loss, FPS comprises the altered spatial configuration of habitat, such that remaining patches are smaller but more numerous. The conservation literature has been focused on fragmentation in general being detrimental to biodiversity (Lawton et al., 2010; Eigenbrod et al., 2017). However, debate continues as to whether the effect of FPS on biodiversity is always negative (Fletcher et al., 2018a), or whether it is insignificant or positive (Fahrig, 2017; Fahrig et al., 2019). In reality, FPS and loss of habitat are intrinsically linked (Fletcher et al., 2018a). Nonetheless, separating the effects of FPS from those of area loss and assessing under what circumstances FPS leads to higher or lower species diversity are important for conservation decisions, such as restoration of habitat networks (Isaac et al., 2018) and the choices made by land managers about where to focus conservation efforts. Decisions include whether to conserve multiple small or fewer large habitat patches (Tulloch et al., 2016) or to allow activities that may limit loss of habitat, but increase fragmentation (Miller-Rushing et al., 2019). While there has been speculation about the mechanisms by which FPS may have positive or negative effects on biodiversity, these mechanisms require theoretical development and testing.

For example, understanding how positive vs. negative effects of FPS on diversity is driven by species' characteristics such as ecological specializations and habitat associations will aid decisions about how to manage specific landscapes. It is often assumed that specialist species and those that are positively associated with the habitat that is becoming fragmented should be negatively affected by FPS (Kosydar et al., 2014; Halstead et al., 2019). If studies report a positive effect of FPS on biodiversity, one explanation given is that species richness and abundance of generalists increases with habitat fragmentation, leading to this rise in diversity (Hu et al., 2012). But in 97% of the studies considered in a review by Fahrig (2017), FPS had a positive effect on the landscape-level (gamma) diversity of apparently specialist, rare, or threatened species. This could be because FPS allows for separation of otherwise competing species among habitat patches within the landscape (Ramiadantsoa et al., 2018). However, this mechanism requires more clarity as to the distinction between specialists and generalists. Specialists and generalists are often defined by an association with a particular habitat or with many, respectively, but this association is open to interpretation (Da Silveira et al., 2016). Being a generalist does not mean the species has no habitat preferences (Townsend et al., 2008; Da Silveira et al., 2016). Chetcuti et al. (2019) analyzed the habitat associations of hundreds of beetle species, showing most species had a positive association with several habitats, and only a few species showed a strong restriction to only a narrow range of habitats. It is also often assumed that specialists are more competitive when in a preferred or more suitable habitat compared to generalists, but the generalists are more competitive on average across multiple habitats (Marvier et al., 2004). Here we assess responses to FPS using clear definitions: specialists are more competitive in their suitable habitats than are generalists, but less competitive than the generalists elsewhere, even where the species may share the same habitat preferences within a landscape.

A further issue concerning mechanisms is relating patch-scale effects to landscape-scale impacts of fragmentation. Long-term manipulation experiments usually show that patch attributes typically associated with fragmentation (e.g., reduced patch size) reduce biodiversity at the scale of an individual patch, i.e., alpha-diversity (Haddad et al., 2015; Fletcher et al., 2018a; Damschen et al., 2019). However, it has been suggested that mechanisms identified in patch-scale studies may not extrapolate to negative effects on biodiversity at the landscape scale (Fahrig, 2017). Indeed, at the landscape scale, containing multiple patches of a habitat, Fahrig (2017) reported that different studies find either a neutral or a positive response of biodiversity (gamma-diversity) to FPS (Fahrig, 2017). By contrast, the species-fragmented area relationship suggests that negative effects of FPS should reduce gamma-diversity compared to that predicted by the species-area relationship alone (Hanski et al., 2013). However, patch-scale studies and the modeling that describes the species-fragmented area relationship do not take into consideration mechanisms that can lead to positive effects of fragmentation, such as increased beta-diversity caused by competitive release and higher habitat diversity (Fahrig et al., 2019; Rybicki et al., 2019). These mechanisms may increase beta-diversity and thus lead to overall increases in gamma-diversity with FPS.

It can be difficult to separate the effects of habitat area loss from those of FPS, as highly fragmented habitats are often in smaller patches (Fahrig, 2003). In general, conducting manipulative landscape-scale studies is difficult and it is often impossible to control for habitat area, which results in a confounding of FPS with habitat loss (Fahrig, 2003; Betts et al., 2019). The effects of area can be isolated statistically (De Camargo et al., 2018; Watling et al., 2020), but in these cases, the change in area can swamp any FPS signal (Fahrig, 2003). Theoretical modeling is a useful way to address contested issues where field data are difficult to collect and are subject to confounding variables. To this end, simulation models have been used to study FPS, which represent individual organisms moving across simulated landscapes (Gunton et al., 2017; Rybicki et al., 2019; Thompson et al., 2019). However, these studies are conducted on binary landscapes with the space between the focal-habitat patches (the habitat type of interest), the matrix, being a single, usually highly unsuitable, habitat type. Obviously, binary landscapes are rarely found in nature, and so using binary habitats likely reduces the relevance and applicability

of these simulation studies (Fardila et al., 2017). With only a single matrix habitat, one possible mechanism of FPS benefits is lost, that of increased habitat diversity. FPS can increase the diversity of other habitats adjacent to focal-habitat patches, and that in turn can increase the beta diversity of inhabitants of the focal-habitat (Fahrig et al., 2019; Rybicki et al., 2019). By having multiple matrix habitats, edge effects, which are typically considered a negative mechanism of FPS due to modification of edge due to changes in microclimate and increased predation from species outside of the habitat, can have a positive effect (Betts et al., 2019; Fahrig et al., 2019). In this study, we model FPS in terms of a single focal-habitat type, as is standard, but represent the matrix as a mix of different habitat types with randomized fragmentation and area (Bender and Fahrig, 2005). The multiple matrix habitats allow landscape-level effects to arise from species differences in their habitat preferences, thus better reflecting species' differences in nature (Bollmann et al., 2005; Betts et al., 2014; Brodie and Newmark, 2019; Chetcuti et al., 2019).

We use an individual based model (IBM) to provide a mechanistic assessment of FPS effects on alpha-, beta- and gamma-diversity, by simulating FPS at the landscape scale. We did this with multiple matrix habitats, and for species with differing specializations for, and dependencies on, the different habitats in the landscape. Dependencies are different from specialization as they relate to the habitats the species would do best in and species preference. Specialization on the other hand is if a species is a specialist or generalist. Even if two species had the same preference for habitats, the specialist would be more competitive than the generalist in the preferred habitat, but less so in less preferred habitats. Having created a complex simulation with multiple species and habitats, we keep the individual species simple, to allow clear conclusions to be drawn. We opt to focus the simulation on the interaction of these different species types with the landscape to assess if these would lead to differing effects of FPS on alpha-, beta- and gamma-diversity. We expect a loss in patch-scale alpha-diversity with FPS, at least for species favoring the focal-habitat, both because the area of patches is lower, and because this smaller area should reduce population viability. Conversely, we predict that FPS will cause beta-diversity to increase because FPS allows the persistence of more species among different focal-habitat patches. Furthermore, FPS will increase the edge-to-area ratio and so the degree to which the focal-habitat interfaces with matrix habitats. This will lead to higher beta-diversity of species for whom the focal-habitat has high suitability due to these species having access to a greater variety of secondary habitat, reflecting the different species-specific habitat preferences. This increased edge will also lead to an influx of species for whom the focal-habitat is less suitable. These tourist species, also called vagrants (Magurran, 2004; Rickert et al., 2012), could potentially counter some of the overall loss in alpha-diversity, but lead to higher competition for species dependent on focal-habitat. Gamma-diversity is a product of alpha- and beta-diversity. So, depending on the rate of decline of alpha-diversity compared to the rate of increase in beta-diversity with FPS, gamma-diversity can increase, decrease or stay the same.

Therefore, we hypothesize that, (1) Increasing fragmentation (FPS) of a habitat causes steeper declines in alpha-diversity of species associated with that habitat due to increased competition from an influx of "tourist" species from the matrix, and therefore the gamma-diversity will decrease with FPS. Focal-habitat dependent specialist species, which are more competitive in the focal-habitat they find most suitable, will be better able to hold out against the influx of tourist species. Generalists will be able to utilize more of the landscape and will coexist with the specialists, but will decline with fragmentation. Therefore, we hypothesize that, (2) Declines in alpha-diversity with fragmentation of a habitat (FPS) will be less steep where the species using that habitat divide into specialists and generalists, due to decreased competition, and so gamma-diversity will be either unaffected or increase with FPS. Additionally, we assess whether results are consistent at high (40%) vs. low (10%) levels of focal-habitat cover, testing the fragmentation threshold hypothesis that FPS should only have a negative effect when habitat amount is low (Fahrig, 2017; De Camargo et al., 2018). We include some of the possible factors influencing the impact of FPS, i.e., reduced competition and higher local habitat diversity surrounding fragmented patches. We also partially include edge effects, through the inherent increase in edge with fragmentation, although we do not include edge effects on the micro-climate.

MATERIALS AND METHODS

We created a multi-species and landscapes IBM simulation to assess the emergent properties arising from multiple individuals and species moving around a landscape containing a high number of patches and habitats (**Figure 1**). Our IBM was built using NetLogo (v6.0.4) (Wilensky, 1999). The model is a discrete-time, discrete-space model with the landscape being a grid of habitat cells. The simulation parameters were set up, run and the outputs analyzed using R version 3.5 (R Core Team, 2018). We describe the model in the **Supplementary Information** following the Overview, Design concepts, Details (ODD), protocol for describing individual-based models (Grimm et al., 2006, 2010), but summarize it here.

Generating Habitats and Landscapes

We generated landscapes containing patches of a focal-habitat and other matrix habitats to allow for exploration of FPS without confounding variables such as focal-habitat area loss that is often present in empirical data. We did this by writing an R script (**Supplementary Material 2**). The simulated landscapes were 1000 × 1000 (=10^6) cells in size and contained habitat patches that were a range of shapes (**Figure 1**). We generated eleven habitat types: the focal-habitat, and ten others comprising the matrix. In keeping with the known complexity of species habitat associations, we allowed species to have a diversity of associations with and use of habitats within the landscape (Betts et al., 2014; Chetcuti et al., 2019). We generated a new landscape for every model run.

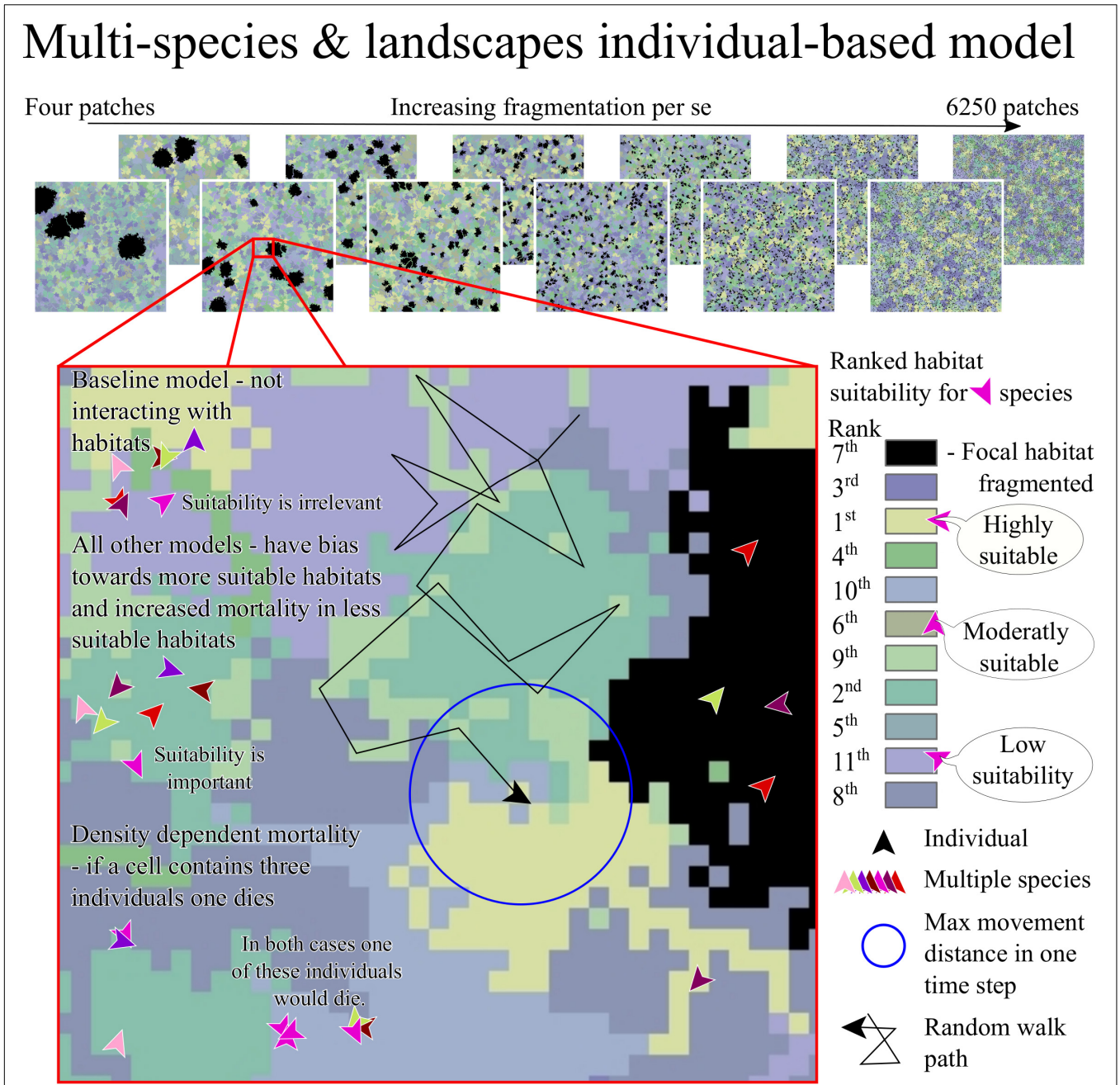


FIGURE 1 | A figurative description of the individual-based model, showing how we represented FPS by increasing the number of patches of the focal-habitat (in black) while keeping its total area the same. We give an example of the ranked suitability for habitats on the right for one species and an example of a random walk in the middle. The simulation used a baseline model, in which the individuals did not have differing mortality or movement bias for different habitats. We simulated two other scenarios in which the individuals interact with the habitats according to their assigned suitability. In the first scenario, the habitat modified mortality and individuals showed biased movement. The second scenario was the same, with the addition that half of the species were specialists and half generalists. We defined specialists and generalists as the former being more competitive in preferred habitats and less competitive in non-preferred habitats compared to generalists.

We generated fragmented landscapes by increasing the number of focal-habitat patches – being contiguous cells of the focal-habitat – while keeping the overall, landscape-level area of focal-habitat constant. We increased the number of focal-habitat patches following a geometric series, starting with four patches (allowing calculation of beta-diversity), up to a maximum number of patches. Patches were spatially separated by at least two cells. The maximum number of patches was defined as when the patches were a minimum size of four cells. We considered fragmentation in landscapes where the focal-habitat covered 10% or 40% of the landscape. When the focal-habitat covered 10% the maximum number of patches was 6250. For 40% cover, we used

a maximum number of patches of 8192, less than the theoretical maximum, but ensuring computational feasibility. For the focal-habitat, we imposed the number of patches and percentage habitat cover. The patches were located in the landscape by generating random coordinates (seeds) for starting locations using the R package "mobsim" (May et al., 2018). Our program then grew patches around the seeds by selecting adjacent cells at random. Patches grew until the area of the focal-habitat ("habitat one") reached the required amount (10% or 40% of the total cells in the habitat). The size of each patch was determined by using a uniform distribution allowing for a range of patch sizes. The program repeated the procedure for habitats two to eleven (in a random order) one at a time to fill remaining space. Each matrix habitat had between one and 200 patches and each covered a uniform random proportion of the matrix. This process was sequential; once patches of a habitat could grow no more, then that habitat was considered complete and the next randomly chosen habitat was grown. The last habitat was distributed differently, so that it filled all the remaining matrix.

Multi-Species Landscape Model Description

Our simulation had much in common with other IBMs, such as random-walking species, density-dependence at a cell scale, random distance of movement up to a maximum distance and random starting locations (Fahrig, 2001). To test our hypotheses, the simulation modeled different species types; which were not based on real species but designed to vary realistically in key characteristics while having the same dispersal ability. In the baseline model (which could be considered a type of neutral model) these species were identical, and were defined only by unique identifiers. In the two more complex models, we varied species in terms of their habitat-biased movement and habitat-modified mortality. To do this, we assigned each species a different set of associations with the eleven habitats. We did this by randomly ranking each of the eleven habitats from one (most suitable) to eleven (least suitable), and doing this anew for each species. Preference, dependence, association or ranked suitability, as used in this paper, are simply the rank value a habitat had for a species.

Individuals of each species (see "Modeled scenarios" concerning assignment of individuals to species) moved with a random walk in the baseline scenario and a habitat-biased random walk in the other scenarios (see "Modeled scenarios"). At each time-step the simulation iterated through individuals in random order so that the simulation did not always assess the same individuals first within each time-step. This random order was important when the population was over the carrying capacity and when assessing density-dependent mortality. To simulate density-dependent mortality, if a cell had more than two individuals after the movement phase in each time step, all but two individuals, chosen at random, died. Where movement and mortality were habitat-biased we used a logistic function defined by a midpoint and slope to determine a multiplier between zero and one for each habitat (**Figure 2**). The multiplier for habitat-modified mortality increased the probability of dying in a time-step for individuals in less suitable habitats. The habitat-biased movement multiplier modified the probability of moving into a cell of a habitat, giving bias toward preferred habitat, but still allowing individuals to move into other habitats. Each individual did this by counting the cells of each habitat in the circle around it defined by the maximum movement distance (**Figure 1**) and multiplying these by the bias multiplier. Each habitat was then assigned a proportion of values between zero and one and a random number generated between zero and one selected a habitat (**Figure 3**). The probability of individuals moving to any point in the circle was equal in the baseline model. In the other scenarios with habitat-biased movement, individuals were more likely to choose to move into the habitat more preferred (those with a higher ranking) by their species within the circle around it.

We chose a maximum movement distance of individuals of five cells per time-step and 5×10^{-4} chance of reproducing during a time-step. These arbitrary values are realistic for different species. For example, based on allometric equations (Sibly et al., 2013) this could relate to: invertebrates if a cell was a meter and the time-step a minute, resulting in ∼5 m per minute and ∼260 offspring a year (525,600 min in a year $\times$ 5 $\times$ 10^{-4} = 260);, or birds or mammals if a cell was a kilometer and the time-step an hour, resulting in ∼5 km per hour and four offspring a year. To stop our simulation from running longer than the 24-h time-limit of the JASMIN HPC cluster LOTUS (Lawrence et al., 2013) we used, we chose a carrying capacity of 4000 individuals in the landscape. We implemented the carrying capacity by increasing the chance of dying for all individuals when numbers were higher than the carrying capacity. We added a bounding area around the edge of the landscape of 10 cells wide with each cell in the area being randomly assigned a different habitat, to prevent species with biased movement from being influenced by the edge of the simulation. This created an invisible edge with individuals remaining or leaving the landscape. Individuals that left the landscape died.

Modeled Scenarios

We generated 400 species per simulation run, starting with ten individuals of each species. In the baseline scenario, all had identical mortalities, fecundities and movement abilities, and no habitat preferences. In the other two models, species' were given ranked habitat suitabilities as described above, and these were generated anew for each simulation run using the R packages "gtools" (Warnes et al., 2018) to permute the order of the vector 1:11 to give a rank for each habitat and "prodlim" (Gerds, 2018) to avoid repeating a particular ranking for >1 species within a simulation run. For the habitat-dependency model, each species had movement and mortality modified by their habitat suitability (**Figure 2**). In the specialization scenario, we compared the effect of FPS on specialists and generalists. In this case, we created 200 of each type of species, using the values in **Figure 2** for the logistic slope for habitat bias and mortality. Specialist species had a slope value of one, and a higher bias toward more suitable habitats but higher mortality

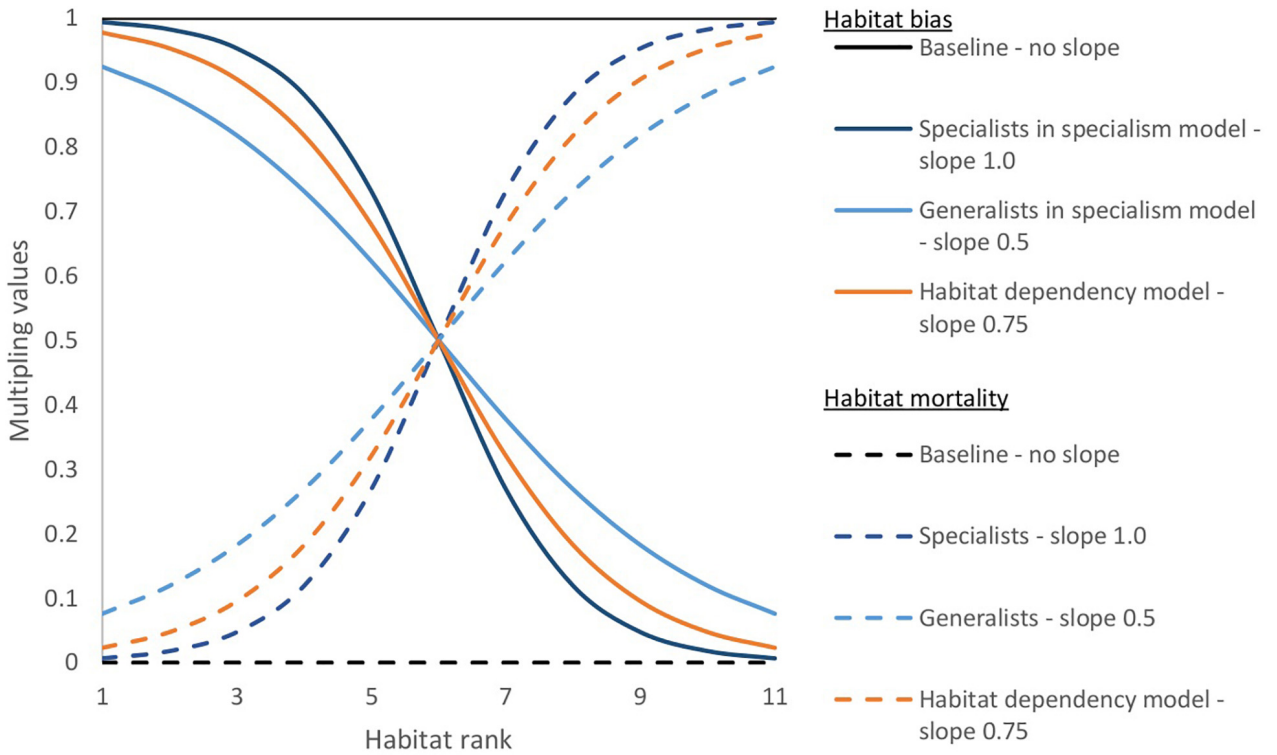


FIGURE 2 | Values used for the logistic slope within each scenario for habitat-biased movement and mortality. The graph shows the effect the slopes have on the multiplying values used to bias the movement toward more suitable habitat and to increase mortality in less suitable habitat. There is mortality due density-dependence and from being over the carrying capacity of the whole simulation. The habitat-modified mortality is additional mortality above the normal levels. To link levels of additional mortality to that of the reproductive rate, the habitat mortality is multiplied by the reproduction rate 5×10^{-4} to give the additional amount of mortality. We used the same scenarios and values for 10 and 40% cover simulations. The specialist species were more competitive in more suitable habitats than the species in the habitat-dependency model and those more so than the generalists. Competitiveness was reversed in less suitable habitat.

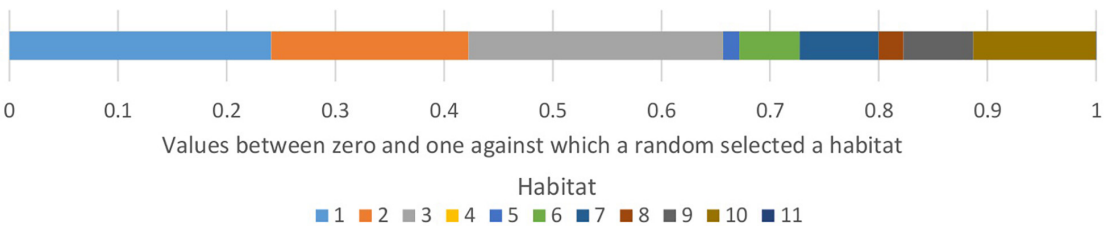


FIGURE 3 | A representation of how each individual chose where to move to in a time-step. It did this by multiplying the proportion of each habitat in a circle around it up to the maximum movement distance, by the bias multiplier. The values were normalized and stacked and then a random number between zero and one was drawn which selected the habitat. The individual then moved to a random cell of that habitat within the maximum movement distance.

in less suitable habitats than the generalists, which were species with a slope of 0.5.

We carried out preliminary simulation runs using four patches of the focal-habitat. We used these runs to calibrate the model, choosing values for habitat movement bias and modified mortality that allowed the simulation to run for 200,000 model time-steps and from these runs we realized the need to include a carrying capacity to limit the population. This number of time-steps allowed the number of species to reduce to close to the equilibrium number of species (i.e., if the model ran until no more species were lost), and allowed time-efficiency in running multiple models. Each scenario and FPS level combination was repeated 50 times. Seventy-one runs failed due to java issues on the clusters, resulting in the minimum number of replicates being 45.

Alpha-, Beta- and Gamma-Diversity

We calculated diversity scores for the focal-habitat only, habitat one, reflecting our focus on impacts of fragmentation of a single habitat type. At the end of the simulation, we counted species within each patch of the focal-habitat. We

then calculated the focal-habitat gamma-diversity, mean alpha-diversity per patch and mean pairwise (between pairs of patches) beta-sim-diversity (Barwell et al., 2015) using the R package "vegan" (Oksanen et al., 2019). We used beta-sim-diversity as it is considered the best metric for presence-absence data and is unaffected by sample size, which could be an issue here as our patches got smaller with FPS and therefore included fewer individuals (Koleff et al., 2003; Barwell et al., 2015). For the habitat-dependency and specialization models, we classified species into three groups: high suitability, those for whom the focal-habitat was highly suitable (rank one to three); low suitability (rank nine to eleven); and moderate suitability (all other species). The moderate suitability was therefore a bigger group and could contain more species.

Analysis of Results

We analyzed the effect of FPS on species diversity using generalized linear models for gamma-diversity (with a Poisson distribution) and alpha-diversity (with a gamma distribution), and beta regression for beta-diversity ("betareg") (values between zero and one) (Cribari-Neto and Zeileis, 2010). FPS was represented by the number of patches (on a log scale in the case of the beta-diversity, see **Supplementary Information**). Differences between pairs of scenarios were tested by including both scenarios and creating interaction terms. Due to the simulation nature of our study, using *p*-values is not advisable as significance can be forced simply running more replicates (White et al., 2014). We instead focus on effect size and 95% confidence intervals. The effect size is usually calculated over an increase of a unit of the independent variable. In our study this would be a single patch but this slight increase is not very informative. It is more appropriate to consider the effect size over the range of FPS simulated. We calculated the effects over the range of FPS using the R package "effects" (Fox, 2003; Fox and Weisberg, 2019).

RESULTS

Considering all species found across the focal-habitat patches, gamma-diversity increased with higher FPS in all models. In the baseline (neutral) model, where habitat did not influence movement or mortality, we observed that individuals became scattered randomly across the landscape. As a result, individual species became more concentrated by chance in different locations through random movement combined with reproduction, and conversely, became vacant from other parts; this resulted in increasing beta-diversity with higher FPS (**Supplementary Figure 1**). Because the species were equivalent in the baseline model, individual species only went extinct through stochasticity. The gamma-diversity, therefore, remained high after the 200,000-time-steps of the simulation. Although there was a positive effect of FPS on gamma-diversity in the baseline model, the 95% confidence interval of the slope included negative values and the effect size was low (**Figure 4**). By contrast, the FPS showed an increasing positive effect on gamma-diversity in the habitat dependency and specialization models; increasing gamma-diversity by 2.7 and 4.5 species, respectively, over the full range of fragmentation. In these habitat dependency and specialization models, the mean pairwise beta-diversity between patches increased faster than mean patch alpha-diversity declined with increasing FPS. When the focal-habitat had low FPS, beta-diversity was low, as the few large patches contained similar sets of species. As FPS increased, beta-diversity increased because there were more patches, and these were possibly in different landscape settings that suited different sets of species. By contrast, in the baseline model, the movement and mortality of species did not differ among the habitats, and so the species distributed across the landscape through stochastic processes only. Therefore, more patches in different landscape settings made no difference to the beta-diversity in the baseline model and gamma-diversity only increased slightly due to the increased number of patches sampling more of the landscape (+ 0.69 species). All results were qualitatively the same for high (40%) and low (10%) overall focal-habitat cover (**Supplementary Tables 1–3**). So, we present results for 10% cover results here, while those for 40% cover are in **Supplementary Tables 4–6**.

The habitat-dependency and specialization models had differences in habitat-dependent mortality and movement bias among species. This led to lower gamma-diversity values than in the baseline model as the species were more rapidly sorted in space and species less suited overall to the specific landscape of a simulation run died out. In these models, we observed that particular species became concentrated in areas of the landscape through their habitat associations (**Supplementary Figure 1**). In many cases, a few species dominated single habitat patches. In the specialization model, the gamma-diversity of the specialists and generalists together summed to give a higher overall gamma-diversity than in the simpler habitat-dependency model (**Figure 4**).

Considering the different species groups in the habitat-dependency model, the focal-habitat gamma-diversity of the species for whom the focal-habitat had low or moderate suitability increased with FPS (**Figure 5**). This was as expected, as the increased edge-to-area ratio under FPS would mean more of these species moved into focal-habitat patches by chance. Interestingly the gamma-diversity of species for whom the focal-habitat had high suitability declined with FPS. The reduction in gamma-diversity over the whole range of FPS was, 2 species, amounting to a 25% reduction. This reduction was also due to a greater amount of edge with higher FPS. In this case this greater edge meant these species were more likely to leave focal-habitat patches and also to be excluded from these patches by the influx of those species for whom the focal-habitat had low or moderate suitability.

In the specialization model, the gamma-diversity of the species for whom the focal-habitat had high suitability did not change with FPS. In contrast to the simpler habitat-dependency model, the specialist species were more competitive in habitats to which they were suited, so they were better able to resist species that found the habitat less suitable and their beta-diversity increased at a rate similar or slightly greater than the

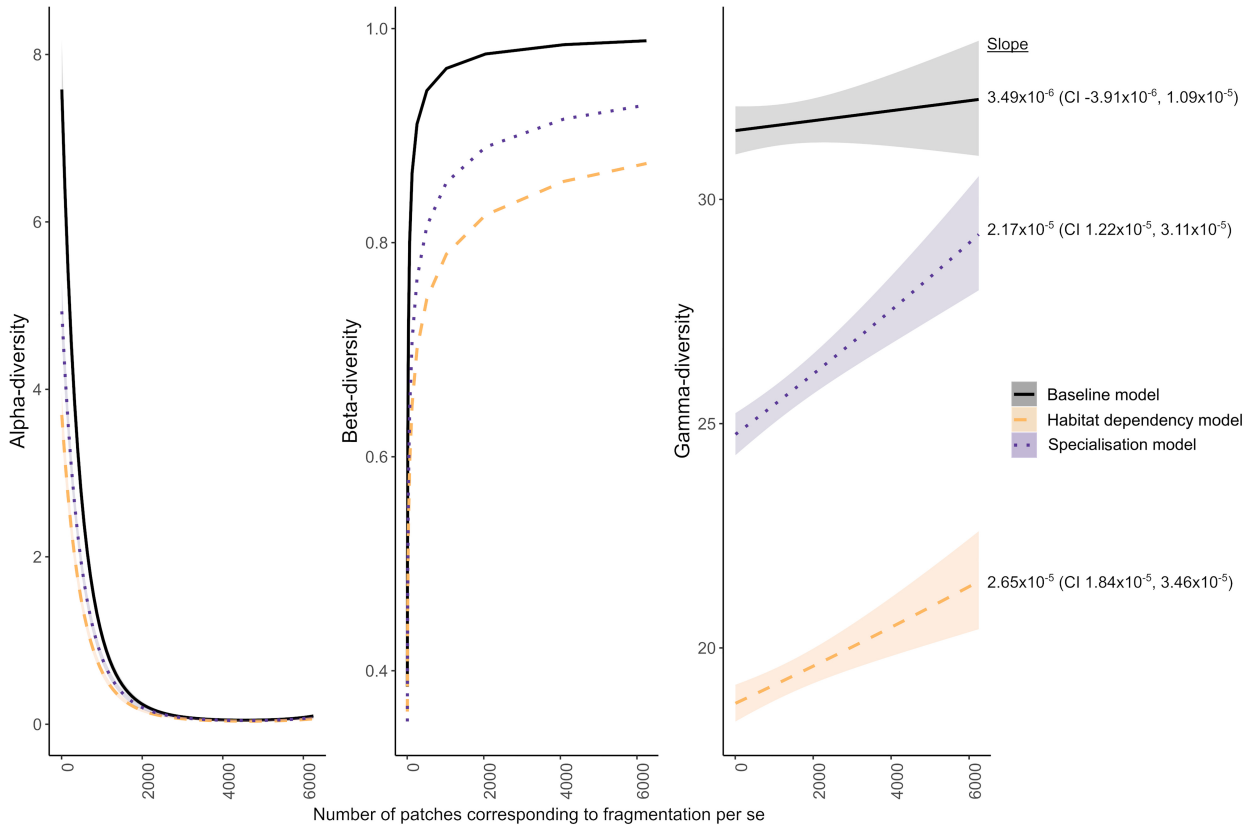


FIGURE 4 | Mean patch scale alpha-diversity, mean pairwise beta-diversity and gamma-diversity for all species in the focal-habitat at 10% cover, with fitted lines and standard errors. Gamma-diversity increased with the number of patches (albeit not greatly for the baseline model), which represents FPS. In all cases, alpha-diversity declined, and beta-diversity increased.

decline in alpha-diversity, so gamma-diversity did not decline (**Figure 6**). The generalist species for whom the focal-habitat had high suitability also did better under high FPS than the species in the habitat-dependency model (which had neither specialists nor generalists), as they were able to use more of the wider landscape.

DISCUSSION

This study helps to inform the debate on the effects of FPS on biodiversity (Fahrig, 2017; Fletcher et al., 2018a; Fahrig et al., 2019; Thompson et al., 2019). FPS had no effect or a positive effect on overall gamma-diversity of the focal-habitat across the patches in a landscape, but the gamma-diversity of species for which the habitat had high suitability could decline with FPS depending on species characteristics with respect to specialization and competitive ability. Our results were consistent when contrasting landscapes with relatively low (10%) to relatively high (40%) cover by the focal-habitat, suggesting some generality.

We found that beta-diversity and gamma-diversity increased overall even without differences among species in habitat specializations. In the baseline the increase in gamma-diversity was negligible, with a possible small increase caused by patches covering more of the landscape with FPS and increasing sampling of species which were aggregated through limited dispersal as predicted by neutral theory (Hubbell, 2011). Despite overall increases, in the habitat-dependency model, the gamma-diversity of species for whom the focal-habitat was highly suitable declined with FPS, while these species did not decline in the specialization model. In the habitat-dependency model, the species for whom the focal-habitat was highly suitable were likely under pressure due to the influx into the more fragmented patches by species for whom the habitat was less suitable, and the beta-diversity increase of these species suited to the focal-habitat did not outweigh the loss in alpha-diversity, so gamma-diversity declined.

In the specialization model, the specialists were more competitive against other species in the focal-habitat and therefore their beta-diversity increased at a similar rate to the decline in alpha-diversity decline with higher FPS, resulting in no change in the specialist species gamma-diversity with FPS. The gamma-diversity of the generalists did not decline, probably due to competitive release, or as they were better able to use multiple habitats outside of the focal-habitat. These findings indicate that the effect of FPS is context dependent. More competitive (specialist) species did not decline with FPS. But, the gamma-diversity of those species suited to the focal-habitat, but which were not more competitive there (in our habitat-dependency

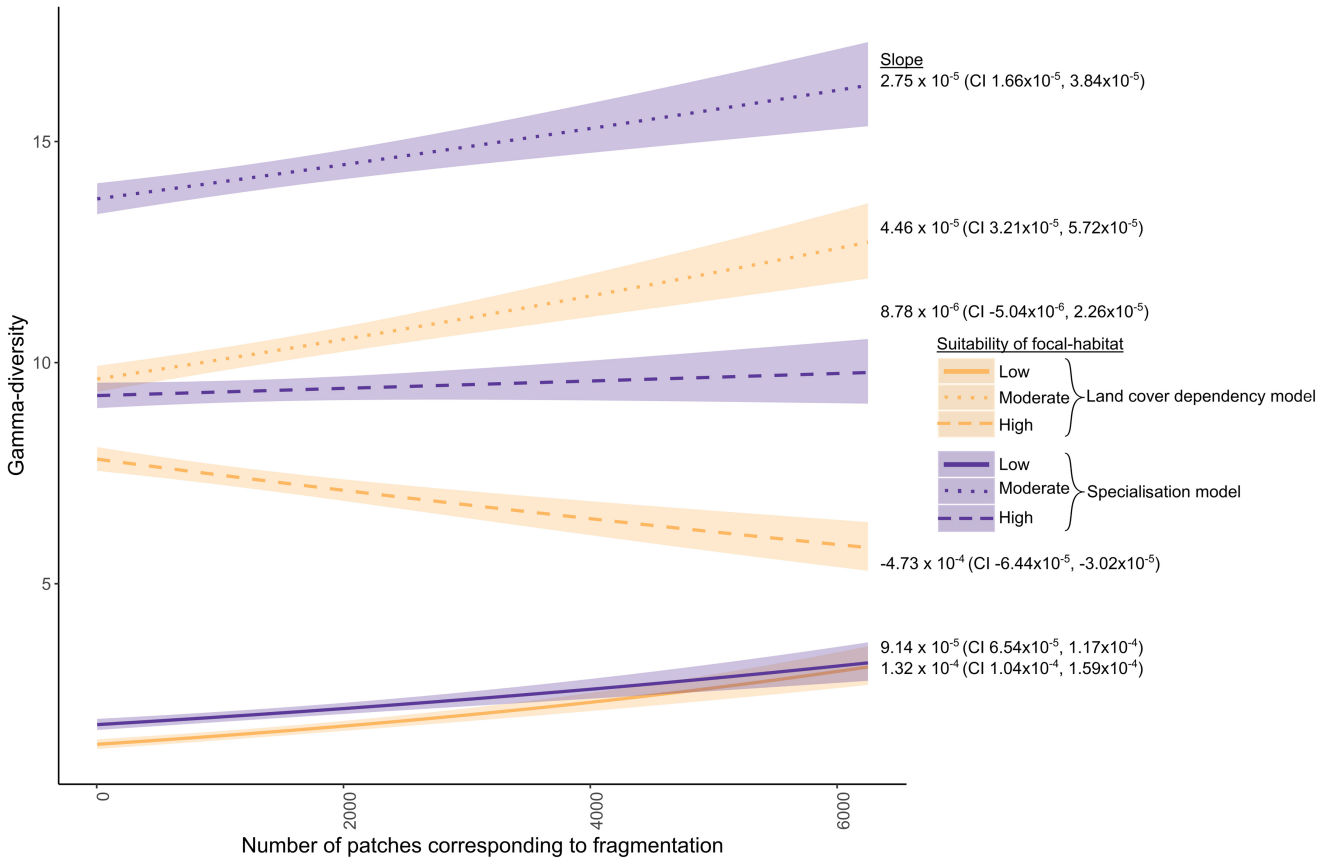


FIGURE 5 | Gamma-diversity for three groups of species – those for whom the focal-habitat had high, moderate, or low suitability – for the habitat dependency model (habitat bias and mortality slope 0.75) and specialization model (habitat bias and mortality slope 1 and 0.5, respectively). Gamma-diversity increased with FPS in both models for the species who for whom the focal-habitat had low or moderate suitability, and those for whom the focal-habitat had high suitability in the specialization model. By contrast, in the habitat dependency model, gamma-diversity declined with increasing FPS for the species for whom the focal-habitat had high suitability.

model) did decline with FPS. Such patterns might arise in nature where environmental change could cause species to become less competitive in their preferred habitat and therefore become more affected by FPS. For example, increased nutrients or climate change can change competitive abilities (Staley et al., 2011; Lancaster et al., 2017). Future studies might conduct simulations considering FPS more explicitly with such environmental change. It is often assumed that less competitive species have increased dispersal efficiency (Bonte et al., 2012). For simplicity, we did not include such differences in dispersal ability. If we had, we suggest a negative effect of FPS would be less likely as the generalists with increased dispersal would be able to better utilize other areas of the landscape. This could be explored in future studies.

Habitat dependency and the degree of specialization of species were very important in changing the direction of the relationship of gamma-diversity to FPS, suggesting information on species' habitat relationships are critical to planning landscape-scale conservation. In terms of conservation, it is often the less competitive species, with high dependencies on specific habitats that are of highest concern and that are the targets for conservation (Carrete et al., 2010; Fletcher et al., 2018b). The effect fragmentation has on these species should, therefore, be assessed and this might determine how the landscape should be managed to conserve these species. Doing so will have consequences for species in other ("matrix") habitats, however, and the resulting trade-offs should be analyzed and considered. Fragmentation *per se* creates smaller patches, which have lower mean alpha-diversity as shown in our modeling. Lower alpha-diversity has a negative effect on ecosystem functioning at the patch scale, but beta-diversity has been suggested as important at a larger scale in supporting multiple ecosystem functions (Mori et al., 2018). Our model used a series of static landscapes. Dynamic landscapes could be simulated in different ways, but if we had simulated dynamic landscapes that changed over time, without species arriving at the edge of the model domain, species diversity could not have increased. As patches are removed, moved or shrunk, some species would be extirpated, possibly with a lag (extinction debt). Without new species, the resulting colonization credit (Jackson and Sax, 2010) would go unfulfilled. A further implication of this, is that if fragmentation is happening over a large enough scale or in an isolated landscape, colonization may not be able to counter the debt.

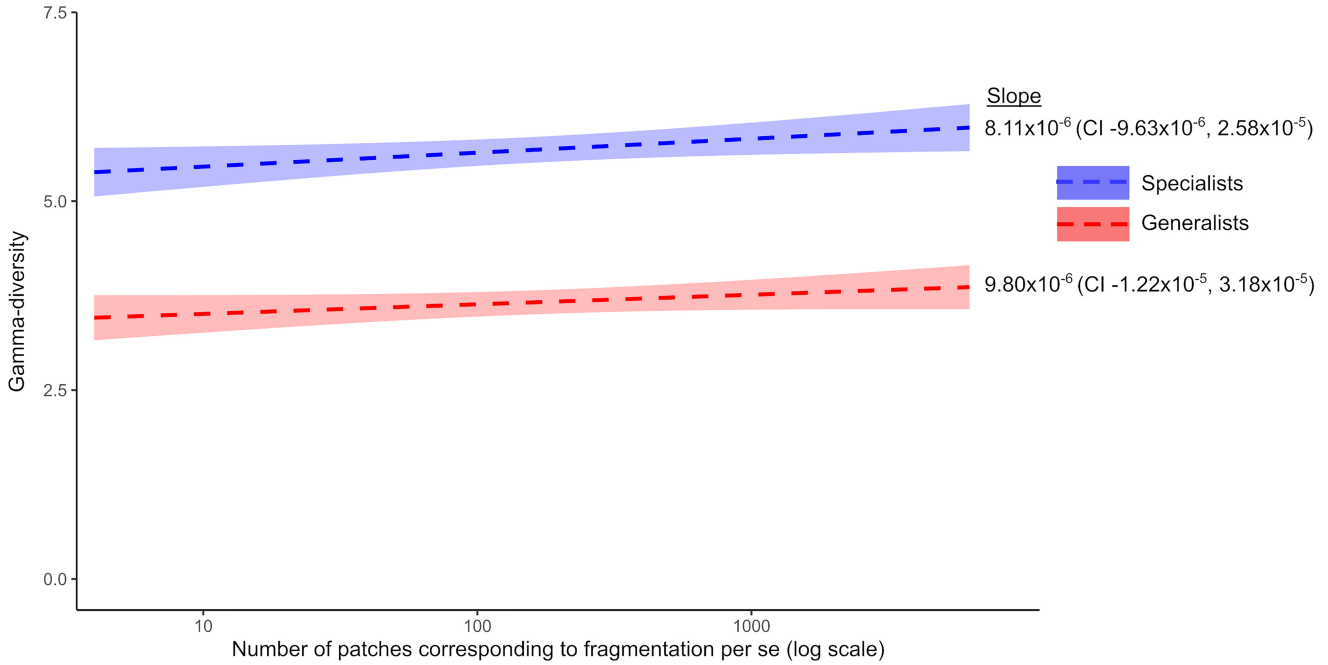


FIGURE 6 | Gamma-diversity for those species for whom the focal-habitat had high suitability for the specialization model (habitat bias and mortality slope 1 and 0.5, respectively) showing specialists and generalists separately. Gamma-diversity increases with FPS for both specialists and generalists in all cases.

If our simulation provided for colonization, we would expect the result of dynamic landscapes to be similar to that of our static landscapes.

We generated species and habitats at random, meaning our results are general and not specific to any real landscape or communities. Our simulations do not show whether any particular species would be retained or, conversely, lost with increasing FPS. We used a large species pool of potential diversity, providing each simulation run with 400 randomly generated species. We also randomly generated the habitat matrix between the patches of the focal-habitat, always having ten other habitats. Given the importance of the intervening habitat matrix in determining what species are in the landscape and how species move between patches (Brodie and Newmark, 2019; Chetcuti et al., 2019), future studies might look at the matrix specifically, non-randomly generating habitat matrices and including different mixes of anthropogenic and semi-natural habitats (Fletcher et al., 2018a). Our baseline model represented movement as a random walk, and we introduced bias based on habitat suitability in the more complex models. In reality, many organisms show complex movement behavior (Gurarie et al., 2016), which is likely to be important in modeling how FPS affects biodiversity and could be a focus of future research. If our results are representative of the ways in which introducing greater ecological complexity (movement, multiple habitat types, habitat dependency and specialization) can affect conclusions about FPS effects on species diversity, then they have important consequences for practical conservation. The habitat dependency model shows fragmentation can have strongly negative effects on (gamma) diversity. But where the biota shows strong division of species into specialists or generalists, gamma-diversity may be unaffected by fragmentation as long as the landscape contains a diversity of habitats. Considering how we define species as specialist or generalist and using analysis of habitat association, conservation efforts could be focused on mitigating fragmentation for those species groups that may be negatively affected by a fragmented landscape. Further, specialization could be looked at from the emergent perspective of species appearing specialist because of the other species in the landscape or the structure of the landscape. In the absence of more competitive species, such as on islands or isolated parts of the landscape, specialist species may utilize a broader range of habitats and benefit further from a heterogeneous landscape of resources and habitats.

This theoretical modeling considered FPS in heterogeneous landscapes, unpicking some of the mechanisms that can cause gamma-diversity to increase or decrease with FPS. Interestingly, we found that FPS could have a positive, negative or no effect on gamma-diversity, suggesting there is no simple answer to the question; is habitat fragmentation good or bad for biodiversity (Fletcher et al., 2018a; Fahrig et al., 2019). Species that were highly suited to the focal-habitat showed declining gamma-diversity with fragmentation, but where we added characteristics separating species into specialists and generalist, both did better under FPS. A key process was species' movement; for example species suited to the focal-habitat declined with FPS in the habitat-dependency model, as they were unable to hold out against increasing influxes of species for whom the focal-habitat had lower suitability. Our research opens new avenues for research into how species demography and movement in relation

to the focal-habitat affect biodiversity responses to FPS. Species' specializations, habitat preferences and demography in different habitats (Chetcuti et al., 2019) should be taken into consideration when planning conservation as well as considering that under some circumstances FPS may lead to the conservation objectives of increased beta-diversity.

DATA AVAILABILITY STATEMENT

The NetLogo simulation can be found at: https://zenodo.org/badge/latestdoi/289454188 and the R package LcvGen at: https://zenodo.org/badge/latestdoi/205820444. All other supporting data supporting the conclusions of this article will be made available by the authors, without undue reservation.

AUTHOR CONTRIBUTIONS

JC did the analysis and wrote the first draft of the manuscript. JB wrote parts of the text. JB and WK provided guidance on the building and parameterization of the simulation and provided significant guidance and editing of the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

JC was funded by a studentship from the NERC SPHERES Doctoral Training Partnership (NE/L002574/1). This work used the JASMIN at RAL STFC (http://jasmin.ac.uk), operated jointly by the centre of environmental data analysis and the scientific computing department. This facility was funded by NERC.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fevo.2020.607619/full#supplementary-material

REFERENCES

Barwell, L. J., Isaac, N. J. B., and Kunin, W. E. (2015). Measuring β - diversity with species abundance data. *J. Anim. Ecol.* 84, 1112–1122. doi: 10.1111/1365-2656.12362

Bender, D. J., and Fahrig, L. (2005). Matrix structure obscures the relationship between interpatch movement and patch size and isolation. *Ecology* 86, 1023–1033. doi: 10.1890/03-0769

Betts, M. G., Fahrig, L., Hadley, A. S., Halstead, K. E., Bowman, J., Robinson, W. D., et al. (2014). A species-centered approach for uncovering generalities in organism responses to habitat loss and fragmentation. *Ecography* 37, 517–527. doi: 10.1111/ecog.00740

Betts, M. G., Wolf, C., Pfeifer, M., Banks-Leite, C., Arroyo-Rodríguez, V., Ribeiro, D. B., et al. (2019). Extinction filters mediate the global effects of habitat fragmentation on animals. *Science* 366, 1236–1239. doi: 10.1126/science.aax9387

Bollmann, K., Weibel, P., and Graf, R. F. (2005). An analysis of central Alpine capercaillie spring habitat at the forest stand scale. *For. Ecol. Manage.* 215, 307–318. doi: 10.1016/j.foreco.2005.05.019

Bonte, D., Van Dyck, H., Bullock, J. M., Coulon, A., Delgado, M., Gibbs, M., et al. (2012). Costs of dispersal. *Biol. Rev.* 87, 290–312. doi: 10.1111/j.1469-185X.2011.00201.x

Brodie, J. F., and Newmark, W. D. (2019). Heterogeneous matrix habitat drives species occurrences in complex, fragmented landscapes. *Am. Nat.* 193, 748–754. doi: 10.1086/702589

Carrete, M., Lambertucci, S. A., Speziale, K., Ceballos, O., Travaini, A., Delibes, M., et al. (2010). Winners and losers in human-made habitats: Interspecific competition outcomes in two Neotropical vultures. *Anim. Conserv.* 13, 390–398. doi: 10.1111/j.1469-1795.2010.00352.x

Chetcuti, J., Kunin, W. E., and Bullock, J. M. (2019). A weighting method to improve habitat association analysis: tested on British carabids. *Ecography* 42, 1395–1404. doi: 10.1111/ecog.04295

Cribari-Neto, F., and Zeileis, A. (2010). Beta Regression in {R}. *J. Stat. Softw.* 34, 1–24.

Da Silveira, N. S., Niebuhr, B. B. S., Muylaert, R. D. L., Ribeiro, M. C., and Pizo, M. A. (2016). Effects of land cover on the movement of frugivorous birds in a heterogeneous landscape. *PLoS One* 11:e0156688. doi: 10.1371/journal.pone.0156688

Damschen, E. I., Brudvig, L. A., Burt, M. A., Fletcher, R. J., Haddad, N. M., Levey, D. J., et al. (2019). Ongoing accumulation of plant diversity through habitat connectivity in an 18-year experiment. *Science* 365, 1478–1480. doi: 10.1126/science.aax8992

De Camargo, R. X., Boucher-Lalonde, V. V. V., and Currie, D. J. (2018). At the landscape level, birds respond strongly to habitat amount but weakly to fragmentation. *Divers. Distrib.* 24, 629–639. doi: 10.1111/ddi.12706

Eigenbrod, F., Adams, W., Hill, C., Macgregor, N., Osborne, P., Clarke, D., et al. (2017).A review of large-scale conservation in England, Scotland and Wales. *Nat. Engl. Joint Publ.* JP019:227.

Fahrig, L. (2001). How much habitat is enough? *Biol. Conserv.* 100, 65–74. doi: 10.1016/S0006-3207(00)00208-1

Fahrig, L. (2003). Effects of Habitat Fragmentation on Biodiversity. *Annu. Rev. Ecol. Evol. Syst.* 34, 487–515. doi: 10.1146/annurev.ecolsys.34.011802.132419

Fahrig, L. (2017). Ecological Responses to Habitat Fragmentation Per Se. *Annu. Rev. Ecol. Evol. Syst.* 48:annurev-ecolsys-110316-122612. doi: 10.1146/annurev-ecolsys-110316-122612

Fahrig, L., Arroyo-Rodríguez, V., Bennett, J. R., Boucher-Lalonde, V., Cazetta, E., Currie, D. J., et al. (2019). Is habitat fragmentation bad for biodiversity? *Biol. Conserv.* 230, 179–186. doi: 10.1016/j.biocon.2018.12.026

Fardila, D., Kelly, L. T., Moore, J. L., and McCarthy, M. A. (2017). A systematic review reveals changes in where and how we have studied habitat loss and fragmentation over 20 years. *Biol. Conserv.* 212, 130–138. doi: 10.1016/j.biocon.2017.04.031

Fletcher, R. J., Didham, R. K., Banks-leite, C., Barlow, J., Ewers, R. M., Rosindell, J., et al. (2018a). Is habitat fragmentation good for biodiversity? *Biol. Conserv.* 226, 9–15. doi: 10.1016/j.biocon.2018.07.022

Fletcher, R. J., Reichert, B. E., and Holmes, K. (2018b). The negative effects of habitat fragmentation operate at the scale of dispersal. *Ecology* 99, 2176–2186. doi: 10.1002/ecy.2467

Fox, J. (2003). Effect Displays in {R} for Generalised Linear Models. *J. Stat. Softw* 8, 1–27.

Fox, J., and Weisberg, S. (2019). *An R Companion to Applied Regression*, 3rd Edn. Thousand Oaks CA: Sage.

Gerds, T. A. (2018). *prodlim: Product-Limit Estimation for Censored Event History Analysis*. Available online at: https://cran.r-project.org/package=prodlim.

Grimm, V., Berger, U., Bastiansen, F., Eliassen, S., Ginot, V., Giske, J., et al. (2006). A standard protocol for describing individual-based and agent-based models. *Ecol. Modell.* 198, 115–126. doi: 10.1016/j.ecolmodel.2006.04.023

Grimm, V., Berger, U., DeAngelis, D. L., Polhill, J. G., Giske, J., and Railsback, S. F. (2010). The ODD protocol: a review and first update. *Ecol. Modell.* 221, 2760–2768. doi: 10.1016/j.ecolmodel.2010.08.019

Gunton, R. M., Marsh, C. J., Moulherat, S., Malchow, A.-K. K., Bocedi, G., Klenke, R. A., et al. (2017). Multicriterion trade-offs and synergies for spatial conservation planning. *J. Appl. Ecol.* 54, 903–913. doi: 10.1111/1365-2664.12803

Gurarie, E., Bracis, C., Delgado, M., Meckley, T. D., Kojola, I., and Wagner, C. M. (2016). What is the animal doing? Tools for exploring behavioural structure in animal movements. *J. Anim. Ecol.* 85, 69–84. doi: 10.1111/1365-2656.12379

Haddad, N. M., Brudvig, L. A., Clobert, J., Davies, K. F., Gonzalez, A., Holt, R. D., et al. (2015). Habitat fragmentation and its lasting impact on Earth's ecosystems. *Sci. Adv.* 1, 1–9. doi: 10.1126/sciadv.1500052

Halstead, K. E., Alexander, J. D., Hadley, A. S., Stephens, J. L., Yang, Z., and Betts, M. G. (2019). Using a species-centered approach to predict bird community responses to habitat fragmentation. *Landsc. Ecol* 34, 1919–1935. doi: 10.1007/s10980-019-00860-5

Hanski, I., Zurita, G. A., Bellocq, M. I., and Rybicki, J. (2013). Species-fragmented area relationship. *Proc. Natl. Acad. Sci. U.S.A.* 110, 12715–12720. doi: 10.1073/pnas.1311491110

Hodgson, J. A., Moilanen, A., Wintle, B. A., and Thomas, C. D. (2011). Habitat area, quality and connectivity: striking the balance for efficient conservation. *J. Appl. Ecol.* 48, 148–152. doi: 10.1111/j.1365-2664.2010.01919.x

Hu, G., Wu, J., Feeley, K. J., Xu, G., and Yu, M. (2012). The effects of landscape variables on the species-area relationship during late-stage habitat fragmentation. *PLoS One* 7:e0043894. doi: 10.1371/journal.pone.0043894

Hubbell, S. P. (2011). *The Unified Neutral Theory of Biodiversity and Biogeography (MPB-32)*. Princeton: Princeton University Press.

IPBES (2018). "The IPBES assessment report on land degradation and restoration," in *Secretariat of the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services*, eds L. Montanarella, R. Scholes, and A. Brainich (Bonn: IPBES).

Isaac, N. J. B., Brotherton, P. N. M., Bullock, J. M., Gregory, R. D., Boehning-Gaese, K., Connor, B., et al. (2018). Defining and delivering resilient ecological networks: Nature conservation in England. *J. Appl. Ecol.* 55, 2537–2543. doi: 10.1111/1365-2664.13196

Jackson, S. T., and Sax, D. F. (2010). Balancing biodiversity in a changing environment: extinction debt, immigration credit and species turnover. *Trends Ecol. Evol.* 25, 153–160. doi: 10.1016/j.tree.2009.10.001

Keil, P., Storch, D., and Jetz, W. (2015). On the decline of biodiversity due to area loss. *Nat. Commun.* 6:8837. doi: 10.1038/ncomms9837

Koleff, P., Gaston, K. J., and Lennon, J. J. (2003). Measuring beta diversity for presence-absence data. *J. Anim. Ecol.* 72, 367–382. doi: 10.1046/j.1365-2656.2003.00710.x

Kosydar, A. J., Conquest, L. L., and Tewksbury, J. J. (2014). Can life histories predict the effects of habitat fragmentation? a meta-analysis with terrestrial mammals. *Appl. Ecol. Environ. Res.* 12, 505–521. doi: 10.15666/aeer/1202_505521

Lancaster, L. T., Morrison, G., and Fitt, R. N. (2017). Life history trade-offs, the intensity of competition, and coexistence in novel and evolving communities under climate change. *Philos. Trans. R. Soc. B Biol. Sci.* 372:20160046. doi: 10.1098/rstb.2016.0046

Lawrence, B. N., Bennett, V. L., Churchill, J., Juckes, M., Kershaw, P., Pascoe, S., et al. (2013). "Storing and manipulating environmental big data with JASMIN," in *Proceedings of the 2013 IEEE International Conference on Big Data, Big Data 6-9 2013*, San Francisco, CA, 68–75. doi: 10.1109/BigData.2013.6691556

Lawton, J. H., Brotherton, P. N. M., Brown, V. K., Elphick, C., Fitter, A. H., Forshaw, J., et al. (2010). Making space for nature: a review of England's wildlife Sites and ecological network. *Rep. to Defra* 107.

MacArthur, R. H., and Wilson, E. O. (1967). *The Theory of Island Biogeography*. Princeton: Princeton University Press.

Magurran, A. (2004). *Measuring Biological Diversity*. Hoboken, NJ: Wiley, doi: 10.1017/CBO9781107415324.004

Marvier, M., Kareiva, P., and Neubert, M. G. (2004). Habitat destruction, fragmentation, and disturbance promote invasion by habitat generalists in a multispecies metapopulation. *Risk Anal.* 24, 869–878. doi: 10.1111/j.0272-4332.2004.00485.x

May, F., Gerstner, K., and McGlinn, D. (2018). MoBiodiv/mobsim: first mobsim release (Version v0.1.1). *bioRxiv[Preprint]*. doi: 10.1101/209502

Miller-Rushing, A. J., Primack, R. B., Devictor, V., Corlett, R. T., Cumming, G. S., Loyola, R., et al. (2019). How does habitat fragmentation affect biodiversity? A controversial question at the core of conservation biology. *Biol. Conserv.* 232, 271–273. doi: 10.1016/j.biocon.2018.12.029

Mori, A. S., Isbell, F., and Seidl, R. (2018). β-Diversity, community assembly, and ecosystem functioning. *Trends Ecol. Evol.* 33, 549-564. doi: 10.1016/j.tree.2018.04.012

Oksanen, J., Blanchet, F. G., Friendly, M., Kindt, R., Legendre, P., McGlinn, D., et al. (2019). *vegan: Community Ecology Package*. Available online at: https://cran.r-project.org/package=vegan.

R Core Team. (2018). *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing.

Ramiadantsoa, T., Hanski, I., and Ovaskainen, O. (2018). Responses of generalist and specialist species to fragmented landscapes. *Theor. Popul. Biol.* 124, 31–40. doi: 10.1016/j.tpb.2018.08.001

Rickert, C., Fichtner, A., van Klink, R., and Bakker, J. P. (2012). α- and β-Diversity in moth communities in salt marshes is driven by grazing management. *Biol. Conserv.* 146, 24–31. doi: 10.1016/j.biocon.2011.11.024

Rybicki, J., Abrego, N., and Ovaskainen, O. (2019). Habitat fragmentation and species diversity in competitive communities. *Ecol. Lett.* 23, 506-517. doi: 10.1111/ele.13450

Sibly, R. M., Grimm, V., Martin, B. T., Johnston, A. S. A., Kulakowska, K., Topping, C. J., et al. (2013). Representing the acquisition and use of energy by individuals in agent-based models of animal populations. *Methods Ecol. Evol.* 4, 151–161. doi: 10.1111/2041-210x.12002

Staley, J. T., Stafford, D. B., Green, E. R., Leather, S. R., Rossiter, J. T., Poppy, G. M., et al. (2011). Plant nutrient supply determines competition between phytophagous insects. *Proc. R. Soc. B Biol. Sci.* 278, 718–724. doi: 10.1098/rspb.2010.1593

Thompson, S. E. D., Chisholm, R. A., and Rosindell, J. (2019). Characterising extinction debt following habitat fragmentation using neutral theory. *Ecol. Lett.* 22, 2087–2096. doi: 10.1111/ele.13398

Townsend, C. R., Begon, M., and Harper, J. L. (2008). *Essentials of Ecology*, 3rd Edn. Hoboken, NJ: Blackwell Publishing.

Tulloch, A. I. T. T., Barnes, M. D., Ringma, J., Fuller, R. A., and Watson, J. E. M. M. (2016). Understanding the importance of small patches of habitat for conservation. *J. Appl. Ecol.* 53, 418–429. doi: 10.1111/1365-2664.12547

Warnes, G. R., Bolker, B., and Lumley, T. (2018). *gtools: Various R Programming Tools*. Available online at: https://cran.r-project.org/package=gtools.

Watling, J., Arroyo-Rodríguez, V., Pfeifer, M., Baeten, L., Banks-Leite, C., Cisneros, L., et al. (2020). Support for the habitat amount hypothesis from a global synthesis of species density studies. *Ecol. Lett.* 23, 674-681.

White, J. W., Rassweiler, A., Samhouri, J. F., Stier, A. C., and White, C. (2014). Ecologists should not use statistical significance tests to interpret simulation model results. *Oikos* 123, 385–388. doi: 10.1111/j.1600-0706.2013.01073.x

Wilensky, U. (1999). *NetLogo. http://ccl.northwestern.edu/netlogo/*. Evanston, IL: Center for Connected Learning and Computer-Based Modeling, Northwestern University.

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Ecology and Evolution

ORIGINAL RESEARCH
published: 14 January 2021
doi: 10.3389/fevo.2020.597542

Specialist Birds Replace Generalists in Grassland Remnants as Land Use Change Intensifies

Ingmar R. Staude[1,2], Gerhard E. Overbeck[3], Carla Suertegaray Fontana[4], Glayson A. Bencke[5], Thaiane Weinert da Silva[4], Anne Mimet[6] and Henrique M. Pereira[1,2]*

[1] German Centre for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig, Leipzig, Germany, [2] Institute of Biology, Martin-Luther University Halle-Wittenberg, Halle, Germany, [3] Department of Botany, Universidade Federal do Rio Grande do Sul, Porto Alegre, Brazil, [4] Laboratório de Ornitologia, Museu de Ciências e Tecnologia, Programa de Pós-Graduação em Ecologia e Evolução da Biodiversidade, Pontifícia Universidade Católica do Rio Grande do Sul, PUCRS, Porto Alegre, Brazil, [5] Museu de Ciências Naturais, Fundação Zoobotânica do Rio Grande do Sul, Porto Alegre, Brazil, [6] Department of Ecology and Ecosystem Management, Technische Universität München, Freising, Germany

The ideal free distribution theory predicts that mobile species distribute themselves among habitat patches so as to optimize their fitness. Changes in land use alter the quality of habitat patches and thereby affect the distribution of species. Following the loss of native habitat, habitat specialists are expected to move to patches where native habitat still remains in order to survive. Competition for resources in habitat remnants should consequently increase. As generalists are able to use other habitats, generalists are expected to gradually disappear in remnants in order to avoid increasing competition with specialists. Here, we test these predictions by studying the response of habitat specialist and generalist birds to land-use change in Brazil's southern grasslands. Using a space-for-time substitution approach, we surveyed bird communities in native grassland sites (~4 ha) in 31 regions (10 × 10 km) with differing levels of conversion to agriculture (1–94%). We found a higher abundance of specialists in native grassland patches with increasing agricultural cover in the region, while the total number of individuals in remnants remained constant. At the same time, the share of generalists in total abundance and total species richness decreased. To gain insights into whether these patterns could be driven by shifts in competition, we tested whether generalists that continued to co-occur with specialists in remnants, had less dietary overlap with specialists. As a consequence of community composition in remnants, a higher proportion of generalists were omnivorous and the average generalist species fed less on seeds, whereas the average specialist species fed more on seeds when agricultural cover was high in the region. Our results, therefore, support predictions of the ideal free distribution theory. Specialists that are assumed to have a low survivorship outside of their specialized habitat, distribute to remnants of this habitat when it is converted elsewhere, while generalists, being able to survive in other habitats, disappear gradually in remnants. Such a process could partly explain the segregation of habitat specialist and generalist birds observed in many agricultural landscapes. Finally, our results suggest that native habitat remnants can be important temporary refugia for specialists.

Keywords: ideal free distribution, trophic niche, habitat loss, biodiversity change, Brazil

OPEN ACCESS

Edited by:
John Maxwell Halley, University of Ioannina, Greece

Reviewed by:
Gabor L. Lovei, Aarhus University, Denmark
Nicole Michel, National Audubon Society, United States

****Correspondence:***
Ingmar R. Staude
ingmar.staude@idiv.de
orcid.org/0000-0003-2306-8780

Specialty section:
This article was submitted to Models in Ecology and Evolution, a section of the journal Frontiers in Ecology and Evolution

Received: *21 August 2020*
Accepted: *07 December 2020*
Published: *14 January 2021*

Citation:
Staude IR, Overbeck GE, Fontana CS, Bencke GA, Silva TWd, Mimet A and Pereira HM (2021) Specialist Birds Replace Generalists in Grassland Remnants as Land Use Change Intensifies. Front. Ecol. Evol. 8:597542. doi: 10.3389/fevo.2020.597542

INTRODUCTION

The population abundance of many bird species has declined at large spatial scales over the last few decades (Rosenberg et al., 2019). Grassland birds, in particular, face high threat levels as their habitats are rapidly being converted for human land uses (Schipper et al., 2016; Stanton et al., 2018; Correll et al., 2019). While abundant empirical evidence demonstrates that generalists dominate bird communities in agricultural habitats (Lockwood et al., 2000; Devictor et al., 2008), it is much less clear how bird communities change in grassland remnants when grassland is increasingly converted to agriculture in the region. Here we study these changes in consideration of the ideal free distribution (IFD) theory (Fretwell, 1970; Tregenza, 1995).

The IFD theory predicts that mobile species, such as birds, should move between patches as habitat quality changes, so as to optimize their fitness (Tregenza, 1995). As native grassland is converted to cultivated fields, opportunities for grassland specialists to survive and reproduce are decreasing. The IFD theory therefore predicts that specialists should move to patches in which native grassland remains (Tregenza, 1995). As a result of the arrival of new specialist individuals, competition for resources in these remnants is expected to increase. Grassland birds are assumedly better at exploiting the resources they specialized on than generalists that are able to use a broad range of habitats and resources (Poisot et al., 2011; Ponge, 2013; Reif et al., 2016). The IFD theory therefore predicts that, if resources are limited, habitat generalists should gradually disappear in remnants, as they are able to use other habitats that are more relaxed from competition with specialists (MacArthur and Pianka, 1966; Tregenza, 1995).

Of all generalist species, those generalists with resource use patterns most similar to specialists are expected to be affected most by increasing specialists (MacArthur and Pianka, 1966; MacArthur and Levins, 1967). This non-randomness should lead to an observable decrease in the share of generalists in total species richness in remnants as habitat conversion increases in the region. Moreover, generalists that still co-occur with specialists in remnants should show a reduced dietary overlap with specialists. Overall, these predictions of the IFD theory would be consistent with the pattern of spatial segregation of specialists and generalists that has been observed in mosaic landscapes of disturbed and stable habitats, where specialists have been found to dominate communities in more stable habitats and generalists to aggregate in more disturbed habitats (Julliard et al., 2006; Devictor et al., 2008).

Here we test whether such predictions based on the IFD theory are supported by empirical evidence from bird communities in Brazil's southern grasslands, which face a very recent and intensive conversion to agriculture (Overbeck et al., 2015; de Oliveira et al., 2017). We used a space-for-time substitution approach and surveyed local bird communities in grassland remnants of 31 regions with differing levels of agricultural cover. First, we test whether total species abundance and the abundance of specialists (i.e., grassland birds) in native grassland remnants changes with increasing agricultural cover in the region. Second, we test for specialist-generalist segregation by assessing whether the share of generalists (i.e., non-grassland birds) in total species richness is lower in remnants when regional agricultural cover is high. Third, we test whether generalist species that continue to co-occur with specialists in remnants, have a reduced dietary overlap with specialists.

METHODS

Study Region

We studied bird communities within the South Brazilian grassland region in Rio Grande do Sul (RS), Brazil's southernmost state. Climate in RS is humid subtropical with warm summers and no pronounced dry seasons (Alvares et al., 2013). In the last four decades, ∼50% of natural grasslands in the region have been converted to agriculture, i.e., mainly to soybeans and rice (Cordeiro et al., 2009; de Oliveira et al., 2017). This rapid conversion rate and lack of sufficient protection give these grasslands the highest Conservation Risk Index of all Brazilian biomes (Overbeck et al., 2015).

Study Sites and Bird Sampling

We identified suitable sampling regions by assessing the spatial distribution of land use/cover types in the entire territory of RS. For this, we georeferenced and visually interpreted Landsat 5 satellite images (from 2009) (Hasenack and Weber, 2010). We adopted a 10 × 10 km grid used by the Brazilian Ministry of the Environment for national forest inventories, and calculated the percentage of different land uses in each grid cell (henceforth referred to as region). We selected 31 regions (**Supplementary Figure 1**; **Supplementary Table 1**) to represent a gradient of regional agricultural cover (1–94% agricultural cover) and the distribution of native grasslands in RS.

In each region, we conducted bird surveys at three local sites within native grassland remnants. Local site selection followed judgement by botanists (presence of native grassland) and operational criteria (accessibility and permission). Each of the three local sites per region was surveyed once with one point count covering an area of ∼ 4 ha (fixed radius of 112 m) for 15 min. This time of detection is sufficient, as it is common for grassland habitats to be sampled with only 5-min point counts (Ralph et al., 1995). We acknowledge, however, that by surveying only three points within remnants per region, we may have missed species and/or undercounted individuals. Mean minimum and maximum distance between the three local sites was 538 and 1,125 m, respectively, where the recommended minimum distance between point counts is 200 m to reduce the probability of the same bird being counted at different points (Sutherland et al., 2004).

All point counts were carried out by the same team of six experienced observers during the Austral spring to summer (i.e., during the breeding season of birds in southern Brazil; October–February). In order to standardize time and reduce temporal variation, all point counts were conducted under similar weather conditions either before 10:00 a.m. or after 16:00 p.m. when birds are most conspicuous (but note that during the breeding season grassland birds can be easily detected over the whole day). Observers recorded the presence of all birds seen and heard (see **Supplementary Material** for more details) and their abundance

(i.e., the maximum number of individuals per species) at each local site. We did not correct for imperfect detection, as we were not able to assign distance bands to all individuals. Our counts therefore present relative and not absolute abundance measures. Sampling of the 31 regions took 3 years starting in 2011 until 2014.

Classification of Birds

Generalist vs. Specialist Species

Grassland bird species were characterized *sensu* Azpiroz et al. (2012). That is, species that are restricted to or make extensive use of grassland habitats—thus thought to be particularly sensitive to native grassland conversion—were classified as grassland specialists (51 of 106 species). The remaining birds were considered as non-grassland specialists and henceforth referred to as generalists (**Supplementary Table 2**, for all species and classification).

Trophic Niches

Foraging attributes were obtained for 96 out of the 106 recorded species (91% of species) from the Elton traits database (Wilman et al., 2014). We focused on two foraging attributes to test for reduced trophic niche overlap between generalists and specialists: seed use and omnivory (**Supplementary Table 2**). We *a priori* expected grassland specialists to be superior competitors for seeds, one of the main resources of grasslands. As a response variable we therefore used the percentage of seeds in a species' diet (column "Diet-Seed" in Elton trait database; esapubs.org/archive/ecol/E095/178/metadata.php) averaged across all generalist species (and separately across all specialist species) per site. We expected a decreasing average seed consumption by generalists and an increasing average seed consumption by specialists in remnants, as agricultural cover increases in the region. Because omnivores with broad trophic niches can opportunistically forage on any available food item as *per capita* resource availability declines, we expected an increasing share of co-occurring generalist species to be omnivorous, as agricultural cover increases in the region. In the Elton trait database, species are defined as omnivores if they have a score of <=50% in all four diet categories (i.e., Plant and Seeds; Fruits and Nectar, Invertebrates; Vertebrates, Fish and Carrion; see column "Diet-5Cat" on esapubs.org/archive/ecol/E095/178/metadata.php). We calculated the proportion of generalist species (and separately specialist species) that are omnivores in each community.

Statistical Analysis

The focal explanatory environmental variable in all following models is regional agricultural cover (%), henceforth referred to as percent agriculture. Response variables are the average of the respective community parameter across the three local sites per region in all subsequent analyses. We model the average because the allocation of the three local sites per region was not standardized across regions due to the aforementioned operational criteria for site selection (i.e., limited accessibility and owner permission requirements prohibited entry to space sites equally across regions).

Change in Total and Specialist Abundance

We used Poisson regression models to predict total abundance and specialist abundance with percent agriculture in the region, including an observation-level random effect to account for over-/underdispersion (Harrison, 2014). Models were fitted using the function "glmer" from the LME4 package (Bates et al., 2014). Two species, *Vanellus chilensis* and *Myiopsitta monachus,* were observed in large flocks of 200 and 94 individuals, respectively. This led to outliers in total and specialist abundance (**Supplementary Figures 2A,B**). We thus leveled those counts to the mean number of individuals across all other sites these species occupied, rounded to the nearest integer (cf. Julliard et al., 2006). Results were robust when the unlevelled data were analyzed (**Supplementary Figure 2C**).

Spatial Segregation of Specialists and Generalists

To test whether changes in land use induce a spatial segregation of specialists from generalists, we tested (1) whether percent agriculture influences total richness, and (2) whether percent agriculture predicts the share of generalists in total species richness. Proportions of generalists were calculated for each local site individually and then averaged for each region. Averaged proportions were regressed on percent agriculture with a beta regression model using the BETAREG package (Zeileis et al., 2010). The relationship between total species richness and percent agriculture was assessed using the same model as for abundance.

Trophic Niche Overlap Between Generalists and Specialists

First, we modeled the percent use of seeds averaged across all generalist species (and separately across all specialist species) with percent agriculture as predictor. Second, we regressed the proportion of generalist and specialist species with an omnivorous diet on percent agriculture. Again, we calculated these response variables as the average of the three local sites. These values were regressed again on percent agriculture using a beta regression model. To accommodate 0/1 values, percent use of seeds was transformed to the open unit interval (0, 1) using a continuity correction $(y^*(n-1)+0.5)/n$, where n is the sample size (Smithson and Verkuilen, 2006; Zeileis et al., 2010).

Influence of Native Grassland in the Immediate Surroundings

We tested whether the amount of native grassland in the immediate surroundings of the three local sites has an effect on community parameters. Within a 1 km radius of the three local sites' centroid, native grassland area was on average 2,500 ha (ranging from 540–3,100 ha). We added native grassland area within this radius to all the above models as an explanatory variable (Pearson correlation of native grassland area and agricultural cover in the region was $\rho = -0.35$). We used the asymptotic likelihood ratio test, as implemented by the function "LRTEST" from the LTEST package (Hothorn et al., 2019) to test whether adding this variable improved the statistical fit of the model.

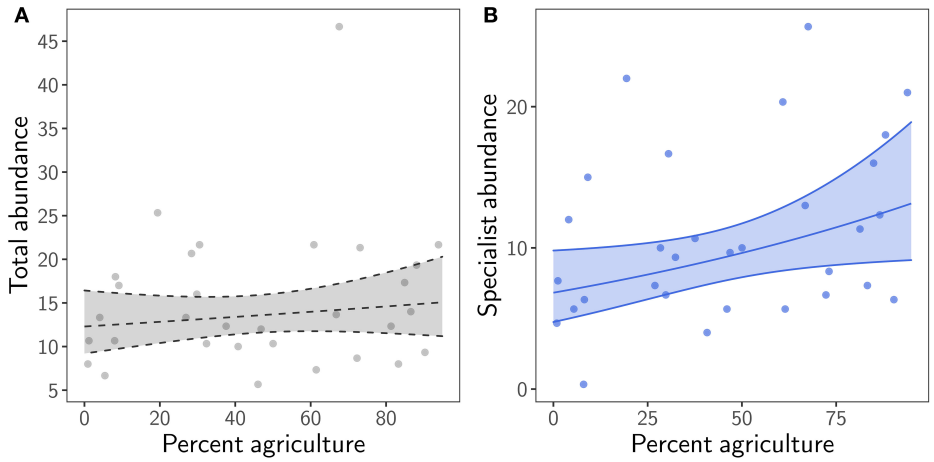


FIGURE 1 | Relationship between percent agriculture in the region and **(A)** total abundance (i.e., the combined number of individuals of specialists and generalists), and **(B)** specialist abundance in grassland remnants. Transparent ribbon represents the 95% confidence interval around the mean. Ribbon and regression line are dashed when $p \geq 0.05$.

Model Validation and Spatial Autocorrelation

Generalized linear mixed effect models were validated with DHARMA scaled residual plots (Hartig, 2017) and beta regression models with BETAREG diagnostic plots (Zeileis et al., 2010). To test for spatial autocorrelation we performed the Moran's I test on all models.

RESULTS

Change in Total and Specialist Abundance

We found that total species abundance in grassland remnants did not change clearly with agricultural cover in the region ($\chi^2 = 0.65$, $p = 0.42$). However, analyzing grassland specialists separately from generalists showed a statistically clear increase in specialist abundance with increasing regional land use ($\chi^2 = 4.4$, $p = 0.036$; **Figure 1**; **Supplementary Table 3**). Correspondingly, the share of generalists in total abundance declined with increasing agriculture in the region ($\chi^2 = 4.58$, $p = 0.032$; **Supplementary Figure 3**).

Spatial Segregation of Specialists and Generalists

While total species richness in remnants remained constant ($\chi^2 = 0.9$, $p = 0.342$), the contribution of generalist species to total species richness decreased by more than 50% with increasing agricultural cover in the region ($\chi^2 = 5$, $p = 0.025$; **Figure 2**; **Supplementary Table 3**).

Trophic Niche Overlap Between Generalists and Specialists

As a consequence of species composition in remnants, the percent use of seeds averaged across generalist species strongly decreased from above 30 to <10% with more agriculture in the region ($\chi^2 = 7.63$, $p = 0.005$) (**Figure 3A; Supplementary Table 4**). In contrast, we found a marginally significant increase in the average percent use of seeds by specialists with more agriculture in the region ($\chi^2 = 3.67$, $p = 0.055$) (**Figure 3B**). The proportion of generalist species with an omnivorous diet doubled ($\chi^2 = 4.17$, $p = 0.041$), while the proportion of omnivorous specialist species remained constant with increasing regional agriculture ($\chi^2 = 0$, $p = 0.86$) (**Figures 3C,D**; **Supplementary Table 4**).

None of the models improved significantly when native grassland amount in the surrounding landscape was added as a second explanatory variable (**Supplementary Table 5**). Residuals did not show spatial autocorrelation in any model (**Supplementary Figure 2**).

DISCUSSION

Using data from bird communities in native grassland remnants, we studied the response of grassland specialist and generalist birds to the increasing conversion of native grasslands to agriculture. While total species abundance and richness was constant, specialist abundance increased and the contribution of generalist species to total species abundance and richness decreased in grasslands remnants with increasing agriculture in the region. Furthermore, generalist species that continued to co-occur with specialists had a smaller dietary overlap with specialists. Together, these results support the predictions of the ideal free distribution theory: specialists that are assumed to have a low survivorship outside of their preferred habitat move away upon habitat conversion to patches where this habitat remains, while habitat generalists that are able to use other habitats, gradually disappear from these remnants.

While our results are in support of the IFD theory, our study has several limitations. We did not directly study temporal changes and instead base our inference on space-for-time substitution. Thus, we do not unequivocally show that bird communities in grassland remnants are indeed changing in

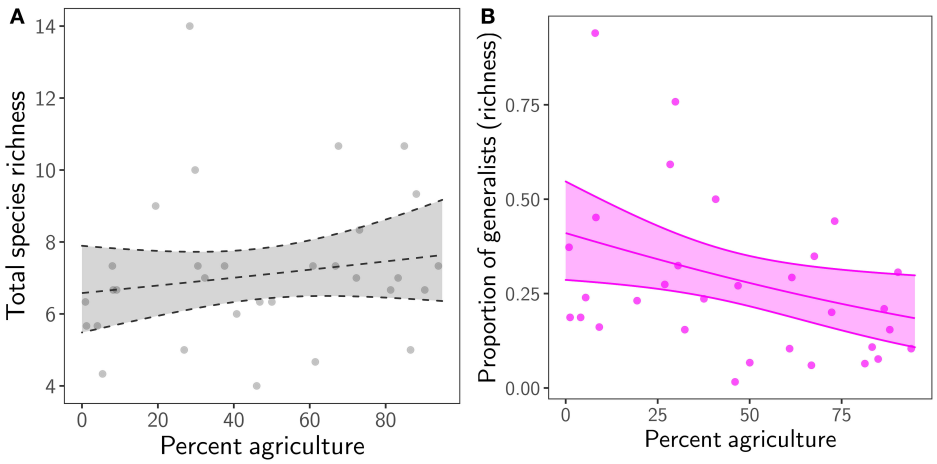


FIGURE 2 | Relationship between percent agriculture in the region and **(A)** total species richness (i.e., specialists and generalists), and **(B)** the proportion of generalist species (to total species richness) in grassland remnants. Transparent ribbon represents the 95% confidence interval around the mean. Ribbon and regression line are dashed when $p \geq 0.05$.

response to regional land-use intensification. Moreover, we were not able to measure what individual birds were eating and whether birds changed their individual diets. We could therefore not unequivocally demonstrate that competition is driving the patterns found in this study (Dhondt, 2011). In addition, biological communities respond slower than the rate at which humans change the environment (Damgaard, 2019). Especially in our study area, land use change is very recent and rapid (Cordeiro et al., 2009), such that the patterns we found here may be transient and not stable in the long-term.

Within these limitations, we found that local species richness and total abundance of birds did not decrease in grassland remnants with increasing regional land-use change. It is likely that communities in grassland remnants show simply a lagged response (Tilman et al., 1994; Kuussaari et al., 2009). That is, species loss in intensive agriculture (Hendershot et al., 2020) and widespread negative population trends of birds (Rosenberg et al., 2019) are likely to translate to declines in richness and abundance in remnants too. However, our results suggest that at least temporarily, native habitat remnants could play an important role in sustaining the populations of some species. We found that the abundance of grassland specialists is increasing in remnants when natural habitat is increasingly converted in the region. Thus, the preservation of native habitat remnants could be an important conservation tool for sustaining populations of species that are facing loss of their specialized habitat.

The stability of total species abundance in remnants could furthermore indicate a carrying capacity and that resources are limited in remnants. The arrival of new specialist individuals is therefore likely to move the remnant beyond carrying capacity and increase the competition for resources. Since specialists are assumedly superior competitors for the resources of their specialized habitat (Reif et al., 2016), their arrival may act to replace generalists. Indeed, we found that the relative abundance and species richness of generalists has declined in remnants. Moreover, we found that within the generalist sub-assemblage, species with similar resource use pattern to specialists gradually disappeared. We therefore hypothesize that competition could indeed play a role in explaining such a replacement pattern.

The replacement of generalists by specialists, which we infer here, contrasts with biotic homogenization, the process by which rare, specialist species are widely replaced by common, generalist species owing to human activities (Lockwood et al., 2000; Mimet et al., 2019). However, given the lack of temporal data in this study, future research on mobile organisms is needed to determine whether this is a more general pattern found in matrices of agriculture and habitat remnants. For example, land-use history can influence processes underlying community change (Isbell et al., 2019; Mimet et al., 2019). In our study region, habitat remnants are still close to a pristine state (i.e., no changes in land-use occurred in the recent past) and thus more likely to support specialists. These conditions may be different in agricultural landscapes of, for example, Europe where land use change has a much longer history (Goldewijk, 2001).

Nonetheless, also in European agricultural landscapes where bird communities should have stabilized in response to land-use history, generalists seem to spatially segregate from specialists in mosaics of perturbed and more stable sites (Julliard et al., 2006; Devictor et al., 2008). While generalists dominate communities in human-modified, disturbed sites, specialists dominate in stable, more natural sites (Devictor et al., 2008). Although the mechanism behind this spatial segregation is still largely unclear, our results suggest this segregation is consistent with predictions of the IFD theory. Both specialists and generalists seek to escape declines in resources; specialists likely in response to native habitat conversion and generalists likely in response to consequent increases in specialists in habitat remnants.

We draw two conclusions from our study. First, our results support predictions of the IFD theory. Specialized birds that are assumed to have little prospect of surviving outside of their specialized habitat, increased in remnants of this habitat when it was converted elsewhere. In accordance with our

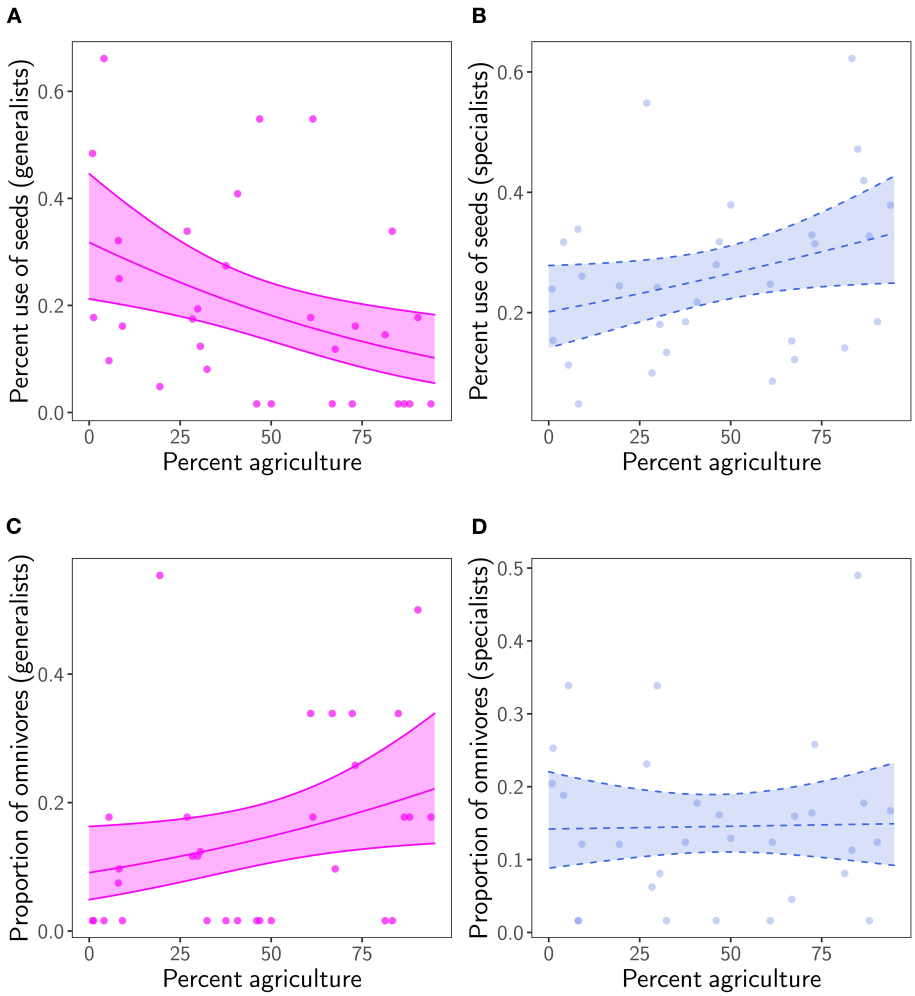


FIGURE 3 | Relationship between percent agriculture in the region and percent use of seeds averaged **(A)** across generalist species and **(B)** specialist species, and with an omnivorous diet in **(C)** generalists and **(D)** specialists. Transparent ribbon represents the 95% confidence interval around the mean. Ribbon and regression line are dashed when $p \geq 0.05$.

hypothesis that this may increase competition in remnants, generalists with similar resource use patterns to specialists gradually disappeared in remnants. The IFD theory could therefore help explain aspects of bird community change, such as specialist-generalist segregation, in response to land-use change. Second, our results indicate that remnants of native grassland embedded in regions devoted primarily to human activities may buy time for the conservation of grassland birds (Silva et al., 2015). Considering that ongoing agricultural intensification may eventually significantly reduce local bird diversity (Hendershot et al., 2020) preserving remnants of native habitat merits more consideration in strategies to conserve biodiversity.

DATA AVAILABILITY STATEMENT

Rmarkdown file with R code for all statistical analyses and data visualization, as well as the data to reproduce the findings of this study are available on figshare at: https://doi.org/10.6084/m9.figshare.13318211.v1.

AUTHOR CONTRIBUTIONS

IS and HP conceived the study. CF, GB, and TS organized and carried out fieldwork. IS analyzed data with input from AM and wrote the manuscript with substantial input from all authors. All authors contributed to the article and approved the submitted version.

FUNDING

We acknowledge funding by CNPq/FAPERGS (grants 563271/2010-8 and 11/2185-0, respectively) within the SISBIOTA program. IS, AM, and HP acknowledge funding of iDiv *via* the German Research Foundation (DFG FZT 118). GO and CF acknowledge CNPq grants (310345/2018-9 and 457475/2012-9, 309438/2016-0, respectively).

ACKNOWLEDGMENTS

We thank Valério D. Pillar, UFRGS, for project coordination and Heinrich Hasenack and Eduardo Vélez-Martin, UFRGS, for remote sensing data and site selection. We wish to thank Diogenes Machado, Maurício Silveira, Gabriel Larré, Maurício Bettio, Felipe Zilio, Eduardo Chiarani, Daniele Cantelli, and Tiago Steffen for help in the field work and/or data digitalization, and all land users for the permission to work on their land.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fevo.2020.597542/full#supplementary-material

REFERENCES

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

Azpiroz, A. B., Isacch, J. P., Dias, R. A., Di Giacomo, A. S., Fontana, C. S., and Palarea, C. M. (2012). Ecology and conservation of grassland birds in southeastern South America: a review. *J. F. Ornithol.* 83, 217–246. doi: 10.1111/j.1557-9263.2012.00372.x

Bates, D., Mächler, M., Bolker, B., and Walker, S. (2014). Fitting linear mixed-effects models using lme4. *arXiv* Preprint arXiv1406.5823. doi: 10.18637/jss.v067.i01

Cordeiro, J. L. P., Hasenack, H., et al. (2009). *Cobertura Vegetal Atual do Rio Grande do Sul. Campos Sulinos: Conservação e uso Sustentável da Biodiversidade*. Brasilia: Ministério do Meio Ambiente, 285–299.

Correll, M. D., Strasser, E. H., Green, A. W., and Panjabi, A. O. (2019). Quantifying specialist avifaunal decline in grassland birds of the northern great plains. *Ecosphere* 10:e02523. doi: 10.1002/ecs2.2523

Damgaard, C. (2019). A critique of the space-for-time substitution practice in community ecology. *Trends Ecol. Evol.* 34, 416–421. doi: 10.1016/j.tree.2019.01.013

de Oliveira, T. E., de Freitas, D. S., Gianezini, M., Ruviaro, C. F., Zago, D., Mércio, T. Z., et al. (2017). Agricultural land use change in the Brazilian pampa biome: the reduction of natural grasslands. *Land Use Policy* 63, 394–400. doi: 10.1016/j.landusepol.2017.02.010

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

Dhondt, A. A. (2011). *Interspecific Competition in Birds*. Oxford; New York, NY: Oxford University Press.

Fretwell, S. D. (1970). On territorial behavior and other factors influencing habitat distribution in birds. *Acta Biotheor.* 19, 45–52. doi: 10.1007/BF01601955

Goldewijk, K. K. (2001). Estimating global land use change over the past 300 years: the HYDE database. *Glob. Biogeochem. Cycles* 15, 417–433. doi: 10.1029/1999GB001232

Harrison, X. A. (2014). Using observation-level random effects to model overdispersion in count data in ecology and evolution. *PeerJ* 2:e616. doi: 10.7717/peerj.616

Hartig, F. (2017). *DHARMa: Residual Diagnostics for Hierarchical (Multi-Level/Mixed) Regression Models*. R Package version 0.15.

Hasenack, H., and Weber, E. (2010). *Base cartográfica vetorial contínua do Rio Grande do Sul-escala 1:50.000*. Porto Alegre: UFRGS Centro de Ecologia 1.

Hendershot, J. N., Smith, J. R., Anderson, C. B., Letten, A. D., Frishkoff, L. O., Zook, J. R., et al. (2020). Intensive farming drives long-term shifts in avian community composition. *Nature* 579, 393–396. doi: 10.1038/s41586-020-2090-6

Hothorn, T., Zeileis, A., Farebrother, R. W., Cummins, C., Millo, G., Mitchell, D., et al. (2019). *Package 'lmtest'*.

Isbell, F., Tilman, D., Reich, P. B., and Clark, A. T. (2019). Deficits of biodiversity and productivity linger a century after agricultural abandonment. *Nat. Ecol. Evol.* 3, 1533–1538. doi: 10.1038/s41559-019-1012-1

Julliard, R., Clavel, J., Devictor, V., Jiguet, F., and Couvet, D. (2006). Spatial segregation of specialists and generalists in bird communities. *Ecol. Lett.* 9, 1237–1244. doi: 10.1111/j.1461-0248.2006.00977.x

Kuussaari, M., Bommarco, R., Heikkinen, R. K., Helm, A., Krauss, J., Lindborg, R., et al. (2009). Extinction debt: a challenge for biodiversity conservation. *Trends Ecol. Evol.* 24, 564–571. doi: 10.1016/j.tree.2009.04.011

Lockwood, J. L., Brooks, T. M., and Mckinney, M. L. (2000). Taxonomic homogenization of the global avifauna. *Anim. Conserv. Forum.* 1, 27–35. doi: 10.1111/j.1469-1795.2000.tb00084.x

MacArthur, R., and Levins, R. (1967). The limiting similarity, convergence, and divergence of coexisting species. *Am. Nat.* 101, 377–385. doi: 10.1086/282505

MacArthur, R. H., and Pianka, E. R. (1966). On optimal use of a patchy environment. *Am. Nat.* 100, 603–609. doi: 10.1086/282454

Mimet, A., Buitenwerf, R., Sandel, B., Svenning, J.-C., and Normand, S. (2019). Recent global changes have decoupled species richness from specialization patterns in North American birds. *Glob. Ecol. Biogeogr.* 28, 1621–1635. doi: 10.1111/geb.12978

Overbeck, G. E., Vélez-Martin, E., Scarano, F. R., Lewinsohn, T. M., Fonseca, C. R., Meyer, S. T., et al. (2015). Conservation in Brazil needs to include non-forest ecosystems. *Divers. Distrib.* 21, 1455–1460. doi: 10.1111/ddi.12380

Poisot, T., Bever, J. D., Nemri, A., Thrall, P. H., and Hochberg, M. E. (2011). A conceptual framework for the evolution of ecological specialisation. *Ecol. Lett.* 14, 841–851. doi: 10.1111/j.1461-0248.2011.01645.x

Ponge, J.-F. (2013). Disturbances, organisms and ecosystems: a global change perspective. *Ecol. Evol.* 3, 1113–1124. doi: 10.1002/ece3.505

Ralph, J. C., Sauer, J. R., and Droege, S. (1995). *Monitoring Bird Populations by Point Counts*. Gen. Tech. Rep. PSW-GTR-149. Albany, CA: US Department of Agriculture; Forest Service; Pacific Southwest Research Station, 149.

Reif, J., Hovrák, D., Krištin, A., Kopsová, L., and Devictor, V. (2016). Linking habitat specialization with species' traits in European birds. *Oikos* 125, 405–413. doi: 10.1111/oik.02276

Rosenberg, K. V., Dokter, A. M., Blancher, P. J., Sauer, J. R., Smith, A. C., Smith, P. A., et al. (2019). Decline of the north american avifauna. *Science* 366, 120–124. doi: 10.1126/science.aaw1313

Schipper, A. M., Belmaker, J., de Miranda, M. D., Navarro, L. M., Böhning-Gaese, K., Costello, M. J., et al. (2016). Contrasting changes in the abundance and diversity of North American bird assemblages from 1971 to 2010. *Glob. Change Biol.* 22, 3948–3959. doi: 10.1111/gcb.13292

Silva, T. W., Dotta, G., and Fontana, C. S. (2015). Structure of avian assemblages in grasslands associated with cattle ranching and soybean agriculture in the Uruguayan savanna ecoregion of Brazil and Uruguay. *Condor Ornithol. Appl.* 117, 53–63. doi: 10.1650/CONDOR-14-85.1

Smithson, M., and Verkuilen, J. (2006). A better lemon squeezer? Maximum-likelihood regression with beta-distributed dependent variables. *Psychol. Methods* 11:54. doi: 10.1037/1082-989X.11.1.54

Stanton, R. L., Morrissey, C. A., and Clark, R. G. (2018). Analysis of trends and agricultural drivers of farmland bird declines in North America: a review. *Agric. Ecosyst. Environ.* 254, 244–254. doi: 10.1016/j.agee.2017.11.028

Sutherland, W. J., Newton, I., and Green, R. (2004). *Bird ecology and Conservation: A Handbook of Techniques*. Oxford: Oxford University Press.

Tilman, D., May, R. M., Lehman, C. L., and Nowak, M. A. (1994). Habitat destruction and the extinction debt. *Nature* 371, 65–66. doi: 10.1038/371065a0

Tregenza, T. (1995). "Building on the ideal free distribution," in *Advances in Ecological Research* (Elsevier), 253–307.

Wilman, H., Belmaker, J., Simpson, J., de la Rosa, C., Rivadeneira, M. M., and Jetz, W. (2014). EltonTraits 1.0: species-level foraging attributes of the world's birds and mammals: ecological archives E095-178. *Ecology* 95:2027. doi: 10.1890/13-1917.1

Zeileis, A., Cribari-Neto, F., Grün, B., and Kos-midis, I. (2010). Beta regression in R. *J. Stat. Softw.* 34, 1–24. doi: 10.18637/jss.v034.i02

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Ecology and Evolution

ORIGINAL RESEARCH
published: 18 February 2021
doi: 10.3389/fevo.2021.531833

Stability Analysis of Delayed Age-Structured Resource-Consumer Model of Population Dynamics With Saturated Intake Rate

*Vitalii V. Akimenko**

Faculty of Computer Science and Cybernetics, Taras Shevchenko National University of Kyiv, Kyiv, Ukraine

This article studies nonlinear n-resource-consumer autonomous system with age-structured consumer population. The model of consumer population dynamics is described by a delayed transport equation, and the dynamics of resource patches are described by ODE with saturated intake rate. The delay models the digestion period of generalist consumer and is included in the calorie intake rate, which impacts the consumer's fertility and mortality. Saturated intake rate models the inhibition effect from the behavioral change of the resource patches when they react to the consumer population growing or from the crowding effect of the consumer. The conditions for the existence of trivial, semi-trivial, and non-trivial equilibria and their local asymptotic stability were obtained. The local asymptotic stability/instability of non-trivial equilibrium of a system with depleted patches is defined by new derived criteria, which relate the demographic characteristics of consumers with their search rate, growth rate of resource in patches, and behavioral change of the food resource when consumer population grows. The digestion period of a generalist consumer does not cause local asymptotical instabilities of consumer population at the semi-trivial and nontrivial equilibria. These theoretical results may be used in the study of metapopulation dynamics, desert locust populations dynamics, prey-predator interactions in fisheries, etc. The paper uses numerical experiments to confirm and illustrate all dynamical regimes of the n-resource-consumer population.

Keywords: age-structured model, saturated intake rate, stability analysis, digestion period, resource-consumer model

OPEN ACCESS

Edited by:
John Maxwell Halley,
University of Ioannina, Greece

Reviewed by:
Leonardo Montagnani,
Free University of Bozen-Bolzano, Italy
Luís Borda-de-Água,
Universidade do Porto, Portugal

***Correspondence:**
Vitalii V. Akimenko
vitaliiakm@gmail.com

Specialty section:
This article was submitted to Models in Ecology and Evolution, a section of the journal Frontiers in Ecology and Evolution

Received: *01 February 2020*
Accepted: *18 January 2021*
Published: *18 February 2021*

Citation:
Akimenko VV (2021) Stability Analysis of Delayed Age-Structured Resource-Consumer Model of Population Dynamics With Saturated Intake Rate. Front. Ecol. Evol. 9:531833. doi: 10.3389/fevo.2021.531833

INTRODUCTION

Competition between several food patches and common consumers has been thoroughly studied in the ecological literature (Holt, 1977, 1984; Holt and Kotler, 1987; Martinez, 1991; Holt and Lawton, 1993; Holt et al., 1994; Wootton, 1997; Abrams et al., 1998; Křivan, 2003, 2014; Williams and Martinez, 2004; Křivan and Eisner, 2006; Vrkoc and Krivan, 2015; Becker and Hall, 2016). Here, increasing biomass of one food patch causes increases in generalist consumer population, thus a

negative impact on other resource patches and vice versa. Thus, apparent competition is similar to exploitative competition (Levin, 1970) and can reduce the number of coexisting resource patches. In traditional unstructured Lotka-Volterra ODE models of population dynamics, many details of life history are neglected (de Ross and Persson, 2013). The more reasonable approach in population dynamics modeling is based on physiologically structured models (Von Foerster, 1959; Gurtin and MacCamy, 1979; Cushing and Saleem, 1982; Elderkin, 1985; Webb, 1985, 2008; Metz and Diekmann, 1986; Cushing, 1998; Bekkal-Brikci et al., 2007; Hritonenko and Yatsenko, 2007; de Ross and Persson, 2013; Mohr et al., 2014; Akimenko and Křivan, 2018). In this article, the apparent competition model of unstructured resource patches with age-structured consumer population is studied. This approach allows us to relate foraging to the life history and demographical characteristics of consumer population (fertility and mortality).

The dynamic interaction between resources and consumers in prey-predator models is described by the consumer's functional response. Beddington (1975) and DeAngelis et al. (1975) introduced and analyzed the functional response with saturation, which is often used now in applied models providing the more realistic description of prey-predator interaction (Capasso and Serio, 1978; Qiu et al., 2004; Wang and Zhao, 2004; Han et al., 2017). The functional response of such "saturated incidence rate" was first introduced into SIR (susceptible-infected-recovered) epidemic models in a study of the cholera epidemic spread in Bari (Capasso and Serio, 1978) and was used later in various epidemic models (Wang and Zhao, 2004; Han et al., 2017) and ecological studies (Essington and Hansson, 2004). The feature of saturated incidence rate is that it tends to saturation when the population of predators (or infectives in epidemic models, parasites in parasite-host model, consumers in resource-consumer models, etc.) gets large and, as a consequence, it prevents the unboundedness of the contact rate between prey and predator. Since this functional response considers the behavioral change of prey (or hosts, susceptibles, resources in patches, etc.) as a reaction on the predator population growing or "crowding effect" of predator, the resource-consumer models with intake rates of such form are more reasonable in comparison with traditional Lotka-Volterra models. The resource consumption in biological and ecological models is often characterized also by the calorie intake rate, which depends linearly from the amount of food resource taken by one consumer per unit of time from all patches. This function depends on the handling time, i.e., the time a consumer needs to handle and digest a unit of resource. This time period is included in a model as a time delay parameter. Thus, the resulting model studied in this article consists of several unstructured resource patches and a single age-structured consumer population that forages in these patches including the saturated intake rate, calorie intake rate, and the digestion period of a generalist consumer as a time delay parameter. The model is formulated in Section "Model".

The conditions of existence of the trivial, semi-trivial, and non-trivial equilibria of autonomous systems are studied in Section "Existence of Stationary Equilibria of the Autonomous System (1)–(5)." The local asymptotic stability of all equilibria is considered in Section "Local Asymptotic Stability of Equilibria of the Autonomous System (1)–(5)." Stability analysis is based on the traditional perturbation theory and linearization of autonomous system and includes the study of impact of the time delay parameter on the asymptotic stability of equilibria (Gourley and Kuang, 2004; Shi, 2013; Mohr et al., 2014; Akimenko, 2017a; Akimenko and Křivan, 2018; Martsenyuk et al., 2018; Liu et al., 2019).

Research shows that the stability indicator of non-linear autonomous age-structured models, partial derivative of basic reproduction number of consumer population by their density used earlier in Cushing (1998); Akimenko and Křivan (2018), can be applied only for non-trivial equilibria of system with non-depleted patches. Local asymptotic stability/instability of non-trivial equilibrium of system with depleted patches is defined in the paper by new derived criteria, which relate the demographic characteristics of consumers with their search rate, growth rate of resource in patches, and behavioral change of the food resource when consumer population grows. We show also that the digestion period of a generalist consumer does not cause local asymptotical instabilities of consumer population at the semi-trivial and nontrivial equilibria. These theoretical results may be used in study of metapopulation dynamics (Nakazawa, 2015; Becker and Hall, 2016), desert locust populations dynamics (Guttal et al., 2012; Akimenko and Piou, 2018), prey-predator interactions in fisheries (Essington and Hansson, 2004; Smith and Smith, 2020), and many others.

The numerical algorithms obtained in earlier works (Akimenko, 2017b,c,d) are used in the Section "Numerical Experiments" for numerical analysis of dynamical regimes of autonomous system that were considered in the previous sections. In the first and second groups of experiments the local asymptotic stability of the trivial and semi-trivial (i.e., resources can only exist at positive densities) equilibria for three resource patches with one generalist consumer is studied. Depending on the reproduction number of consumers, trajectories of system are unstable, oscillate in the vicinities of the trivial and semi-trivial equilibria, or converge asymptotically to the semi-trivial equilibrium. The further increasing of consumer's basic reproduction number or time delay parameter leads to the consumer population outbreaks in the form of pulse sequence, which are classified in the quantitative population ecology as the populations with cyclical eruption dynamics (Abbott and Dwyer, 2007; Akimenko and Anguelov, 2017). The results of simulations illustrated the properties of the outbreak solutions are presented in Section "The Trivial and Semi-Trivial Equilibria."

The next group of experiments focuses on the study of asymptotic behavior of solutions in the vicinity of the non-trivial equilibrium with one non-depleted and two depleted resource patches (1st–3rd experiments), three non-depleted resource patches (4th experiment), and one non-trivial consumer population. The results of simulations confirm and illustrate the statements of theorems and exhibit the different dynamical regimes of system with unstable and asymptotically stable trajectories for the selected parameters of the model.

Several concluding remarks are given in Section "Conclusion and Discussion."

MODEL

In this article, we study an apparent competition food web module that consists of n resource patches and consumers that move freely between these patches. Resource density of i-th patch is denoted as $y_i(t)$, $i = 1, ..., n$. The resource dynamics in each patch is described by the logistic model with constant growth rate $r_i > 0$, and environmental carrying capacity $K_i > 0$. The age-specific density of consumer population at age a and time t is denoted by $w(a, t)$, the quantity of consumers in population is $W_0(t) = \int_0^{a_d} w(a, t)da$ (where $a_d > 0$ is the maximum consumer's life-span), and a weighted quantity of consumers at the fixed time t is $\hat{W}(t) = \int_0^{a_d} \gamma(a)w(a, t)da$ (where $\gamma(a)$ is an age-specific consumer's preferences in food resource). The interaction strength between resources and consumers is a product of the saturated intake rate $y_i(t)g_i(\hat{W}(t))$, where $g_i(\hat{W}(t))$ evolves to a saturation level when $\hat{W}(t)$ gets large, i.e., $g_i(\hat{W}(t)) = \beta_i\hat{W}(t)\left(1 + \alpha_i\hat{W}(t)\right)^{-1}$. Functions $g_i(\hat{W}(t))$ have a form of the Beddington–DeAngelis type of functional responses (Beddington, 1975; DeAngelis et al., 1975; Qiu et al., 2004) under assumption that handling time of predator is effectively zero [Eq. (12) in Beddington, 1975]. Constant $\beta_i > 0$ is a search rate of resource $i = 1, ..., n$. Saturation coefficient $\alpha_i \geq 0$ is proportional to the rate of encounter between consumers, related both to their speed of movement and the range at which they sense each other and the time wasted by consumer per one encounter (Beddington, 1975). On the other hand, this coefficient can consider also the behavioral change of the food resource when consumer population grows [like in epidemic models for pair susceptibles-infectives (Capasso and Serio, 1978; Wang and Zhao, 2004; Han et al., 2017)]. The greater the coefficient α_i, the greater the activity of the food resource in i-th patch and vice versa. When $\alpha_i = 0$ the saturated intake rate is a bilinear form of Lotka-Voltera functional response, which considers the inactive food resource without behavioral reaction on the consumer population changes. For our convenience we introduce the food resource classification: the higher activity resource with $\alpha_i \geq r_i^{-1}\beta_i$, the lower activity resource with $0 < \alpha_i < r_i^{-1}\beta_i$ and non-active resource with $\alpha_i = 0$. These assumptions lead to the following resource population dynamics.

$$\frac{dy_i}{dt} = r_i y_i(t)\left(1 - y_i(t)K_i^{-1}\right) - y_i(t)\beta_i\hat{W}(t) \times \left(1 + \alpha_i\hat{W}(t)\right)^{-1}, t \in (0, T], \quad (1)$$

Consumer population dynamics $w(a, t)$ are governed by the delayed McKendrick-Von Foerster's age-structured model (Von Foerster, 1959; Gurtin and Maccamy, 1974; Gurtin and MacCamy, 1979):

$$\frac{\partial w}{\partial t} + \frac{\partial w}{\partial a} = -s(a, C(t - \tau))w(a, t), (a, t) \in Q, \quad (2)$$

where $Q = \{(a,\ t)\,|a \in (0, a_d],\ t \in (0, T]\}$, $\overline{Q} = \{(a,\ t)|a \in [0, a_d]\}$, $t \in [0, T]$. Eqs (1, 2) are completed by the following initial and boundary conditions:

$$y_i(t) = y_{0_i}(t),\ t \in [-\tau - 0], \quad (3)$$

$$w(a, 0) = \varphi(a), a \in [0, a_d], \quad (4)$$

$$w(0, t) = \int_{a_r}^{a_m} \theta(a, C(t - \tau))w(a, t)da, t \in (0, T], \quad (5)$$

where $a_r > 0$ is an age of maturation, $a_m > 0$ is a maximum age of reproduction, and $\varphi(a)$ is an initial density of consumers. Functions $s(a, C(t - \tau))$ in Eq. (2) and $\theta(a, C(t - \tau))$ in Eq. (5) are age and calorie intake rate dependent consumer's death and fertility rates, respectively. Consumption of food resources by one consumer per unit of time is measured by calorie intake rate $C(t)$. This function is used in Eqs (2, 5) with the time delay parameter $\tau > 0$, which is a handling time, i.e., the time a consumer needs to handle and digest a unit of resource. We assume that the calorie intake rate is a linear function of the amount of food resource taken by one consumer per unit of time from all patches and is defined through the resource intake rate:

$$C(y(t), \hat{W}(t)) = \sum_{i=1}^{n} C_i(y_i(t), \hat{W}(t)) = \sum_{i=1}^{n} e_i y_i(t)\beta_i\left(1 + \alpha_i\hat{W}(t)\right)^{-1}, \quad (6)$$

where $C_i \geq 0$ is a calorie intake rate for i-th resource patch, $e_i > 0$ is an efficiency with which the consumed resource i is transformed to energy. We impose the following natural restrictions on the consumer's death and fertility rates, preferences in food resource:

$$\theta(a, C) \in C^1([0, a_d] \times R_{\geq 0}), \theta(a, C) \geq 0, \theta(a, 0) = 0, \frac{\partial\theta}{\partial C} > 0, \quad (7)$$

$$s(a, C) \in C^1([0, a_d] \times R_{\geq 0}), s(a, C) > 0, \frac{\partial s}{\partial C} < 0,$$

$$\gamma(a) \in (0, 1], \gamma(a) \in L_2([0, a_d]). \quad (8)$$

where $C^1(X)$ is a space of continuously differentiable functions defined in domain X, $L_2([0, a_d])$ is a space of square-integrable functions on an interval $[0, a_d]$, $R_{\geq 0}$ is a set of non-negative real numbers (Kolmogorov and Fomin, 1999).

Equations (7, 8) mean that decreasing of calorie intake rate corresponds to the critical foraging or starvation, and increasing it corresponds to the sufficient foraging and satiation with increasing resource intake rate. Increasing of calorie intake rate provides also maximum comfortable conditions for reproduction of consumers that corresponds to increasing of birth rate, and decreasing of calorie intake rate provides the most poor and unfavorable conditions for reproduction of consumers, decreasing birth rate.

The basic reproduction number of age-structured model of consumer population dynamics is a calorie intake rate depending function:

$$R(C) = \int_{a_r}^{a_m} \theta(a, C) \exp\left(-\int_0^a s(\xi, C)d\xi\right) da. \quad (9)$$

Derivation of Eq. (9) is traditional and is given in works (Hoppensteadt, 1975; Metz and Diekmann, 1986; Webb, 2008; de Ross and Persson, 2013; Akimenko and Křivan, 2018). The novelty of function (9) lies in using the calorie intake rate depending birth and death rates of consumer population in integral. This formal substitution of calorie intake rate emphasizes the foraging-depending demographical processes of consumer population.

EXISTENCE OF STATIONARY EQUILIBRIA OF THE AUTONOMOUS SYSTEM (1)–(5)

We consider the equilibria $y^* = (y_1^*, ..., y_n^*)$, $w^*(a)$ of the autonomous system (1)–(5). Trivial equilibrium $y^* = (0, ..., 0)$, $w^*(a) \equiv 0$ means that all food patches are depleted and consumer population is empty. Semi-trivial equilibrium $y^* = (K_1, ..., K_n)$, $K = (K_1, ..., K_n)$, $w^*(a) \equiv 0$ corresponds to the abundant food patches (with saturated food density) and empty consumer population. It is easy to verify that trivial and semi-trivial equilibria of autonomous system (1)–(5) with coefficients satisfied Eqs (7, 8) always exist.

Non-trivial equilibrium means that there exist a nonempty set of non-depleted food patches with positive and bounded equilibrium densities $0 < y_i^* < K_i$, $i \in I_+$, while the remaining patches are depleted $y_i^* = 0$, $i \in I_0$, and the consumer population is not empty with nonnegative equilibrium density $w^*(a) \geq 0$, positive equilibrium quantity and weighted quantity of consumers $W_0^* = \int_0^{a_d} w^*(a)da > 0$, $\hat{W}^* = \int_0^{a_d} \gamma(a)w^*(a)da > 0$. Symbols I_0 and I_+ denote here the bounded non-overlapping sets of integer indexes such that their union contains the indexes of all patches: $I_0 \cap I_+ = \emptyset$, $I_0 \cup I_+ = \{i \,|\, i \in N, i = 1, ..., n\}$. This type of equilibria corresponds to the stationary state of system in which consumer population coexists with several or all non-depleted food patches. In this section we study the conditions of existence of such nontrivial equilibria. Equilibrium y_i^*, $i = 1, ..., n$, satisfies the equation:

$$y_i^* \left(r_i \left(1 - y_i^* K_i^{-1}\right) - \beta_i \hat{W}^* \left(1 + \alpha_i \hat{W}^*\right)^{-1} \right) = 0. \quad (10)$$

Equation (10) has at most two nonnegative solutions:

$$y_i^* = K_i \left(1 - r_i^{-1}\beta_i \hat{W}^* \left(1 + \alpha_i \hat{W}^*\right)^{-1}\right) > 0, \quad \text{only if } \hat{W}^*(r_i^{-1}\beta_i - \alpha_i) < 1; \quad (11)$$

$$y_i^* = 0. \quad (12)$$

Hence, the nontrivial equilibrium of food web contains the nonempty set of non-depleted patches that necessarily satisfy condition $\hat{W}^*(r_i^{-1}\beta_i - \alpha_i) < 1$ with indexes $i \in I_+ = \left\{ i \,\middle|\, 1 \leq i \leq n,\ \hat{W}^*(r_i^{-1}\beta_i - \alpha_i) < 1 \right\}$ and the set (empty or not) of depleted patches that can satisfy or not the condition $\hat{W}^*(r_i^{-1}\beta_i - \alpha_i) \geq 1$ with indexes $i \in I_0$.

From Eq. (11) we obtain the positive equilibria $\hat{W}^* > 0$:

$$\hat{W}^* = \left(1 - y_i^* K_i^{-1}\right) \left(r_i^{-1}\beta_i - \alpha_i \left(1 - y_i^* K_i^{-1}\right)\right)^{-1} = \text{const} > 0, \quad \text{for all } i \in I_+, \quad (13)$$

or

$$\hat{W}^* = n_0^{-1} \sum_{i \in I_+} \left(1 - y_i^* K_i^{-1}\right) \left(r_i^{-1}\beta_i - \alpha_i \left(1 - y_i^* K_i^{-1}\right)\right)^{-1}, \quad (14)$$

where n_0 is a number of patches of set $I_+ \neq \emptyset$, $1 \leq n_0 \leq n$. Substituting Eqs (11) and (13) in Eq. (6) we obtain the equilibrium calorie intake rate C^*:

$$C^* = \sum_{i \in I_+} e_i y_i^* \beta_i \left(1 + \alpha_i \hat{W}^*\right)^{-1}. \quad (15)$$

The second equation of equilibrium is obtained from Eqs (2, 5):

$$\frac{dw^*}{da} = -s(a, C^*)w^*(a), \quad (16)$$

$$w^*(0) = \int_{a_r}^{a_m} \theta(a, C^*)w^*(a)da. \quad (17)$$

The general solution of Eq. (16) is $w^*(a) = w^*(0) \exp\left(-\int_0^a s(\upsilon, C^*)d\upsilon\right)$. Substituting Eq. (17) in this solution yields the integral equation for the equilibrium density of consumer $w^*(a)$:

$$w^*(a) = \exp\left(-\int_0^a s(\upsilon, C^*)d\upsilon\right) \int_{a_r}^{a_m} \theta(\upsilon, C^*)w^*(\upsilon)d\upsilon. \quad (18)$$

Integrating Eq. (18) with respect to a from 0 to a_d we obtain expression with the equilibrium quantity of consumers W_0^*:

$$\int_{a_r}^{a_m} \theta(\upsilon, C^*)w^*(\upsilon)d\upsilon = W_0^* \left(\int_0^{a_d} \exp\left(-\int_0^a s(\upsilon, C^*)d\upsilon\right) da\right)^{-1}. \quad (19)$$

Multiplying both sides of Eq. (18) by $\gamma(a)$, integrating them with respect to a from 0 to a_d and substituting in obtained equation the left side of Eq. (19) we obtain W_0^*:

$$W_0^* = \hat{W}^* \int_0^{a_d} \exp\left(-\int_0^a s(\upsilon, C^*)d\upsilon\right) da \times \left(\int_0^{a_d} \gamma(a) \exp\left(-\int_0^a s(\upsilon, C^*)d\upsilon\right) da\right)^{-1}. \quad (20)$$

By analogy with Theorem 1 from Hritonenko and Yatsenko (2007) formulated for harvesting problem, we obtain

Theorem 1. Let coefficients of system (1)–(5) satisfy conditions (7)–(8). System (11), (12), (18), possess a nontrivial equilibrium $y_i^* > 0$ ($i \in I_+$), $y_i^* = 0$ ($i \in I_0$), and $w^*(a) \geq 0$,

$W_0^* = \int_0^{a_d} w^*(a)da > 0$, if and only if there exists the positive solution $\hat{W}^* > 0$ of equation $R(C(y^*(\hat{W}^*), \hat{W}^*)) = 1$ with restrictions $\hat{W}^*(r_i^{-1}\beta_i - \alpha_i) < 1$, $i \in I_+$. The basic reproduction number $R(C^*)$, equilibrium y_i^*, $i \in I_+$, equilibrium calorie intake rate C^* and equilibrium quantity of consumers W_0^* are given by Eqs (9, 11, 15, 20), respectively. The equilibrium distribution of consumer's density $w^*(a) \in C^1([0, a_d])$ is defined by:

$$w^*(a) = W_0^* \exp\left(-\int_0^a s(\upsilon, C^*)d\upsilon\right) \times \left(\int_0^{a_d} \exp\left(-\int_0^a s(\upsilon, C^*)d\upsilon\right) da\right)^{-1}. \quad (21)$$

Proof. Multiplying both sides of Eq. (18) by $\theta(a, C^*)$, integrating them with respect to a from a_r to a_m after a little algebra we arrive to the equation $R(C^*(y^*(\hat{W}^*), \hat{W}^*)) = 1$ [see Eq. (9)]. If this equation has solution $\hat{W}^* > 0$ satisfied $\hat{W}^*(r_i^{-1}\beta_i - \alpha_i) < 1$ [$i \in I_+$, see Eq. (11)], $y_i^* = 0$ ($i \in I_0$), we can obtain y_i^*, $i \in I_+$, [Eq. (11)], C^* [Eq.(15)], and W_0^* [Eq. (20)]. Conversely, if equation $R(C(y^*(\hat{W}^*), \hat{W}^*)) = 1$ does not have solution $\hat{W}^* > 0$ satisfied $\hat{W}^*(r_i^{-1}\beta_i - \alpha_i) < 1$, $i \in I_+$, $y_i^* = 0$, $i \in I_0$, the stationary solution of problem (11), (12), (18), (20) does not exist.

Substituting the left-hand side of Eq. (19) in Eq. (18) we obtain the equilibrium distribution of consumer's density (21). Since coefficients of system (1)–(5) satisfy conditions (7)–(8) the equilibrium $w^*(a) \in C^1([0, a_d])$. Theorem 1 is proved.

Corollary 1. If some patches have higher activity resources with $\alpha_i \geq r_i^{-1}\beta_i$, condition $\hat{W}^*(r_i^{-1}\beta_i - \alpha_i) < 1$ holds for them and such patches always have positive equilibria $y_i^* > 0$ defined by Eq. (11) (i.e., non-depleted patches).

Theorem 1 imposes the restriction on the basic reproduction number of consumer population $R(C(y^*, \hat{W}^*))$ at the equilibrium y^*, $\hat{W}^*$, taking into account the impact of foraging on the consumer fertility and mortality. The condition of existence of nontrivial balance between food resource growing and consumer demographical processes (nontrivial equilibrium) is given in the form of transcendental integral equation $R(C(y^*(\hat{W}^*), \hat{W}^*)) = 1$. Implementation of such condition in biological applications is difficult from the technical point of view. In the next theorem we provide the sufficient conditions for existence of the nontrivial equilibrium in the simpler form of restrictions on the coefficients of the system (1)–(5).

Theorem 2. Let the sets of indexes of the lower and higher activity resources of non-depleted patches are $\breve{I}_+ = \{i \,|\, i \in I_+, \alpha_i < r_i^{-1}\beta_i\}$ and $\widehat{I}_+ = \{i \,|\, i \in I_+, \alpha_i \geq r_i^{-1}\beta_i\}$, respectively, constant $\hat{W}_c = \min_{i \in \breve{I}_+}(r_i^{-1}\beta_i - \alpha_i)^{-1}$,

$$\tilde{y}_i = \begin{cases} K_i\left(1 - r_i^{-1}\beta_i\hat{W}_c(1 + \alpha_i\hat{W}_c)^{-1}\right), & \text{if } i \in I_+ \text{ and } \breve{I}_+ \neq \emptyset, \\ K_i\left(1 - (r_i\alpha_i)^{-1}\beta_i\right), & \text{if } i \in I_+ \text{ and } \breve{I}_+ = \emptyset, \end{cases} \quad (22)$$

$$\tilde{C}_i = \begin{cases} e_i\beta_i K_i\left(1 - r_i^{-1}\beta_i\hat{W}_c\left(1 + \alpha_i\hat{W}_c\right)^{-1}\right)\left(1 + \alpha_i\hat{W}_c\right)^{-1}, \\ \text{if } i \in I_+ \text{ and } \breve{I}_+ \neq \emptyset, \\ 0, \quad \text{if } i \in I_+ \text{ and } \breve{I}_+ = \emptyset, \end{cases} \quad (23)$$

and coefficients of system (1)–(5) satisfy Eqs (7, 8). Then, for existence of at least one non-trivial solution of stationary problem (11), (12), (18) equilibrium $y_i^*(\hat{W}^*) \in (\tilde{y}_i, K_i)$, $i \in I_+$, $y_i^* = 0$ ($i \in I_0$), and $w^*(a) \geq 0$, $w^*(a) \in C^1([0, a_d])$, $W_0^* > 0$ it is sufficient that $R(C_{\text{sup}}^*) > 1$ and $R(C_{\text{inf}}^*) < 1$ where the infimum and supremum of equilibrium calorie intake rate are:

$$C_{\text{inf}}^* = \sum_{i \in I_+} \tilde{C}_i, \; C_{\text{sup}}^* = \sum_{i \in I_+} e_i\beta_i K_i. \quad (24)$$

The proof of Theorem 2 is given in **Supplementary Appendix A**.

Corollary 1. If all non-depleted patches have only lower activity resources with $\alpha_i < r_i^{-1}\beta_i$, when $\breve{I}_+ = \emptyset$, from Eq. (23) it follows that the infimum of equilibrium intake rate $C_{\text{inf}}^* = \sum_{i \in I_+} \tilde{C}_i = 0$, $\theta(a, 0) = 0$ [Eq.(7)] and, consequently, $R(0) = 0$. In this case condition $R(C_{\text{inf}}^*) < 1$ is always satisfied and can be omitted in Theorem 2.

LOCAL ASYMPTOTIC STABILITY OF EQUILIBRIA OF THE AUTONOMOUS SYSTEM (1)–(5)

The conditions of local asymptotic stability of the trivial and semi-trivial equilibria are addressed in Theorem 3.

Theorem 3.

(i) The trivial equilibrium $y_i^* = 0$, $i = 1, ..., n$, $w^*(a) \equiv 0$ of system (1)–(5) is unstable for all $\tau > 0$.

(ii) The semi-trivial equilibrium $y_i^* = K_i$, $i = 1, ..., n$, $w^*(a) \equiv 0$ is unconditionally (i.e., for all $\tau > 0$) locally asymptotically stable if the consumer's basic reproduction number $R(K) < 1$ whereas it is unstable for all $\tau > 0$ if $R(K) \geq 1$.

The proof of Theorem 3 is given in **Supplementary Appendix B**.

Remark 1. If consumer population is fully extinct and cannot renew the reproduction, the Eq. (47) (**Supplementary Appendix B**) has only the trivial solution $\tilde{\xi}(a) \equiv 0$. The roots of Eq. (46) (**Supplementary Appendix B**) are always negative $\lambda^* = -r_i < 0$, that is the perturbations $\zeta_i(t) \to 0$ and the semi-trivial equilibrium is locally asymptotically stable. The examples of the nonlinear age-structured models of population dynamics in the form of a single pulse–population outbreak with following extinction were obtained in works (Akimenko, 2017d; Akimenko and Anguelov, 2017).

The conditions of local asymptotic stability of the nontrivial equilibrium are addressed in Theorem 4.

Theorem 4. Let coefficients of the system (1)–(6) satisfy conditions (7), (8), the nontrivial equilibrium $y_i^* > 0$ ($i \in I_+ \neq \emptyset$), $y_i^* = 0$ ($i \in I_0$), $w^*(a) \geq 0$, $W_0^* > 0$, $\hat{W}^* > 0$ of autonomous system (1)–(5) is a solution of the equation

$R(C(y^*(\hat{W}^*), \hat{W}^*)) = 1$ with restrictions $\hat{W}^*(r_i^{-1}\beta_i - \alpha_i) < 1$ ($i \in I_+$) satisfied Eqs (11, 12, 20, 21).

This equilibrium is unstable for all $\tau > 0$ if $I_0 \neq \emptyset$ and at least one of two following statements is true:

(i) $\exists i \in I_0$ for which $\alpha_i \geq r_i^{-1}\beta_i$, or

(ii) $\alpha_i < r_i^{-1}\beta_i$ for all $i \in I_0$, and $\hat{W}^* \leq \max\limits_{i \in I_0} \left(r_i^{-1}\beta_i - \alpha_i\right)^{-1}$.

This equilibrium is unconditionally locally asymptotically stable (for all $\tau > 0$) if one of two following statements is true:

(iii) $I_0 \neq \emptyset$, $\alpha_i < r_i^{-1}\beta_i$ for all $i \in I_0$ and $\hat{W}^* > \max\limits_{i \in I_0} \left(r_i^{-1}\beta_i - \alpha_i\right)^{-1}$, or

(iv) $I_0 = \emptyset$.

Digestion period of generalist consumer τ does not cause local asymptotical instabilities of consumer population at the nontrivial equilibria.

The proof of Theorem 4 is given in **Supplementary Appendix C**.

Remark 1. According to the statement (iv) of Theorem 4 the nontrivial equilibrium with non-depleted food patches ($I_0 = \emptyset$) is always locally asymptotically stable. That means that there exists a balance between resource growing, demographical process of consumer population and their consumption regime (within the framework of the considered model), which guarantees the steady coexistence of all non-depleted resource patches with non-empty consumer population. The local asymptotic stability of the nontrivial equilibria of nonlinear age-structured models with density dependent fertility and death rates is well predicted by the partial derivative of basic reproduction number (Cushing, 1998; Akimenko and Křivan, 2018). Such stability indicator of non-trivial equilibrium with non-depleted patches of system (1)–(5) has the form:

$$\frac{\partial R(y^*(\hat{W}^*), \hat{W}^*)}{\partial \hat{W}^*} = -\int_{a_r}^{a_m} \exp\left(-\int_0^a s(\upsilon, C^*)d\upsilon\right) \times \left(\frac{\partial \theta(a, C^*)}{\partial C^*} - \theta(a, C^*)\int_0^a \frac{\partial s(\upsilon, C^*)}{\partial C^*}d\upsilon\right)da \times \sum_{i=1}^{n} e_i\beta_i K_i \left(1 + \alpha_i \hat{W}^*\right)^{-3} \times \left(r_i^{-1}\beta_i + \alpha_i\left(1 - \hat{W}^*\left(r_i^{-1}\beta_i - \alpha_i\right)\right)\right) < 0. \quad (25)$$

The negative value of this expression indicates the local asymptotic stability of the nontrivial equilibrium (see Cushing, 1998; Akimenko and Křivan, 2018) with non-depleted patches that confirms the statement (iv) of Theorem 4. On the other hand, this stability indicator cannot be used for analysis of local asymptotic stability of nontrivial equilibria with depleted patches [cases (i), (ii), (iii) of Theorem 4], because equilibrium calorie intake rate C^* [Eq. (15)] depends only from the equilibrium of non-depleted patches $y_i^* \in I_+$ and stability indicator (25) is negative for all non-trivial equilibria. That is why we will use in the next section the conditions of local asymptotic stability of non-trivial equilibria given in Theorem 4 instead of indicator (25).

NUMERICAL EXPERIMENTS

Parameterization of the Autonomous System (1)–(5)

We assume that the consumer fertility rate is increasing with a saturation monotone function and the death rate is a decreasing with extinction monotone function of calorie intake rate satisfied Eqs (7, 8). They are defined on the parametrized classes of algebraic functions:

$$s(a, C(t)) = 0.5s_0\big(0.5 + \pi^{-1}\arctan(4(a - a_m)) + 4\exp(-4C(t))\big), a \in [0, a_d], \quad (26)$$

$$\theta(a, C(t)) = \theta_0\left(1 - \exp(-0.5\ C(t))\right), a \in [a_r, a_m], \quad (27)$$

where s_0, θ_0 are given constants. The generalist consumer population is partitioned into three age-structured groups with young, matured, and senile individuals with a number of individuals in each group $W_r(t) = \int_0^{a_r} w(a, t)dt$, $W_m(t) = \int_{a_r}^{a_m} w(a, t)dt$, and $W_d(t) = \int_{a_m}^{a_d} w(a, t)dt$, respectively. We use the following piece-constant function of resource intake weighted coefficient among age-structured groups [$\gamma(a) \in L_2([0, a_d])$, Eq. (8)]:

$$\gamma(a) = \begin{cases} \gamma_r, & \text{if } a \in [0, a_r], \\ \gamma_m, & \text{if } a \in [a_r, a_m], \\ \gamma_d, & \text{if } a \in [a_m, a_d]. \end{cases} \quad (28)$$

where $0 < \gamma_r < \gamma_d < \gamma_m < 1$ the set of constant dimensionless weights of resource intake for young, senile, and matured consumers, respectively. The biggest value of γ_m in comparison with γ_d and γ_r means that one matured consumer takes more resource biomass than young or senile consumer.

For illustration of theoretical results obtained in Theorems 3 and 4, we consider the minimal set of three food resource patches in all experiments in the vicinities of stationary equilibria. The values of coefficients of Eqs (1, 6), and initial values (3, 4) vary in each experiment depending on the conditions of Theorems 3, 4.

The Trivial and Semi-Trivial Equilibria

The numerical method based on the method of characteristics (Akimenko, 2017b,c,d) is used here for study the dynamical regimes of autonomous system (1)–(5) in the vicinities of all equilibria considered in "Local Asymptotic Stability of Equilibria of the Autonomous System (1)–(5)."

In the first group of experiments we study the asymptotic behavior of solutions in the vicinity of the trivial and semi-trivial equilibria with fixed small value of time delay $\tau = 0.001a_d$. The dynamics of mean resource density $Y(t) = n^{-1}\sum_{i=1}^{n} y_i(t)$, the quantity of consumers $W_0(t) = \int_0^{a_d} w(a, t)da$, and the basic reproduction number $R(t)$ are shown in **Figure 1**. The numerical simulations illustrate the results obtained in Theorem 3.

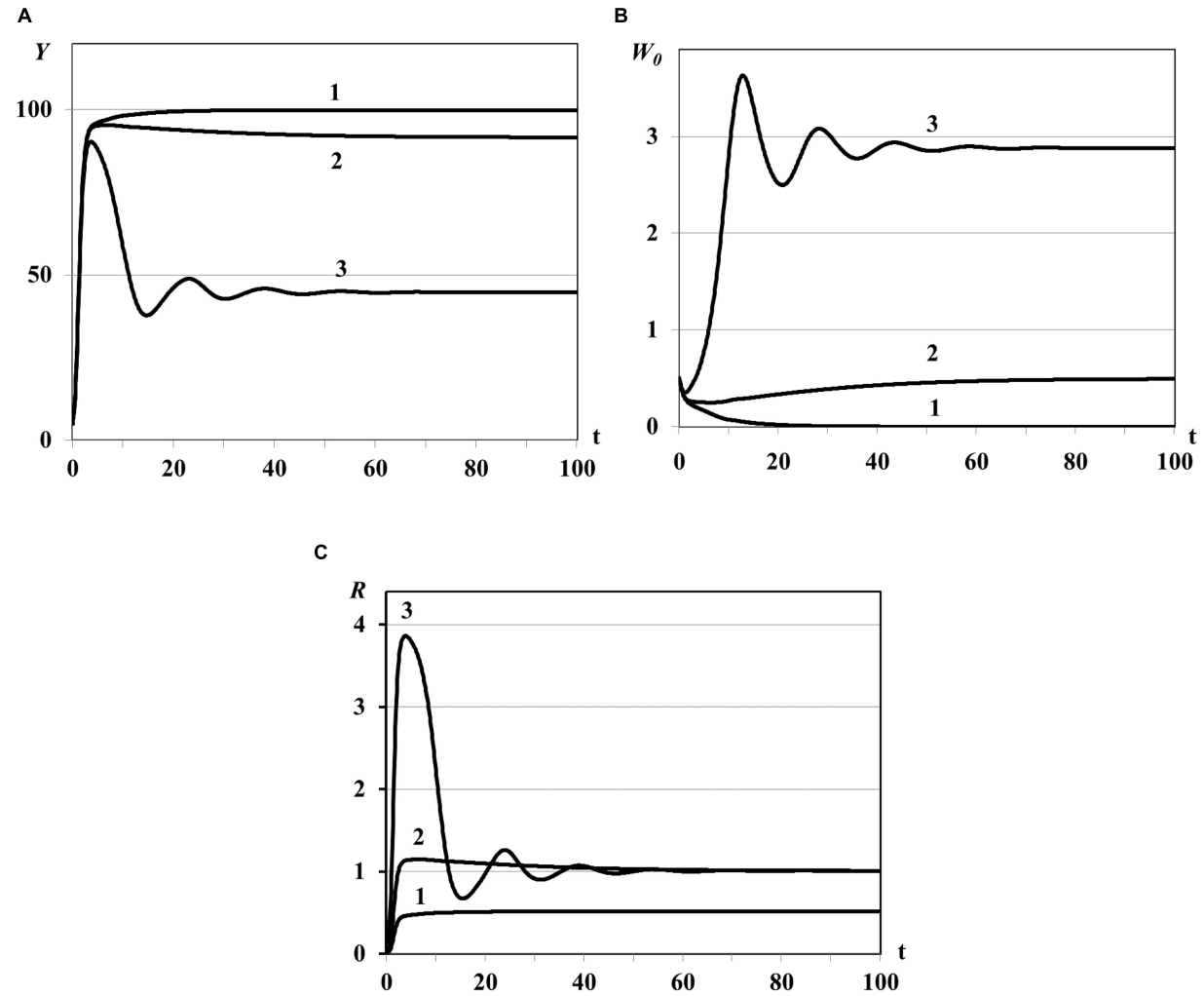


FIGURE 1 | Graphs of asymptotic convergence of solutions with $R(K) < 1$ (curve 1), $R(K) > 1$ (curve 2), and $R(K) >> 1$ (curve 3). $Y(t)$ - mean resource density **(A)**, $W_0(t)$ - quantity of consumers **(B)**, $R(t)$ - consumer's basic reproduction number **(C)**.

Case (i) of Theorem 3: in this experiment the trivial equilibrium is unstable (all curves in **Figure 1**), $Y(t)$ and $W_0(t)$ evolve to the semi-trivial equilibrium (curves 1 in **Figures 1A,B**, consumer population becomes extinct while the resource biomass in all patches saturates, $R(K) < 1$) or evolve to the nontrivial equilibrium (curves 2 and 3 in **Figures 1A,B**, $R(K) > 1$).

Case (ii) of Theorem 3: in this experiment the semi-trivial equilibrium is asymptotically stable with $R(K) < 1$ (curves 1 in **Figure 1**) whereas it is unstable with $R(K) > 1$ (graphs 2 in **Figure 1**).

For the very large basic reproduction number $R(K) >> 1$ obtained with large value of θ_0 [Eq. (86)] we observe the oscillatory regime of the system with asymptotic convergence of solution to the steady state (curve 3 in **Figure 1**). The existence of such periodic solutions of some Lotka-Volterra prey-predator models was proved in theoretical work (Xu et al., 2004) and was observed in numerical experiments in Akimenko (2017d) and Akimenko and Anguelov (2017).

Further increasing of basic reproduction number by parameter θ_0 causes the consumer population outbreaks (special dynamical regimes of population, see Abbott and Dwyer, 2007; Akimenko and Anguelov, 2017; Akimenko and Piou, 2018). The pulse sequence or sequence of outbreaks (**Figure 2**) of consumers population and resource densities describe the quasi-periodic dynamical regime in the vicinities of the trivial and semi-trivial equilibria. The fast growing of consumer population is accompanied by huge resource consumption, and as a consequence, by resource extinction and following decreasing of consumer population density to minimal but not critical values. Although this minimal value cannot be seen on the graphs due to their small scale, we observe a quasi-periodic recovery and renewal of customer population that would be impossible with the complete disappearance of reproductive individuals in the population.

The system moves to the trivial equilibrium from the vicinity of unstable semi-trivial equilibrium ($R(K) >> 1$, statement

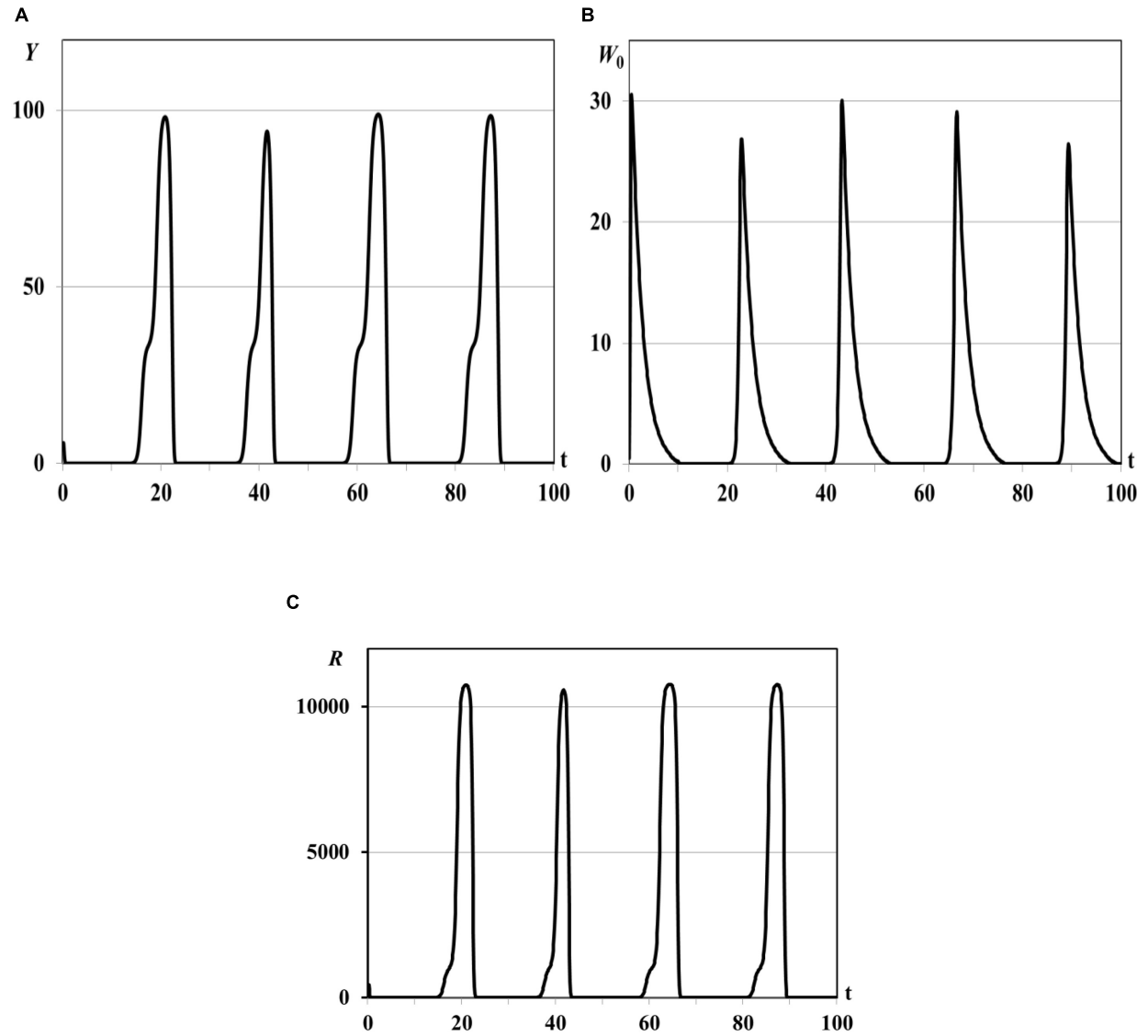


FIGURE 2 | Graphs of periodic population outbreaks with $R(K) >> 1$, $\tau = 0.001a_d$. $Y(t)$ - mean resource density **(A)**, $W_0(t)$ - quantity of consumers **(B)**, $R(t)$ - consumer's basic reproduction number **(C)**.

(ii) of Theorem 2). But, since the trivial equilibrium is unstable too (statement (i) of Theorem 2), and the minimal number of consumers is sufficient for the following renewal of population, system moves to the semi-trivial equilibrium again. This process is repeated at quasi-periodic time intervals and results in the pulse sequence of consumer population and resource densities. The same regimes were obtained in work (Akimenko, 2017d) for the nonlinear age-structured model of population dynamics with density-dependent delayed death rate only for the big values of delay parameter and/or for the periodic time-dependent death and fertility rates. Since in this experiment the impact of the time delay parameter is insignificant and all coefficients of the model are time-independent the dynamical regimes of periodic outbreaks are result of the repeating dynamical interaction between total resource consumption and its renewing from the one hand and consumer population growth and extinction from the other hand.

Further increasing of parameter θ_0 leads to the consumer population outbreaks of the single pulse form (**Figures 3A,B**). The same rapid consumer population growth like in the previous experiment is accompanied by huge resource consumption and resource extinction but with following decreasing of consumer population density up to critical values when population is not able to renew the reproduction and becomes fully extinct (see Remark 1 to Theorem 3). The food resources in all patches saturate with time and system evolves eventually to the asymptotically stable semi-trivial equilibrium (**Figure 3A**).

In the second group of experiments we study the dynamical regimes of autonomous system (1)–(5) in the vicinity of the

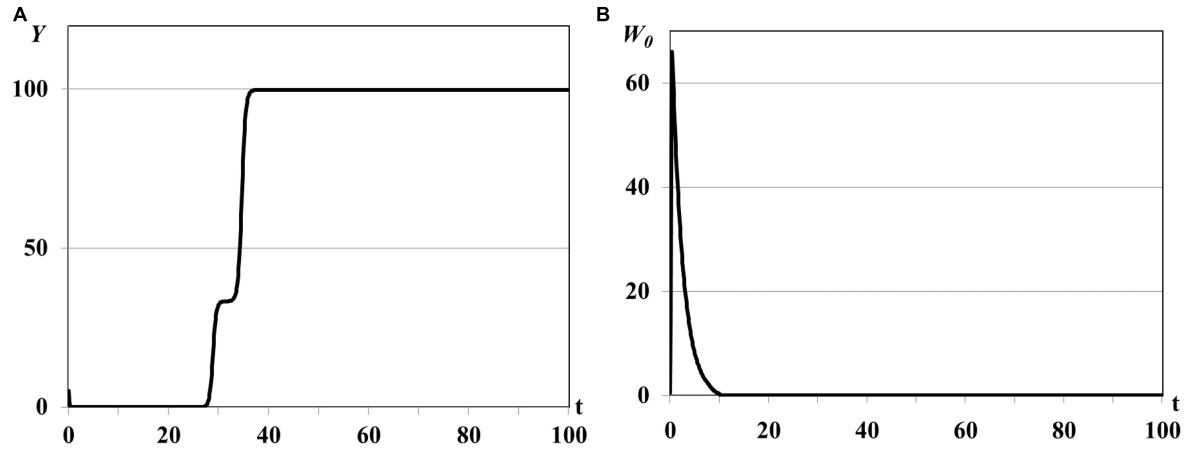


FIGURE 3 | Graphs of population single outbreak and extinction, $\tau = 0.001a_d$. $Y(t)$ - mean resource density **(A)**, $W_0(t)$ - quantity of consumers **(B)**.

trivial and semi-trivial equilibria with different values of time delay from interval $\tau = 0.001a_d, \ldots, \tau = 0.03a_d$, and large basic reproduction number $R(K) >> 1$. Solution of the autonomous system (1)–(5) oscillates with bounded magnitude in the vicinity of the semi-trivial equilibrium in all experiments with different value of time delay (**Figures 4–6**). For the small value of $\tau = 0.001a_d$ the trajectories of system have the magnitude with exctincted oscillations and converge to the positive equilibrium (curve 3 in **Figure 1**). The bigger value of $\tau = 0.005a_d$ causes the periodic dynamics of $Y(t)$ and $W_0(t)$ with bigger magnitudes, shown in **Figure 4**. Further increasing of time delay leads to the periodic outbreakes of $Y(t)$ and $W_0(t)$ with $\tau = 0.015a_d$ (**Figure 5**) and single outbreak of $W_0(t)$ with saturating $Y(t)$ with $\tau = 0.03a_d$ (**Figure 6**).

In all experiments of the second group the graphs of mean resource density $Y(t)$ and quantity of consumers $W_0(t)$ oscillate in antiphase. This is a traditional form of dynamics of prey-predator interaction. It means that increasing biomass of food patches stimulates the reproduction of consumers and decreasing of their mortality and leads to increasing size of consumer population. In this case resource consumption is rapidly increasing and, eventually, leads to the depletion of food patches. The growing deficit of food resource causes the decreasing consumer reproduction, increasing their mortality, and, eventually, decreasing size of consumer population up to the complete disappearance. Since the model considers the renewable food resource the food biomass in patches starts for growing again after consumer population decreasing or disappearance. The different growth rates r_i and carrying capacities K_i in three food patches cause the rounded smooth steps on the graphs of $Y(t)$ in **Figures 4A**, **5A**, **6A**. If the consumer population does not extinct this process repeats again and we obtain the quasiperiodical dynamics of consumer population [graphs of $W_0(t)$ in **Figures 4B**, **5B**]. Otherwise we observe the consumer population outbreak in the form of single pulse with following extinction (**Figure 6B**). In some insect and animal populations the cannibalism can be a common response to nutritional deficiency when the food patches are limited in resources or depleted (Richardson et al., 2010; Guttal et al., 2012). For example, if some population of locusts cannot migrate to new food patch and suffers from the lack of food resource in depleted patches the gregarious locusts can cannibalize each other (Guttal et al., 2012). The survivor and solitarious locusts in this population eventually die of starvation. The graphs shown in **Figures 4**, **5** correspond to the regimes of population outbreaks obtained in the previous experiments (**Figures 2**, **3**). Similar dynamical regimes were obtained and described in works (Akimenko, 2017d; Akimenko and Anguelov, 2017) for the age-structured model with density-dependent delayed death rate and discussed in work (Akimenko and Piou, 2018) for the two-compartment age-structured model of locust population dynamics.

Comparison of graphs of $Y(t)$, $W_0(t)$ on **Figures 4**, **5** reveals that an increasing value of digestion period of generalist consumer τ leads to the growth of consumer population size and increase in the time period of its outbreaks. Consumers with bigger digestion periods extend the pause between foraging that allows patches to use this time to increase their resource up to satiation. Increasing of food resource in patches leads eventually to the growth in consumer population size and resource food consumption. Larger consumer population consumes more resource emptying food patches, resulting in nutritional deficiency and increasing in the mortality of the consumer population (starvation, cannibalism, etc.). If consumers of reproductive age survive, they recover and renew the consumer population with a larger time period (graphs of $W_0(t)$ in **Figures 4B**, **5B**), otherwise consumer population cannot recover band extinct (graph of $W_0(t)$ in **Figure 6B**).

The results of this group of experiments illustrate also that digestion period of generalist consumer τ does not lead to the high-frequency oscillations of the solution in the vicinity of the semi-trivial equilibrium known as deterministic chaos.

The Nontrivial Equilibria

The third group of numerical experiments focuses on the study of asymptotic behavior of solutions in the vicinity of

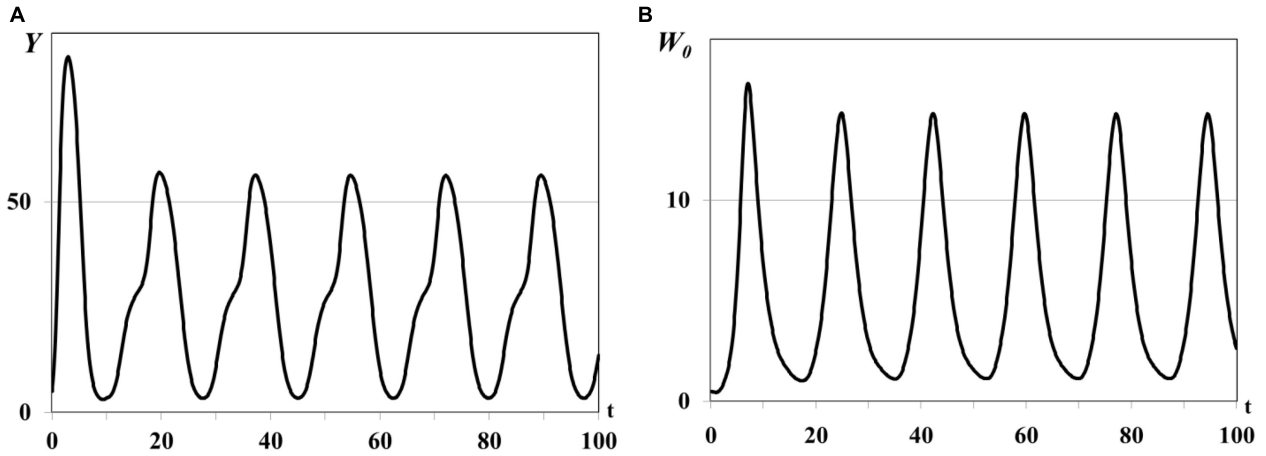


FIGURE 4 | Graphs of periodic dynamics of $Y(t)$ and $W_0(t)$, $\tau = 0.005a_d$. $Y(t)$ - mean resource density **(A)**, $W_0(t)$ - quantity of consumers **(B)**.

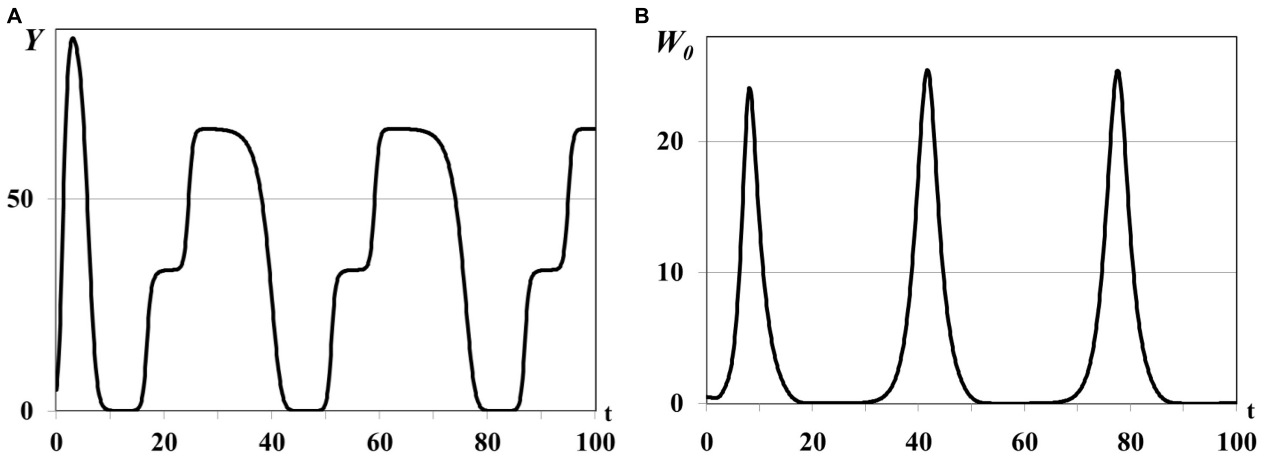


FIGURE 5 | Graphs of periodic outbreaks of $Y(t)$ and $W_0(t)$, $\tau = 0.015a_d$. $Y(t)$ - mean resource density **(A)**, $W_0(t)$ - quantity of consumers **(B)**.

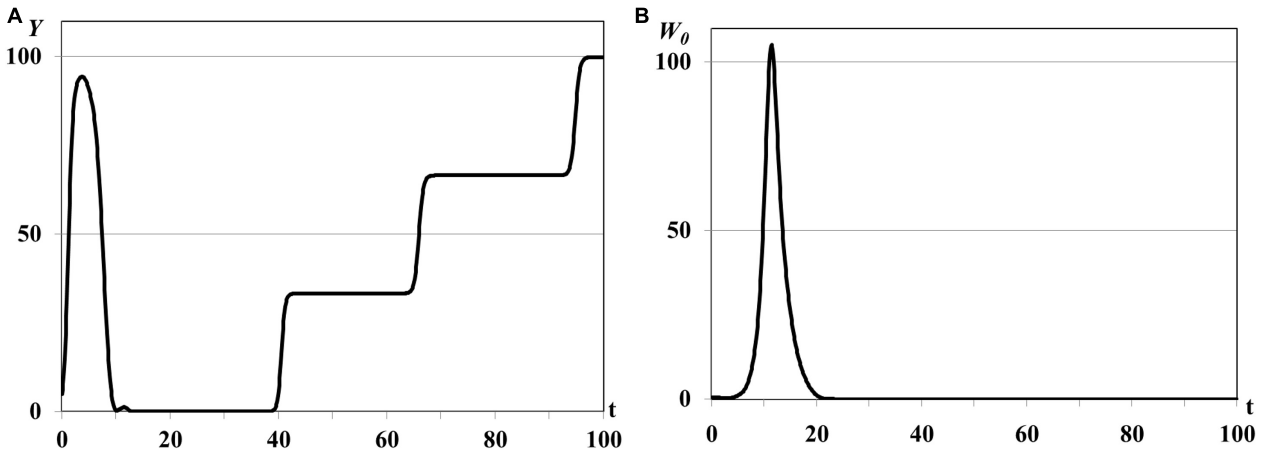


FIGURE 6 | Graphs of asymptotic saturation of $Y(t)$ **(A)**, single outbreak of $W_0(t)$ **(B)**, $\tau = 0.03a_d$.

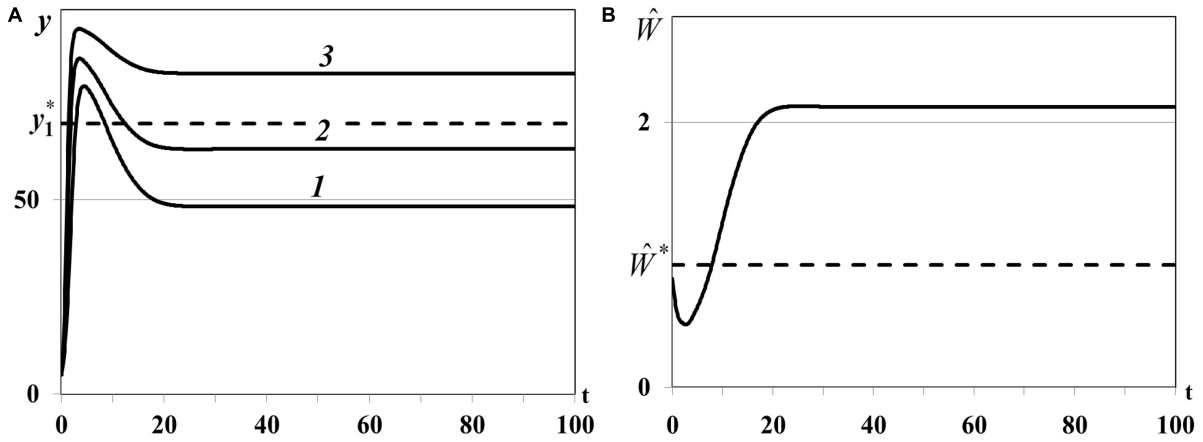


FIGURE 7 | (A) Case (i). Resource densities $y_i(t)$ $(i = 1, 2, 3)$. $y_1{}^*$ is a computed equilibrium of the first resource. **(B)** Weighted number of consumers $\hat{W}(t)$. $\hat{W}^*$ is a computed equilibrium of $\hat{W}(t)$.

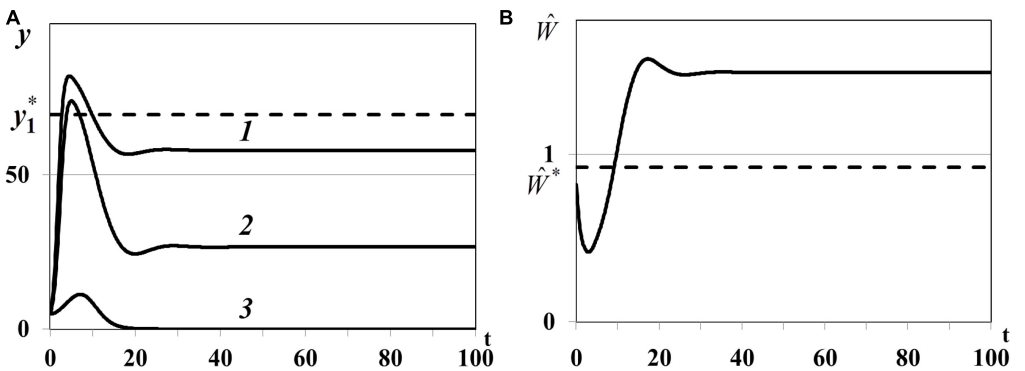


FIGURE 8 | (A) Case (ii). Resource densities $y_i(t)$ $(i = 1, 2, 3)$. $y_1{}^*$ is a computed equilibrium of the first resource. **(B)** Weighted number of consumers $\hat{W}(t)$. $\hat{W}^*$ is a computed equilibrium of $\hat{W}(t)$.

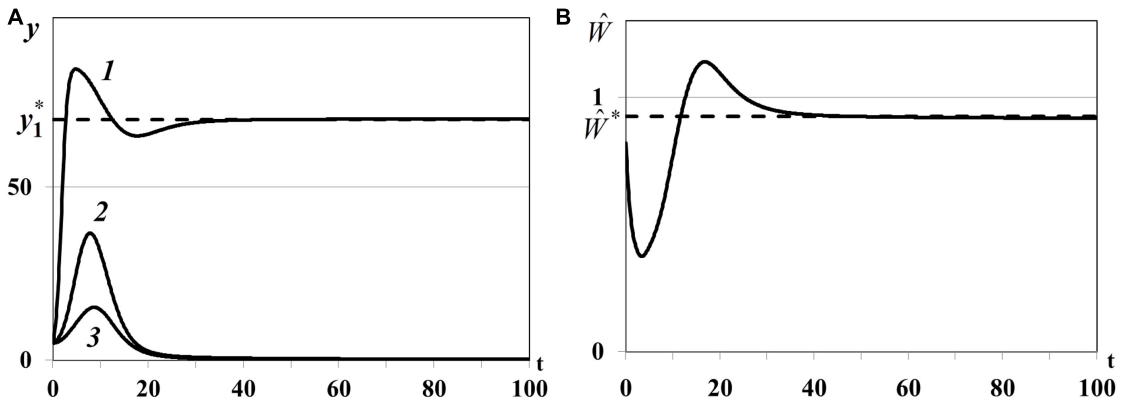


FIGURE 9 | (A) Case (iii). Resource densities $y_i(t)$ $(i = 1, 2, 3)$. $y_1{}^*$ is a computed equilibrium of the first resource. **(B)** Weighted number of consumers $\hat{W}(t)$. $\hat{W}^*$ is a computed equilibrium of $\hat{W}(t)$.

the non-trivial equilibrium where some patches are depleted, i.e., $y_i^* > 0$ $(i = 1, ..., n_0)$, $y_i^* = 0$ $(i = n_0 + 1, ..., n)$, $n_0 \leq n$, $w^*(a) \geq 0$, $W_0^* > 0$. In all experiments $n = 3$, but n_0 varies in each experiment depending from the condition of Theorem 4. For the fixed values of n, n_0 the root of transcendental equation $R(C(y^*(\hat{W}^*), \hat{W}^*)) = 1$ is defined numerically by the

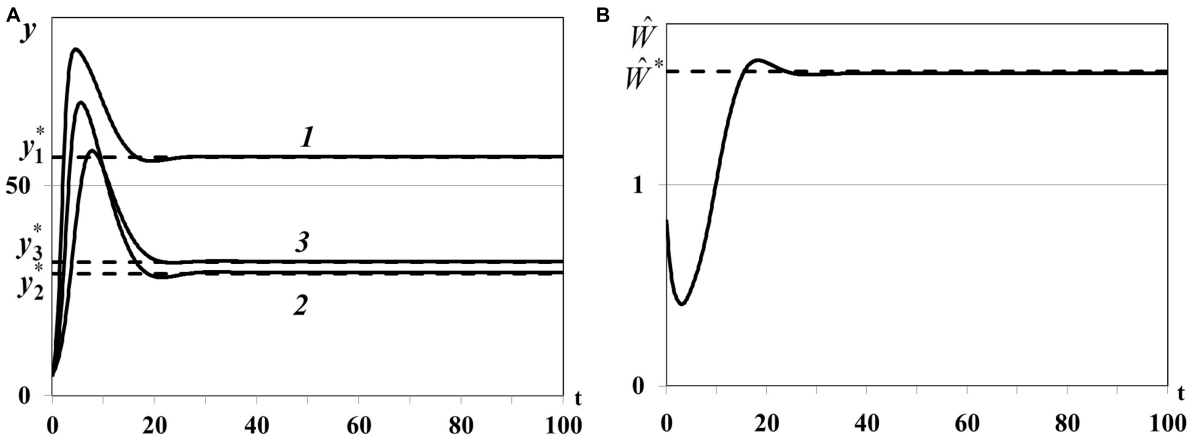


FIGURE 10 | (A) Case (iv). Resource densities $y_i(t)$ ($i = 1, 2, 3$). $y_1{}^*$ is a computed equilibrium of the i-th resource, $i = 1, 2, 3$. **(B)** Weighted number of consumers $\hat{W}(t)$. $\hat{W}^*$ is a computed equilibrium of $\hat{W}(t)$.

bisection method with restrictions $\hat{W}^*(r_i^{-1}\beta_i - \alpha_i) < 1$, $i \in I_+$. The equilibrium values y_i^*, W_0^*, $w^*(a)$ are defined by Eqs (11, 20, 21) respectively. The value of $\hat{W}^*$ is used also in figures for illustration of asymptotic convergence of trajectories to the equilibrium.

Case (i) of Theorem 4. In the first experiment we study the positive equilibrium with one non-depleted patch, less active and more active depleted patches: $n_0 = 1$, $\alpha_1 < r_1^{-1}\beta_1$, $\alpha_2 > r_2^{-1}\beta_2$, $\alpha_3 > r_3^{-1}\beta_3$. In equation $R(C(y^*(\hat{W}^*), \hat{W}^*)) = 1$ we use $y_1^* > 0$, $y_2^* = 0$, $y_3^* = 0$, for which $\hat{W}^*(r_1^{-1}\beta_1 - \alpha_1) < 1$, $\hat{W}^*(r_2^{-1}\beta_2 - \alpha_2) < 1$, $\hat{W}^*(r_3^{-1}\beta_3 - \alpha_3) < 1$. The dynamics of resource densities $y_i(t)$ $(i = 1, 2, 3)$ and consumer weighted quantity $\hat{W}(t)$ are shown in **Figure 7**. The density of first food patch $y_1(t)$ does not evolve to the equilibrium y_1^* (curve 1 in **Figure 7A**), patches $y_2(t)$ and $y_3(t)$ are not depleted (curves 2 and 3 in **Figure 7A**), and consumer weighted quantity $\hat{W}(t)$ does not evolve to the equilibrium $\hat{W}^*$ (**Figure 7B**), i.e., the nontrivial equilibrium is unstable.

Case (ii) of Theorem 4. In the second experiment we study positive equilibrium with one non-depleted patch and two less active depleted patches with $\hat{W}^* \le \max\limits_{i \in I_0} \left(r_i^{-1}\beta_i - \alpha_i\right)^{-1}$: $n_0 = 1$, $\alpha_1 < r_1^{-1}\beta_1$, $\alpha_2 < r_2^{-1}\beta_2$, $\alpha_3 < r_3^{-1}\beta_3$. In equation $R(C(y^*(\hat{W}^*), \hat{W}^*)) = 1$ we use $y_1^* > 0$, $y_2^* = 0$, $y_3^* = 0$, for which $\hat{W}^*(r_1^{-1}\beta_1 - \alpha_1) < 1$, $\hat{W}^*(r_2^{-1}\beta_2 - \alpha_2) < 1$, $\hat{W}^*(r_3^{-1}\beta_3 - \alpha_3) > 1$. The density of first food patch $y_1(t)$ does not evolve to the equilibrium y_1^* (curve 1 in **Figure 8A**), second patch is not depleted [$y_2(t)$, curve 2 in **Figure 8A**], $y_3(t)$ converges to the trivial equilibrium and becomes depleted [curve 3 in **Figure 8A**], and consumer weighted quantity $\hat{W}(t)$ does not evolve to the equilibrium $\hat{W}^*$ (**Figure 8B**), i.e., the nontrivial equilibrium of food web is unstable.

Case (iii) of Theorem 4. In the third experiment we study the positive equilibrium with one non-depleted patch, two less active depleted patches with $\hat{W}^* > \max\limits_{i \in I_0} \left(r_i^{-1}\beta_i - \alpha_i\right)^{-1}$: $n_0 = 1$, $\alpha_1 < r_1^{-1}\beta_1$, $\alpha_2 < r_2^{-1}\beta_2$, $\alpha_3 < r_3^{-1}\beta_3$. In equation $R(C(y^*(\hat{W}^*), \hat{W}^*)) = 1$ we use $y_1^* > 0$, $y_2^* = 0$, $y_3^* = 0$, for which $\hat{W}^*(r_1^{-1}\beta_1 - \alpha_1) < 1$, $\hat{W}^*(r_2^{-1}\beta_2 - \alpha_2) > 1$, $\hat{W}^*(r_3^{-1}\beta_3 - \alpha_3) > 1$. The density of first food patch $y_1(t)$ in this case evolves to the positive equilibrium y_1^* (curve 1 in **Figure 8A**), the densities of the other patches $y_2(t)$ and $y_3(t)$ evolve to the trivial equilibrium ($y_2^* = 0$, $y_3^* = 0$), and become depleted (curves 2 and 3 in **Figure 8A**). Consumer weighted quantity $\hat{W}(t)$ evolves to the positive equilibrium $\hat{W}^*$ (**Figure 8B**), i.e., the nontrivial equilibrium with one non-depleted patch and two depleted patches is locally asymptotically stable.

Case (iv) of Theorem 4. In the last fourth experiment we study the positive equilibrium with three non-depleted patches and nontrivial consumer population: $n_0 = 3$, $\alpha_1 < r_1^{-1}\beta_1$, $\alpha_2 < r_2^{-1}\beta_2$, $\alpha_3 < r_3^{-1}\beta_3$. In equation $R(C(y^*(\hat{W}^*), \hat{W}^*)) = 1$ we use $y_1^* > 0$, $y_2^* > 0$, $y_3^* > 0$, for which $\hat{W}^*(r_1^{-1}\beta_1 - \alpha_1) < 1$, $\hat{W}^*(r_2^{-1}\beta_2 - \alpha_2) < 1$, $\hat{W}^*(r_3^{-1}\beta_3 - \alpha_3) < 1$. The densities of food patches $y_i(t)$ evolve to the corresponding positive equilibria y_i^*, $(i = 1, 2, 3)$ (curves 1, 2, 3 in **Figure 10A**), consumer weighted quantity $\hat{W}(t)$ evolves to the positive equilibrium $\hat{W}^*$ (**Figure 10B**), i.e., the nontrivial equilibrium with all non-depleted patches is locally asymptotically stable.

In the third experiment the nontrivial equilibrium with one non-depleted patch and two depleted patches ($I_0 \neq \emptyset$) and in the fourth experiment the nontrivial equilibrium with three non-depleted patches ($I_0 = \emptyset$) are locally asymptotically stable (**Figures 9**, **10**). Overall, we can conclude that stability indicators obtained in Theorem 3 and 4 correctly predict the asymptotic stability or instability of trivial, semi-trivial and non-trivial equilibria of system (1)–(5) in all numerical experiments.

CONCLUSION AND DISCUSSION

In this work an autonomous system was studied–a resource-consumer model in a heterogenous environment consisting of several food patches with active resource. Food resources do

not disperse between patches, while consumers do disperse. The model of food resources is unstructured while the model of consumer population is age-structured. The relationship between the consumed food resource and consumer demographic parameters (fertility and death rates) is modeled by means of a calorie intake rate that describes the amount of energy obtained by consumer at a given age from all food patches per unit of time. In biological applications calorie intake rate can be obtained from the observations, or foraging experiments focusing on age-structured consumer behavior. The consumer calorie intake rate is proportional to the saturated intake rate (where the coefficient of saturation is a behavioral parameter of food resource) and depends on the time period a consumer needs to handle and digest a unit of resource (delayed parameter). Thus, the model considered in this paper extends the classic apparent competition models (Holt, 1984; Holt and Kotler, 1987; Holt and Lawton, 1993; Holt et al., 1994; Křivan, 2014) to a structured consumer population with time delay and active food resources.

All types of possible equilibria: trivial (depleted all n resource patches and extinct consumer population), semi-trivial (abundant all n resource patches with satiated density and extinct consumer population), and non-trivial equilibria (food web with at least one non-depleted patch, at most $(n-1)$ depleted patches and consumer population) were studied. All theorems used a new condition of sign-preserving partial derivatives of calorie intake rate-dependent fertility and mortality rates of consumer: $\frac{\partial\theta}{\partial C} > 0$, $\frac{\partial s}{\partial C} < 0$. The trivial and semi-trivial equilibria of the nonlinear autonomous system always exist while the non-trivial equilibria exist if and only if the basic reproduction number of the consumer population $R = 1$. Since this condition leads actually to a complex nonlinear equation, derivation of conditions for the existence of non-trivial equilibria in practice is not an easy problem. For facilitation of this problem, the additional sufficient conditions for existence of the nontrivial equilibria were obtained in this paper in the form of simple constraints on the coefficients of the autonomous system.

The conditions of local asymptotic stability/instability of trivial and semi-trivial equilibrium were formulated in terms of the consumer's basic reproduction number. These conditions were derived on the basis of the perturbation theory and linearization methods. Unfortunately, the well-known stability indicator of equilibria of nonlinear age-structured models–partial derivative of density-dependent basic reproduction number of consumer population (Cushing, 1998; Akimenko and Křivan, 2018) cannot be used for stability analysis of nontrivial equilibria of resource-consumer model with depleted patches. Instead of it, the conditions of instability/local asymptotic stability of nontrivial equilibria with several or without depleted resource patches were obtained in the form of additional restrictions on coefficients of system. As expected, the time-delay parameter, the consumer's digestion period does not cause local asymptotical instabilities of consumer population at the trivial, semi-trivial or nontrivial equilibria.

The dynamical regimes of autonomous system with the different values of time-delay parameter were studied in numerical experiments for illustration of the obtained theoretical results. In the 1st and 2nd groups of experiments (**Figures 1–6**) the local asymptotic stability/instability of the trivial and semi-trivial equilibria, consumer population outbreaks, extinct and non-extinct quasi-periodic dynamic regimes were obtained for the different values of the time delay parameter. The processes of resource handle and food digestion are inherent for all biological organisms, although the value of handling and digestion period can significantly differ among species. Numerical experiments showed that digestion period of generalist consumer τ does not cause the local asymptotical instabilities or high-frequency oscillations (deterministic chaos) of consumer population in the vicinity of semi-trivial equilibrium.

In the 3rd and 4th groups of experiments (**Figures 7–10**) we study the local asymptotic stability/instability of the nontrivial equilibria of system with one generalist consumer and one non-depleted and two depleted resource patches (3rd group), three non-depleted resource patches (4th group). The numerical results showed that if there exists the non-trivial equilibrium positive solution of equation $R(C(y^*(\hat{W}^*), \hat{W}^*)) = 1$, which satisfies condition $\hat{W}^*(r_i^{-1}\beta_i - \alpha_i) \geq 1$ for all depleted patches $(i \in I_0)$ and $\hat{W}^*(r_i^{-1}\beta_i - \alpha_i) < 1$ for all non-depleted patches $(i \in I_+)$ such equilibrium is always locally asymptotically stable. The coefficient of saturation (behavioral characteristic of a food resource) α_i plays an important role in this criterion: if $\alpha_i > r_i^{-1}\beta_i$ the corresponding i-th patch cannot be depleted in the asymptotically stable equilibrium of a food web. Thus, the numerical experiments illustrated and confirmed all theoretical results obtained in paper.

The theoretical results obtained in this work can be applied, for example, to study: (i) metapopulation dynamics that include the prey-predator interactions (Nakazawa, 2015; Becker and Hall, 2016); (ii) desert locust–food resource population dynamics with non-active resource ($\alpha_i = 0$) (Guttal et al., 2012; Akimenko and Piou, 2018); (iii) the predator-dependent functional responses and prey-predator interaction strengths in a natural food web in fisheries (Essington and Hansson, 2004; Smith and Smith, 2020), and many others.

DATA AVAILABILITY STATEMENT

The datasets generated for this study are available on request to the corresponding author.

AUTHOR CONTRIBUTIONS

The author confirms being the sole contributor of this work and has approved it for publication.

ACKNOWLEDGMENTS

I would like to acknowledge Vlastimil Krivan, Cyril Piou, and referees for the helpful and valuable comments and suggestions

to the topic of manuscript. This manuscript has been released as a pre-print at bioRxiv (Akimenko, 2020) (VA, Stability Analysis of Delayed Age-structured Resource-Consumer Model of Population Dynamics with Saturated Intake Rate. BioRxiv.doi.org/10.1101/2020.01.10.901629).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fevo.2021.531833/full#supplementary-material

REFERENCES

Abbott, K. C., and Dwyer, G. (2007). Food limitation and insect outbreaks: complex dynamics in plant–herbivore models. *J. Anim. Ecol.* 76, 1004–1014. doi: 10.1111/j.1365-2656.2007.01263.x

Abrams, P. A., Holt, R. D., and Roth, J. D. (1998). Apparent competition or apparent mutualism? Shared predation when population cycle. *Ecology* 79, 201–212. doi: 10.1890/0012-9658(1998)079[0201:acoams]2.0.co;2

Akimenko, V. V. (2017a). An age-structured SIR epidemic model with the fixed incubation period of infection. *Comput. Math. Appl.* 73, 1485–1504. doi: 10.1016/j.camwa.2017.01.022

Akimenko, V. V. (2017b). Asymptotically stable states of non-linear age-structured monocyclic cell population model I. Travelling wave solution. *Math. Comput. Simul.* 133, 2–23. doi: 10.1016/j.matcom.2015.06.004

Akimenko, V. V. (2017c). Asymptotically stable states of non-linear age-structured monocyclic cell population model II. Numerical simulation. *Math. Comput. Simul.* 133, 24–38. doi: 10.1016/j.matcom.2015.06.003

Akimenko, V. V. (2017d). Nonlinear age-structured models of polycyclic population dynamics with death rates as a power functions with exponent n. *Math.Comput. Simul.* 133, 175–205. doi: 10.1016/j.matcom.2016.08.004

Akimenko, V. V. (2020). Stability analysis of delayed age-structured resource-consumer model of population dynamics with saturated intake rate. *BioRxiv [Preprint]* doi: 10.1101/2020.01.10.901629

Akimenko, V. V., and Anguelov, R. (2017). Steady states and outbreaks of two-phase nonlinear age-structured model of population dynamics with discrete time delay. *J. Biol. Dynam.* 11, 75–101. doi: 10.1080/17513758.2016.1236988

Akimenko, V. V., and Křivan, V. (2018). Asymptotic stability of delayed predator age-structured population models with an Allee effect. *Math. Biosci.* 306, 170–179. doi: 10.1016/j.mbs.2018.10.001

Akimenko, V. V., and Piou, C. (2018). Two-compartment age-structured model of solitarious and gregarious locust population dynamics. *Math. Meth. Appl. Sci.* 41, 8636–8672. doi: 10.1002/mma.4947

Becker, D. J., and Hall, R. J. (2016). Heterogeneity in patch quality buffers metapopulations from pathogen impacts. *Theor. Ecol.* 9, 197–205. doi: 10.1007/s12080-015-0284-6

Beddington, R. (1975). Mutual interference between parasites or predators and its effect on searching efficiency. *J. Animal Ecol.* 44, 331–340. doi: 10.2307/3866

Bekkal-Brikci, F., Boushaba, K., and Arino, O. (2007). Nonlinear age structured model with cannibalism, Discrete Contin. *Dyn. Syst. Ser. B* 7, 201–218. doi: 10.3934/dcdsb.2007.7.201

Capasso, V., and Serio, G. (1978). A generalization of the Kermack-Mckendrick deterministic epidemic model. *Math. Biosci.* 42, 43–61. doi: 10.1016/0025-5564(78)90006-8

Cushing, J. M. (1998). *An Introduction to Structured Population Dynamics*. Philadelphia, PA: SIAM.

Cushing, J. M., and Saleem, M. (1982). A consumer prey model with age structure. *J. Math. Biol.* 14, 231–250.

de Ross, A. M., and Persson, L. (2013). *Population and Community Ecology of Ontogenetic Development*. Princeton: Princeton University Press.

DeAngelis, D. L., Goldstein, R. A., and O'Neil, R. V. (1975). A model for trophic interaction. *Ecology* 56, 881–892.

Elderkin, R. H. (1985). Nonlinear, globally age-dependent population models: some basic theory. *J. Math. Anal. Applic.* 108, 546–562. doi: 10.1016/0022-247x(85)90044-7

Essington, T. E., and Hansson, S. (2004). Predator-dependent functional responses and interaction strengths in a natural food web. *Can. J. Fish. Aquat. Sci.* 61, 2215–2226. doi: 10.1139/f04-146

Gourley, S. A., and Kuang, Y. (2004). A stage structured consumer-prey model and its dependence on maturation delay and death rate. *J. Math. Biol.* 49, 188–200.

Gurtin, M. E., and Maccamy, R. C. (1974). Nonlinear age-dependent population dynamics. *Arch. Ration. Mech. Anal.* 54, 281–300.

Gurtin, M. E., and MacCamy, R. C. (1979). Some simple models for nonlinear age-dependent population dynamics. *Math. Biosci.* 43, 199–211. doi: 10.1016/0025-5564(79)90049-x

Guttal, V., Romanczuk, P., Simpson, S. J., Sword, G. A., and Couzin, I. D. (2012). Cannibalism can drive the evolution of behavioural phase polyphenism in locusts. *Ecol. Lett.* 15, 1158–1166. doi: 10.1111/j.1461-0248.2012.01840.x

Han, Q., Chen, L., and Jiang, D. (2017). A note on the stationary distribution of stochastic SEIR epidemic model with saturated incidence rate. *Sci. Rep.* 7:3996.

Holt, R. D. (1977). Predation, apparent competition, and the structure of prey communities. *Theor. Popul. Biol.* 12, 197–229. doi: 10.1016/0040-5809(77)90042-9

Holt, R. D. (1984). Spatial heterogeneity, indirect interactions, and the coexistence of prey species. *Amer. Natur.* 124, 377–406. doi: 10.1086/284280

Holt, R. D., Grover, J., and Tilman, D. (1994). Simple rules to interspecific dominance in system with exploitative and apparent competition. *Amer. Natur.* 144, 741–771. doi: 10.1086/285705

Holt, R. D., and Kotler, B. P. (1987). Short-term apparent competition. *Amer.Natur.* 130, 412–430. doi: 10.1086/284718

Holt, R. D., and Lawton, J. H. (1993). Apparent competition and enemy-free space in insect host-parasitoid communities. *Amer. Natur.* 142, 623–645. doi: 10.1086/285561

Hoppensteadt, F. (1975). "Mathematical theory of population: demographics, genetics and epidemics," in *Proceedings of the CBMS-NSF Regional Conference Series in Applied Mathematics*, (Philadelphia, PA: SIAM).

Hritonenko, N., and Yatsenko, Y. (2007). The structure of optimal time- and age-dependent harvesting in the Lotka–McKendrik population model. *Math. Biosci.* 208, 48–62. doi: 10.1016/j.mbs.2006.09.008

Kolmogorov, A. N., and Fomin, S. V. (1999). *Elements of the Theory of Functions and Functional Analysis*. New York, NY: Dover Publications.

Křivan, V. (2003). Competitive coexistence caused by adaptive consumers. *Evol. Ecol. Res.* 5, 1163–1182.

Křivan, V. (2014). Competition in di- and tri-trophic food web modules. *J. Theor. Biol.* 343, 127–137. doi: 10.1016/j.jtbi.2013.11.020

Křivan, V., and Eisner, J. (2006). The effect of the Holling type II functional response on apparent competition. *Theor. Popul. Biol.* 70, 421–430. doi: 10.1016/j.tpb.2006.07.004

Levin, S. A. (1970). Community equilibria and stability: an extension of the competitive exclusion principle. *Am. Nat.* 104, 413–423. doi: 10.1086/282676

Liu, Z., Guo, C., Li, H., and Zhao, L. (2019). Analysis of a nonlinear age-structured tumor cell population model. *Nonlinear Dyn.* 98, 283–300.

Martinez, N. D. (1991). Artefacts or attributes? Effects of resolution on the little rock lake food web. *Ecol. Monogr.* 61, 367–392. doi: 10.2307/2937047

Martsenyuk, V., Klos-Witkowska, A., and Sverstiuk, A. (2018). Stability, bifurcation and transition to chaos in a model of immunosensor based on lattice differential equation with delay, Electron. *J. Qual. Theory Differ. Equ.* 27, 1–31. doi: 10.14232/ejqtde.2018.1.27

Metz, J. A. J., and Diekmann, O. (1986). *The Dynamics of Physiologically Structured Populations*. Berlin: Springer-Verlag.

Mohr, M., Barbarossa, M. V., and Kuttler, C. (2014). Consumer-prey interactions, age structures and delay equations. *Math. Model. Nat. Phenom.* 9, 92–107. doi: 10.1051/mmnp/20149107

Nakazawa, T. (2015). Introducing stage-specific spatial distribution into the Levins metapopulation model. *Sci. Rep.* 5:7871. doi: 10.1038/srep07871

Qiu, Z. P., Yu, J., and Zou, Y. (2004). The asymptotic behavior of a chemostat model with the Beddington-DeAngelis functional response. *Math. Biosci.* 187, 175–187. doi: 10.1016/j.mbs.2003.10.001

Richardson, M. L., Mitchell, R. F., Reagel, P. F., and Hanks, L. M. (2010). Causes and consequences of cannibalism in noncarnivorous insects. *Annu. Rev. Entomol.* 55, 39–53. doi: 10.1146/annurev-ento-112408-085314

Shi, J. (2013). "Absolute stability and conditional stability in general delayed differential equations," in *Advances in Interdisciplinary Mathematical Research*, ed. B. Toni (New York, NY: Springer), 117–131. doi: 10.1007/978-1-4614-6345-0_5

Smith, B. E., and Smith, L. A. (2020). Multispecies functional responses reveal reduced predation at high prey densities and varied responses among and within trophic groups. *Fish. Fish.* 21, 891–905. doi: 10.1111/faf.12468

Von Foerster, H. (1959). "Some remarks on changing populations," in *The Kinetics of Cellular Proliferation*, ed. F. Stohlman (New York, NY: Grune and Stratton), 382–407.

Vrkoc, I., and Krivan, V. (2015). Asymptotic stability of tri-trophic food chains sharing a common resource. *Math. Biosci.* 270, 90–94. doi: 10.1016/j.mbs.2015.10.005

Wang, W., and Zhao, X.-Q. (2004). An epidemic model in a patchy environment. *Math. Biosci.* 190, 97–112. doi: 10.1016/j.mbs.2002.11.001

Webb, G. F. (1985). *Theory of Nonlinear Age-Dependent Population Dynamics*. New York, NY: CRC Press.

Webb, G. F. (2008). "Population models structured by age, size and spatial position," in *Structured Population Models in Biology and Epidemiology, Lecture Notes in Mathematics*, eds P. Magal and S. Ruan (Berlin: Springer), 1–49. doi: 10.1007/978-3-540-78273-5_1

Williams, R. J., and Martinez, N. D. (2004). Stabilization of chaotic and non-permanent food-web dynamics. *Eur. Phys. J. B* 38, 297–303. doi: 10.1140/epjb/e2004-00122-1

Wootton, J. T. (1997). Estimates and tests of per capita interaction strength: diet, quantity, and impact of intertidally foraging birds. *Ecol. Monographs* 67, 45–64. doi: 10.1890/0012-9615(1997)067[0045:eatopc]2.0.co;2

Xu, R., Chaplain, M. A. J., and Davidson, F. A. (2004). Periodic solution of a Lotka–Volterra consumer–prey model with dispersion and time delays. *Appl. Math. Comput.* 148, 537–560. doi: 10.1016/s0096-3003(02)00918-9

Conflict of Interest: The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Ecology and Evolution

ORIGINAL RESEARCH
published: 22 February 2021
doi: 10.3389/fevo.2021.607177

Spatial Scaling Patterns of Functional Diversity

Saeid Alirezazadeh[1,2†], Paulo A. V. Borges[3], Pedro Cardoso[3,4], Rosalina Gabriel[3], François Rigal[3,5] and Luís Borda-de-Água[1,2]*

[1] CIBIO/InBio, Centro de Investigação em Biodiversida de e Recursos Genéticos, Universidade do Porto, Vairão, Portugal, [2] CIBIO/InBio, Centro de Investigação em Biodiversida de e Recursos Genéticos, Laboratório Associado, Instituto Superior de Agronomia, Universidade de Lisboa, Tapada da Ajuda, Lisbon, Portugal, [3] cE3c–Centre for Ecology, Evolution and Environmental Changes/Azorean Biodiversity Group, Faculdade de Ciências e Engenharia do Ambiente, Universidade dos Açores, Angra do Heroísmo, Portugal, [4] Laboratory for Integrative Biodiversity Research (LIBRe), Finnish Museum of Natural History Luomus, University of Helsinki, Helsinki, Finland, [5] CNRS—Université de Pau et des Pays de l'Adour—E2S UPPA, Institut Des Sciences Analytiques et de Physico Chimie pour L'environnement et les Materiaux UMR 5254, Pau, France

OPEN ACCESS

Edited by:
Ruscena P. Wiederholt, Everglades Foundation, United States

Reviewed by:
Buntarou Kusumoto, University of the Ryukyus, Japan
Cong Liu, Harvard University, United States

***Correspondence:**
Luís Borda-de-Água
lbagua@gmail.com

†Present address:
Saeid Alirezazadeh, C4–Cloud Computing Competence Centre (C4-UBI), UBIMedical, Covilhã, Portugal

Specialty section:
This article was submitted to Models in Ecology and Evolution, a section of the journal Frontiers in Ecology and Evolution

Received: *16 September 2020*
Accepted: *29 January 2021*
Published: *22 February 2021*

Citation:
Alirezazadeh S, Borges PAV, Cardoso P, Gabriel R, Rigal F and Borda-de-Água L (2021) Spatial Scaling Patterns of Functional Diversity. Front. Ecol. Evol. 9:607177. doi: 10.3389/fevo.2021.607177

Ecology, biogeography and conservation biology, among other disciplines, often rely on species identity, distribution and abundance to perceive and explain patterns in space and time. Yet, species are not independent units in the way they interact with their environment. Species often perform similar roles in networks and their ecosystems, and at least partial redundancy or difference of roles might explain co-existence, competitive exclusion or other patterns reflected at the community level. Therefore, considering species traits, that is, the organisms' functional properties that interact with the environment, might be of utmost importance in the study of species relative abundances. Several descriptive measures of diversity, such as the species-area relationship (SAR) and the species abundance distribution (SAD), have been used extensively to characterize the communities and as a possible window to gain insight into underlying processes shaping and maintaining biodiversity. However, if the role of species in a community is better assessed by their functional attributes, then one should also study the SAR and the SAD by using trait-based approaches, and not only taxonomic species. Here we merged species according to their similarity in a number of traits, creating functional units, and used these new units to study the equivalent patterns of the SAR and of the SAD (functional units abundance distributions - FUADs), with emphasis on their spatial scaling characteristics. This idea was tested using data on arthropods collected in Terceira island, in the Azorean archipelago. Our results showed that diversity scales differently depending on whether we use species or functional units. If what determines species communities' dynamics is their functional diversity, then our results suggest that we may need to revaluate the commonly assumed patterns of species diversity and, concomitantly, the role of the underlying processes.

Keywords: functional diversity, moments, functional species abundance distribution, functional species-area relationship, traits combination

INTRODUCTION

Species diversity, or simply biodiversity, encompasses several scales, from the genetic (phylogenetic diversity at the species level) to that of species (taxonomic diversity), populations and ecosystem functions (functional diversity) (Wilcox, 1984; United Nations Environment Programme [UNEP], 1992). One of the major goals of ecology is to describe this diversity at different spatial and temporal scales and seek for the processes shaping and maintaining it. However, species do not exist in isolation, they interact with each other and are influenced, and influence, abiotic processes in the environments where they exist. Therefore, a major challenge is to connect the description of taxonomic diversity with the processes by which species interact with their environments. Functional diversity has the potential to link these two views of a community (Asner et al., 2017). Our main purpose here is to study patterns of functional diversity using tools commonly used to assess patterns of species taxonomic diversity with an emphasis on its scaling properties.

By functional diversity we mean: "the values and range in the values, for the species present in an ecosystem, of those organismal traits that influence one or more aspects of the functioning of an ecosystem" (Tilman, 2001). As (Petchey and Gaston, 2006) pointed out, this is a relatively narrower definition than that found in original studies, and that emphasizes the importance of "measuring functional trait diversity, where functional traits are components of an organism's phenotype that influence ecosystem level processes" (Petchey and Gaston, 2006). An important aspect highlighted by functional diversity is that species often overlap in their traits, leading communities with different species (taxonomic) composition to having a similar functional composition (Cardoso et al., 2014). Therefore, because the emphasis is no longer on taxonomic differences, as recently highlighted by Wong et al. (2019).

The main novelty of our work is its emphasis on the spatial scaling attributes of functional diversity by looking at the equivalent to the species-area relationship (hereafter SAR) that we will call the functional units-area relationship (hereafter FUAR), and the equivalent to the species abundance distribution (hereafter SAD), that we will call the functional units abundance distribution (hereafter FUAD) and where we adopted the following definition of functional units: a set of organisms sharing the same unique combination of attributes. Concerning the FUAD, as we elaborate later, our purpose is not to discuss which distribution gives the best fit at one spatial scale, but how the distributions change as a function of scale. To do this we will use the (raw) moments of the distributions. This is done with an aim to providing new tools and, importantly, to discerning new quantitative patterns associated with functional diversity.

One the most studied patterns of species taxonomic richness is the species-area relationship (e.g., Rosenzweig, 1995). The species-area relationship describes how the number of species changes as a function of area size. However, as shown by Scheiner (2003), there are different types of SARs depending on how data are collected and presented. For example, some SARs describe how the number of species change as a function of the size of islands in an archipelago (Whittaker et al., 2014), while others describe how the number of species accumulates as sampled (nested or non-nested) areas increase (Matthews et al., 2016); in this study we look at the SAR of non-nested areas. An important aspect of SAR studies is that they implicitly assume a scaling approach and try to identify a pattern. Which function best describes this scaling pattern has been a topic of contention among ecologists, as reviewed by Rosenzweig (1995). One of the most popular functions is a power law of the form, $S = cA^z$, where S is the number of species, A the area size, and c and z constants to be determined. In our analyses we will compute the equivalent to the SAR but using species functional units instead of species and will compare functional units-area relationships, the FUAR, and taxonomic-based SAR curves of the same dataset. Smith et al. (2013) have already introduced the concept of functional-diversity-area relationship, to explore the scaling properties of individual traits alone, and when extending the concept to multiple traits they used the concept of convex hull model introduced by Cornwell et al. (2006), while here we group taxonomic species based on their traits to form functional units.

Although species richness and how it changes as a function of sample or area size are important attributes of a community, they do not provide information on the species relative abundances. There are several ways to account for species abundance, but probably the most intuitive, and the one we will use here, is the number of individuals. Typically, species relative abundances are depicted using a histogram, called a species abundance distribution, relating the number of species (the y-axis) with the number of individuals (the x-axis), with the latter usually expressed on a logarithmic scale of base 2 (e.g., McGill et al., 2007); logarithms are used in order to accommodate the wide range of abundances usually encountered in a community sample. In other words, a SAD answers the question of how many species exist within a given range of number of individuals. In line with our previous delineated approach of functional units, we will analyze SADs based on species functional units, the functional units abundance distributions, or FUAD, and, again, we will compare these distributions with the taxonomic SADs of the same dataset.

In contrast with the SAR, that looks at the number of species at different scales, the SAD is often computed and studied at one single scale (e.g., McGill et al., 2007). However, Borda-de-Água et al. (2012, 2017) argued that SADs should also be studied as a function of scale, because histograms (i.e., the distributions) change considerably depending on the size of a sample (or area size). The visual description by Preston (1948) of a veil-line progressively revealing more classes of the distribution as sample size increases, provides a good visual representation of the way the distributions change: for small sample sizes SADs are usually monotonically decreasing functions with the maximum occurring for the singleton class, while for larger sample sizes the distributions become bell-shaped. Once we realize that the SADs change as a function of sample size, an important task is to characterize that change, that is, the scaling properties of the distributions. Borda-de-Água et al. (2012, 2017) used the raw moments of the distribution to describe the scaling of the SADs (see section "Materials and Methods" for the definition of raw moments). The characterization of the distributions at

different scales based on their moments, enables the SADs to be extrapolated to larger areas, whose sizes pose sampling problems because of economical or other practical reasons. Here, and for the same reasons, we will study how SAD based on functional groups instead of species (i.e., functional units abundance distributions) change as a function of scale and will use the raw moments to describe their scaling properties.

In summary, the rationale of our study is the following. If functional units provide a more accurate description of how species interact among them and the environment, then the above scaling patterns of species richness and relative abundance should be reformulated in terms of the species' functional units and not of the species taxonomical identities. As we will see, using data on arthropods collected in the Azores archipelago, describing the scaling proprieties of the diversity of a community using species' functional units leads to quantitatively different results from those obtained when using taxonomic species.

MATERIALS AND METHODS

Study Sites and Sampling Strategy

We used data on arthropod species collected in Terceira island, Azores (**Figure 1**). The Azores archipelago (37°–40° N; 25°–31° W) consists of nine islands and it is one of the world's most isolated archipelagos. Terceira is the island for which we have the most data (see Borda-de-Água et al., 2017), hence we will focus our analysis on this island. Terceira is the third largest island of the archipelago with an area of 402 km^2 and it is currently estimated to be 0.40 million years old (Calvert et al., 2006). We gathered data in native forest reserves from 39 transects, each 150 m long and 5 m wide. We sampled all transects with the same sampling effort using standardized methods; for more details see Borges et al. (2005) or Rigal et al. (2018). We have identified 181 arthropod species for Terceira island, and for each species we have also recorded the number of individuals sampled.

Species Identification

The arthropods collected were identified in the laboratory at the species level for the taxa Araneae, Opiliones, Pseudoscorpiones, Diplopoda, Chilopoda and Insects (excluding Collembola, Diplura, Diptera and Hymenoptera). Taxonomic identification was performed in two steps: (i) trained parataxonomists sorted samples into morphospecies (i.e., recognizable taxonomic units, sensu Oliver and Beattie, 1996) using a non-complete reference collection; (ii) experienced taxonomists assisted in the identification of the morphospecies. All species were classified as indigenous or exotic. Indigenous species comprise Azorean endemics and other native non-endemics. Exotic species are those considered to have colonized via human mediation, many of which have a cosmopolitan distribution (Borges et al., 2010). As in some of our previous studies (e.g., Borges et al., 2005; Gaspar et al., 2008), we dealt with unidentified morphospecies as follows. When other species in the same genus, subfamily or family were present in the archipelago and all belonged to the same colonization category (according to Borges et al., 2010), the unknown morphospecies were classified similarly. If no information was available, we assumed the species to be native since exotics are usually widespread and easier to identify (Borges et al., 2010).

Functional Data

Different species interact with the environment and other organisms in different ways. Yet, standardization across taxa is not only possible but provides important insights (e.g., Chichorro et al., 2019, Chichorro et al., 2020). Often a trait-based approach has to be adapted to accommodate such differences, by using different proxies that are equivalent for contrasting taxa. In our case, as we only studied arthropods, which share many characteristics, standardization was relatively easy. In fact, in this last decade recent efforts have been made to develop trait database for arthropods as a whole (See Schweiger et al., 2005; Simons et al., 2016; Rigal et al., 2018). Functional traits were then selected based on the known life-history of the taxa analyzed as representing different ways the organisms can be affected in their probability of survival and hence abundance in the studied forests. This way we are implicitly studying mechanisms that affect species abundances in the native forests of the Azores and how such abundances change depending on the spatial scale.

For all arthropod species, we collated data for five traits namely, body size, dispersal abilities, trophic level, microhabitat and origin. Body size was measured for the individuals sampled in this study and was coded using ordered size classes delimited by the intervals (0, 0.5), (0.5, 1.5) ... (40.5, 41.5). Both dispersal abilities and trophic level were collected from an extensive literature search, including manuscripts with the first descriptions of the species, first species records for the Azores, brief notes and ecological studies. Information was also obtained from experts who have identified the specimens or from experts of a given taxonomic group when information for a particular species was not available. Functional information was assigned to each species according to their adult characteristics, except for Lepidoptera, where traits were assigned with reference to the larval stage. For the unidentified morphospecies, we assigned functional traits of the nearest taxonomic resolution (genus, family), except for body size which was measured directly from the individuals.

Specifically, for the trait "dispersal ability," the species were categorized into high and low dispersal classes based on ecological attributes and morphological characteristics. This could be for instance, the presence of active wings for beetles (Coleoptera) and bugs (Hemiptera), ballooning for spiders and evidence of flying ability for endemics and general natural history guides for the other species. To be considered as a good disperser, a species has to be able to disperse between fragments of native forest and surpass the current matrix of man-made habitats (e.g., pastures). For origin we identified three classes: "endemic," "introduced," or "native non-endemic." One can argue that the "origin" of a species is not a relevant trait to how a species functions within a community. We use this category because introduced species may have (new) traits that are not present among the native species and, for a similar reason, we identify endemic species because they may have unique traits adapted to

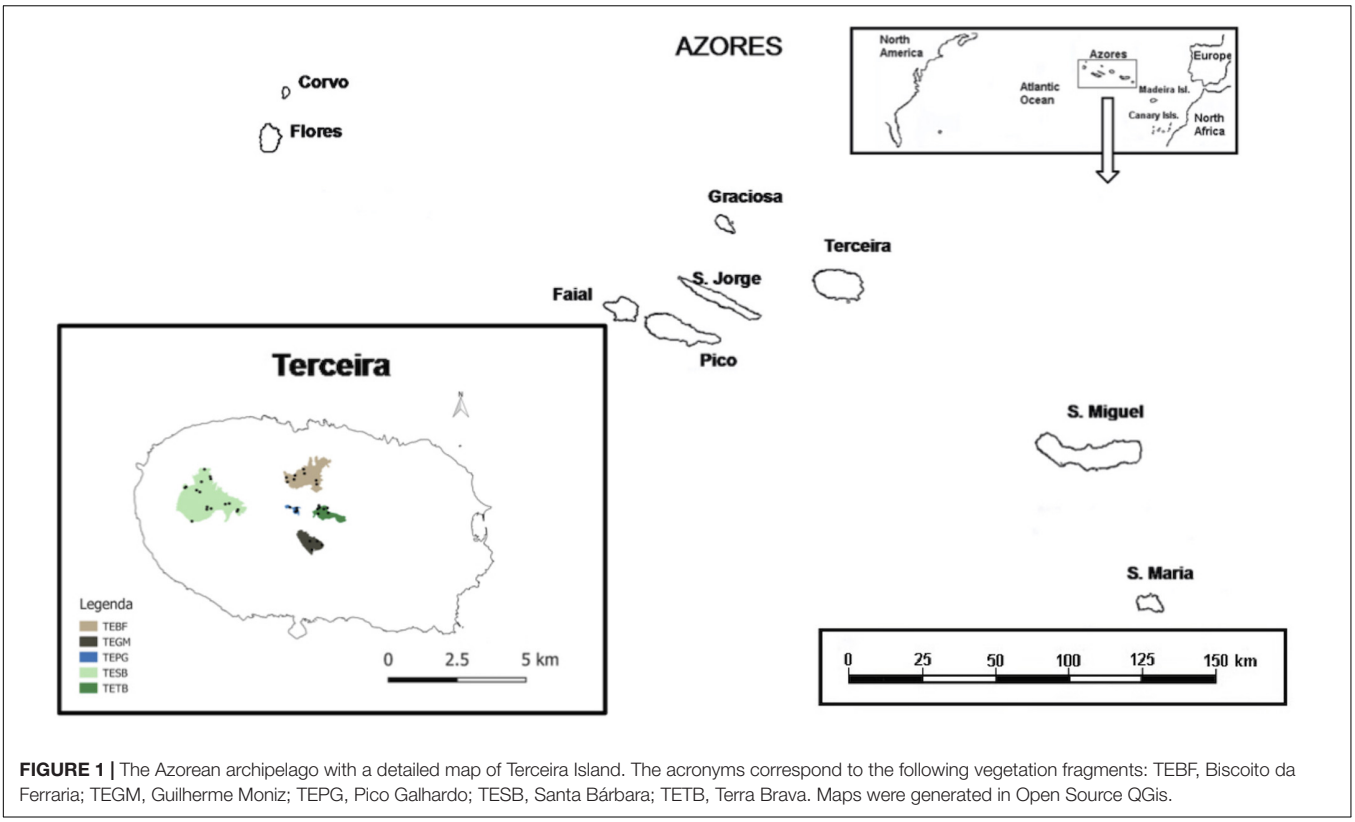


FIGURE 1 | The Azorean archipelago with a detailed map of Terceira Island. The acronyms correspond to the following vegetation fragments: TEBF, Biscoito da Ferraria; TEGM, Guilherme Moniz; TEPG, Pico Galhardo; TESB, Santa Bárbara; TETB, Terra Brava. Maps were generated in Open Source QGis.

particular conditions of the archipelago or, in the case of single island endemics, of an island (see also Rigal et al., 2018).

Thus, each species was categorized based on five traits:

- *Average body size* classes;
- *Dispersal ability*, with two classes: "high" or "low";
- *Trophic level*, with two classes: "herbivore" or "predator";
- *Microhabitat*, with two classes: "canopy" or "pitfall";
- *Origin*, with three classes: "endemic," "introduced," or "native non-endemic."

For each species we determined the combination arising from the five above traits. For instance, a species could have the following combination (average body size = 3.5, dispersal = "low," location = "canopy," origin = "native," trophic level = "herbivore"). From a functional perspective, we think of all species with the same traits' combination as being just one functional unit. Thus, we identified all species with the same traits' combination and added their number of individuals to form a functional unit; obviously, the number of functional units is necessarily smaller or equal to the number of species.

Statistical Procedures

In order to obtain the species and functional units-area relationships (FUAR) we started from a random transect and added the species of the nearest transect, and then of the next nearest transect until reaching the desired number of transects. By adding the nearest transect we better replicate a spatial sampling process that avoids overinflating the effect of beta diversity that would happen if sites were added, irrespectively, of spatial distance. The final SAR and FUAR curves are the average of the curves obtained from different starting transects. This procedure leads to one of the six types of species-area relationship described by Scheiner (2003), and it has the advantage of retaining information of the relative distance between the transects.

We construct histograms of functional units abundance distributions (FUAD) as the classical taxonomic species abundance distributions (SADs). **Figure 2** shows some examples. On the *y*-axis we plot the number of functional units and on the *x*-axis the number of individuals on a logarithmic scale of base 2, with the classes of the number of individuals defined as 1, 2–3, 4–7, *et seq.* When a species abundance distribution resulted from the accumulation of data from two or more plots, we always added the data from the nearest transect, as we did for the SARs. Also, as before, the final SAD and FUAD curves are the average of the distributions obtained from different starting transects.

In order to characterize the scaling properties of the FUADs, we use their raw moments (occasionally, when there is not risk of confusion, we call them simply "moments") as we have done previously for SADs (Borda-de-Água et al., 2012, 2017). If a community is made of S_{FG} functional units and $x_j = \log_2(X_j)$, where X_j is the number of individuals of the j functional unit, then the raw moment of order n, M_n, is calculated from

$$M_n = \frac{1}{S_{FG}} \sum_{j=1}^{S_{FG}} x_j^n; \quad (1)$$

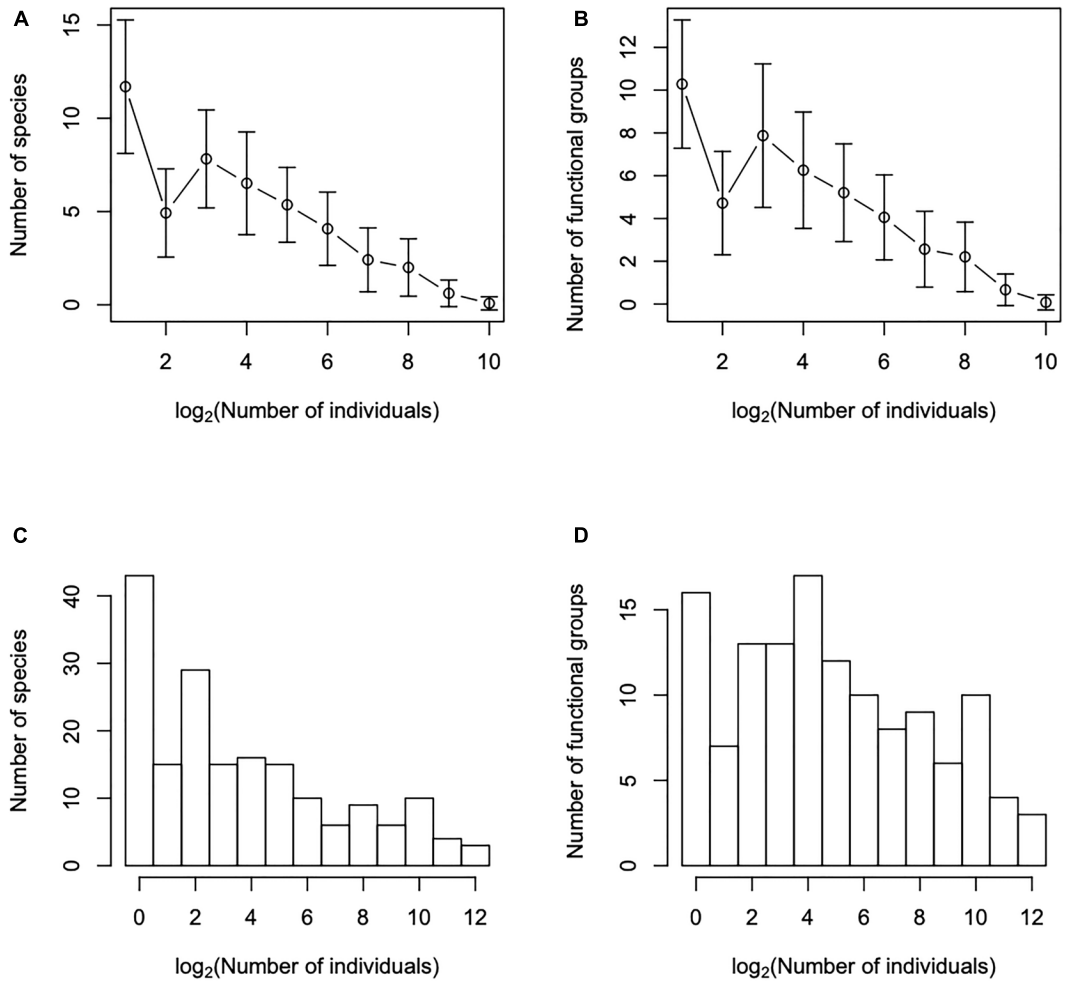


FIGURE 2 | Species abundance distributions, plots **(A,C)**, and functional units abundance distributions, plots **(B,D)**; functional units are the functional entities obtained by aggregating species with the same trait combinations. Plots **(A,B)** are obtained after averaging the abundance distributions of all transects and the bars correspond to two standard deviations, while plots **(C,D)** are obtained after adding all transects. Notice that the shape of the distributions changed when we consider the average of all transects and when we all add all transects.

notice that M_0 is equal to 1, and M_1 corresponds to the average. Other, more familiar moments, such as the variance (the second central moment), can be obtained as combinations of the raw moments.

The Extrapolation Procedure

Here we used an improved method to extrapolate the distributions based on that described in Borda-de-Água et al. (2012, 2017). The basic idea consists of estimating the moments, using formula (1), up to a given order n, plotting these values and fitting the curves of each order moment and then, using these curves, estimating the moment values for the new areas. In order to reproduce the abundance distribution based on the extrapolated moments, we used an improved version (Mukundan, 2004) of the scale discrete Tchebichef moments and polynomials first introduced by Mukundan et al. (2001). We refer the details to the **Supplementary Material**.

RESULTS

Based on the traits described in the section "Materials and Methods," there are 648 possible trait combinations. Since, there are 181 arthropod species identified in the Terceira arthropod dataset, there are, at the maximum, 181 observable combinations. From this we only observe 117 trait combinations, or 117 functional units. Notice that from all the classes ascribed to the average body size, species only occurred in 27 (from classes 0 to 22 and then in classes 25, 33, 37, 40, and 41).

We start by presenting the SAR and the FUAR (**Figure 3**); because all transects have the same dimensions, we equate the number of transects with area size. The obvious difference between the two curves is that the FUAR in a log-log plot starts exhibits a plateau, showing clear signs of starting to "saturate," while the traditional species-area curve reveals the typical (almost) straight line in a log-log plot, that is the number of taxonomic species keeps increasing until the very last transect

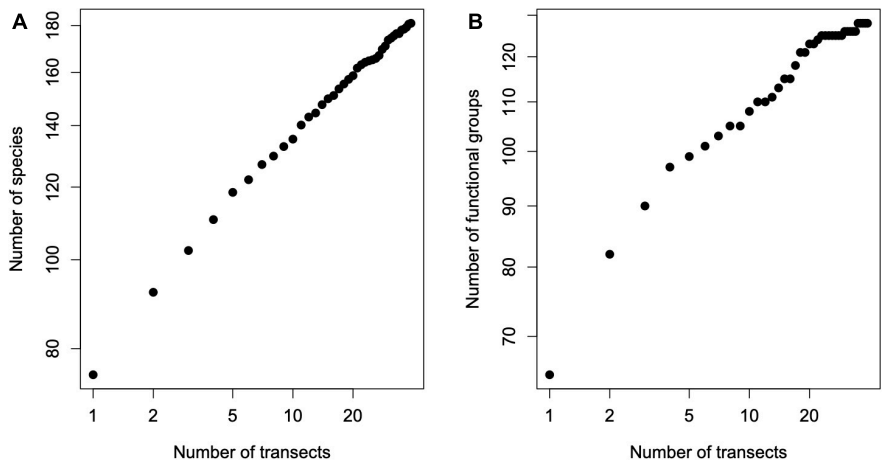


FIGURE 3 | The species-area relationship (SAR), plot **(A)**, and the functional units area relationship (FUAR), plot **(B)**. Recall that all transects have the same dimensions, thus we equate the number of transects with area size. Both plots are on double logarithmic scales. Notice that the SAR exhibits the typical linear relationship when plotted on double logarithmic scales (i.e., it is approximately a power law relationship) while the FUAR shows some signs of reaching a plateau.

considered. In a sense, given that the number of functional units is a subset of the number of species, one could expect this to happen, but notice that the possible plateau only starts to appear at a large number of transects.

We showed in **Figure 2** the SADs and FUADs: plots a and c were obtained after averaging the distributions of individual transects and plots b and d when we added the data from all the transects. Notice that at the level of one transect (plots A and B) the SAD and FUAD are very similar, being almost monotonically decreasing curves. From here on, we will describe curves that have such a monotonically decreasing shape as "logseries-like," but we do not ascertain that the logseries distribution would provide the best fit, it is, hence, just a terminology to describe the general shape of the curve. On the other hand, when all transects are added together they are very different (plots C and D). Specifically, while the SAD of the 39 transects retains a logseries-like shape, the FUAD has its maximum at an intermediate class and, although it still has a large number of singletons, overall it has a bell shape. In other words, the FUAD evolves faster from a logseries-like distribution to a bell-shaped distribution than the species-based SAD. From here on, we will describe curves with a bell-shape distribution, thus with the maximum at intermediate classes, as "lognormal-like" without implying that the lognormal would provide the best fit; as with the terminology for the "logseries-like" distributions, it just a general description of the shape of the distribution.

We estimated the moments of the FUAD using Eq. 1, **Figure 4**. The overall behavior of the moments of the functional units as a function of the number of transects is not very different from those of taxonomical diversity (Borda-de-Água et al., 2017), that is, above a certain number of transects (or area) the moments are approximately linear in log-log scales. Using the procedure described in the **Supplementary Material**, we extrapolate these moments for a larger number of transects and used them to reconstruct the FUAD (**Figure 5**); for sake of comparison we applied the same procedure to the traditional taxonomical SAD. The SAD evolves from an almost monotonically decreasing function to a curve with a clear peak when we extrapolate the distribution for 70 transects (the green line in **Figure 5A**). However, the appearance of a peak for the SAD is slow compared to that of the FUAD, whose peak for intermediate classes is already clear in the FUAD for 39 transects. When we increase the number of transects the maximum peak for the extrapolated curves shifts to higher abundance classes, and there is a reduction in the number of singletons.

DISCUSSION

Under the premise that functional units have a more meaningful correspondence with the underlying processes of maintenance and generation of species diversity (Tilman, 2001), we applied several procedures commonly used on studies of communities but using, instead of taxonomical species, functional units, that is, entities resulting from aggregating species with identical traits. Common criticisms often leveled at studies of functional diversity can also be raised here, namely, that the results depend on the choice of the traits, on the way traits are coded (e.g., number of categories identified) and on the scales used to identify traits (e.g., Ricotta, 2005). Undoubtedly, we would have obtained quantitatively different results if we had chosen a different set of traits or a different scale for the average body size, for instance. However, we describe this work as a first exploratory attempt to determine how diversity scaling changes depending on whether we use species or functional units.

These exploratory results showed that the scaling behavior of the species and of the functional units is quantitatively different. For instance, the FUAR relationship does not have a linear relation when plotted on log-log scales and shows a clear plateau, and contrary to the SAR, when area increases the FUAD develops

a peak earlier than the corresponding SAD. Concerning the FUAR we acknowledge that its shape is a function of the number of traits chosen. For instance, if we had chosen fewer traits, the plateau would have been reached even earlier, because there would have been fewer functional units. Although it is outside the scope of the work, a pattern that is worth analysing in the future is how the FUAR shape changes in a community as a function of the number of traits used. Furthermore, to overcome the problem of trait dimension in driving FUAD patterns, we could also use a null model approach to test whether FUAR was simply a product of species richness (see Whittaker et al., 2014, for example).

To understand the faster transition in a FUAD, it is important to understand the transitions in a typical SAD. Using computer simulations, Borda-de-Água et al. (2007) conclude that the SADs undergo three regimes when area increases (their Fig. 17.9). For very small sample sizes, the distributions are monotonically decreasing curves, thus the maximum occurs for the singleton class, and they can be described as logseries-like distributions. When the sample size increases the number of singleton species, and those of other rare classes, decreases and the distributions develop a peak for intermediate abundance classes, having a lognormal like shape, although with some singletons and other rare species still being present. This shift in the shape of the species abundance distribution can be observed in real data, such as the data on tropical tree species of a 50 ha plot in Barro Colorado Island, Panama; see for example Borda-de-Água et al. (2012; **Figure 1**). The explanation for such a transition is the following: (i) when the number of individuals is small and several species are present, the number of individuals per species is small and most species are rare; (ii) once more individuals are collected some, if not most, belong to species already present, thus, species accumulate more individuals and they move away from the rarity classes toward those classes corresponding to more abundant species, therefore, overall, intermediate abundant classes start having more species and the curve becomes more bell shaped. When sample size increases further, the distribution becomes again monotonically decreasing, resembling again a logseries distribution. The explanation for such transition is the following: most species are rare for very small sample sizes, but only a few are really rare in the (meta)community, therefore, those that are not truly rare move to more abundant classes when sample sizes increase, but those that are truly rare remain in the classes of the rare species. For very large samples, we collect species from different regions of the community, each region containing rare species, thus at the level of the entire community there are a large number of rare species that are added to the rare abundance classes. The presence of a large number of rare species can also be the result of the temporal evolution of community. In fact, hyperdominance (i.e., few species having most of the individuals) may happen due to few winners dominating the community and the addition of mostly rare species to the regional pool with speciation operating on long timescales (McGill et al., 2019). As new species are added at a regional scale through evolution, hyper-rarity and hyperdominance are generated. Consequently, in time, the number of such (truly) rare species can be very high and the SAD is a monotonically decreasing function. We suggest that the latter transition from a lognormal to a logseries is difficult to observe in real data because it requires large sample sizes. However, in addition to the evidence from simulations (Borda-de-Água et al., 2007), theoretical results based on Neutral theory (Hubbell, 2001) and data on Amazon tree genera do suggest that the SAD for very large communities is logseries-like (Ter Steege et al., 2006; Hubbell et al., 2008; Hubbell, 2013; Ter Steege et al., 2013).

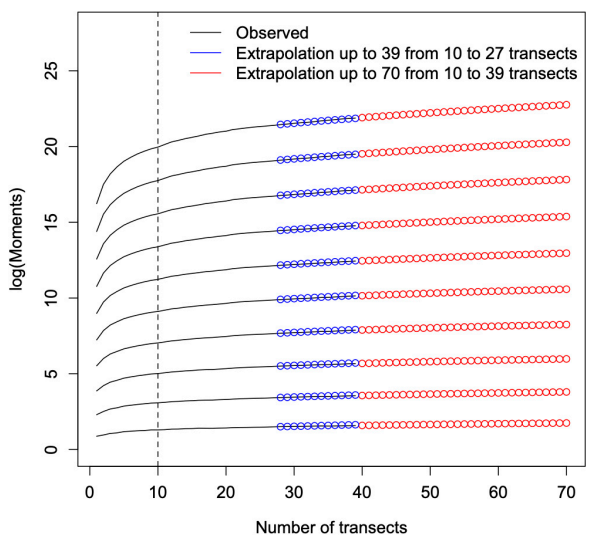


FIGURE 4 | The raw moments up to order 10 as a function of the number of transects. The full lines are the values calculated from the data, the blue circles are the values obtained by extrapolating the observed values from 10 to 27 transects and are, then, used to assess the method of extrapolation, and, finally, the red circles are the values obtained by extrapolating the observed values from 10 to 39 transects. The dashed vertical line corresponds to the threshold above which data are used to extrapolate the abundance distributions.

The previous description for the three regimes observed for SADs may not apply when we look at the abundance distributions of functional units. The first transition from the logseries-like to the lognormal-like may still be present, as revealed by **Figure 5**, though it occurs faster than the corresponding SAD. This is not surprising given that functional units are collections of species, thus a faster increase in the number of intermediate abundance functional units and a decrease in the number of rare functional units is not surprising. However, we speculate that the transition from lognormal-like to logseries-like for large scales may never occur for FUADs, because, though we expect more rare species to enter the sample, these may have the same set of traits of other species already present in the community, thus the number of functional units in the rare classes will not increase. In fact, we expect the number of functional units in rare abundance classes to decrease. In other words, although the number of species is likely to increase when the number of sampled individuals increase, the number of trait combinations, and thus that of functional units (but not their abundance as measured by the number of individuals), is likely to remain approximately

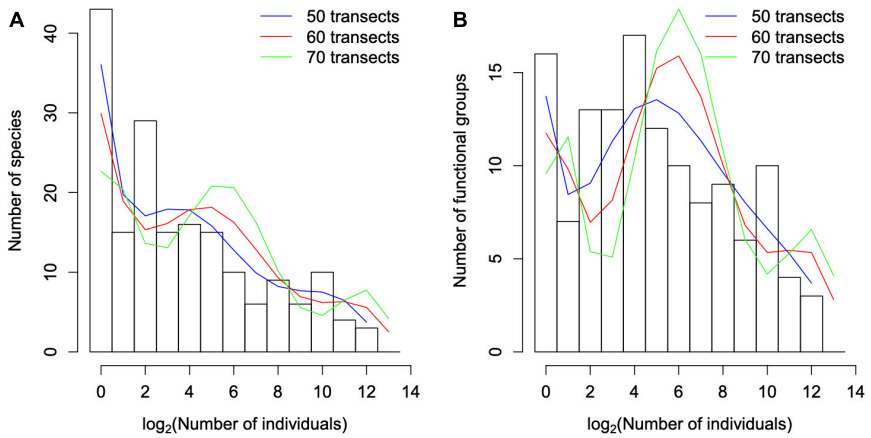


FIGURE 5 | The histograms correspond to the species **(A)** and functional units **(B)** abundance distributions of the arthropod species for Terceira's 39 transects, and as a reference, the extrapolations to 50, 60, and 70 transects are plotted in the blue, red, and green lines, respectively. Notice, in both cases, the reduction in the number of singletons when the number of transects increases. For the species abundance distribution there is a clear development of a peak for intermediate classes, while for the functional units such a peak is already present for 39 transects, but it increases when the number of transects increases.

constant. If this is indeed the case, and if the functioning of the community is (partially) determined by the relative abundance of the functional units, then communities at large scales are mainly characterized by lognormal-like distributions and not logseries-like distributions. This has implications for the investigation of the mechanisms and patterns that govern communities. The important mechanisms shaping a community may lead to FUADs that are inherently lognormal-like.

We finalize with an observation that points to future work. Once we group species based on their similarity, it is implicit that some species may be redundant. Such a view, however, should be taken carefully because we only identified a subset of all possible traits. It is possible that some non-identified traits may correspond to a species role in the ecosystem that is irreplaceable, and whose absence could lead to profound transformations in the community, including the loss of other species. In this respect it is worth noting that several works have shown that communities exhibit approximately the same number of species, *S*, and individuals, *N*, over time, that is, there is a regulation of temporal trends of species diversity (Brown et al., 2001; Gotelli et al., 2017; Blowes et al., 2019). This is a remarkable result, although the mechanisms of such regulation are not (fully) understood. One question that remains is whether such regulatory trends also occur if temporal community diversity is analyzed in terms of functional diversity. In fact, some studies have shown that functional diversity changes over time in some communities (e.g., Mendez et al., 2012; White et al., 2018; Dolezal et al., 2020). However, if regulation is even stronger when we consider functional units, this suggests that one possible way to identify functional units, would be to consider several trait combinations, and identify which trait combinations lead to the smallest variation of the number of functional units and their abundances over time. This approach could provide an objective choice of the traits and respective functional entities. Finally, we define functional units using categorical traits but some traits are usually recorded as continuous variables (e.g., morphological traits), therefore requiring alternative measures of functional units such as clustering methods (Villéger et al., 2012).

DATA AVAILABILITY STATEMENT

Publicly available datasets were analyzed in this study. This data can be found here: https://doi.org/10.3897/BDJ.4.e10948.

AUTHOR CONTRIBUTIONS

LB-Á, PC, and PB designed the original research. SA did most of the analytic work. All the authors contributed to writing up the article.

FUNDING

SA was financed by the Fundação para a Ciência e a Tecnologia project PTDC/BIA-BIC/5558/2014. LB-Á was financed by the Portuguese national funds through FCT—Fundação para a Ciência e a Tecnologia, I.P., under the Norma Transitória—DL57/2016/CP1440/CT0022.

ACKNOWLEDGMENTS

We thank H. M. Pereira for discussions related to functional diversity metrics.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fevo.2021.607177/full#supplementary-material

REFERENCES

Asner, G. P., Martin, R. E., Knapp, D. E., Tupayachi, R., Anderson, C. B., Sinca, F., et al. (2017). Airborne laser-guided imaging spectroscopy to map forest trait diversity and guide conservation. *Science* 355, 385–389. doi: 10.1126/science.aaj1987

Blowes, S. A., Supp, S. R., Antão, L. H., Bates, A., Bruelheide, H., Chase, J. M., et al. (2019). The geography of biodiversity change in marine and terrestrial assemblages. *Science* 366, 339–345.

Borda-de-Água, L., Borges, P. A., Hubbell, S. P., and Pereira, H. M. (2012). Spatial scaling of species abundance distributions. *Ecography* 35, 549–556. doi: 10.1111/j.1600-0587.2011.07128.x

Borda-de-Água, L., Hubbell, S. P., and He, F. (2007). *Scaling Biodiversity Under Neutrality. Scaling Biodiversity*. (Cambridge: Cambridge University Press), 347–375.

Borda-de-Água, L., Whittaker, R. J., Cardoso, P., Rigal, F., Santos, A. M., Amorim, I. R., et al. (2017). Dispersal ability determines the scaling properties of species abundance distributions: a case study using arthropods from the Azores. *Sci. Rep.* 7:3899.

Borges, P. A., Aguiar, C., Amaral, J., Amorim, I. R., André, G., Arraiol, A., et al. (2005). Ranking protected areas in the Azores using standardised sampling of soil epigean arthropods. *Biodivers. Conserv.* 14, 2029–2060. doi: 10.1007/s10531-004-4283-y

Borges, P. A., Vieira, V., Amorim, I. R., Bicudo, N., Fritzén, N., Gaspar, C., et al. (2010). "List of arthropods (Arthropoda)," in *A List of the Terrestrial and Marine Biota from the Azores*, eds P. A. V. Borges, A. Costa, R. Cunha, R. Gabriel, V. A. Gonçalves, F. Martins, et al. (Cascais: Princípia), 179–246.

Brown, J. H., Ernest, S. M., Parody, J. M., and Haskell, J. P. (2001). Regulation of diversity: maintenance of species richness in changing environments. *Oecologia* 126, 321–332. doi: 10.1007/s004420000536

Calvert, A. T., Moore, R. B., McGeehin, J. P., and da Silva, A. M. R. (2006). Volcanic history and 40Ar/39Ar and 14C geochronology of Terceira Island, Azores, Portugal. *J. Volcanol. Geotherm. Res.* 156, 103–115. doi: 10.1016/j.jvolgeores.2006.03.016

Cardoso, P., Rigal, F., Borges, P. A., and Carvalho, J. C. (2014). A new frontier in biodiversity inventory: a proposal for estimators of phylogenetic and functional diversity. *Methods Ecol. Evol.* 5, 452–461. doi: 10.1111/2041-210x.12173

Chichorro, F., Juslén, A., and Cardoso, P. (2019). A review of the relation between species traits and extinction risk. *Biol. Conserv.* 237, 220–229. doi: 10.1016/j.biocon.2019.07.001

Chichorro, F., Urbano, F., Teixeira, D., Väre, H., Pinto, T., Brummitt, N., et al. (2020). Species traits predict extinction risk across the tree of life. *bioRxiv* [Preprint]. doi: 10.1101/2020.07.01.183053

Cornwell, W. K., Schwilk, D. W., and Ackerly, D. D. (2006). A trait-based test for habitat filtering: convex hull volume. *Ecology* 87, 1465–1471. doi: 10.1890/0012-9658(2006)87[1465:attfhf]2.0.co;2

Dolezal, J., Fibich, P., Altman, J., Leps, J., Uemura, S., Takahashi, K., et al. (2020). Determinants of ecosystem stability in a diverse temperate forest. *Oikos* 129, 1692–1703. doi: 10.1111/oik.07379

Gaspar, C., Borges, P. A. V., and Gaston, K. J. (2008). Diversity and distribution of arthropods in native forests of the Azores archipelago. *Arquipélago Life Mar. Sci.* 25, 1–30.

Gotelli, N. J., Shimadzu, H., Dornelas, M., McGill, B., Moyes, F., and Magurran, A. E. (2017). Community-level regulation of temporal trends in biodiversity. *Sci. Adv.* 3:e1700315. doi: 10.1126/sciadv.1700315

Hubbell, S. P. (2001). *The Unified Neutral Theory of Biodiversity and Biogeography (MPB-32)*. Princeton, NJ: Princeton University Press.

Hubbell, S. P. (2013). Tropical rain forest conservation and the twin challenges of diversity and rarity. *Ecol. Evol.* 3, 3263–3274.

Hubbell, S. P., He, F., Condit, R., Borda-de-Água, L., Kellner, J., and ter Steege, H. (2008). How many tree species are there in the Amazon and how many of them will go extinct? *Proc. Natl. Acad. Sci. U.S.A.* 105(Suppl. 1), 11498–11504. doi: 10.1073/pnas.0801915105

Matthews, T. J., Triantis, K. A., Rigal, F., Borregaard, M. K., Guilhaumon, F., and Whittaker, R. J. (2016). Island species–area relationships and species accumulation curves are not equivalent: an analysis of habitat island datasets. *Glob. Ecol. Biogeogr.* 25, 607–618. doi: 10.1111/geb.12439

McGill, B. J., Chase, J., Hortal, J., Overcast, I., Rominger, A., Rosindell, J., et al., (2019). Unifying Macroecology and Macroevolution to answer fundamental questions about biodiversity. *Glob. Ecol. Biogeogr.* 28, 1925–1936. doi: 10.1111/geb.13020

McGill, B. J., Etienne, R. S., Gray, J. S., Alonso, D., Anderson, M. J., Benecha, H. K., et al. (2007). Species abundance distributions: moving beyond single prediction theories to integration within an ecological framework. *Ecol. Lett.* 10, 995–1015. doi: 10.1111/j.1461-0248.2007.01094.x

Mendez, V., Gill, J. A., Burton, N. H., Austin, G. E., Petchey, O. L., and Davies, R. G. (2012). Functional diversity across space and time: trends in wader communities on British estuaries. *Divers. Distrib.* 18, 356–365. doi: 10.1111/j.1472-4642.2011.00868.x

Mukundan, R. (2004). Some computational aspects of discrete orthonormal moments. *IEEE Trans. Image Process.* 13, 1055–1059. doi: 10.1109/tip.2004.828430

Mukundan, R., Ong, S. H., and Lee, P. A. (2001). Image analysis by Tchebichef moments. *IEEE Trans. Image Process.* 10, 1357–1364. doi: 10.1109/83.941859

Oliver, I., and Beattie, A. J. (1996). Invertebrate morphospecies as surrogates for species: a case study. *Conserv. Biol.* 10, 99–109. doi: 10.1046/j.1523-1739.1996.10010099.x

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

Preston, F. W. (1948). The commonness, and rarity, of species. *Ecology* 29, 254–283. doi: 10.2307/1930989

Ricotta, C. (2005). Through the jungle of biological diversity. *Acta Biotheor.* 53, 29–38. doi: 10.1007/s10441-005-7001-6

Rigal, F., Cardoso, P., Lobo, J. M., Triantis, K. A., Whittaker, R. J., Amorim, I. R., et al. (2018). Functional traits of indigenous and exotic ground-dwelling arthropods show contrasting responses to land-use change in an oceanic island, Terceira, Azores. *Divers. Distrib.* 24, 36–47. doi: 10.1111/ddi.12655

Rosenzweig, M. L. (1995). *Species Diversity in Space and Time*. Cambridge: Cambridge University Press.

Scheiner, S. M. (2003). Six types of species-area curves. *Glob. Ecol. Biogeogr.* 12, 441–447. doi: 10.1046/j.1466-822x.2003.00061.x

Schweiger, O., Maelfait, J. P., Wingerden, W. V., Hendrickx, F., Billeter, R., Speelmans, M., et al. (2005). Quantifying the impact of environmental factors on arthropod communities in agricultural landscapes across organizational levels and spatial scales. *J. Appl. Ecol.* 42, 1129–1139. doi: 10.1111/j.1365-2664.2005.01085.x

Simons, N. K., Weisser, W. W., and Gossner, M. M. (2016). Multi-taxa approach shows consistent shifts in arthropod functional traits along grassland land-use intensity gradient. *Ecology* 97, 754–764.

Smith, A. B., Sandel, B., Kraft, N. J., and Carey, S. (2013). Characterizing scale-dependent community assembly using the functional-diversity–area relationship. *Ecology* 94, 2392–2402. doi: 10.1890/12-2109.1

Ter Steege, H., Pitman, N. C., Phillips, O. L., Chave, J., Sabatier, D., Duque, A., et al. (2006). Continental-scale patterns of canopy tree composition and function across Amazonia. *Nature* 443, 444–447. doi: 10.1038/nature05134

Ter Steege, H., Pitman, N. C., Sabatier, D., Baraloto, C., Salomão, R. P., Guevara, J. E., et al. (2013). Hyperdominance in the Amazonian tree flora. *Science* 342:1243092.

Tilman, D. (2001). Functional diversity. *Encycl. Biodivers.* 3, 109–120.

United Nations Environment Programme [UNEP], (1992). *Convention on Biological Diversity*. Nairobi: United Nations Environment Programme.

Villéger, S., Miranda, J. R., Hernandez, D. F., and Mouillot, D. (2012). Low functional β-diversity despite high taxonomic β-diversity among tropical estuarine fish communities. *PLoS One* 7:e40679. doi: 10.1371/journal.pone.0040679

White, H. J., Montgomery, W. I., Pakeman, R. J., and Lennon, J. J. (2018). Spatiotemporal scaling of plant species richness and functional diversity in a temperate semi-natural grassland. *Ecography* 41, 845–856. doi: 10.1111/ecog.03111

Whittaker, R. J., Rigal, F., Borges, P. A., Cardoso, P., Terzopoulou, S., Casanoves, F., et al. (2014). Functional biogeography of oceanic islands and the scaling of functional diversity in the Azores. *Proc. Natl. Acad. Sci. U.S.A.* 111, 13709–13714. doi: 10.1073/pnas.1218036111

Wilcox, B. A. (1984). *In situ Conservation of Genetic Resources: Determinants of Minimum Area Requirements. National parks, Conservation, and Development: The Role of Protected Areas in Sustaining Society*. (Washington, DC: Smithsonian Institution Press), 825.

Wong, M. K., Guénard, B., and Lewis, O. T. (2019). Trait-based ecology of terrestrial arthropods. *Biol. Rev.* 94, 999–1022. doi: 10.1111/brv.12488

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Ecology and Evolution

ORIGINAL RESEARCH
published: 19 March 2021
doi: 10.3389/fevo.2021.620062

Identifying the Drivers of Spatial Taxonomic and Functional Beta-Diversity of British Breeding Birds

Joseph P. Wayman[1], Jonathan P. Sadler[1], Thomas A. M. Pugh[1,2], Thomas E. Martin[3], Joseph A. Tobias[4] and Thomas J. Matthews[1,5]*

[1] School of Geography, Earth and Environmental Sciences, and Birmingham Institute of Forest Research, University of Birmingham, Birmingham, United Kingdom, [2] Department of Physical Geography and Ecosystem Science, Lund University, Lund, Sweden, [3] Operation Wallacea, Spilsby, United Kingdom, [4] Department of Life Sciences, Imperial College London, Ascot, United Kingdom, [5] Centre for Ecology, Evolution and Environmental Changes-Azorean Biodiversity Group and Universidade dos Açores–Depto de Ciências Agráriase Engenharia do Ambiente, Angra do Heroísmo, Portugal

OPEN ACCESS

Edited by:
John Maxwell Halley, University of Ioannina, Greece

Reviewed by:
William Kunin, University of Leeds, United Kingdom
Joana Raquel Vicente, Centro de Investigacao em Biodiversidade e Recursos Geneticos (CIBIO-InBIO), Portugal

***Correspondence:**
Joseph P. Wayman
jpw830@student.bham.ac.uk

Specialty section:
This article was submitted to Models in Ecology and Evolution, a section of the journal Frontiers in Ecology and Evolution

Received: *21 October 2020*
Accepted: *25 February 2021*
Published: *19 March 2021*

Citation:
Wayman JP, Sadler JP, Pugh TAM, Martin TE, Tobias JA and Matthews TJ (2021) Identifying the Drivers of Spatial Taxonomic and Functional Beta-Diversity of British Breeding Birds. Front. Ecol. Evol. 9:620062. doi: 10.3389/fevo.2021.620062

Spatial variation in community composition may be driven by a variety of processes, including environmental filtering and dispersal limitation. While work has been conducted on the relative importance of these processes on various taxa and at varying resolutions, tests using high-resolution empirical data across large spatial extents are sparse. Here, we use a dataset on the presence/absence of breeding bird species collected at the 10 km × 10 km scale across the whole of Britain. Pairwise spatial taxonomic and functional beta diversity, and the constituent components of each (turnover and nestedness/richness loss or gain), were calculated alongside two other measures of functional change (mean nearest taxon distance and mean pairwise distance). Predictor variables included climate and land use measures, as well as a measure of elevation, human influence, and habitat diversity. Generalized dissimilarity modeling was used to analyze the contribution of each predictor variable to variation in the different beta diversity metrics. Overall, we found that there was a moderate and unique proportion of the variance explained by geographical distance *per se*, which could highlight the role of dispersal limitation in community dissimilarity. Climate, land use, and human influence all also contributed to the observed patterns, but a large proportion of the explained variance in beta diversity was shared between these variables and geographical distance. However, both taxonomic nestedness and functional nestedness were uniquely predicted by a combination of land use, human influence, elevation, and climate variables, indicating a key role for environmental filtering. These findings may have important conservation implications in the face of a warming climate and future land use change.

Keywords: avifauna, beta-diversity, community composition, GDM, spatial, climate

INTRODUCTION

Biodiversity is currently facing a multitude of global-scale threats from human activity (Dirzo et al., 2014; McGill et al., 2015). As the human footprint on the natural world grows, it is becoming increasingly important to understand how these factors are impacting ecological communities in order to inform conservation efforts and make predictions about impacts under future scenarios (Newbold, 2018; Soininen et al., 2018). Analyzing spatial variation in species diversity is a powerful means of assessing the impact of different environmental factors on biodiversity as it provides us with information on what is currently limiting species ranges and occupation of sites. The analysis of spatial variation in biodiversity generally involves focusing on taxonomic changes between sites in the form of alpha (α) or, to a lesser extent, beta (β) diversity (Field et al., 2009; Calderón-Patrón et al., 2016; Soininen et al., 2018).

Comparing the alpha diversity of two communities separated in space provides a measure of the difference in the number of species between these sites but ignores species replacement/turnover (i.e., a species being extirpated from a site and another species colonizing) and can therefore mask biodiversity change (Gonzalez et al., 2016). In contrast, beta-diversity provides a measure of community dissimilarity between sites (Whittaker, 1960; Koleff et al., 2003). Various metrics have been proposed to measure beta-diversity, which can be grouped into variance-based approaches (the focus of the present study) and diversity-partitioning approaches (Legendre and de Cáceres, 2013; Matthews et al., 2019). Recently, several variance-based metrics (e.g., the Sørensen index) have been partitioned into constituent components, such as species replacement/turnover (that is independent of richness differences) and species richness differences or 'nestedness' (Baselga, 2010). It has been argued that the study of these partitions provides insight into the drivers of compositional difference between sites (Baselga and Leprieur, 2015). Nestedness in this context is not 'true' nestedness [e.g., as measured by the nestedness metric based on overlap and decreasing fill (NODF)], but rather nestedness resultant dissimilarity that allows for the separation of dissimilarity due to turnover from that of nestedness (Baselga, 2012). For ease, we henceforth use the term 'nestedness' to describe nestedness resultant dissimilarity.

Standard beta-diversity metrics (herein termed 'taxonomic beta-diversity') assume all species are the same in terms of the role they play within an ecosystem (Sekercioglu, 2006), thereby ignoring the vital role that functional diversity plays in assemblage dynamics (Devictor et al., 2010; Eskildsen et al., 2015). Recently, several beta-diversity metrics have been expanded to incorporate functional information (Cardoso et al., 2014; Baselga and Leprieur, 2015) and can be used to shed light on the biotic/abiotic factors that cause variation in functional diversity between sites (Villéger et al., 2013; Cardoso et al., 2014). This evidence can be used to inform conservation activities such as protected area design and biological corridor selection (Socolar et al., 2016), and help protect ecosystem services (Şekercioğlu et al., 2004; Cardinale et al., 2012; Galetti et al., 2013). For example, bird communities play an important role in providing services such as pollination, pest control, and carrion removal (Whelan et al., 2008; Wenny et al., 2011), which are critical services for humans and other taxa. Using a beta-diversity measure that incorporates species traits, and hence functionality, is therefore essential to gain a better understanding of biodiversity change and its consequences (Devictor et al., 2010; Jarzyna and Jetz, 2018; Tobias and Pigot, 2019; Carvalho et al., 2020).

Biological communities are predicted to vary spatially in the absence of anthropogenic influences (i.e., the natural pattern of distance decay in species similarity; Nekola and White, 1999) due to both dispersal limitation and niche filtering, among other factors (Lomolino et al., 2010). Dispersal limitation is hypothesized to impact the spatial variation in community composition by restricting the range of species through distance alone (Dambros et al., 2017), independent of environmental differences between the communities (Hubbell, 2001). Niche-filtering occurs when environmental gradients constrain communities to those species adapted and able to persist in local conditions (Weiher and Keddy, 1999; Cornwell et al., 2006). In addition, human-induced change (e.g., land-use change and climate change) are likely important drivers of both spatial taxonomic and functional beta-diversity (Devictor et al., 2007; Davey et al., 2012; Barnagaud et al., 2017).

The effect of dispersal on spatial beta-diversity can be assessed through the analysis of the geographic distance between sites, while the effect of environment can be tested by evaluating measures of habitat and land use types (hereafter 'land use') and climate (Luck et al., 2013; Wieczynski et al., 2019; Fluck et al., 2020). However, due to the spatial structuring of environmental gradients (i.e., a distance decay in environmental conditions), it is difficult to partition the unique effects of each (Leibold and Chase, 2017). Relating current land use and climate to spatial variation in community composition will also enable inferences to be made on how increases in the relative intensity of these drivers may impact spatial variation in the future (Barbet-Massin and Jetz, 2015).

The effect of land use on community composition is mostly a result of niche filtering, where species which are adapted to a specific land use type are unable to survive in contrasting land use types (Weiher and Keddy, 1999; Cornwell et al., 2006). Rapidly growing human populations (Tratalos et al., 2007) have facilitated considerable land use changes via increasing urbanization in some areas of the world (Seto et al., 2012), and conversion of natural land to agriculture, as well as an overall intensification of agricultural practices (Zabel et al., 2019). There is strong evidence that these practices have disrupted communities, leading to pools of generalist species in heavily modified habitats via the exclusion, through filtering, of species with narrower environmental requirements (i.e., specialists) (McKinney, 2006; Vellend et al., 2007; Clavel et al., 2011; Flohre et al., 2011; Barnagaud et al., 2017; Hagen et al., 2017). Thus, in anthropogenic landscapes (such as those that occur across much of the United Kingdom), turnover is predicted to be low across large spatial scales, and communities are predicted to become more nested, with high redundancy in functional diversity (Liang et al., 2019; Weideman et al., 2020).

Climate is also an important environmental filter and can drive high spatial beta-diversity between regions due to differences in energy availability (energy richness hypothesis; Hutchinson, 1959; Currie, 1991; Hurlbert and Haskell, 2003), variation in species tolerance (physiological tolerance hypothesis; Root, 1988), and variation in speciation rates (the speciation rates hypothesis; Currie et al., 2004; Hua and Wiens, 2013). While there is mixed support for these hypotheses, substantial evidence exists showing that differences in species composition between sites are often correlated with climatic gradients (Currie et al., 2004). Additional filtering impacts may occur where land use change interacts with climate (Auffret and Thomas, 2019).

Better understanding the role of climate in driving spatial beta-diversity is essential in order to accurately predict the effects of future climate change on community composition. For example, species populations have been found to expand or contract their ranges in response to changing climatic conditions (Fox et al., 2014; Batt et al., 2017). As well as range shifts, shifting phenologies across the trophic web can lead to disruption of communities through cascade effects due to altered species interactions (Bell et al., 2019). Other impacts of a changing climate, such as more frequent severe weather events, are also increasingly recognized as significant drivers of spatial beta-diversity (Maxwell et al., 2019).

There are thus many potential drivers of spatial beta-diversity. However, few studies exist assessing the relative roles of these different drivers (e.g., land use, climate, human impacts) in terms of both taxonomic and functional beta-diversity, the aim of the present study. We use generalized dissimilarity models (GDMs) in combination with a dataset containing presence/absence data of British breeding birds collected at the 100 km^2 scale over the entirety of the British Isles. We aim to (1) test the effect of geographic distance and a range of environmental variables (e.g., land use type, climate) on the spatial taxonomic and functional beta-diversity of British breeding bird communities and (2) evaluate the role of human influence on spatial beta-diversity patterns. While we do not set out to test the niche-filtering and dispersal-limitation hypotheses directly, we interpret increasing dissimilarity due to geographical distance as an indication that dispersal limitation may play a role in the structuring of communities. In contrast, increasing dissimilarity due to climate, land use, or human influence would point to a role for niche filtering.

MATERIALS AND METHODS

Data Collection

Species Composition Data

Data showing the summer (breeding) distributions of the British avifauna (Gillings et al., 2019) were collected during April-July over the period 2008–2011 (BA2010) by volunteers on behalf of the British Trust of Ornithology (BTO) and the Scottish Ornithologists' Club (SOC). Some fieldwork effort was permitted out of this field season, with specific instructions given on what evidence was permitted (see Gillings et al., 2019, for further information).

FIGURE 1 | Map displaying the study location (Great Britain) and its location within Europe. The lower proportion of the island is gridded with the quadrats used to sample the avifauna (only a sample is shown here and the whole island was sampled).

The dataset summarizes the presence/absence of British bird species within 10 km × 10 km (100 km^2) quadrats covering the British Isles on a continuous grid (**Figure 1**). Only species designated as being "confirmed" or "probable" breeders (Gillings et al., 2019) were retained here. Vagrant and pelagic species were excluded, but we retained introduced breeding species for the analysis (McInerny et al., 2018). While some introduced species' occupied ranges may reflect in part their initial introduction sites, many of these species are now established, so their presence exerts an influence on community structure, resource use, and competition (Lennon et al., 2000). Species under threat from human persecution (particularly hunting or egg-collecting) were also removed from the analyses as data for such species were provided at larger spatial grains (i.e., larger than 100 km^2) or their locations were omitted entirely (Gillings et al., 2019). All quadrats with less than 50% land and all island regions that were considered disconnected from the mainland were removed. A total of 2257 100 km^2 quadrats remained with a species pool containing 169 species (**Supplementary Table 1**).

Trait Data

Continuous morphometric variables were measured from museum specimens or extracted from literature and used to characterize the functional diversity of each community (defined as all the species present in each quadrat). We selected eight morphological traits to represent the functional role of birds: two estimates of beak length (culmen from tip-to-skull and tip-to-nares), beak width, beak depth, tarsus length, wing chord length, tail length, and body mass, with evidence showing all of these traits provide useful information about avian dietary niche, locomotion and ecological function (Trisos et al., 2014; Tobias and Pigot, 2019; Pigot et al., 2020). Further information on measurements, including

sampling per species and methods, are published separately (Pigot et al., 2020).

A principal components analysis (PCA) was performed using all eight traits, with the full eight axes extracted. All the axes were then standardized to a mean of 0 and a standard deviation of 1, producing a trait matrix (species x traits) with eight trait axes for each species. To test for the effect of raw trait variability, we also standardized the traits prior to running the PCA. The results using both the raw traits and the standardized traits in the PCAs were comparable, so we report only the results using the raw traits within the PCA here. All eight axes were included as it has been shown that all the axes provide useful information, with even minor axes capturing significant variation in traits (Pigot et al., 2020).

Climate Data

Monthly temperature and precipitation data were downloaded from the United Kingdom Met Office, which provides climate data interpolated from local weather stations onto a 1 km × 1 km grid across the United Kingdom (Hollis et al., 2019). Data were downloaded from the period 2000–2011. For the breeding season (defined as the start of May to the end of July, as many arriving migrants in April will not yet be breeding), key climate variables were selected *a priori,* and averages calculated. Precipitation (mm) was summed for each 100 km^2 quadrat over the breeding season for each year. The average temperature (°C) was calculated as the daily average temperature across the quadrat and the breeding season. The range in temperature was also calculated as the average mean maximum daily temperature over the breeding season minus the average mean minimum daily temperature. The mean of each of these climatic variables was then calculated over the 2000 - 2011 period to reduce the influence of yearly variation, leaving three measures of climate ($Tavg_{mean}$, $Prec_{mean}$, and $Range_{mean}$). Climate averages were also constructed over a more extended period (1960–2011) to test if birds were responding to longer-term climate variation. The majority of the GDMs fitted using the shorter period had a better fit, and thus we only report the results using 2000–2011 climate here.

Land Use Data

Data on land use were obtained from the EDINA environment digimap service for 2007 (Land Cover Map, 2007). These data provide land cover (23 land use classes) for the British Isles at a 25 m scale. From these data, the percentage cover for each land use within each 10 × 10 km quadrat was calculated. The woodland classes (coniferous woodland and broadleaved woodland) were grouped into one variable named 'forest,' as were 'grasslands' (grouped from the improved grasslands and semi-natural grasslands categories), and 'urban areas' (grouped from the suburban and urban categories). Arable land was also included as a predictor variable. Shannon's diversity index was calculated for each quadrat as a measure of land-use heterogeneity (hereafter called Shan).

Human Influence Index (HII)

The Human Influence Index (HII) was used to assess the contribution of human impact on the variation in community composition (Wildlife Conservation Society-WCS, and Center for International Earth Science Information Network-CIESIN-Columbia University, 2005). The HII is derived from multiple data sources on population density, infrastructure (railroads, urban development, night-time lighting), and landcover ranging in date collected 1994–2005 (although in this version about half of the measures were collected around 2000 instead of 1995, as was the case for the first version). The measures are each weighted differently in the methodology and then standardized giving a measure of human impact ranging from zero (no human impact) to 100 (maximum human impact possible using the methodology). HII values were extracted from each of the 1 km^2 grid squares within each 100 km^2 quadrat. The average was taken over these values to obtain the mean HII within each quadrat. It is important to note here that there is some temporal disparity in the period the HII was developed over (see above) and the period the atlas was conducted (2008–2011). However, even with this small disparity, HII should provide a robust indication of the impact human influences have on spatial variation in taxonomic and functional composition.

Elevation Data

Elevation data were obtained from the shuttle radar topography mission (SRTM; Jarvis et al., 2008). For each 100 km^2 quadrat, data were extracted using 400 equally spaced points. The mean ($Mean_{elev}$) and the standard deviation (SD_{elev}) were then calculated from these data as measures of elevation and variability in the elevation across the area.

Testing for Multicollinearity

Pearson's and Spearman's correlations were used to test for multicollinearity between the predictor variables. SD_{elev} was removed due to the variable being strongly correlated with multiple other variables (Elevation, $Tavg_{mean}$, and $Prec_{mean}$). The climatic variables were found to be collinear with one another, and with other variables (**Supplementary Figures 1a,b**). Therefore, the climatic variables were combined using a PCA. The PCA yielded three axes that explained all the variation of the original three climatic variables [hereafter; Climate 1 (81.90%), Climate 2 (13.54%), and Climate 3 (4.56%)]. All the axes were retained in the models to capture all the variability that could be explained by climate. Scatter plots and correlations between the PCA axes and the raw climate variable showed that Climate 1 was positively correlated with average temperature and negatively correlated with average precipitation (**Supplementary Figure 2**). Climate 2 and Climate 3 had less clear and more complex correlations with the original variables (**Supplementary Figure 2**).

After substituting the climate variables with the PCA axes, all variables had correlations < 0.70, with two exceptions: Elevation and Climate 1, and urban land use and HII (**Supplementary Figures 1c,d**). As a result, urban land use was removed from the analysis and the human influence index, which is a composite

measure including urban land use, was retained. Both elevation and Climate 1 were retained because (1) both variables are known to be significant predictors of spatial variation in breeding avian communities, (2) the correlation was still below 0.8, and 3) GDM is known to be robust to multicollinearity to a certain degree (Glassman et al., 2017). A variance inflation factor test (VIF) was also performed, with all remaining variables having VIF values < 5 (Neter et al., 1983; Gareth et al., 2013).

Measuring Spatial Dissimilarity in Community Composition

Spatial Taxonomic and Functional Beta-Diversity

To assess taxonomic dissimilarity between the assemblages, pairwise taxonomic beta diversity was calculated for each 100 km^2 quadrat using the function beta.pair from the package 'betapart' (Baselga and Orme, 2012). This function computes the dissimilarity (here measured using Sørensen's dissimilarity index, βsor; Baselga, 2010; Koleff et al., 2003) between an assemblage and every other assemblage present in the dataset to create a pairwise dissimilarity matrix.

Using Sørensen's dissimilarity, total beta-diversity can then be partitioned into its two constituent components: dissimilarity due to turnover (BD_{TURN}) and nestedness resultant dissimilarity (BD_{NEST}), with $BD_{TOTAL} = BD_{TURN} + BD_{NEST}$. Turnover is the proportion of dissimilarity due to species replacement between two assemblages, whereas nestedness is the proportion of the dissimilarity due to one assemblage being a nested subset of another assemblage through either species loss or gain (Baselga, 2010). It is important to note that, unlike BD_{TURN}, BD_{TOTAL} and BD_{NEST} are not independent of species richness changes, as the measurements are dependent upon species richness gradients (Baselga and Leprieur, 2015).

A measure of functional beta-diversity was then calculated using Sørensen's dissimilarity index and Baselga's (2010) partitioning framework (Phylosor). For this approach, a global functional dendrogram was created containing all the United Kingdom breeding species, using a Euclidean trait distance matrix and the agglomerative hierarchical clustering method (UPGMA). This method produces a rooted tree with a constant weight assumption (i.e., where the distance between the root to all tips is equal), and this then describes the functional relationships between species (Petchey and Gaston, 2002). The phylo.sor function in the 'betapart' package (Baselga and Orme, 2012) was used to calculate functional dissimilarity based on the shared branch length of the functional dendrogram between each assemblage and every other assemblage (hereafter called FD_{TOTAL}). Although this method is usually used on phylogenies, here it is used on a functional dendrogram to give a functional measure analogous to taxonomic beta-diversity, allowing for straightforward comparison. In addition, using a convex hull approach (the standard Baselga functional metric) was not possible here due to the size of the dataset and the computational demands of such an approach. FD_{TOTAL} was also partitioned into its constituent components of nestedness resultant dissimilarity (FD_{NEST}) and turnover (FD_{TURN}).

A Pearson's correlation was performed between the Euclidean distances (in the trait distance matrix) and the cophenetic distances (in the dendrogram) (Villéger et al., 2017). The resultant correlation was high (Pearson's $r = 0.97$), indicating that the dendrogram provides an adequate measure of the functional distances between species.

MNTD (Mean Nearest Taxon Distance) and MPD (Mean Pairwise Distance)

As an alternative to Baselga's functional beta-diversity framework, mean nearest taxon distance (MNTD) and mean pairwise distance (MPD) were calculated. While MNTD is also sensitive to species richness differences, MPD is a measure that is mostly independent of species differences between sites (Miller et al., 2017). MNTD represents the mean distance (smallest non-diagonal value) between species in a community and is most sensitive to changes at the 'tips' of a dendrogram (Webb et al., 2002). MPD is a similar measure but is calculated as the mean between all non-diagonal elements between species within a community (Webb, 2000; Webb et al., 2008), and so it is more sensitive to changes at the roots of the functional dendrogram. Here, the beta-diversity versions of MNTD and MPD, that calculate the same measures but between assemblages are used. For ease, we refer to these as MNTD and MPD (Miller et al., 2017). MPD and MNTD were calculated using the comdist and comdistnt functions, respectively, in the R package 'picante' (Webb et al., 2008). MNTD and MPD were standardized by dividing each pairwise measure by the largest pairwise measure to produce dissimilarity bounded between 0 and 1. Standardizing MPD and MNTD in this way allowed the measures to be modeled using GDMs.

Modelling Variation in Spatial Beta-Diversity

As a first step, multidimensional scaling (MDS) was applied to each of the pairwise measures. The first axes from the MDS were taken and plotted. These were then assessed visually for any pattern in the dissimilarity/similarity between assemblages.

Generalized dissimilarity modeling (GDM) was then used to model functional and taxonomic beta-diversity. GDM is a statistical technique (an extension of matrix regression) that can be used for assessing the relationship between environmental gradients and variation in community composition (Ferrier et al., 2007). The modeling accommodates non-linearity that is present in ecological datasets over large extents (Ferrier, 2002; Ferrier et al., 2007) and can also incorporate geographical distance. This is vital to include, as dispersal limitation modulated by distance is known to be an important driver of community composition (Keil et al., 2012). GDM can deal with higher multicollinearity among predictor variables than many commonly used regression models (Glassman et al., 2017), and uses monotonic I-splines that constrain the coefficients of the regressions to be positive for a non-decreasing fit and non-positive for a non-increasing fit. The I-splines allow the evaluation of predictor effects on the dissimilarity metrics through the height and slope. The maximum height represents the total deviance explained by the predictor while holding all the other predictors constant, while

the slope displays the rate of compositional change across the predictor's range (Fitzpatrick et al., 2013; Fitzpatrick and Keller, 2015). We applied a modeling framework using GDM, aimed at assessing the unique and shared roles of both geographic distance and environmental factors on our measures of beta diversity (**Supplementary Figure 3**).

A separate GDM was fitted with each of the taxonomic and functional beta-diversity metrics (i.e., BD_{TOTAL}, BD_{TURN}, BD_{NEST}, FD_{TOTAL}, FD_{TURN}, FD_{NEST}, MNTD, and MPD) as response variables using the 'gdm' package in R (Fitzpatrick et al., 2020). As a first step, matrix permutation was used to assess the model significance and variable importance using the *gdm.varImp* function. Due to the large amount of memory and time required to model the full site-pair combinations (N = 2,545,896) only a subset of the data could be used for the matrix permutation and variable importance process (as recommended for datasets with a large number of sites when calculating variable importance; Fitzpatrick et al., 2020). First, 60% of the sites were randomly removed, leaving 407,253 site-pairs. Another 60% of the site-pair combinations were then removed from the remaining site-pairs, leaving a total of 162,901 site-pairs for analysis. The removal of the site-pairs, after the initial site removal, removes site-pairs randomly but does not remove all site-pair combinations (Fitzpatrick et al., 2020). The predictor variables were not scaled, which allows assessment of the impacts each of the predictors has along actual environmental gradients (e.g., Fitzpatrick et al., 2013; Heino et al., 2019).

For the matrix permutation process, a GDM was first run using all of the predictor variables. The rows of the environmental data were then permuted, and a GDM model was fitted to those data. Significance was then evaluated by comparing the deviance of the model with unpermuted data to the model using permuted data. Variable importance was assessed by permuting each of the variables in turn while holding the other variables constant (unpermuted). Variable importance was then assessed as the difference in deviance explained using the permuted and unpermuted variable, with more important variables explaining a larger proportion of the deviance when unpermuted. The process was then repeated after dropping the least important variable (backward elimination) with variable importance and significance recalculated. All variable importance scores reported in the text were calculated at this point. The first model where all variables were significant ($p < 0.05$) was identified, and this model was then fit using all the sites (i.e., the full dataset) (final model). One hundred permutations were used for model significance testing and variable importance scores (Ferrier et al., 2007; Heino et al., 2019). Uncertainty in the I-splines was then evaluated using a bootstrapping approach (Shryock et al., 2015). A total of 50% of the sites were first removed randomly from the dataset. From these data, a GDM model was fit, and the I-spline coefficients extracted. For the bootstrapping, a further 80% of the sites were randomly removed, and a GDM model fit. The process was repeated 100 times. The I-splines were then plotted with error bands showing the standard deviation from the permutation process.

Variance shared between geographical distance and the environmental variables was calculated for each model in turn using the formula:

$$V_s = V_{full} - (V_{full} - V_g) - (V_{full} - -V_e),$$

where V_s is the explained variance shared between the geographic predictor and the environmental variables, V_{full} is the variance explained by the full model, V_g is the variance explained by the geographic distance only model, and V_e is the variance explained by the model containing only the environmental variables (Ray-Mukherjee et al., 2014).

RESULTS

Taxonomic and Functional Beta-Diversity of the British Avifauna

BD_{TOTAL} was higher on average (0.313 ± 0.142) than FD_{TOTAL} (0.285 ± 0.126). In both cases, overall beta-diversity was determined mostly by the turnover component (0.225 ± 0.121 (71.88% of the total) and 0.195 ± 0.108 (68.42% of the total), for BD_{TURN} and FD_{TURN}, respectively). Nestedness was responsible for a smaller proportion on average for both measures (**Figures 2A,B**). Average MPD (0.799 ± 0.045) was higher than MNTD (0.255 ± 0.125) (**Figure 2C**), as is to be expected.

Spatial Variation in Taxonomic and Functional Beta Diversity

Heat maps of the community metrics showed clear spatial patterns in the different beta diversity metrics. A north-south divide was present for BD_{TOTAL} (**Figure 3A**) and FD_{TOTAL} (**Figure 3E**), and for the turnover component of each (**Figures 3B,F**, respectively). Alternatively, this could also be interpreted as a divide along the classic 'Tees-exe line' that roughly divides the uplands and lowlands of Britain (Prakash and Rumsey, 2018). The west of Wales was more similar to the north of England and parts of Scotland than the south of England for total beta diversity and turnover for both taxonomic and functional metrics, representing elevation changes between upland and lowland regions (**Figures 3A,B,E,F**). The eastern coastal regions in Scotland were more congruent with southern regions than with inland and west coast Scottish assemblages (**Figures 3A,B,E,F**). Nestedness patterns generally mirror the other metrics but the patterns were patchier (i.e., the divides between regions (north/south and Tees-exe line) was less delineated). The south-west of England and the west coast of Wales were a closer match with most of Scotland than they were with the majority of southern England (**Figures 3C,G**), again closely matching the 'Tees-exe line' (Prakash and Rumsey, 2018).

The pattern for MNTD was largely the same as that found for total taxonomic and functional beta-diversity and turnover (**Figure 3D**). MPD showed a pattern similar to that of nestedness for both taxonomic and functional beta diversity (**Figure 3H**).

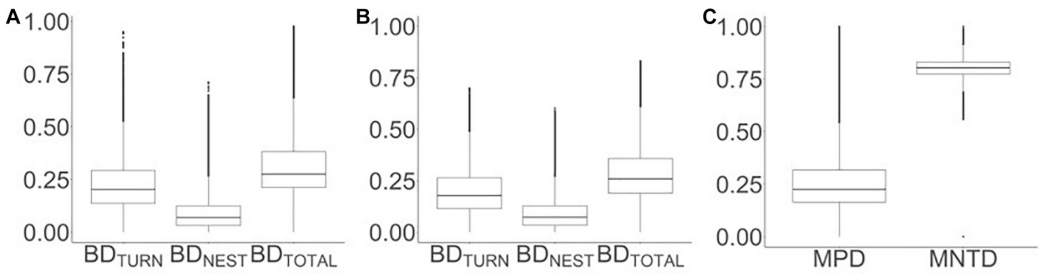


FIGURE 2 | Boxplots of pair-wise dissimilarity measures of community composition for British breeding birds. These are all on the same row **(A)** displays taxonomic spatial dissimilarity (turnover, nestedness resultant dissimilarity, and total), **(B)** is functional pairwise dissimilarity (turnover, nestedness resultant dissimilarity, and total), and **(C)** shows standardised mean pairwise distance (MPD) and standardised mean nearest taxon distance (MNTD). The horizontal line within the box represents the median, the box indicates the inter-quartile range (IQR), and the whiskers show data 1.5 times the IQR. Points highlight outliers.

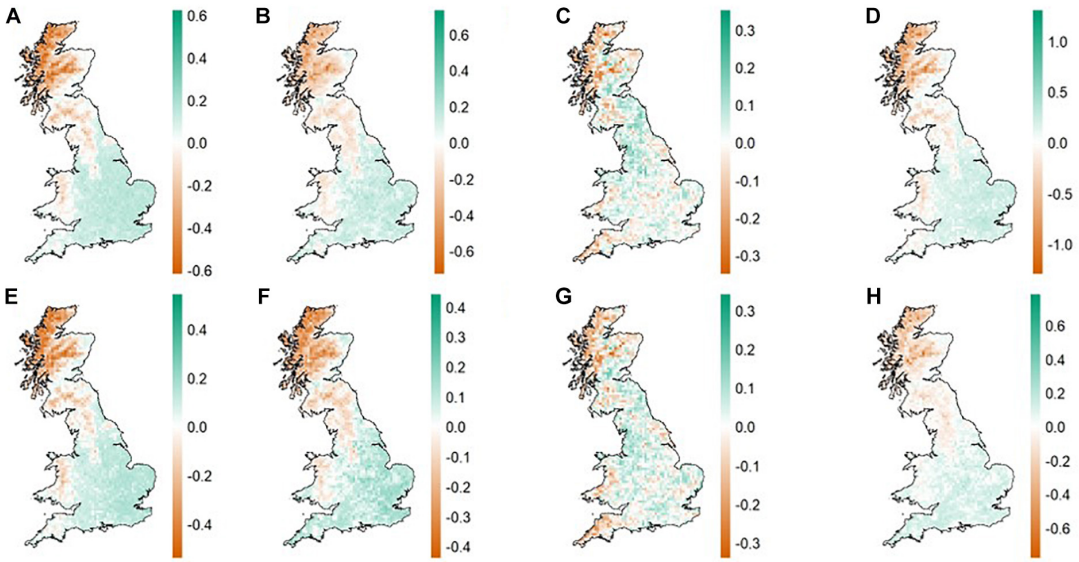


FIGURE 3 | Heat maps of community dissimilarity in British breeding birds based on the first axis from a Principal co-ordinate analysis analysis (PCoA) for taxonomic and functional beta diversity, turnover, and nestedness resultant dissimilarity. Colors represent the ordering scores obtained from the PCoA, with areas displaying similar colors more similar and areas with differing colors less similar in terms of community composition. The first three maps on the first row **(A–C)** are for taxonomic beta-diversity (BD_{TOTAL}, BD_{TURN}, and BD_{NEST}, respectively) and the first three maps on the second row **(E–G)** are for functional beta-diversity (FD_{TOTAL}, FD_{TURN}, and FD_{NEST}, respectively). Mean nearest taxon distance **(D)** and mean pairwise distance **(H)** are the last maps on each row, respectively.

Modelling the Drivers of Spatial Beta-Diversity

Overall, the variance explained by all the final models (bar those for MPD, FD_{NEST}, and BD_{NEST}) was high, with between 55.61% and 68.45% variance explained (**Table 1**). The variance explained by the final GDM models for MPD, FD_{NEST}, and BD_{NEST} was lower (between 9.60% and 12.09%) (**Table 1**). The deviance explained for the functional metrics was lower than for the taxonomic metrics (**Table 1**).

For all final models, the deviance explained by GDM models using only geographical distance as a predictor of dissimilarity was lower than the deviance explained by the models run using only the significant environmental variables (**Table 1**). However, there was overlap in the deviance explained by geographical distance and the environmental variables. Between 19.31% and 23.42% of the variance explained was shared within the MNTD, taxonomic, and functional beta-diversity models, excluding the FD_{NEST} and BD_{NEST} models (**Table 1**). For the nestedness components, the geographical distance between sites explained a low percentage of deviance (**Table 1**). Geographic distance was non-significant in the MPD model.

Drivers of Taxonomic Beta-Diversity

The following results (and those in the 'Drivers of functional beta-diversity' section) relate to the final models (i.e., the models that have been simplified using the permutation approach described in the methods).

BD_{TOTAL} was most impacted by geographical distance, Climate 1, Elevation, and HII (in order of decreasing importance, **Table 1**). BD_{TOTAL} rose gradually with geographical distance and

TABLE 1 | Results of generalized dissimilarity models (GDMs) analyzing the spatial taxonomic and functional beta-diversity of British breeding birds as a function of environmental variables and geographical distance.

	Taxonomic			Functional				
	Total	**Turn**	**Nest**	**Total**	**Turn**	**Nest**	**MPD**	**MNTD**
GDM Deviance	74977	80731	143637	68867	85707	143167	25588	74664
Null Deviance	237649	213424	163388	196464	193057	159253	28306	202198
Variance Explained (%)	68.45	62.17	12.09	64.95	55.61	10.10	9.60	62.97
Intercept	0.13	0.06	0.06	0.12	0.05	0.06	1.53	0.10
Geographic Only (%)	28.04	29.72	1.43	27.95	28.17	1.22	0.00	26.30
Environment Only (%)	61.33	55.87	12.00	56.68	49.08	10.05	9.60	55.98
Shared (%)	20.92	23.42	1.34	19.68	21.64	1.17	0.00	19.31
Variable Importance								
Geographic	10.83	9.95	0.04	9.16	11.42	1.87	0.00	9.68
Climate 1	6.57	3.71	25.05	5.56	3.48	31.64	7.01	10.72
Climate 2	0.00	0.35	0.00	0.27	0.45	0.00	0.00	0.00
Climate 3	0.00	0.00	0.00	0.38	0.13	4.39	0.00	0.26
Arable	1.55	2.88	0.00	0.95	2.83	0.00	0.00	1.00
Forest	1.46	1.82	6.92	0.58	0.74	6.40	0.00	1.09
Grass	0.95	0.52	0.00	0.82	0.46	0.00	0.00	0.90
Shan	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Elevation	5.64	2.14	2.43	1.38	1.07	0.00	37.81	1.34
HII	1.61	1.19	4.51	2.60	1.48	5.12	15.22	3.48

Also included are the results from models analyzing mean pairwise distance (MPD) and mean nearest taxon distance (MNTD) between sites. Variable importance scores were calculated using a subset of the data through permutation of each one of the predictors in turn, while holding all other predictors constant. The variable importance is then the mean difference in variation described by the model including the non-permuted variable and the permuted variable. Therefore, the higher the importance score the more important that variable is to explaining variation in the community dissimilarity metric. For individual variable descriptions, see the main text.

Elevation, with a sharper rise observed for Climate 1 initially, followed by a leveling out (**Figure 4A**). HII had an initial effect on total beta-diversity, with a small but sharp increase observed over initial environmental dissimilarity, but then leveled off and remained relatively constant (**Supplementary Figure 1**). Arable, forest, and grass cover were also in the final BD_{TOTAL} model (**Supplementary Figure 4**).

Geographical distance, Climate 1, and arable land cover (in order of decreasing importance) had similar impacts on BD_{TURN} as they did on BD_{TOTAL} (**Figure 4B**). Elevation and forest cover were the two other most important variables in regard to this response variable, with turnover increasing sharply with forest cover initially before leveling off, and a near-linear increase with elevation (**Supplementary Figure 5**). HII, grass cover, and Climate 2 were also included in the final BD_{TURN} model (**Supplementary Figure 5**).

Geographical distance was relatively unimportant for predicting BD_{NEST}, although it was included in the final model. The best predictor (the predictors with the highest variable importance values; **Table 1**) for BD_{NEST} was Climate 1, and its relationship was similar to that found for BD_{TOTAL} (**Figure 4C**). Forest cover, HII, and elevation were also in the final model (**Supplementary Figure 6**).

Drivers of Functional Beta-Diversity

FD_{TOTAL} had a similar relationship with geographic distance, Climate 1 and HII as was observed with BD_{TOTAL}, and they were also the three most important predictors (**Figure 5A**). Elevation had an almost linear relationship with FD_{TOTAL} (**Supplementary Figure 7**). Arable, grass, and forest cover, along with the other two climate axes (Climate 2 and Climate 3), were also included in the final FD_{TOTAL} model (**Supplementary Figure 7**).

The main difference between the FD_{TURN} and BD_{TURN} final models was the inclusion of Climate 3 for FD_{TURN}, and a slight difference in the ordering of the variable importance (**Table 1**, **Figure 5B** and **Supplementary Figure 8**). FD_{NEST}, as with BD_{NEST}, was also unaffected by geographical distance (**Table 1**). FD_{NEST} was impacted by Climate 1, forest cover, HII, climate 3, and geographical distance (**Figure 5C** and **Supplementary Figure 9**).

MNTD was mainly impacted by Climate 1, geographical distance, HII, Elevation, and forest cover (**Figure 6A** and **Supplementary Figure 10**). Overall, the final MNTD model contained the same significant predictors as in the final models for BD_{TOTAL} and FD_{TOTAL}, and similar relationships were found with all the predictors, although the order was slightly different (**Supplementary Figure 10**). MPD was the only measure not predicted by geographical distance (**Table 1**). MPD was found to sharply increase with Elevation before leveling off around 100 m (**Figure 6B**). HII was the next best predictor of MPD and its relationship was different to that found with the other measures, with a curvilinear increase observed midway through the gradient (**Figure 6B**). The relationship between Climate 1 and MPD was also different to that found between the predictor and the other response variables, with a threshold effect found near the tail end of the gradient (**Figure 6B**).

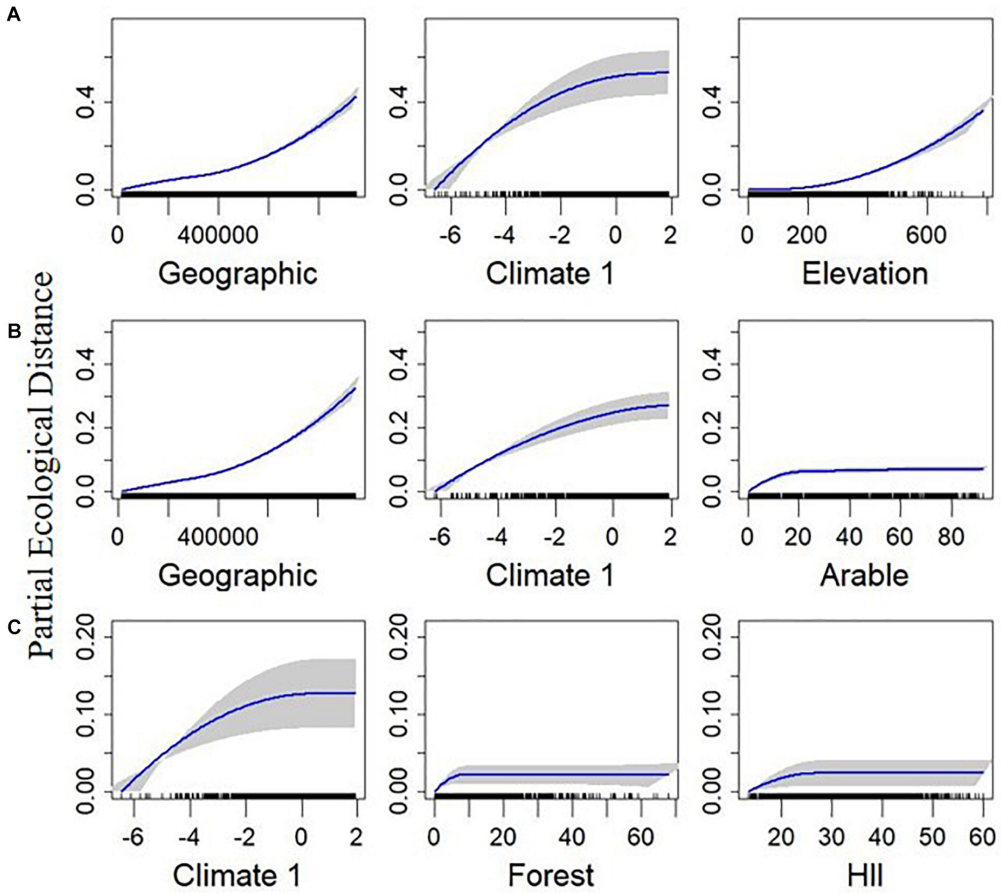


FIGURE 4 | Plotted I-splines of the three most important variables (determined from the variable importance score) from generalized dissimilarity models analyzing the relationship between environmental and geographic gradients, and spatial community composition in British breeding birds. Plots on row **(A)** are for total Sorensen's beta-diversity (BD_{TOTAL}), **(B)** are for the turnover component (BD_{TURN}), and **(C)** are for the nestedness resultant dissimilarity (BD_{NEST}). Climate 1 is the first axis from a principal component analysis calculated from the average temperature, range of temperature, and precipitation over the months May to July, across the 2000–2011 period. Elevation is the average elevation across each 100 km^2 quadrat. Forest and Arable are the percentage of each land use within each quadrat. The human influence index (HII) is the average human influence across each quadrat. Geographic is the geographic distance between sites. Curves show the relationship between the gradients and community dissimilarity obtained using I-splines. The most important variables are on the left with decreasing variable importance to the right. Blue lines show the I-Spline correlations, with standard deviation (gray shaded area) calculated through bootstrapping (100 permutations) on a portion of the dataset. A rug plot on the x axis shows the spread of the data. Note the varying y-axis for each measure.

It should be noted that a large majority of the predictors were close in their variable importance scores across the models (**Table 1**). The geographical distance between sites was the exception across the models, as it was the most important predictor by a large proportion in all but the two nestedness, and MPD/MNTD models. It is also important to note that although multicollinearity was assessed, geographic structuring of certain predictors may mean that the permutation of the variables may not have resulted in large importance values if geographic distance was retained in the model.

DISCUSSION

We found that variation in overall spatial functional and taxonomic beta-diversity (and their turnover components) of British breeding avian assemblages is driven by a combination of geographical distance *per se* and environmental gradients. The environment-only models explained more deviance than the geographic distance only models, but there was also a relatively large, shared variance explained component. Previous studies have suggested that this shared component be considered as indirect effects of climate (climate distance, Mazel et al., 2017; Qian et al., 2020). This may suggest that geographic distance plays a comparatively small role in predicting compositional differences between assemblages in contrast to environment. This aligns with previous studies that have shown that spatial variation in community composition, for a wide range of taxa, can be attributed to a combination of deterministic (including contemporary and historical), and stochastic factors (Steinitz et al., 2006; Melo et al., 2009; Dobrovolski et al., 2012; Vicente et al., 2014; Baselga et al., 2015; Glassman

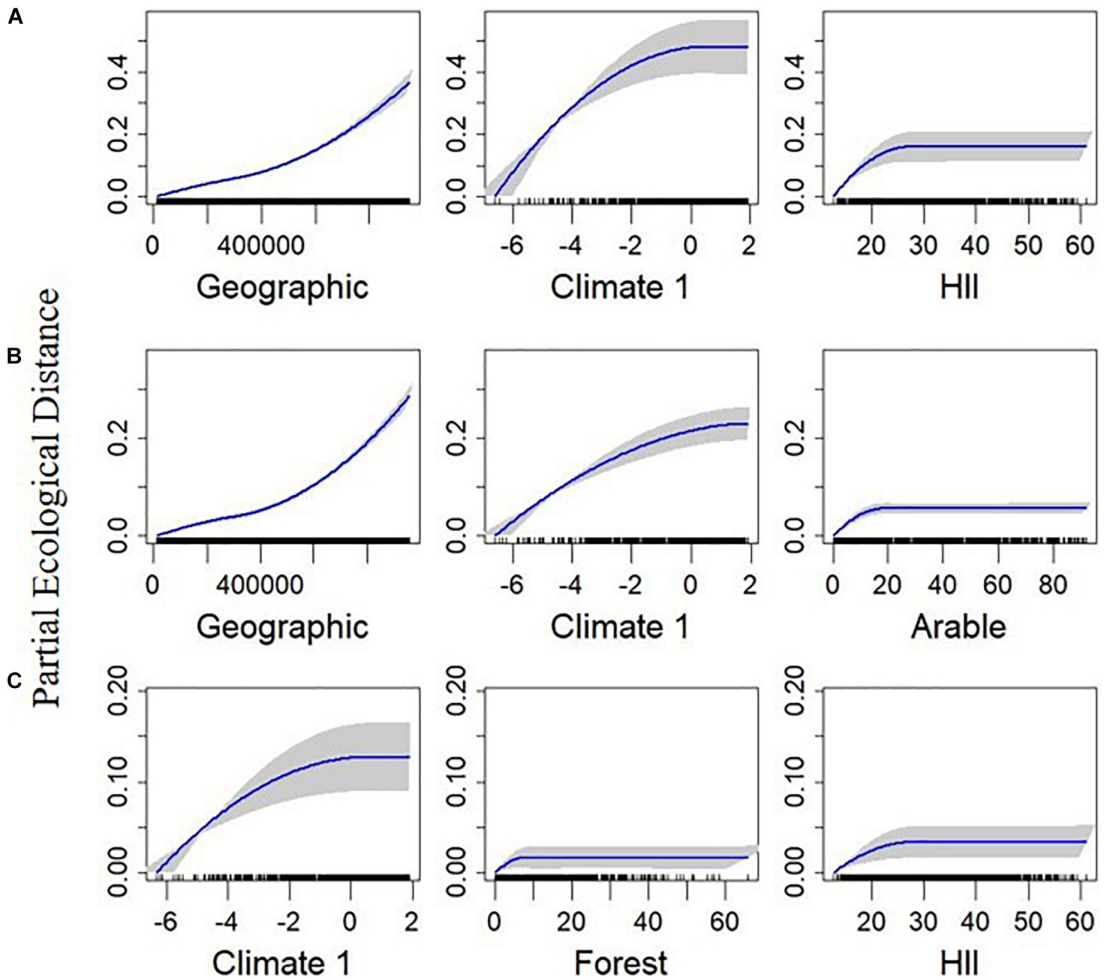


FIGURE 5 | Plotted I-splines of the three most important variables (determined from the variable importance score) from generalized dissimilarity models analyzing the relationship between environmental and geographic gradients, and spatial community composition in British breeding birds. Plots on row **(A)** are for total functional beta-diversity (FD_{TOTAL}), **(B)** are for the turnover component (FD_{TURN}), and **(C)** are for the nestedness resultant dissimilarity (FD_{NEST}). Climate 1 is the first axis from a principal component analysis calculated from the average temperature, range of temperature, and precipitation over the months May to July, across the 2000–2011 period. Forest and Arable are the percentage of each land use within each quadrat. The human influence index (HII) is the average human influence across each quadrat. Geographic is the geographical distance between sites. Curves show the relationship between the gradients and community dissimilarity obtained using I-splines. The most important variables are on the left with decreasing variable importance to the right. Blue lines show the I-Spline correlations, with standard deviation (gray shaded area) calculated through bootstrapping (100 permutations) on a portion of the dataset. A rug plot on the x axis shows the spread of the data. Note the varying y-axis for each measure.

et al., 2017; Carvalho et al., 2020). However, we also found that the differences in species loss/gain between sites (as measured by nestedness resultant dissimilarity) were explained mostly by differences in climate, and to a lesser extent, land use. Overall variation explained by the nestedness models was, however, much lower than for the total beta-diversity and turnover models. The lower explanatory power of the nestedness models implies that other drivers not included here may be impacting dissimilarity between sites due to species loss and gains.

Geographical Distance

Both BD_{TOTAL} and FD_{TOTAL} were driven mostly by the turnover component, highlighting that compositional differences between communities in Britain are mainly a result of different species being replaced. This is consistent with other beta-diversity studies, across multiple taxa (see Soininen et al., 2018). A slow initial increase followed by a steeper rise in the slope of the I-splines of geographical distance in respect to turnover highlights a north-south divide, or a divide between the 'Tees-exe' line, in community dissimilarity (**Figure 1B**). The observed impact of distance could be due to dispersal limitation, geographical barriers, or historical factors (Nekola and White, 1999; Soininen et al., 2007; Dobrovolski et al., 2012; Barnagaud et al., 2017). Here, we expect it is a combination of all these factors, as well as an intertwining of distance with climate and land use.

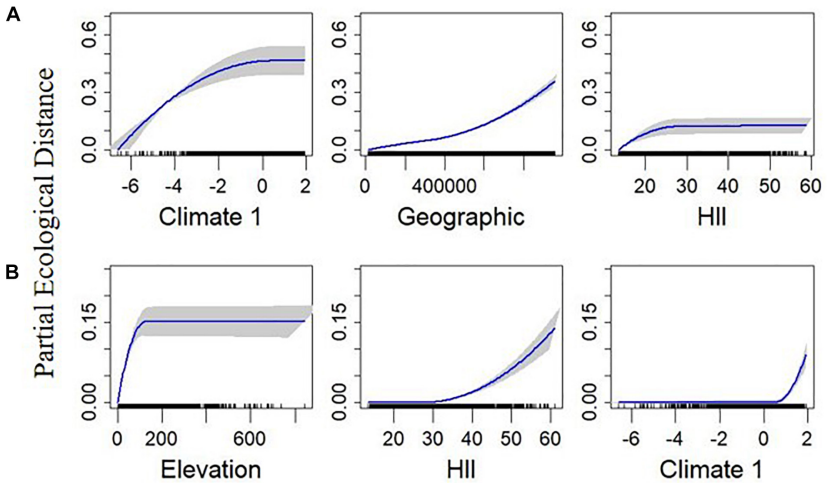


FIGURE 6 | Plotted I-splines of the three most important variables (determined from the variable importance score) from generalized dissimilarity models analyzing the relationship between environmental and geographic gradients, and spatial community composition in British breeding birds. Plots on row **(A)** are for standardised mean nearest taxon distance (MNTD) and **(B)** are for standardised mean pairwise distance (MPD). Climate 1 is the first axis from a principal component analysis calculated from the average temperature, range of temperature, and precipitation over the months May to July, across the 2000 – 2011 period. Elevation is the average elevation across each 100 km^2 quadrat. The human influence index (HII) is the average human influence across each quadrat. Geographic is the geographical distance between sites. Curves show the relationship between the gradients and community dissimilarity obtained using I-splines. The most important variables are on the left with decreasing variable importance to the right. Blue lines show the I-Spline correlations, with standard deviation (gray shaded area) calculated through bootstrapping (100 permutations) on a portion of the dataset. A rug plot on the x axis shows the spread of the data. Note the varying y-axis for each measure.

FD_{TURN} was lower than BD_{TURN}, indicating that the species that are being turned over across communities share some functional traits. Petchey et al. (2007) found that many co-occurring species in British breeding birds were functionally similar, indicating low functional diversity across Britain as a whole and within individual communities. The low functional alpha diversity observed by Petchey et al. (2007) may partly explain the lower functional turnover compared to taxonomic turnover observed here.

Nestedness made up a lower proportion of the total beta-diversity for both the taxonomic and functional metrics. Functional nestedness was mostly on par with taxonomic nestedness (**Figure 2**), highlighting that species lost/gained between assemblages were not functionally redundant (Petchey et al., 2007). The functional distinction of species lost/gained spatially between communities' highlights that there is a significant difference between northern/southern communities. The observed pattern is partially congruent with elevational peaks in Britain (the Tees-exe line; Prakash and Rumsey, 2018). The pattern is also supported by MPD, a measure independent of species richness (**Figure 3H**). It is interesting to note that of the variables that were significant in the MPD model, elevation was the most important and had a threshold effect with any communities over ∼100 m being dissimilar to those at lower elevations (**Figure 6B**). The results found here align with previous evidence that indicate the importance of elevational gradients as environmental filters (Sanders and Rahbek, 2012; Pigot et al., 2016).

Climate

Climate1 was included in the top three most important variables for all the final models (**Table 1**). The impact of climate was similar across all final models (except for the MPD model), with an almost linear relationship observed with dissimilarity before a leveling off. As the correlations with the original climate variables showed a positive correlation with average temperature and a negative correlation with precipitation, this relates to a divide between warmer, drier, and wetter, cooler regions. The near linear relationship between partial ecological distance and Climate 1 for many of the metrics therefore shows a divide in terms of assemblage composition between these two types of regions. Climate was also the best predictor of nestedness by a large margin. This may point to climatic filtering of certain species, with a divide between north and south Britain. However, the overall pattern for both measures was patchy (**Figures 2C,G**). With global warming predicted to change the seasonality and severity of rainfall within Britain in the coming decades (Watts et al., 2015), these results indicate that drier and hotter summers in the future could lead to shifts in community composition.

Land Use and the Human Influence Index

While the overall impact of the different land-use types differed, the form of the relationship between the various land-use predictors and all the measures of beta-diversity were similar. After an initial sharp increase in community dissimilarity over the first small portion of the range of the environmental gradient, the dissimilarity leveled out. This impact over the initial environmental dissimilarity highlights the difference in

community composition between areas with none of that land cover, and those with just a small percentage.

HII impacted spatial variation in communities in much the same way as the land-use predictors. After an initial increase in dissimilarity between communities with increasing HII, there was a leveling off. There are several reasons why this could be the case. For example, Tratalos et al. (2007) found that richness of all species initially increased faster with household densities (one of the measures included in the HII) than urban adapted species, but these then declined significantly after peaking at a very low density. For urban areas, captured within HII, this could highlight the initial homogenization impact of urbanization, with urban areas similar after this initial disturbance driven by an increase in generalists and loss of specialists (Davey et al., 2012). However, as HII is a composite measure, this cannot be conclusively confirmed from these results alone. The human influence captured in this measure may also be masked at this scale as it is likely that a remnant of suitable habitat with lower human disturbance exists within a quadrat, or near to a quadrat (Fattorini et al., 2016). As this study is also only considering presence/absence, it also makes no inferences about abundance, which is likely more sensitive to human influence, and which could be very different between quadrats (Tratalos et al., 2007).

Future Climate and Land-use Change

The importance of both climate and land-use variables points toward potential future disruption of community composition if these drivers increase in intensity as is expected (Seto et al., 2012; Watts et al., 2015). The latitudinal divide between northern and southern (or highland and lowland) community composition also indicates that future warming could see species extirpations/extinctions from the colder, northern regions of the United Kingdom (Tayleur et al., 2016). This would also likely see an increased similarity of the northern communities with southern communities as southern species extend their ranges northward (Hickling et al., 2006). However, species extending their ranges into the UK from Europe could obscure this impact, so studies approaching this question should focus on species identities. Indeed, this has already been observed in the study of bird abundance within England, with resident and short-distance migrants increasing in abundance through time potentially at the expense of long-distance migrants, habitat specialists and cold-associated species (Pearce-Higgins et al., 2015). With many rare species also dispersal limited, future changes in land-use and climate could potentially extirpate some of the few rare bird species Britain has (Baur, 2014).

Limitations

While we consider a range of predictors, we have not included measures of biotic interactions. Competitive interactions and predator presence/abundance can all have an impact on spatial beta-diversity (Wittwer et al., 2015; Koròan and Svitok, 2018). Abundance differences between sites were also not directly considered due to a lack of appropriate data. Given that a species' abundance can be an important determinant of that species' influence on ecosystem functioning (Winfree et al., 2015; Gaston et al., 2018), future studies aiming to analyze spatial variation in community and functional composition should attempt to analyze measures of population size.

CONCLUSION

Spatial variation in both the taxonomic and functional composition of United Kingdom breeding birds is driven mainly by species turnover, which can be explained through a combination of geographical distance *per se* and environmental gradients. The unique variance explained by distance alone could reflect an important role for dispersal limitation in driving these patterns, but more work is needed, as this variance component could also be due to a process not considered here. In contrast, species loss/gain, observed through nestedness between sites, was driven mainly by environmental factors. Future climate warming and land-use change could lead to an increase in the loss of species, particularly cold-adapted or rare and dispersal-limited, from communities, particularly in the north and in the uplands. With turnover driving these patterns, broad-ranging conservation efforts would be preferable to conservation focused on target areas (Si et al., 2016).

Future work should look for potential synergies between climate and land-use in order to assess if future increases in both could potentially have larger than expected impacts on biodiversity based on the individual effects of each in isolation (Brook et al., 2008; de Chazal and Rounsevell, 2009; Mantyka-pringle et al., 2012; Frishkoff et al., 2016). Future comparison between the results presented here and results of similar tests from areas in different climate regions or in less disturbed regions than the United Kingdom would also be informative. For example, in tropical systems it would be expected that, as the United Kingdom is a post-perturbation system and rates of habitat loss will be higher in the tropics, land use will play a larger role in driving community composition dissimilarity between sites than found here (Hansen et al., 2013). Another potential future research direction could be to assess whether other taxonomic groups within the British Isles have similar patterns of functional or taxonomic beta-diversity, or whether the patterns observed here are bird-specific.

DATA AVAILABILITY STATEMENT

Publicly available datasets were analyzed in this study. These data can be found here: http://www.bto.org/datasets.

AUTHOR CONTRIBUTIONS

JW, JS, TP, and TJM conceived the original idea. JT collected the trait data. JW performed the analysis with input from JS, TP, and TJM. JW wrote the manuscript with significant contributions from TJM. All authors contributed to the writing and commented on the draft.

FUNDING

Central OA block grant processed through the University of Birmingham.

ACKNOWLEDGMENTS

We thank British Trust of Ornithology (BTO) volunteers for their hard work in collecting the data. We are also grateful to the Sir Stanley Stapley Trust for financial support (to JW). The computations described in this manuscript were performed using the University of Birmingham's BlueBEAR HPC service.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fevo.2021.620062/full#supplementary-material

REFERENCES

Auffret, A. G., and Thomas, C. D. (2019). Synergistic and antagonistic effects of land use and non-native species on community responses to climate change. *Glob. Change Biol.* 25, 4303–4314. doi: 10.1111/gcb.14765

Barbet-Massin, M., and Jetz, W. (2015). The effect of range changes on the functional turnover, structure and diversity of bird assemblages under future climate scenarios. *Glob. Change Biol.* 21, 2917–2928. doi: 10.1111/gcb.12905

Barnagaud, J. Y., Kissling, W. D., Tsirogiannis, C., Fisikopoulos, V., Villéger, S., Sekercioglu, C. H., et al. (2017). Biogeographical, environmental and anthropogenic determinants of global patterns in bird taxonomic and trait turnover. *Glob. Ecol. Biogeogr.* 26, 1190–1200. doi: 10.1111/geb.12629

Baselga, A. (2010). Partitioning the turnover and nestedness components of beta diversity. *Glob. Ecol. Biogeogr.* 19, 134–143. doi: 10.1111/j.1466-8238.2009.00490.x

Baselga, A. (2012). The relationship between species replacement, dissimilarity derived from nestedness, and nestedness. *Glob. Ecol. Biogeogr.* 21, 1223–1232. doi: 10.1111/j.1466-8238.2011.00756.x

Baselga, A., Bonthoux, S., and Balent, G. (2015). Temporal beta diversity of bird assemblages in agricultural landscapes: land cover change vs. stochastic processes. *PLoS One* 10:e0127913. doi: 10.1371/journal.pone.0127913

Baselga, A., and Leprieur, F. (2015). Comparing methods to separate components of beta diversity. *Methods Ecol. Evol.* 6, 1069–1079. doi: 10.1111/2041-210X.12388

Baselga, A., and Orme, C. D. L. (2012). Betapart: an R package for the study of beta diversity. *Methods Ecol. Evol.* 3, 808–812. doi: 10.1111/j.2041-210X.2012.00224.x

Batt, R. D., Morley, J. W., Selden, R. L., Tingley, M. W., and Pinsky, M. L. (2017). Gradual changes in range size accompany long-term trends in species richness. *Ecol. Lett.* 20, 1148–1157. doi: 10.1111/ele.12812

Baur, B. (2014). Dispersal-limited species - A challenge for ecological restoration. *Basic Appl. Ecol.* 15, 559–564. doi: 10.1016/j.baae.2014.06.004

Bell, J. R., Botham, M. S., Henrys, P. A., Leech, D. I., Pearce-Higgins, J. W., Shortall, C. R., et al. (2019). Spatial and habitat variation in aphid, butterfly, moth and bird phenologies over the last half century. *Glob. Change Biol.* 25, 1982–1994. doi: 10.1111/gcb.14592

Brook, B. W., Sodhi, N. S., and Bradshaw, C. J. A. (2008). Synergies among extinction drivers under global change. *Trends Ecol. Evol.* 23, 453–460. doi: 10.1016/j.tree.2008.03.011

Calderón-Patrón, J. M., Goyenechea, I., Ortiz-Pulido, R., Castillo-Cerón, J., Manriquez, N., Ramírez-Bautista, A., et al. (2016). Beta diversity in a highly heterogeneous area: disentangling species and taxonomic dissimilarity for terrestrial vertebrates. *PLoS One* 11:e0160438. doi: 10.1371/journal.pone.0160438

Cardinale, B. J., Duffy, J. E., Gonzalez, A., Hooper, D. U., Perrings, C., Venail, P., et al. (2012). Biodiversity loss and its impact on humanity. *Nature* 486, 59–67. doi: 10.1038/nature11148

Cardoso, P., Rigal, F., Carvalho, J. C., Fortelius, M., Borges, P. A. V., Podani, J., et al. (2014). Partitioning taxon, phylogenetic and functional beta diversity into replacement and richness difference components. *J. Biogeogr.* 41, 749–761. doi: 10.1111/jbi.12239

Carvalho, J. C., Malumbres-Olarte, J., Arnedo, M. A., Crespo, L. C., Domenech, M., and Cardoso, P. (2020). Taxonomic divergence and functional convergence in Iberian spider forest communities: insights from beta diversity partitioning. *J. Biogeogr.* 47, 288–300. doi: 10.1111/jbi.13722

Clavel, J., Julliard, R., and Devictor, V. (2011). Worldwide decline of specialist species: toward a global functional homogenization? *Front. Ecol. Environ.* 9:222–228. doi: 10.1890/080216

Cornwell, W. K., Schwilk, D. W., and Ackerly, D. D. (2006). A trait-based test for habitat filtering: convex hull volume. *Ecology* 87, 1465–1471. doi: 10.1890/0012-9658(2006)87[1465:attfhf]2.0.co;2

Currie, D. J. (1991). Energy and large-scale patterns of animal and plant species richness. *Am. Nat.* 137, 27–49. doi: 10.1086/285144

Currie, D. J., Mittelbach, G. G., Cornell, H. V., Field, R., Guégan, J. F., Hawkins, B. A., et al. (2004). Predictions and tests of climate-based hypotheses of broad-scale variation in taxonomic richness. *Ecol. Lett.* 7, 1121–1134. doi: 10.1111/j.1461-0248.2004.00671.x

Dambros, C. S., Morais, J. W., Azevedo, R. A., and Gotelli, N. J. (2017). Isolation by distance, not rivers, control the distribution of termite species in the Amazonian rain forest. *Ecography* 40, 1242–1250. doi: 10.1111/ecog.02663

Davey, C. M., Chamberlain, D. E., Newson, S. E., Noble, D. G., and Johnston, A. (2012). Rise of the generalists: evidence for climate driven homogenization in avian communities. *Glob. Ecol. Biogeogr.* 21, 568–578. doi: 10.1111/j.1466-8238.2011.00693.x

de Chazal, J., and Rounsevell, M. D. A. (2009). Land-use and climate change within assessments of biodiversity change: a review. *Glob. Environ. Change* 19, 306–315. doi: 10.1016/j.gloenvcha.2008.09.007

Devictor, V., Julliard, R., Couvet, D., Lee, A., and Jiguet, F. (2007). Functional homogenization effect of urbanization on bird communities. *Conserv. Biol.* 21, 741–751. doi: 10.1111/j.1523-1739.2007.00671.x

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

Dirzo, R., Young, H. S., Galetti, M., Ceballos, G., Isaac, N. J. B., and Collen, B. (2014). Defaunation in the Anthropocene. *Science* 345, 401–406. doi: 10.1126/science.1251817

Dobrovolski, R., Melo, A. S., Cassemiro, F. A. S., and Diniz-Filho, J. A. F. (2012). Climatic history and dispersal ability explain the relative importance of turnover and nestedness components of beta diversity. *Glob. Ecol. Biogeogr.* 21, 191–197. doi: 10.1111/j.1466-8238.2011.00671.x

Eskildsen, A., Carvalheiro, L. G., Kissling, W. D., Biesmeijer, J. C., Schweiger, O., and Høye, T. T. (2015). Ecological specialization matters: long-term trends in butterfly species richness and assemblage composition depend on multiple functional traits. *Divers. Distrib.* 21, 792–802. doi: 10.1111/ddi.12340

Fattorini, S., Galassi, D. M. P., and Strona, G. (2016). When human needs meet beetle preferences: tenebrionid beetle richness covaries with human population on the Mediterranean islands. *Insect Conserv. Divers.* 9, 369–373. doi: 10.1111/icad.12170

Ferrier, S. (2002). Mapping spatial pattern in biodiversity for regional conservation planning: where to from here? *Syst. Biol.* 51, 331–363. doi: 10.1080/10635150252899806

Ferrier, S., Manion, G., Elith, J., and Richardson, K. (2007). Using generalized dissimilarity modelling to analyse and predict patterns of beta diversity in regional biodiversity assessment. *Divers. Distrib.* 13, 252–264. doi: 10.1111/j.1472-4642.2007.00341.x

Field, R., Hawkins, B. A., Cornell, H. V., Currie, D. J., Diniz-Filho, J. A. F., Guégan, J. F., et al. (2009). Spatial species-richness gradients across scales: a meta-analysis. *J. Biogeogr.* 36, 132–147. doi: 10.1111/j.1365-2699.2008.01963.x

Fitzpatrick, M. C., and Keller, S. R. (2015). Ecological genomics meets community-level modelling of biodiversity: mapping the genomic landscape of current and future environmental adaptation. *Ecol. Lett.* 18, 1–16. doi: 10.1111/ele.12376

Fitzpatrick, M. C., Sanders, N. J., Normand, S., Svenning, J. C., Ferrier, S., Gove, A. D., et al. (2013). Environmental and historical imprints on beta diversity: insights from variation in rates of species turnover along gradients. *Proc. R. Soc. B Biol. Sci.* 280:20131201. doi: 10.1098/rspb.2013.1201

Fitzpatrick, M. C., Mokany, K., Manion, G., Lisk, M., Ferrier, S., and Nieto-Lugilde, D. (2020). *gdm: Generalised Dissimilarity Modeling. R Package Version 1.4.2*. Avaliable at: https://CRAN.R-project.org/package=gdm.

Flohre, A., Fischer, C., Aavik, T., Bengtsson, J., Berendse, F., Bommarco, R., et al. (2011). Agricultural intensification and biodiversity partitioning in European landscapes comparing plants, carabids, and birds. *Ecol. Appl.* 21, 1772–1781. doi: 10.1890/10-0645.1

Fluck, I. E., Cáceres, N., Hendges, C. D., Brum, M., do, N., and Dambros, C. S. (2020). Climate and geographic distance are more influential than rivers on the beta diversity of passerine birds in Amazonia. *Ecography* 43, 860–868. doi: 10.1111/ecog.04753

Fox, R., Oliver, T. H., Harrower, C., Parsons, M. S., Thomas, C. D., and Roy, D. B. (2014). Long-term changes to the frequency of occurrence of British moths are consistent with opposing and synergistic effects of climate and land-use changes. *J. Appl. Ecol.* 51, 949–957. doi: 10.1111/1365-2664.12256

Frishkoff, L. O., Karp, D. S., Flanders, J. R., Zook, J., Hadly, E. A., Daily, G. C., et al. (2016). Climate change and habitat conversion favour the same species. *Ecol. Lett.* 19, 1081–1090. doi: 10.1111/ele.12645

Galetti, M., Guevara, R., Cortes, M. C., Fadini, R., von Matter, S., Leite, A. B., et al. (2013). Functional extinction of birds drives rapid evolutionary changes in seed size. *Science* 340, 1086–1090. doi: 10.1126/science.342.6164.1316-c

Gareth, J., Daniela, W., Trevor, H., and Robert, T. (2013). *An Introduction to Statistical Learning: With Applications in R*. Berlin: Spinger.

Gaston, K. J., Cox, D. T. C., Canavelli, S. B., García, D., Hughes, B., Maas, B., et al. (2018). Population abundance and ecosystem service provision: the case of birds. *BioScience* 68, 264–272. doi: 10.1093/biosci/biy005

Gillings, S., Balmer, D. E., Caffrey, B. J., Downie, I. S., Gibbons, D. W., Lack, P. C., et al. (2019). Breeding and wintering bird distributions in Britain and Ireland from citizen science bird atlases. *Glob. Ecol. Biogeogr.* 28, 866–874. doi: 10.1111/geb.12906

Glassman, S. I., Wang, I. J., and Bruns, T. D. (2017). Environmental filtering by pH and soil nutrients drives community assembly in fungi at fine spatial scales. *Mol. Ecol.* 26, 6960–6973. doi: 10.1111/mec.14414

Gonzalez, A., Cardinale, B. J., Allington, G. R. H., Byrnes, J., Endsley, K. A., Brown, D. G., et al. (2016). Estimating local biodiversity change: a critique of papers claiming no net loss of local diversity. *Ecology* 97, 1949–1960. doi: 10.1890/15-1759.1

Hagen, E. O., Hagen, O., Ibáñez-álamo, J. D., Petchey, O. L., and Evans, K. L. (2017). Impacts of urban areas and their characteristics on avian functional diversity. *Front. Ecol. Evol.* 5:84. doi: 10.3389/fevo.2017.00084

Hansen, M. C., Potapov, P. V., Moore, R., Hancher, M., Turubanova, S. A., Tyukavina, A., et al. (2013). High-resolution global maps of 21st-century forest cover change. *Science* 342, 850–853. doi: 10.1126/science.1244693

Heino, J., Alahuhta, J., Fattorini, S., and Schmera, D. (2019). Predicting beta diversity of terrestrial and aquatic beetles using ecogeographical variables: insights from the replacement and richness difference components. *J. Biogeogr.* 46, 304–315. doi: 10.1111/jbi.13485

Hickling, R., Roy, D. B., Hill, J. K., Fox, R., and Thomas, C. D. (2006). The distributions of a wide range of taxonomic groups are expanding polewards. *Glob. Change Biol.* 12, 450–455. doi: 10.1111/j.1365-2486.2006.01116.x

Hollis, D., McCarthy, M., Kendon, M., Legg, T., and Simpson, I. (2019). HadUK-Grid—A new UK dataset of gridded climate observations. *Geosci. Data J.* 6, 151–159. doi: 10.1002/gdj3.78

Hua, X., and Wiens, J. J. (2013). How does climate influence speciation? *Am. Nat.* 182, 1–12. doi: 10.1086/670690

Hubbell, S. P. (2001). *The Unified Neutral Theory of Biodiversity and Biogeography*. Princeton: Princeton University Press.

Hurlbert, A. H., and Haskell, J. P. (2003). The effect of energy and seasonality on avian species richness and community composition. *Am. Nat.* 161, 83–97. doi: 10.1086/345459

Hutchinson, G. E. (1959). Homage to Santa Rosalia or why are there so many kinds of animals? *Am. Nat.* 93, 145–159. doi: 10.1086/282070

Jarvis, A., Reuter, H. I., Nelson, A., and Guevara, E. (2008). *Hole-Filled SRTM for the Globe VERSION 4*, Vol. 15. 25–54. Avaliable at:http://srtm.csi.cgiar.org (accessed February 3, 2020).

Jarzyna, M. A., and Jetz, W. (2018). Taxonomic and functional diversity change is scale dependent. *Nat. Commun.* 9:2565. doi: 10.1038/s41467-018-04889-z

Keil, P., Schweiger, O., Kühn, I., Kunin, W. E., Kuussaari, M., Settele, J., et al. (2012). Patterns of beta diversity in Europe: the role of climate, land cover and distance across scales. *J. Biogeogr.* 39, 1473–1486. doi: 10.1111/j.1365-2699.2012.02701.x

Koleff, P., Gaston, K. J., and Lennon, J. J. (2003). Measuring beta diversity for presence-absence data. *J. Anim. Ecol.* 72, 367–382. doi: 10.1046/j.1365-2656.2003.00710.x

Koròan, M., and Svitok, M. (2018). Pairwise null model analyses of temporal patterns of bird assemblages contradict the assumptions of competition theory. *Basic Appl. Ecol.* 31, 72–81. doi: 10.1016/j.baae.2018.07.001

Land Cover Map (2007). *[TIFF Geospatial data], Scale 1:10000000, Tiles: GB, Updated: 18 July 2008, CEH, Using: EDINA Environment Digimap Service*. Avaliable at: https://digimap.edina.ac.uk (accessed May 26, 2019).

Legendre, P., and de Cáceres, M. (2013). Beta diversity as the variance of community data: dissimilarity coefficients and partitioning. *Ecol. Lett.* 16, 951–963. doi: 10.1111/ele.12141

Leibold, M. A., and Chase, J. M. (2017). *Metacommunity Ecology*, Vol. 59. Princeton, NJ: Princeton University Press.

Lennon, J. J., Greenwood, J. J. D., and Turner, J. R. G. (2000). Bird diversity and enviromnental gradients in Britain: a test of the species-energy hypothesis. *J. Anim. Ecol.* 69, 581–598. doi: 10.1046/j.1365-2656.2000.00418.x

Liang, C., Yang, G., Wang, N., Feng, G., Yang, F., Svenning, J. C., et al. (2019). Taxonomic, phylogenetic and functional homogenization of bird communities due to land use change. *Biol. Conserv.* 236, 37–43. doi: 10.1016/j.biocon.2019.05.036

Lomolino, M. V., Riddle, B. R., and Brown, J. H. (2010). *Biogeography*. Sunderland, MA: Sinauer Associates Inc.

Luck, G. W., Carter, A., and Smallbone, L. (2013). Changes in bird functional diversity across multiple land uses: interpretations of functional redundancy depend on functional group identity. *PLoS One* 8:e0063671. doi: 10.1371/journal.pone.0063671

Mantyka-pringle, C. S., Martin, T. G., and Rhodes, J. R. (2012). Interactions between climate and habitat loss effects on biodiversity: a systematic review and meta-analysis. *Glob. Change Biol.* 18, 1239–1252. doi: 10.1111/j.1365-2486.2011.02593.x

Matthews, T. J., Aspin, T. W. H., Ulrich, W., Baselga, A., Kubota, Y., Proios, K., et al. (2019). Can additive beta diversity be reliably partitioned into nestedness and turnover components? *Glob. Eco. Biogeogr.* 28, 1146–1154. doi: 10.1111/geb.12921

Maxwell, S. L., Butt, N., Maron, M., McAlpine, C. A., Chapman, S., Ullmann, A., et al. (2019). Conservation implications of ecological responses to extreme weather and climate events. *Divers. Distrib.* 25, 613–625. doi: 10.1111/ddi.12878

Mazel, F., Wüest, R. O., Lessard, J.-P., Renaud, J., Ficetola, G. F., Lavergne, S., et al. (2017). Global patterns of ß-diversity along the phylogenetic time-scale: the role of climate and plate tectonics. *Glob. Ecol. Biogeogr.* 26, 1211–1221. doi: 10.1111/geb.12632

McGill, B. J., Dornelas, M., Gotelli, N. J., and Magurran, A. E. (2015). Fifteen forms of biodiversity trend in the anthropocene. *Trends Ecol. Evol.* 30, 104–113. doi: 10.1016/j.tree.2014.11.006

McInerny, C. J., Musgrove, A. J., Stoddart, A., Harrop, A. H. J., Dudley, S. P., Balmer, D., et al. (2018). The British list: a checklist of birds of Britain (9th edition). *Ibis* 160, 190–240. doi: 10.1111/ibi.12536

McKinney, M. L. (2006). Urbanization as a major cause of biotic homogenization. *Biol. Conserv.* 127, 247–260. doi: 10.1016/j.biocon.2005.09.005

Melo, A. S., Rangel, T. F. L. V. B., and Diniz-Filho, J. A. F. (2009). Environmental drivers of beta-diversity patterns in New-World birds and mammals. *Ecography* 32, 226–236. doi: 10.1111/j.1600-0587.2008.05502.x

Miller, E. T., Farine, D. R., and Trisos, C. H. (2017). Phylogenetic community structure metrics and null models: a review with new methods and software. *Ecography* 40, 461–477. doi: 10.1111/ecog.02070

Nekola, J. C., and White, P. S. (1999). The distance decay of similarity in biogeography and ecology. *J. Biogeogr.* 26, 867–878. doi: 10.1046/j.1365-2699.1999.00305.x

Neter, J., Kutner, M. H., Nachtsheim, C. J., and Wasserman, W. (1983). *Applied Linear Regression Models*. Burr Ridge, IL: Irwin.

Newbold, T. (2018). Future effects of climate and land-use change on terrestrial vertebrate community diversity under different scenarios. *Proc. R. Soc. B Biol. Sci.* 285:20180792. doi: 10.1098/rspb.2018.0792

Pearce-Higgins, J. W., Eglington, S. M., Martay, B., and Chamberlain, D. E. (2015). Drivers of climate change impacts on bird communities. *J. Anim. Ecol.* 84, 943–954. doi: 10.1111/1365-2656.12364

Petchey, O. L., Evans, K. L., Fishburn, I. S., and Gaston, K. J. (2007). Low functional diversity and no redundancy in British avian assemblages. *J. Anim. Ecol.* 76, 977–985. doi: 10.1111/j.1365-2656.2007.01271.x

Petchey, O. L., and Gaston, K. J. (2002). Functional diversity (FD), species richness and community composition. *Ecol. Lett.* 5, 402–411. doi: 10.1046/j.1461-0248.2002.00339.x

Pigot, A. L., Sheard, C., Miller, E. T., Bregman, T. P., Freeman, B. G., Roll, U., et al. (2020). Macroevolutionary convergence connects morphological form to ecological function in birds. *Nat. Ecol. Evol.* 4, 230–239. doi: 10.1038/s41559-019-1070-4

Pigot, A. L., Trisos, C. H., and Tobias, J. A. (2016). Functional traits reveal the expansion and packing of ecological niche space underlying an elevational diversity gradient in passerine birds. *Proc. R. Soc. B Biol. Sci.* 283, 1–9. doi: 10.1098/rspb.2015.2013

Prakash, R. O., and Rumsey, F. (2018). "Biodiversity in the united kingdom," in *Global Biodiversity: Selected Countries in Europe, Vol. 2* (Apple Academic Press), 379.

Qian, H., Jin, Y., Leprieur, F., Wang, X., and Deng, T. (2020). Patterns of phylogenetic beta diversity measured at deep evolutionary histories across geographical and ecological spaces for angiosperms in China. *J. Biogeogr.* 1–12. doi: 10.1111/jbi.14036

Ray-Mukherjee, J., Nimon, K., Mukherjee, S., Morris, D. W., Slotow, R., and Hamer, M. (2014). Using commonality analysis in multiple regressions: a tool to decompose regression effects in the face of multicollinearity. *Methods Ecol. Evol.* 5, 320–328. doi: 10.1111/2041-210x.12166

Root, T. (1988). Energy constraints on avian distributions and abundances. *Ecol. Soc. Am.* 69, 330–339. doi: 10.2307/1940431

Sanders, N. J., and Rahbek, C. (2012). The patterns and causes of elevational diversity gradients. *Ecography* 35, 1–3. doi: 10.1111/j.1600-0587.2011.07338.x

Sekercioglu, C. H. (2006). Increasing awareness of avian ecological function. *Trends Ecol. Evol.* 21, 464–471. doi: 10.1016/j.tree.2006.05.007

Şekercioğlu, ÇH., Daily, G. C., and Ehrlich, P. R. (2004).). Ecosystem consequences of bird declines. *Proc. Natl. Acad. Sci. U.S.A.* 101, 18042–18047. doi: 10.1073/pnas.0408049101

Seto, K. C., Güneralp, B., and Hutyra, L. R. (2012). Global forecasts of urban expansion to 2030 and direct impacts on biodiversity and carbon pools. *Proc. Natl. Acad. Sci. U.S.A.* 109, 16083–16088. doi: 10.1073/pnas.1211658109

Shryock, D. F., Havrilla, C. A., DeFalco, L. A., Esque, T. C., Custer, N. A., and Wood, T. E. (2015). Landscape genomics of Sphaeralcea ambigua in the Mojave Desert: a multivariate, spatially-explicit approach to guide ecological restoration. *Conserv. Genet.* 16, 1303–1317. doi: 10.1007/s10592-015-0741-1

Si, X., Baselga, A., Leprieur, F., Song, X., and Ding, P. (2016). Selective extinction drives taxonomic and functional alpha and beta diversities in island bird assemblages. *J. Anim. Ecol.* 85, 409–418. doi: 10.1111/1365-2656.12478

Socolar, J. B., Gilroy, J. J., Kunin, W. E., and Edwards, D. P. (2016). How should beta-diversity inform biodiversity conservation? *Trends Ecol. Evol.* 31, 67–80. doi: 10.1016/j.tree.2015.11.005

Soininen, J., Heino, J., and Wang, J. (2018). A meta-analysis of nestedness and turnover components of beta diversity across organisms and ecosystems. *Global Ecol. Biogeogr.* 27, 96–109. doi: 10.1111/geb.12660

Soininen, J., McDonald, R., and Hillebrand, H. (2007). The distance decay of similarity in ecological communities. *Ecography* 30, 3–12. doi: 10.1111/j.2006.0906-7590.04817.x

Steinitz, O., Heller, J., Tsoar, A., Rotem, D., and Kadmon, R. (2006). Environment, dispersal and patterns of species similarity. *J. Biogeogr.* 33, 1044–1054. doi: 10.1111/j.1365-2699.2006.01473.x

Tayleur, C. M., Devictor, V., Gaüzère, P., Jonzén, N., Smith, H. G., and Lindström, Å (2016). Regional variation in climate change winners and losers highlights the rapid loss of cold-dwelling species. *Divers. Distrib.* 22, 468–480. doi: 10.1111/ddi.12412

Tobias, J. A., and Pigot, A. L. (2019). Integrating behaviour and ecology into global biodiversity conservation strategies. *Philos. Trans. R. Soc. B Biol. Sci.* 374:20190012. doi: 10.1098/rstb.2019.0012

Tratalos, J., Fuller, R. A., Evans, K. L., Davies, R. G., Newson, S. E., Greenwood, J. J. D., et al. (2007). Bird densities are associated with household densities. *Glob. Change Biol.* 13, 1685–1695. doi: 10.1111/j.1365-2486.2007.01390.x

Trisos, C. H., Petchey, O. L., and Tobias, J. A. (2014). Unraveling the interplay of community assembly processes acting on multiple niche axes across spatial scales. *Am. Nat.* 184, 593–608. doi: 10.1086/678233

Vellend, M., Verheyen, K., Flinn, K. M., Jacquemyn, H., Kolb, A., van Calster, H., et al. (2007). Homogenization of forest plant communities and weakening of species-environment relationships via agricultural land use. *J. Ecol.* 95, 565–573. doi: 10.1111/j.1365-2745.2007.01233.x

Vicente, J. R., Pereira, H. M., Randin, C. F., Goncalves, J., Lomba, A., Alves, P., et al. (2014). Environment and dispersal paths override life strategies and residence time in determining regional patterns of invasion by alien plants. *Perspect. Plant Ecol. Evol. Syst.* 16, 1–10. doi: 10.1016/j.ppees.2013.10.003

Villéger, S., Grenouillet, G., and Brosse, S. (2013). Decomposing functional β-diversity reveals that low functional β-diversity is driven by low functional turnover in European fish assemblages. *Glob. Ecol. Biogeogr.* 22, 671–681. doi: 10.1111/geb.12021

Villéger, S., Maire, E., and Leprieur, F. (2017). On the risks of using dendrograms to measure functional diversity and multidimensional spaces to measure phylogenetic diversity: a comment on Sobral et al. (2016). *Ecol. Lett.* 20, 554–557. doi: 10.1111/ele.12750

Watts, G., Battarbee, R. W., Bloomfield, J. P., Crossman, J., Daccache, A., Durance, I., et al. (2015). Climate change and water in the UK – past changes and future prospects. *Prog. Phys. Geogr.* 39, 6–28. doi: 10.1177/0309133314542957

Webb, C. O. (2000). Exploring the phylogenetic structure of ecological communities: An example for rain forest trees. *Am. Nat.* 156, 145–155. doi: 10.1086/303378

Webb, C. O., Ackerly, D. D., and Kembel, S. W. (2008). Phylocom: software for the analysis of phylogenetic community structure and trait evolution. *Bioinformatics* 24, 2098–2100. doi: 10.1093/bioinformatics/btn358

Webb, C. O., Ackerly, D. D., McPeek, M. A., and Donoghue, M. J. (2002). Phylogenies and community ecology. *Annu. Rev. Ecol. Syst.* 33, 475–505. doi: 10.1146/annurev.ecolsys.33.010802.150448

Weideman, E. A., Slingsby, J. A., Thomson, R. L., and Coetzee, B. T. W. (2020). Land cover change homogenizes functional and phylogenetic diversity within and among African savanna bird assemblages. *Landsc. Ecol.* 35, 145–157. doi: 10.1007/s10980-019-00939-z

Weiher, E., and Keddy, P. A. (1999). Relative abundance and evenness patterns along diversity and biomass gradients. *Oikos* 87, 335–361.

Wenny, D. G., DeVault, T. L., Johnson, M. D., Kelly, D., Sekercioglu, C. H., Tomback, D. F., et al. (2011). Perspectives in ornithology the need to quantify ecosystem services provided by birds. *Auk* 128, 1–14. doi: 10.1525/auk.2011.10248

Whelan, C. J., Wenny, D. G., and Marquis, R. J. (2008). Ecosystem services provided by birds. *Ann. N. Y. Acad. Sci.* 1134, 25–60. doi: 10.1196/annals.1439.003

Whittaker, R. H. (1960). Vegetation of the Siskiyou Mountains. Oregon and California. *Ecol. Monogr.* 30, 279–338. doi: 10.2307/1943563

Wieczynski, D. J., Boyle, B., Buzzard, V., Duran, S. M., Henderson, A. N., Hulshof, C. M., et al. (2019). Climate shapes and shifts functional biodiversity in forests worldwide. *Proc. Natl. Acad. Sci. U.S.A.* 116, 587–592. doi: 10.1073/pnas.1813723116

Wildlife Conservation Society-WCS, and Center for International Earth Science Information Network-CIESIN-Columbia University (2005). *Last of the Wild Project, version 2, 2005 (LWP-2): Global Human Influence index (HII) Dataset (geographic)*. New York, NY: Wildlife Conservation Society.

Winfree, R., Fox, J. W., Williams, N. M., Reilly, J. R., and Cariveau, D. P. (2015). Abundance of common species, not species richness, drives delivery of a real-world ecosystem service. *Ecol. Lett.* 18, 626–635. doi: 10.1111/ele.12424

Wittwer, T., O'Hara, R. B., Caplat, P., Hickler, T., and Smith, H. G. (2015). Long-term population dynamics of a migrant bird suggests interaction of climate change and competition with resident species. *Oikos* 124, 1151–1159. doi: 10.1111/oik.01559

Zabel, F., Delzeit, R., Schneider, J. M., Seppelt, R., Mauser, W., and Václavík, T. (2019). Global impacts of future cropland expansion and intensification on agricultural markets and biodiversity. *Nat. Commun.* 10, 1–10. doi: 10.1038/s41467-019-10775-z

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Ecology and Evolution

BRIEF RESEARCH REPORT
published: 29 March 2021
doi: 10.3389/fevo.2021.617356

Ability of Current Phylogenetic Clustering to Detect Speciation History

Athanasios S. Kallimanis, Maria Lazarina, Mariana A. Tsianou, Aristi Andrikou-Charitidou and Stefanos Sgardelis*

Department of Ecology, Aristotle University of Thessaloniki, Thessaloniki, Greece

Phylogenetic diversity aims to quantify the evolutionary relatedness among the species comprising a community, using the phylogenetic tree as the metric of the evolutionary relationships. Could these measures unveil the evolutionary history of an area? For example, in a speciation hotspot (biodiversity cradle), we intuitively expect that the species in the community will be more phylogenetically clustered than randomly expected. Here, using a theoretical simulation model, we estimate the ability of phylogenetic metrics of current diversity to detect speciation history. We found that, in the absence of dispersal, if the incipient species do not coexist in the region of speciation (as expected under allopatric speciation), there was no clear phylogenetic clustering and phylogenetic diversity failed to detect speciation history. But if the incipient species coexisted (sympatric speciation), metrics such as standardized effect size of Faith's Phylogenetic Diversity (PD) and of Mean Nearest Taxon Distance (MNTD) were able to identify areas of high speciation, while Mean Pairwise Distance (MPD) was a poor indicator. PD systematically outperformed MNTD. Dispersal was a game-changer. It allowed species to expand their range, colonize areas, and led to the coexistence of the incipient species originating from a common ancestor. If speciation gradient was spatially contiguous, dispersal strengthened the associations between phylogenetic clustering and speciation history. In the case of spatially random speciation, dispersal blurred the signal with phylogenetic clustering occurring in areas of low or no speciation. Our results imply that phylogenetic clustering is an indicator of speciation history only under certain conditions.

Keywords: phylogenetic diversity metrics, speciation history, simulation model, dispersal, extinction

OPEN ACCESS

Edited by:
Paulo A. V. Borges, University of the Azores, Portugal

Reviewed by:
David Eme, Unité Écologie et Modèles pour l'Halieutique, Institut Français de Recherche pour l'Exploitation de la Mer (IFREMER), France
Vanessa Pontara, Universidade Estadual de Mato Grosso do Sul, Brazil

****Correspondence:***
Athanasios S. Kallimanis
kalliman@bio.auth.gr

Specialty section:
This article was submitted to Models in Ecology and Evolution, a section of the journal Frontiers in Ecology and Evolution

Received: *14 October 2020*
Accepted: *10 March 2021*
Published: *29 March 2021*

Citation:
Kallimanis AS, Lazarina M, Tsianou MA, Andrikou-Charitidou A and Sgardelis S (2021) Ability of Current Phylogenetic Clustering to Detect Speciation History. Front. Ecol. Evol. 9:617356. doi: 10.3389/fevo.2021.617356

INTRODUCTION

Phylogenetic diversity is a facet of biodiversity that is worthy of investigation as it incorporates evolutionary history in community analyses (Webb, 2000; Mazel et al., 2016). Phylogenetic diversity of natural communities has been used to illuminate the processes governing the assembly and coexistence of species in ecological communities (Webb, 2000; Hardy and Senterre, 2007) based on the debated assumption that phylogenetic relatedness reflects ecological similarity (Tucker et al., 2018). Phylogenetic diversity has been used as a tool in the selection of conservation targets and designation of protected areas since the 1990s (Faith, 1992; Mazel et al., 2017).

This led to the development of spatial phylogenetics, with a line of research focusing on identifying centers of endemism and distinguishing between neo- and paleo-endemism (Mishler et al., 2014; Kougioumoutzis et al., 2021). This approach led to the use of phylogenetic diversity measures to infer biodiversity cradles (areas where species diversify into new species often neo-endemics) and museums (areas where species persist over a long-term evolutionary time) (Dagallier et al., 2020). Despite the intuitive appeal of this approach, the reliability of current phylogenetic diversity data to identify such historic signal of speciation in areas has received limited attention, with studies focusing mainly on the structure of phylogenetic tree (Mazel et al., 2016; Eme et al., 2020), while the factors that affect the efficiency of this surrogacy have not received critical evaluation.

Here, we aim to test if current phylogenetic diversity can detect speciation history of a region. In simpler terms, is there an association between phylogenetic clustering of the present and a community's history of speciation events? Speciation is difficult (if not impossible) to be examined empirically, since it occurs over evolutionary time frames and many other variables impact the phenomenon in complex ways. Therefore, we developed a theoretical simulation model to test if there is a relationship between phylogenetic diversity indices and history of speciation events, and if so to quantify the relative efficiency of different phylogenetic diversity indices (reflecting either the total phylogenetic history of the area or the species evolutionary pairwise distance). Since phylogenetic diversity and species richness are closely linked, we examined not the indices *per se* but their Standardized Effect Sizes, i.e., their deviation from random expectations given the species richness, indicating phylogenetic clustering. We simulated speciation, local extinction and dispersal in a theoretical model, to assess how different processes (like dispersal and local extinctions) affect the ability of phylogenetic diversity to detect speciation history. Our study focused on how these processes shape the spatial patterns of phylogenetic diversity.

METHODS

The Model

We developed a simple theoretical model simulating the processes of speciation, local extinctions, and dispersal to explore the relationship between speciation and phylogenetic diversity after 300 time steps. The model was employed to a hypothetical one-dimensional landscape, consisting of 90 cells. Cells were divided into three groups with different speciation rate reflecting a "gradient speciation landscape": (i) high speciation rate (first 30 cells), where two speciation events occurred in one randomly selected cell per time step, i.e., 20 events per cell after 300 time steps on average, (ii) low speciation rate (next 30 cells), where one speciation events occurred in a randomly selected cell, i.e., 10 events per cell after 300 time steps on average, and (iii) no speciation (last 30 cells). Speciation was spatially autocorrelated. Note that due to stochasticity, the number of speciation events varied among the cells of each group. In this landscape, we ran different sets of simulations by varying initial local species richness, size of the initial species pool, local extinction rate, and dispersal. For simplicity, we examined each variable independently keeping the values of the remaining variables fixed (50 repetitions per set). We also examined the case of "random speciation hotspots" where we had 30 cells of high speciation, 30 cells of low speciation and 30 cells of no speciation, but the spatial arrangements of these cells were random, i.e., the speciation level of a cell provided no information about the speciation level of its neighbors.

In the first step, we generated a random coalescent phylogenetic tree, using *rcoal* function of package *ape* (Paradis, 2012). Pilot studies showed that using different models for generating random phylogenetic trees influenced our results quantitatively but not qualitatively, possibly because our analysis did not focus on the structure of the phylogenetic tree but on the spatial patterns of phylogenetic diversity (**Supplementary Material**). From the initial tree we estimated the minimum branch length, and at each time step we added one fifth of this length to the branch length of all tips of our tree. All the species in the initial tree defined as the original species pool were randomly assigned to cells (local species richness), considering that all species were represented with equal probability and all cells were equally probable to host a species. In the simulations exploring the effect of the species pool size, total species richness ranged from 100 up to 1,600 species, and when fixed, the value was set to 400 species. The initial local species richness was a fraction of the original species pool and ranged between 20 and 200 species (default value = 40). All cells started off with similar number of species, and similar levels of phylogenetic diversity. Each species consisted of a number of local populations (i.e., the species presence in the different cells). All processes were simulated to operate at the local scale, i.e., speciation, extinction and dispersal affected these local populations.

The focus of our simulation is the spatial patterns of speciation and how they are reflected in the spatial patterns of phylogenetic diversity. Following the generation of the phylogenetic tree and assembly of local communities of similar levels of phylogenetic and taxonomic diversity, we simulated random speciation events by selecting one of the species in a cell and adding a new tip in the existing one of the tree with the *bind.tip* function of package *phytools* (Revell, 2012). The probability of each species in the cell to be selected was equal. In other words, as the local species richness increases the probability for a given species to undergo speciation decreases. These events occurred in a cell, but the tree was the same for all cells. Based on the classic definition, allopatric speciation corresponds to the isolation of two initial populations belonging to the same ancestral species, given that isolation is long enough to ensure reproductive isolation, results in the emergence of two new incipient species with disjunct distributions (Mayr, 1942; Gavrilets, 2003; Harrison, 2012). In our simulations, we considered as the "allopatric" speciation mode the case where isolation of a local population in one cell led to a speciation event where one of the incipient species was found in that cell while the other incipient species was found only outside that cell. The "allopatric" mode of speciation can occur effectively, only when the ancestral species occurred in at least two distinct cells. Sympatric speciation is "speciation

without geographic isolation" (Mayr, 1942). Here, we considered as "sympatric" speciation mode the case when both new incipient species occurred in the cell where the speciation event took place, with the one incipient species occurring only there and the other occurring throughout the ancestral species' extent.

Local extinction was defined as the probability of a local population to go extinct, i.e., the presence of the species in other cells was not affected by the local extinction event. At each time step, all local populations faced the same probability of extinction. Therefore, species present in many cells had lower probability to go extinct from the entire landscape than species present in only few cells. We simulated a range of local extinction probability values from 0.05% up to 1% (default value = 0.1%).

Similarly, dispersal was simulated as a local scale phenomenon. At each time step, each local population may disperse to one of its adjacent cells with a probability that was equal for all species and set at 5%. We tested other dispersal rates (5, 10, 15, and 20% probability that a local population will disperse to an adjacent cell), but the results were only regarding the role of the spatial pattern of speciation hotspots (i.e., "gradient speciation landscape" vs "random speciation hotspots"). Given that our hypothetical landscape is a torus, there was no effect of the margins. Dispersal probability equal to zero represented the case of no dispersal, in which at least one of the new incipient species was local endemic of the region where they appeared.

After 300 time steps, we pruned the phylogenetic tree to remove the extinct from the landscape species, and for each cell we estimated three phylogenetic diversity indices quantifying different facets of phylogenetic diversity: Faith's (1992) phylogenetic diversity metric (PD), mean pairwise distance (MPD), and mean nearest taxon distance (MNTD). Faith's (1992) PD reflects the cumulative evolutionary history of all species in a community, and was calculated as the sum of the branch lengths of a phylogenetic tree connecting species of each cell. MPD is a divergence index defined as the mean pairwise phylogenetic distance of all species of each cell (Webb, 2000). The MNTD quantifies the phylogenetic distance to the nearest relative of each species in the community (Webb, 2000). Given that the phylogenetic diversity indices are correlated to species richness (Tucker et al., 2018), we estimated the Standardized Effect Size (SES) of all indices using the null model *taxa.labels* that shuffles the location of each species in the phylogenetic tree retaining the community structure (199 runs). In our null models, if a certain species was not present at any cell at time step 300, it was removed from the phylogenetic tree and the relevant branch lengths were not used to compute the Standardized Effect Sizes. The calculations were performed with the *picante* package (Kembel et al., 2010).

Analysis

We evaluated the ability of phylogenetic diversity indices at time step 300 to reflect the speciation history by exploring the relationship between the SES of phylogenetic indices and the number of the speciation events that occurred in each cell with linear model analysis. The species richness in each cell is the product of speciation, extinction and dispersal, but we were interested in quantifying the ability of the index to detect speciation history of each cell. The indices' efficiency was quantified by the goodness of fit of the linear model coefficient of determination (R^2). We considered that the index did not detect speciation events in the case of non-significant relationship. To estimate the importance of the different factors in shaping the efficiency of each index we used linear regression with the index efficiency as dependent variable and the value of the different factors as independent when all other factors were fixed to their default values.

RESULTS

In the absence of dispersal, the standardized effect sizes of phylogenetic diversity indices PD and MNTD detected the speciation hotspots when speciation was simulated as "sympatric" (**Figure 1**; for PD $p < 0.0001$, $R^2 = 0.67$; for MNTD $p < 0.0001$, $R^2 = 0.57$; and for MPD $p = 0.18$, $R^2 < 0.01$), while they failed to detect the speciation signal in the case of "allopatric" speciation (**Figure 2** left column). Among the three indices analyzed, the SES of PD systematically outperformed the SES of MNTD, whereas SES of MPD was weaker and tended to give insignificant associations. Therefore, we will focus our results on the SES of PD and of MNTD.

In the absence of dispersal and the "sympatric" speciation mode, the initial local species richness (all else equal) had a conspicuous effect on the ability of phylogenetic clustering to detect speciation (for SES of PD **Figure 2B**, $p < 0.0001$, $R^2 = 0.62$, for SES of MNTD $p < 0.0001$, $R^2 = 0.54$). Given that the speciation rate (the number of speciation events) is fixed for each cell, regardless of the initial species richness of the cell, or the initial species pool size, the probability for a given species to undergo speciation is in fact decreasing for an increasing initial local species richness. On the other hand, the size of the initial species pool (all other factors fixed) did not greatly affect the efficiency of phylogenetic indices (for SES of PD **Figure 2D**, $p = 0.0007$, $R^2 = 0.03$, for SES of MNTD $p = 0.0041$, $R^2 = 0.02$), highlighting that the ratio of original species richness to speciation events matters. Perhaps, this is due to the simulation method of speciation process followed here, and using a proportional speciation rate (all species keep the same probability to speciate a given time) might lead to different results. Regarding local extinctions, given that local populations of all species had equal probability to go extinct, the effect of local extinctions on the ability of phylogenetic diversity indices to detect speciation history was significant but weak, with increased local extinction probability being associated with lower efficiency (for SES of PD **Figure 2F**, $p < 0.0001$, $R^2 = 0.17$, for SES of MNTD $p = 0.0013$, $R^2 = 0.03$).

Dispersal was a game-changer. With dispersal, the efficiency of phylogenetic clustering to detect speciation hotspots strengthened considerably (**Figure 3**) and the "allopatric" mode of speciation displayed as strong associations as the "sympatric" mode of speciation. The effect of initial local species richness (for "allopatric" mode of speciation: SES of PD **Figure 3A**, $p = 0.0040$, $R^2 = 0.04$, for SES of MNTD $p = 0.0002$, $R^2 = 0.07$; for

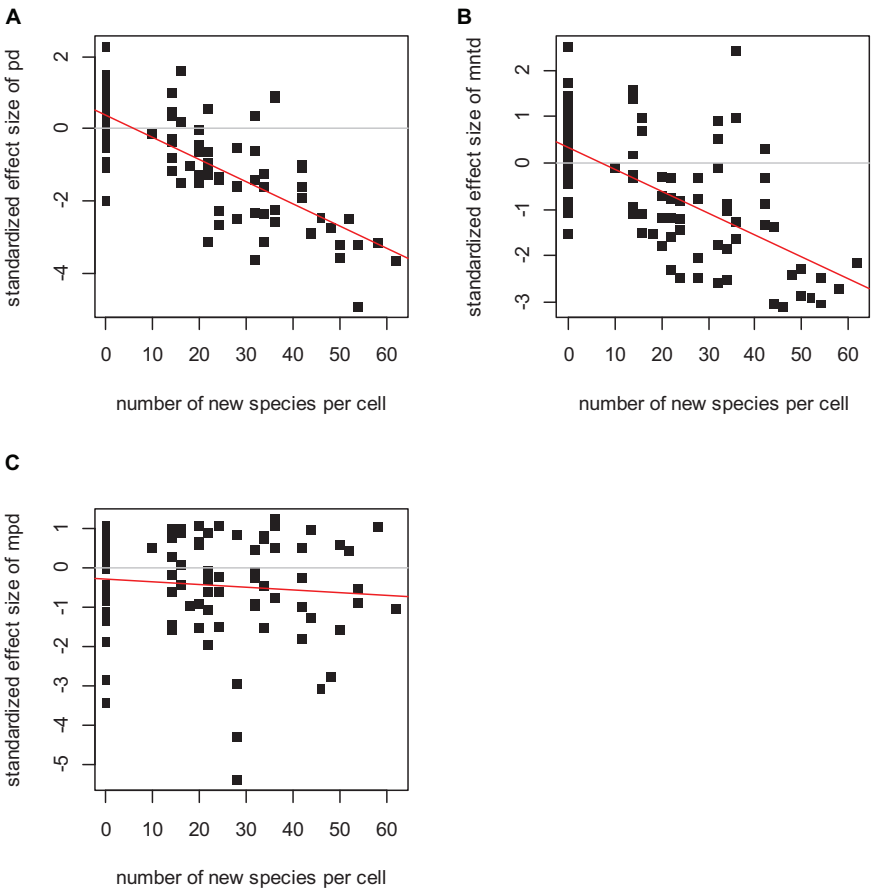


FIGURE 1 | The relationship between the standardized effect size of PD **(A)**, MNTD **(B)**, and MPD **(C)** and the number of speciation events per cell. The linear regression was significant in the case of the standardized effect size of PD ($p < 0.0001$, $R^2 = 0.67$), and of MNTD ($p < 0.0001$, $R^2 = 0.57$), but not significant in the case of MPD ($p = 0.18$, $R^2 < 0.01$). Initial species pool 400 species, initial local species richness 40 species, local population extinction probability 0.1%, no dispersal and sympatric speciation mode.

"sympatric" mode of speciation: SES of PD **Figure 3B**, $p = 0.55$, $R^2 < 0.01$, for SES of MNTD $p = 0.0275$, $R^2 = 0.02$) and initial local species pool weakened (for "allopatric" mode of speciation: SES of PD **Figure 3C**, $p = 0.0515$, $R^2 = 0.01$, for SES of MNTD $p = 0.0004$, $R^2 = 0.04$; for "sympatric" mode of speciation: SES of PD **Figure 3D**, $p = 0.55$, $R^2 < 0.01$, for SES of MNTD $p < 0.0001$, $R^2 = 0.07$). The role of local populations extinction probability did not appear to alter compared to no dispersal, with increased extinctions being associated with lower efficiency (for "allopatric" mode of speciation: SES of PD **Figure 3E**, $p < 0.0001$, $R^2 = 0.05$, for SES of MNTD $p = 0.0003$, $R^2 = 0.04$; for "sympatric" mode of speciation: SES of PD **Figure 3F**, $p < 0.0001$, $R^2 < 0.10$, for SES of MNTD $p < 0.0001$, $R^2 = 0.05$).

Dispersal plays an important role on whether phylogenetic clustering reflects speciation. Dispersal allows the two incipient species originating from the same ancestor to coexist. But dispersal could allow this coexistence to take place in areas where no speciation events occurred. So to elucidate the role of dispersal, we examined two cases when speciation hotspots were spatially contiguous, and there it became apparent that dispersal increased the effectiveness of phylogenetic clustering to detect speciation hotspots up to a point, in our case ∼5% (**Figures 4A,C**). But more illuminating is the case where speciation hotspots were randomly distributed across the landscape, where even low levels of dispersal meant that coexistence of related incipient species could occur in all areas irrespective of their speciation history, and thus phylogenetic clustering failed to detect the speciation history (**Figures 4B,D**).

DISCUSSION

Our simulations support the use of current phylogenetic clustering as indicator of a region's speciation history, but only under specific conditions. The intuitive expectation that speciation events in an area may lead to declines in phylogenetic diversity (after accounting for the species richness, i.e., phylogenetic clustering) (Davies and Buckley, 2011) applies when both the incipient species coexist in the same region of high speciation and not elsewhere. This occurs in our model when speciation is sympatric and there is no dispersal, or when there is a spatially contiguous speciation gradient and local dispersal

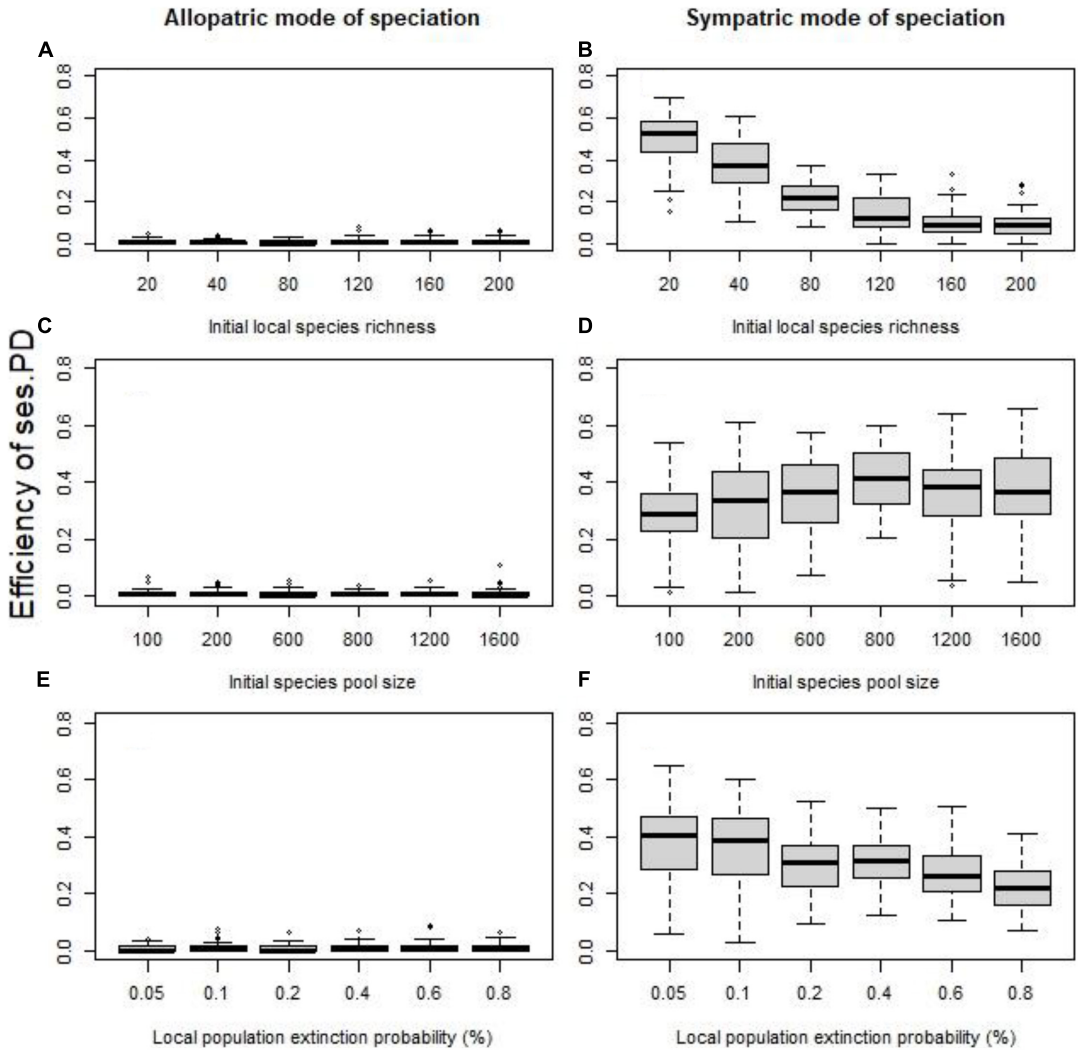


FIGURE 2 | Without dispersal, the effect of the initial local species richness for fixed species pool **(A,B)**, the size of the initial species pool for fixed species richness per cell **(C,D)**, and the probability of local extinctions **(E,F)** on the efficiency of the standardized effect size of PD as indicator of speciation history (quantified using the coefficient of determination R^2). The allopatric speciation mode results are in the left column, while the sympatric speciation mode results in the right column. To assess the role of each factor independently we kept all the other values fixed at their default values (initial local richness = 40 species, species pool size = 400 species, probability of local extinction = 0.1%.

allows the introduction of the second incipient species in the region of the first incipient species and vice versa.

In allopatric speciation, geographical isolation of a species in a region leads to local differentiation that over sufficient evolutionary time results in reproductive isolation from the initial species and hence, the ancestral species leads to two new incipient species with disjunct distributions (Harrison, 2012). In this case, the two incipient species do not coexist in the same region (Wen et al., 2014). Therefore, the total phylogenetic diversity of the community is not significantly affected, and the phylogenetic diversity indices cannot detect the speciation gradient, unless the second incipient species is introduced in the region. On the other hand, the two incipient species coexist in the area in the sympatric speciation (Skeels and Cardillo, 2019). Thus, overall phylogenetic diversity of the area is significantly lower than random assemblages with the same species richness and SES of PD and MNTD phylogenetic diversity indices capture the speciation signal. The PD systematically (albeit slightly) outperformed MNTD in terms of efficiency, while MPD index weakly detected the speciation signal and was an unreliable indicator. This difference in efficiency may be due to MPD being more strongly influenced by the deep branching structure of the tree and thus by basal clustering rather than what happens in the tips of the phylogenetic tree, in contrast to PD and MNTD that are less sensitive to internal branching and therefore more sensitive to patterns occurring at the tips of the tree, like recent speciation events (Mazel et al., 2016).

In our model, extinction was simulated as a local phenomenon that did not differentiate between species and was spatially

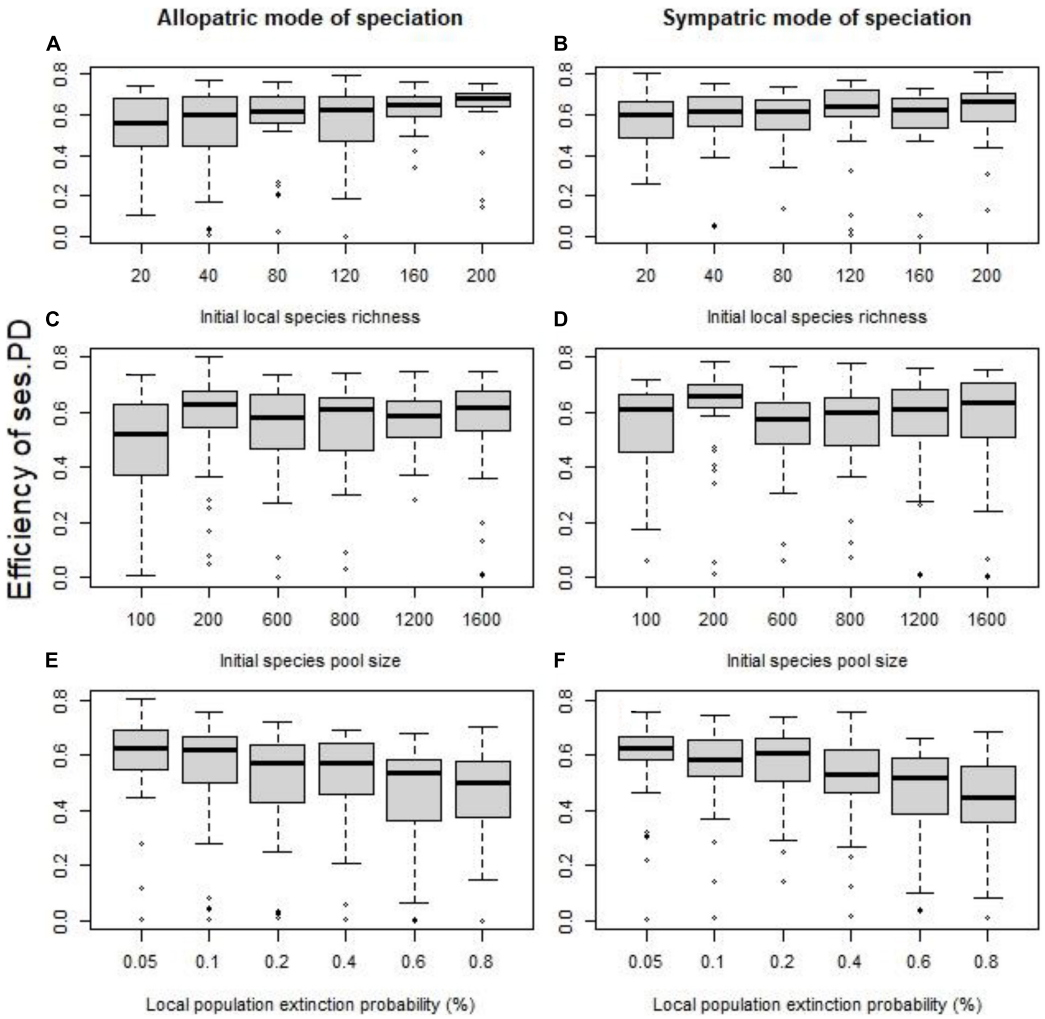


FIGURE 3 | With a 5% probability of dispersal for each local population to its nearest neighbor cell, the effect of the initial local species richness for fixed species pool **(A,B)**, the size of the initial species pool for fixed species richness per cell **(C,D)**, and the probability of local extinctions **(E,F)** on the efficiency of the standardized effect size of PD as indicator of speciation history (quantified using the coefficient of determination R^2). The allopatric speciation mode results are in the left column, while the sympatric speciation mode results in the right column. To assess the role of each factor independently we kept all the other values fixed at their default values (initial local richness = 40 species, species pool size = 400 species, probability of local extinction = 0.1%).

uniform (i.e., we did not include in our simulations extinction gradients or biodiversity museums). Spatially uniform extinction lowered the total species richness of each region and decreased the size of the total species pool, making deviations from random expectations more difficult to detect due to small numbers. So, the efficiency of phylogenetic diversity to detect speciation declined but the effect was not very strong.

In the absence of dispersal, the efficiency of phylogenetic clustering to detect speciation gradients becomes more pronounced as the ratio of speciation events to original species richness increases. It is this ratio, rather than the size of the total species pool throughout the landscape, that significantly affects the efficiency of phylogenetic indices. At the community level, if speciation events are few compared to the community's total species richness, their contribution to phylogenetic diversity patterns is very small. Therefore, the standardized effect size of indices is closer to random expectations.

Dispersal plays an important role in this phenomenon changing the picture drastically. Dispersal provides a rescue effect to counteract local extinctions, as well as means for range expansion, thus nullifying the effect of initial species richness. In allopatric speciation, the new incipient species originating from the same ancestral species do not coexist, and thus they do not affect phylogenetic clustering. But even moderate levels of dispersal, will allow the introduction of the absent incipient species in the speciation hotspot leading to phylogenetic clustering. But simultaneously, dispersal will allow the incipient species to expand their range and colonize new areas. This reinforces the possibility that the incipient species originating from the same ancestral species will coexist in areas with no speciation, and thus phylogenetic clustering may

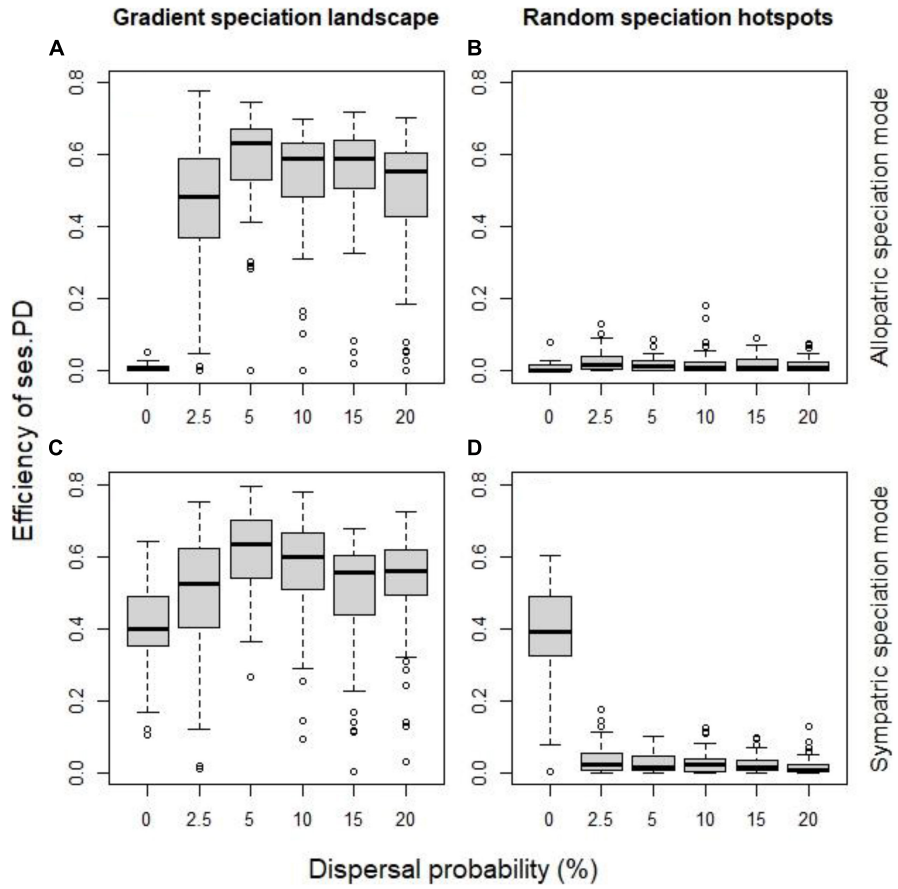


FIGURE 4 | The effect of dispersal probability on the efficiency of the standardized effect size of PD as indicator of speciation history, for allopatric mode of speciation **(A,B)** and the sympatric mode **(C,D)**, when speciation hotspots were spatially continuous **(A,C)** and when they were random **(B,D)**. All other factor fixed at their default values (initial local richness = 40 species, species pool size = 400 species, probability of local extinction = 0.1%).

be observed in areas of low or no speciation. This renders the role of spatial pattern of speciation very important. If speciation occurs in a spatially autocorrelated way (e.g., along geographic latitude), then areas adjacent to speciation hotspots will also be speciation hotspots (albeit for other species), and thus the phylogenetic clustering in a speciation hotspot will increase, even if this increase is not due to speciation *per se*. If speciation occurs in a spatially random way the results will be drastically different, and dispersal will blur the signal of speciation leading to phylogenetic clustering in areas of no speciation, making phylogenetic clustering an unreliable indicator of past speciation history.

Our model is the simplest model, i.e., the model with the fewer assumptions that allowed us to disentangle the effect of speciation. All other processes (initial biodiversity, extinction, dispersal) were spatially uniform. Our model did not incorporate any spatial structure in the original distribution of the species (but dispersal led to spatially autocorrelated species ranges). Also, species traits (e.g., extinction risk, dispersal, speciation probability) were equal among species and had no phylogenetic signal. Future work will look into relaxing these simplifying assumptions and investigate their importance.

Concluding, the use of phylogenetic clustering of current biodiversity as an indicator of the regions' speciation history should be done with caution, since allopatric speciation does not lead to phylogenetic clustering in the absence of dispersal; and if there is dispersal and past speciation is not spatially autocorrelated, phylogenetic clustering may occur in regions of low or no speciation.

DATA AVAILABILITY STATEMENT

The datasets presented in this article are not readily available because the data are the results of theoretical simulations. Requests to access the datasets should be directed to AK, kalliman@bio.auth.gr.

AUTHOR CONTRIBUTIONS

AK planned and executed the experiments, analyzed data, and wrote the manuscript draft. ML, MT, AA-C, and SS assisted in

manuscript draft. All authors contributed to the article and approved the submitted version.

FUNDING

This research was supported by the Hellenic Foundation for Research and Innovation (H.F.R.I.) under the "First Call for H.F.R.I. Research Projects to support Faculty members and Researchers and the procurement of high-cost research equipment grant" (Project Number: HFRI-FM17-2024 Mapping Functional Diversity Drivers, Impacts and Threats—MAPFUN).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fevo.2021.617356/full#supplementary-material

REFERENCES

Dagallier, L. P. M., Janssens, S. B., Dauby, G., Blach-Overgaard, A., Mackinder, B. A., Droissart, V., et al. (2020). Cradles and museums of generic plant diversity across tropical Africa. *New Phytol.* 225, 2196–2213. doi: 10.1111/nph.16293

Davies, T. J., and Buckley, L. B. (2011). Phylogenetic diversity as a window into the evolutionary and biogeographic histories of present-day richness gradients for mammals. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 366, 2414–2425. doi: 10.1098/rstb.2011.0058

Eme, D., Anderson, M. J., Myers, E. M. V., Roberts, C. D., and Liggins, L. (2020). Phylogenetic measures reveal eco-evolutionary drivers of biodiversity along a depth gradient. *Ecography* 43, 689–702. doi: 10.1111/ecog.04836

Faith, D. P. (1992). Conservation evaluation and phylogenetic diversity. *Biol. Conserv.* 61, 1–10. doi: 10.1016/0006-3207(92)91201-3

Gavrilets, S. (2003). Perspective: models of speciation: What have we learned in 40 years? *Evolution* 57, 2197–2215. doi: 10.1111/j.0014-3820.2003.tb00233.x

Hardy, O. J., and Senterre, B. (2007). Characterizing the phylogenetic structure of communities by an additive partitioning of phylogenetic diversity. *J. Ecol.* 95, 493–506. doi: 10.1111/j.1365-2745.2007.01222.x

Harrison, R. G. (2012). The language of speciation. *Evolution* 66, 3643–3657. doi: 10.1111/j.1558-5646.2012.01785.x

Kembel, S. W., Cowan, P. D., Helmus, M. R., Cornwell, W. K., Morlon, H., Ackerly, D. D., et al. (2010). Picante: R tools for integrating phylogenies and ecology. *Bioinformatics* 26, 1463–1464. doi: 10.1093/bioinformatics/btq166

Kougioumoutzis, K., Kokkoris, I. P., Panitsa, M., Kallimanis, A., Strid, A., and Dimopoulos, P. (2021). Plant endemism centres and biodiversity hotspots in Greece. *Biology* 10:72. doi: 10.3390/biology10020072

Mayr, E. (1942). *Systematics and the Origin of Species*. New York, NY: Columbia University Press.

Mazel, F., Davies, T. J., Gallien, L., Renaud, J., Groussin, M., Münkemüller, T., et al. (2016). Influence of tree shape and evolutionary time-scale on phylogenetic diversity metrics. *Ecography* 39, 913–920. doi: 10.1111/ecog.01694

Mazel, F., Mooers, A. O., Riva, G. V. D., and Pennell, M. W. (2017). Conserving phylogenetic diversity can be a poor strategy for conserving functional diversity. *Syst. Biol.* 66, 1019–1027. doi: 10.1093/sysbio/syx054

Mishler, B. D., Knerr, N., González-Orozco, C. E., Thornhill, A. H., Laffan, S. W., and Miller, J. T. (2014). Phylogenetic measures of biodiversity and neo-and paleo-endemism in Australian Acacia. *Nat. Commun.* 5:4473.

Paradis, E. (2012). *Analysis of Phylogenetics and Evolution with R*, Second Edn. New York, NY: Springer.

Revell, L. J. (2012). phytools: R package for phylogenetic comparative biology (and other things). *Methods Ecol. Evol.* 3, 217–223. doi: 10.1111/j.2041-210X.2011.00169.x

Skeels, A., and Cardillo, M. (2019). Reconstructing the geography of speciation from contemporary biodiversity data. *Am. Nat.* 19, 240–255. doi: 10.1086/701125

Tucker, C. M., Davies, T. J., Cadotte, M. W., and Pearse, W. D. (2018). On the relationship between phylogenetic diversity and trait diversity. *Ecology* 99, 1473–1479. doi: 10.1002/ecy.2349

Webb, C. O. (2000). Exploring the phylogenetic structure of ecological communities: an example for rain forest trees. *Am. Nat.* 156, 145–155. doi: 10.1086/303378

Wen, J., Zhang, J., Nie, Z. L., Zhong, Y., and Sun, H. (2014). Evolutionary diversifications of plants on the Qinghai-Tibetan Plateau. *Front. Genet.* 5:4. doi: 10.3389/fgene.2014.00004

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Ecology and Evolution

ORIGINAL RESEARCH
published: 07 April 2021
doi: 10.3389/fevo.2021.626730

The Shape of Species Abundance Distributions Across Spatial Scales

Laura H. Antão[1,2,3]†, Anne E. Magurran[1†] and Maria Dornelas[1†]*

[1] Centre for Biological Diversity, University of St Andrews, St Andrews, United Kingdom, [2] Department of Biology and CESAM, Universidade de Aveiro, Aveiro, Portugal, [3] Organismal and Evolutionary Biology Research Programme, Research Centre for Ecological Change, University of Helsinki, Helsinki, Finland

Species abundance distributions (SADs) describe community structure and are a key component of biodiversity theory and research. Although different distributions have been proposed to represent SADs at different scales, a systematic empirical assessment of how SAD shape varies across wide scale gradients is lacking. Here, we examined 11 empirical large-scale datasets for a wide range of taxa and used maximum likelihood methods to compare the fit of the logseries, lognormal, and multimodal (i.e., with multiple modes of abundance) models to SADs across a scale gradient spanning several orders of magnitude. Overall, there was a higher prevalence of multimodality for larger spatial extents, whereas the logseries was exclusively selected as best fit for smaller areas. For many communities the shape of the SAD at the largest spatial extent (either lognormal or multimodal) was conserved across the scale gradient, despite steep declines in area and taxonomic diversity sampled. Additionally, SAD shape was affected by species richness, but we did not detect a systematic effect of the total number of individuals. Our results reveal clear departures from the predictions of two major macroecological theories of biodiversity for SAD shape. Specifically, neither the Neutral Theory of Biodiversity (NTB) nor the Maximum Entropy Theory of Ecology (METE) are able to accommodate the variability in SAD shape we encountered. This is highlighted by the inadequacy of the logseries distribution at larger scales, contrary to predictions of the NTB, and by departures from METE expectation across scales. Importantly, neither theory accounts for multiple modes in SADs. We suggest our results are underpinned by both inter- and intraspecific spatial aggregation patterns, highlighting the importance of spatial distributions as determinants of biodiversity patterns. Critical developments for macroecological biodiversity theories remain in incorporating the effect of spatial scale, ecological heterogeneity and spatial aggregation patterns in determining SAD shape.

Keywords: spatial scale, biodiversity, community structure, multimodality, macroecology, maximum entropy theory of ecology, neutral theory of biodiversity and biogeography

OPEN ACCESS

Edited by:
Luís Borda-de-Água, Universidade do Porto, Portugal

Reviewed by:
Tom Matthews, University of Birmingham, United Kingdom
David Storch, Charles University, Czechia

****Correspondence:***
Laura H. Antão
laura.antao@helsinki.fi

†ORCID:
Laura H. Antão
orcid.org/0000-0001-6612-9366
Anne E. Magurran
orcid.org/0000-0002-0036-2795
Maria Dornelas
orcid.org/0000-0003-2077-7055

Specialty section:
This article was submitted to Models in Ecology and Evolution, a section of the journal Frontiers in Ecology and Evolution

Received: *06 November 2020*
Accepted: *08 March 2021*
Published: *07 April 2021*

Citation:
Antão LH, Magurran AE and Dornelas M (2021) The Shape of Species Abundance Distributions Across Spatial Scales. Front. Ecol. Evol. 9:626730. doi: 10.3389/fevo.2021.626730

INTRODUCTION

Species Abundance Distributions (SADs) describe the relative abundance of the species within a community. Looking at the whole distribution allows accounting for different aspects that univariate metrics measure separately and readily integrating concepts such as rarity and dominance (Magurran, 2004; McGill et al., 2007). SADs are thus a synthetic measure of biodiversity

and community structure, and a key component of biodiversity theory and research. The two most common distributions used to describe SADs are the logseries (Fisher et al., 1943) and the lognormal (Preston, 1948). These distributions differ mainly in the proportion of rare species they predict; specifically, the logseries has no internal mode, with singletons as the modal class, whereas an unveiled lognormal has one internal mode for species with intermediate abundances (Fisher et al., 1943; Preston, 1948). Additionally, multimodal SADs, i.e., with more than one mode of abundance can also occur (Gray et al., 2006; Dornelas and Connolly, 2008). While this pattern had been mostly disregarded, a recent empirical meta-analysis showed that not only is multimodality more common than previously recognized, it is also more likely to occur for communities encompassing larger spatial extents or with higher taxonomic diversity (Antão et al., 2017). Despite decades of research, multiple hypotheses and theories proposed to explain SADs, a thorough understanding of what drives SAD shape is still lacking (Hubbell, 2001; Connolly et al., 2005; Green and Plotkin, 2007; McGill et al., 2007); this gap is particularly apparent for large spatial scales (Enquist et al., 2019; Fukaya et al., 2020). A systematic empirical assessment of how SAD shape varies with scale is needed to improve our understanding of the mechanisms underpinning SAD shape, and consequently community structure. Given the pivotal role of SADs in biodiversity research, these insights will further facilitate assessments of ongoing biodiversity change.

Understanding how SAD shape varies with sampling scale is a long standing question in ecology (Fisher et al., 1943; Preston, 1948; McGill et al., 2007; Zillio and He, 2010). Both sampling effects and spatial scale are known to affect SAD shape (Fisher et al., 1943; Preston, 1948; Pielou, 1977; Hubbell, 2001; Connolly et al., 2005; Green and Plotkin, 2007). Two approaches to understanding the scaling properties of SADs can be used: downscaling—i.e., try to predict the shape of SADs at smaller scales from the regional scale (e.g., Hubbell, 2001; Green and Plotkin, 2007), and upscaling—i.e., try to infer the SAD for larger spatial scales by pooling smaller scale SADs (Šizling et al., 2009b; Zillio and He, 2010; Borda-de-Água et al., 2012). However, analyzing SADs across different spatial scales has remained largely unexplored (but see Rosindell and Cornell, 2013). As there is no *single ideal* scale for studying SADs (Wiens, 1989; Levin, 1992), systematically assessing SADs along spatial scale gradients allows us to make stronger inferences about patterns of commonness and rarity across scales, and can potentially provide insights to help disentangle which processes are relevant at different scales.

Generally, SAD studies have relied on sampling theory approaches to address the relationship between the large regional community and local scale SADs. For instance, Fisher et al. (1943) viewed the logseries as resulting from a random sample from a gamma distribution (i.e., SADs at smaller scales are random subsamples of SADs at larger scales). The "veil line" proposed by Preston (1948) was a first attempt to explain that the absence of rare species in small samples would lead to a truncation of the "true" underlying lognormal distribution, which would gradually disappear with increasing sampling. The lognormal has since been a particularly prominent SAD model, emerging as the statistical expectation of the Central Limit Theorem (i.e., SADs result from random multiplicative processes acting on species abundances), and from population dynamics and niche partitioning models (Preston, 1948; May, 1975; Magurran, 2004; McGill et al., 2007). However, unveiling does not simply reveal the left-end of the distribution by rigidly moving the veil, but the shape of the overall distribution also changes (Pielou, 1977; Dewdney, 1998; McGill, 2003b), which can affect the overall proportion of the rarest species. For instance, McGill (2003b) showed that pooling repeated autocorrelated small samples can lead to the log-left-skew reported in many empirical SADs, i.e., the existence of more rare species than predicted by a lognormal distribution (Hubbell, 2001). This phenomenon can nonetheless be driven by biological mechanisms, where SAD shape reflects changes in community structure, such as the signature of core-transient species temporal dynamics (Magurran and Henderson, 2003).

Crucially, SAD shape is affected by how species are distributed in space. One of the fundamental patterns in ecology is that individuals are not randomly distributed in space (Condit et al., 2000; McGill, 2010), with both theoretical and empirical studies evaluating these effects. Spatial aggregation patterns and turnover affect SAD shape when upscaling from smaller areas (Šizling et al., 2009a). Green and Plotkin (2007) developed a statistical sampling theory for SAD incorporating conspecific spatial aggregation patterns. They showed that when sampling from regional SADs with randomly distributed populations (i.e., Poisson sampling), the sampled SADs would exhibit the same functional form as the regional SAD. In contrast, using a more realistic description of species spatial aggregation patterns (i.e., negative binomial sampling), sampled SADs diverged from the regional SAD. Specifically, this conspecific spatial aggregation led to sampled SADs skewed toward both rare and more abundant species (Green and Plotkin, 2007). Interspecific differences in aggregation rates were also suggested to produce bimodal abundance distributions (Alonso et al., 2008), which in turn partially explain the existence of multiple modes (two, but not the three modes) in a coral SAD (Dornelas and Connolly, 2008). Subsequently, and taking a completely different approach while attempting to upscale SADs, Borda-de-Água et al. (2012) similarly predicted a bimodal SAD for larger scales without including any information on species aggregation patterns. In this case, bimodality arises from an increase in the number of rare species with area (Borda-de-Água et al., 2002, 2012), with one mode occurring for the singletons class and another mode for intermediate abundance classes.

Two unified theories of biodiversity in particular make explicit predictions for SAD shape at different scales—the Neutral Theory of Biodiversity (NTB; Hubbell, 2001) and the Maximum Entropy Theory of Ecology (METE; Harte et al., 2008). Both theoretical frameworks can be thought of as null models, as they assume demographic equivalence between individuals (NTB), or do not include explicit ecological mechanisms (METE). While not the only theories of SADs, both provide null expectations for what SAD shape should emerge at different scales. Systematic discrepancies between empirical data and theoretical predictions can help identify important mechanisms that need to be

accounted if we are to improve our ability to make stronger inferences about the processes driving SAD shape. Below, we provide only a brief outline of both theories' assumptions relevant for our analysis, given the extensive literature devoted to them.

NTB assumes all individuals in an assemblage have equal demographic rates of birth, death, dispersal and speciation, irrespective of species identity (Hubbell, 2001), with stochastic drift and dispersal limitation as the processes explaining patterns of species occurrence and abundance. The spatially implicit NTB model includes two distinct spatial scales: a local community that consists of a dispersal-limited sample, and the metacommunity (larger scale) from which this sample is taken. In the original model, the metacommunity follows a logseries distribution and the local community follows a zero-sum multinomial distribution (ZSM), which includes fewer rare species than the logseries and resembles a left-skewed lognormal distribution (Hubbell, 2001; Rosindell et al., 2011). This latter distribution is controversial, with numerous studies comparing ZSM and lognormal performances on fitting empirical SADs (Hubbell, 2001; McGill, 2003a; Volkov et al., 2003, 2007; Dornelas et al., 2006). Subsequent model developments also yield a logseries SAD for the largest scale [e.g., spatially explicit model (Rosindell and Cornell, 2013); protracted speciation mode (Rosindell et al., 2010)].

METE is a spatially explicit theory of biodiversity based on the principle of maximization of information entropy (MaxEnt). METE only requires knowledge on four state variables to describe ecological communities: the area of an ecosystem, species richness, the total number of individuals, and total metabolic rate (the latter has been disregarded when analyzing SADs; Harte et al., 2008; Harte and Newman, 2014). MaxEnt rationale is that the least-biased inference of the shape of a probability distribution is as smooth and flat as possible, given the constraints (Harte et al., 2008). Using only these four state variables and without incorporating any specific ecological mechanisms, the most likely distributions for several macroecological patterns are found by maximizing information entropy. The logseries is the SAD distribution that emerges across spatial scales (Harte et al., 2008; Harte and Newman, 2014). Therefore, assessing the performance of the logseries for large scale SADs, as well as for different taxa, provides a relevant test both on current neutral models and a critical assessment of METE's expectation of a scale-independent SAD. On the other hand, neither theory accounts for multimodal SADs.

Here, we systematically assessed empirical SADs shape across a gradient in spatial scale spanning several orders of magnitude and for different taxonomic groups spanning both marine and terrestrial realms. We assessed the effect of sampled area, taxonomic diversity (species and family richness) and total abundance on SAD shape. We aimed to compare predictions for SAD shape from macroecological theories, with sampling theory predictions and empirical patterns. Specifically, we contrasted expectations of a better fit of the logseries as outlined above following NTB or METE predictions (either for larger scales or invariant with scale, respectively) (1), with the expected better performance of lognormal distributions for larger samples (2), and further with recent empirical results showing that multimodality occurs with higher prevalence for larger areas or more diverse communities (3).

MATERIALS AND METHODS

Empirical Data

We analyzed 11 datasets sampled over large extents for different taxa, namely trees, birds, fish and benthos (**Table 1**). We selected datasets from the BioTIME database (Dornelas et al., 2018) with spatial extent larger than 150,000 km^2 and for which the unique sampling locations were distributed across the study area so that the random splitting of the total extent would not result in portions without sampling locations (see below). Importantly, for all these datasets, samples were consistently collected using standardized methods (e.g., plots, transects or tows), where each sample consists of records of species and their abundance in a given time and location (more information can be found in each dataset's original metadata, or from the BioTIME database). For each dataset, we used data corresponding to one year of sampling only, selecting the year with the most and more evenly distributed sampling locations. We further analyzed the Forest Inventory and Analysis Database (FIA; USDA Forest Service, 2010; Woudenberg et al., 2010)[1], as we wanted to include a tree community data in our analysis to ensure the results are robust across a broad range of taxonomic groups. We obtained the latter data via the EcoData Retriever (Morris and White, 2013; McGlinn and White, 2015)[2], and selected data from 2013 only. For each dataset, information of the taxonomic family corresponding to each species was retrieved. These empirical datasets cover a wide range of sampling grains (0.0001 to 25.4 km^2) and total spatial extents (167,455 to 16,663,141 km^2).

Implementing the Scale Gradient

All analyses were performed in the statistical software R (R Core Team, 2017). We established a scale gradient by using the fixed extent of the study area of each dataset and systematically partitioning this area into smaller portions as follows. We drew a circle encompassing all the sampling locations from each dataset and centered on the centroid of the sampling locations. A random point from the circle was selected to split the circle into halves, thirds, quarters, eights and sixteenths, using the initial random point from the bisection as reference (**Figure 1**). This process ensures the spatial relationship between the sections and the sampling locations is maintained. Further details on implementing the scale gradient can be found in Antão et al. (2019). Within each section, species abundances were pooled across all the individual samples to build the species abundance distributions. Thus, one SAD was calculated for the total extent (i.e., the largest level in our scale gradient), two SADs for the bisection level, three for the third level, and so forth. At each level, each section was annotated with the area, species richness (S), total abundance (N), and total number of families, to assess the effect of taxonomic diversity on SAD shape. The areas sampled

[1] http://fia.fs.fed.us/

[2] http://data-retriever.org

TABLE 1 | Community data used and data sources.

Dataset title	Abbreviation	Taxon	Year/usage notes	Number of species	Number of samples	References
Forest inventory analysis	FIA	Trees	2013; excluded Alaska	305	19,427	USDA Forest Service, 2010; Woudenberg et al., 2010
North American breeding bird survey	BBS	Birds	2015; USA data only (excluding Alaska)	521	2,420	Pardieck et al., 2016
Maritimes breeding bird atlas	MBBA	Birds	2009	163	3,243	NatureCounts, 2012a
Landbird monitoring program	LBMP	Birds	2004	229	5,107	USFS, 2012
Ontario breeding bird atlas	OBBA	Birds	2003	233	19,611	NatureCounts, 2012b
East Coast North America strategic assessment	ECNASAP	Fish	1994	110	2,101	Brown et al., 2005
Reef life survey: global reef fish dataset	RLS_F	Fish	Spatial subset around Australia	1,847	6,666	Edgar and Stuart-Smith, 2014a,b
ICES North sea international bottom trawl survey for commercial fish species	NSIBT	Fish	2011	131	688	DATRAS, 2010a
Irish ground fish survey for commercial fish species	IGFS	Fish	2004	100	163	DATRAS, 2010b
Reef life survey: invertebrates	RLS_I	Invertebrates	Spatial subset around Australia	1,013	6,817	Edgar and Stuart-Smith, 2008, 2014b
Snow crab research trawl survey database	SCRT	Benthos	2009	32	354	Wade, 2011

For each community the taxon, species richness and number of unique sample locations are shown.

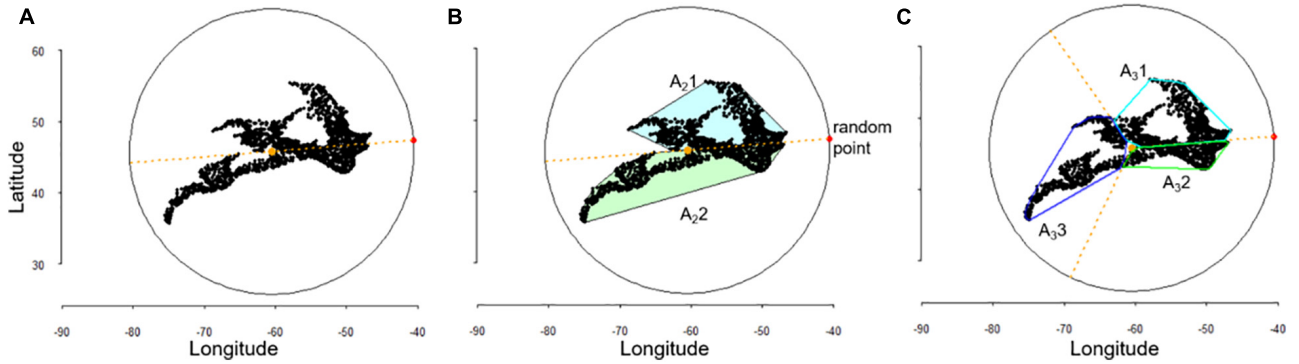


FIGURE 1 | Schematic representation of the scale gradient, showing an example of how the encompassing circle was drawn to include all the sampling locations and a random point was selected to establish bisections **(A,B)** and thirds **(C)**.

were calculated using convex hull polygons encompassing the sampling locations within each section, using package rgeos (Bivand and Rundel, 2016; **Figure 1**).

Model Fitting and Analysis

Along the scale gradient and for each section, each SAD was fitted with four alternative models, following the method implemented in Antão et al. (2017). Specifically, we employed maximum likelihood methods to explicitly compare the fit of logseries distributions (Fisher et al., 1943) and of mixtures of 1, 2, and 3 Poisson Lognormal distributions (1PLN, 2PLN, and 3PLN, respectively), where 2PLN and 3PLN are multimodal distributions (Pielou, 1969; Bulmer, 1974). We did not include the ZSM distribution in our comparisons, since our aim was to evaluate changes in SADs overall shape, rather than explicitly test NTB models or focus on detailed comparisons between alternative models fitting. This is further justified given: (1) the plethora neutral models, implementations and assumptions, without a clear way forward (Hubbell, 2001; McGill, 2003b; Etienne, 2005, 2007, 2009; Dornelas et al., 2006; Rosindell et al., 2011), (2) PLN mixtures can provide suitable alternatives to ZSM (Gray et al., 2005), and (3) ZSM is usually associated with scales overall much smaller than the lowest levels in our scale gradient. For each model above, best-fit parameters were found by minimizing the negative log-likelihood; parameter estimation was performed using the optimization routine nlminb(), with searches initialized from multiple starting points due to the possibility of several local maxima for more complex distributions (Dornelas and Connolly, 2008; Antão et al., 2017). All the fitting routines were run on non-binned data, and we only binned data for plotting purposes—SADs are traditionally plotted as histograms of the number of species as a function of abundance on a log2 scale, conveying the intuitive approach of doubling classes of abundance, called octaves (Preston, 1948; Gray et al., 2006). The second order Akaike's information criterion for small sample

sizes (AIC_c, Burnham and Anderson, 2002) was used for model selection. AICc was used throughout as it converges to AIC when the sample size is large (Burnham and Anderson, 2002), and previous work with simulated communities testing this *PLN-mixture* method has shown that BIC is too conservative and can be insensitive to deviations in SAD shape (Antão et al., 2017). Furthermore, because we were not interested in detecting multimodality *per se*, but rather in detecting changes in SAD shape, we assumed the best model to be the one selected by AIC_c, regardless of support level. We plotted smoothed density estimates of the model selected for each section as a function of the relevant variables across the scale gradient, i.e., sampling level, area sampled, species richness, total abundance and number of families, using the package ggplot2 (Wickham, 2009). These assessments were built for each community individually and for all the SADs together to provide an overview of how these variables affect SAD shape across the different datasets analyzed. In addition to using AIC_c as a model selection criterion, we quantified the deviations between the empirical SADs and the predictions of each model, by comparing the observed and expected number of species per octave.

RESULTS

We built 374 SADs, which included 4,684 species and over 142.5 million individuals. Overall, there was a higher prevalence of multimodal SADs for larger areas and for more taxonomically diverse datasets (**Figures 2**, **3** and **Supplementary Figure 1**), although some smaller areas or less diverse communities were also selected as multimodal. The logseries was never selected as best model for the total extent SADs, and was only selected for much smaller areas, and when species richness or number of families were proportionally much smaller (**Figures 2**, **3**). That is, only 1PLN, 2PLN, or 3PLN provided adequate fit for the total extent SADs, with a better performance for multimodal models. As area sampled decreases, both species richness and total number of individuals are also expected to decrease. However, while species richness showed a similar effect to that of area on model selection, there was no clear pattern for total number of individuals across the datasets analyzed (**Figure 3**).

For the SADs selected as multimodal at the total extent, multimodal models most often provided the best fit across the scale gradient. This was the case for the FIA tree inventory, the BBS and OBBA bird, and the RLS fish datasets (**Figure 2** and **Supplementary Figure 1**). The average ΔAIC_c for multimodality vs. non-multimodality across the scale gradient was 5.53 for FIA, 11.01 for BBS, 6.39 for OBBA and 9.77 for RLS_Fish (calculated as ($\min AIC_{c\ 2PLN/3PLN} - \min AIC_{c\ 1PLN/logser}$) across all sections). Additionally, some datasets exhibited the expected pattern of progressing from multimodality to either 1PLN or logseries as sampled area decreased (MBBA and LBMP birds and IGFS fish datasets; **Figure 2** and **Supplementary Figure 1**). The datasets that were better fit by 1PLN at the total extent showed some variability in the best fit models as area decreased. For these, 1PLN was still frequently selected as best model, but both logseries and multimodal models were selected for smaller scales (ECNASAP and NSIBT fish, SCRT benthos and RLS invertebrate datasets; **Figure 2** and **Supplementary Figure 1**).

Deviations between the empirical SADs and each model's predictions support the results above. For the datasets with consistent support for multimodality across the scale gradient, logseries consistently and severely overestimated the number of singletons and rare species (i.e., octaves 1–2) across the scale gradient, while 1PLN often underestimated them. In addition, both logseries and 1PLN either over- or underestimated the number of species with intermediate to high abundances (**Supplementary Figure 2**). On average, for these assemblages, deviations were smaller for 2 or 3PLN at every scale (FIA, BBS, OBBA, and RLS_Fish; **Supplementary Figure 2**). For the remaining multimodal SADs at the total extent, logseries again overestimated the number of rare species, while the PLN mixtures exhibited large deviations between the observed number of species and the models' predictions across the distribution and across the scale gradient (**Supplementary Figure 2**). For the datasets better fit by 1PLN at the total extent, for RLS invertebrates, deviations are much smaller on average for 2PLN at every scale, while both logseries and 1PLN underestimated the number of rare species. For the remaining SADs, there was no clear pattern, but logseries was systematically unable to accurately predict the rarest and intermediate abundance classes (**Supplementary Figure 2**).

DISCUSSION

Our systematic assessment of empirical SADs shape across a wide scale gradient and taxa showed consistent variation in SAD shape. Furthermore, our results support previous findings of higher prevalence of multimodal SADs for larger areas or more taxonomically diverse communities (Antão et al., 2017), while the logseries never provided an adequate fit for larger and more diverse communities. In addition, we revealed a clear effect of area and taxonomic diversity in determining SAD shape, and a non-directional pattern for total abundance (for the aggregate communities' results). Our findings clearly depart from two important macroecological theories predictions for the SAD. We discuss the implications of our results in turn below.

Variability in SAD Shape Across Spatial Scales

We compared the performance of different models to describe the SAD across a wide scale gradient to assess how the relative abundance patterns depend on spatial scale and also on taxonomic diversity (measured as species and family richness). For the communities selected as multimodal at the total extent, multimodality was consistently and strongly selected across the scale gradient, even for intense sampling effects of area and diversity. This indicates that multimodality is a robust feature of SADs and is indeed reflecting the structure of the underlying communities, rather than being a sampling (c.f. Barabás et al., 2013) or scaling artifact. The sections created across the scale gradient for these spatially large and taxonomically diverse datasets (FIA tree inventory, BBS and

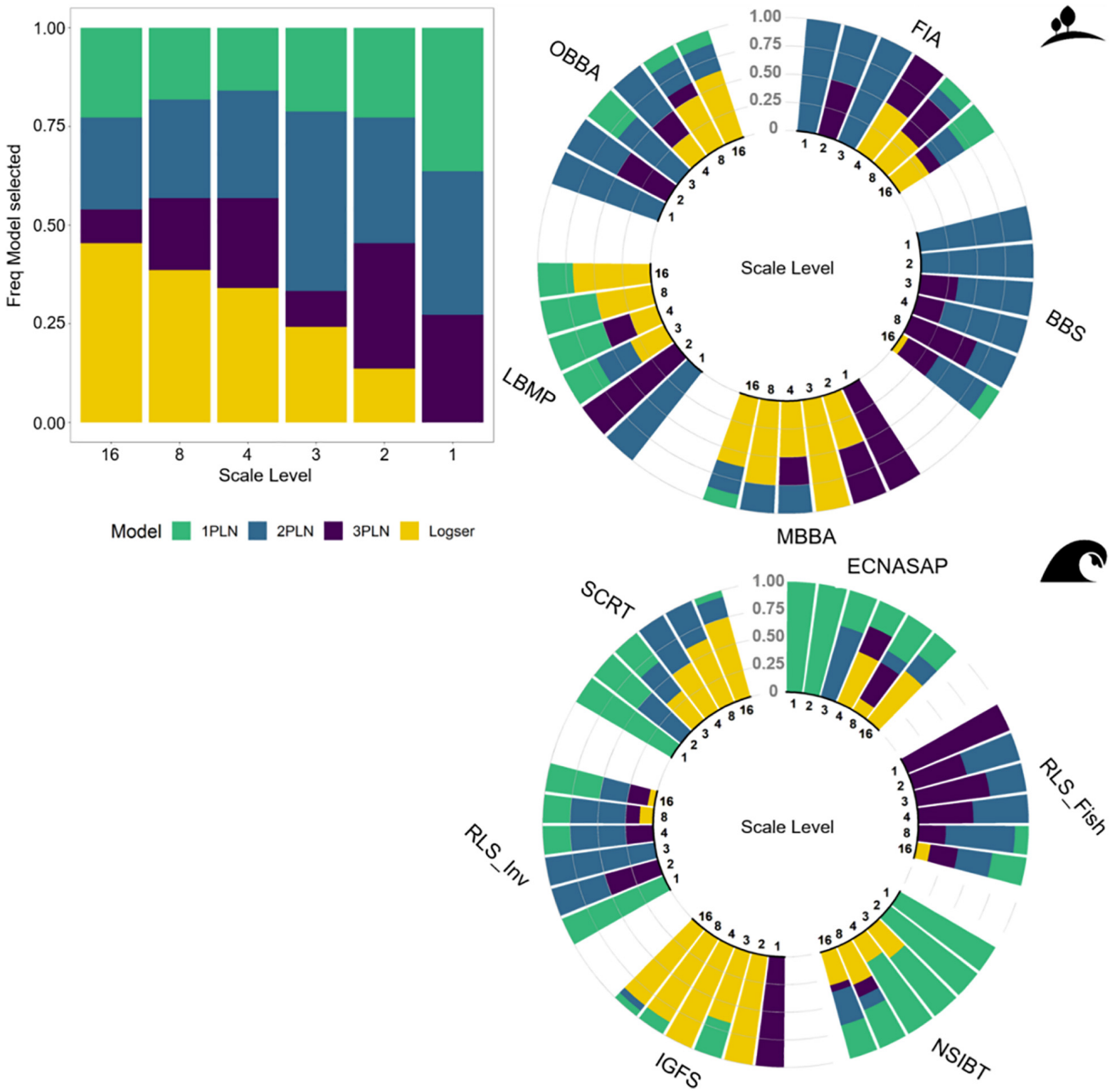


FIGURE 2 | Effect of the sampling level on SAD shape across the scale gradient, showing the aggregated pattern across all the communities (barplot) and for each individual community in the circular plots for terrestrial (top) and marine (bottom) datasets. The colors represent the best model selected by AICc, i.e., either 1PLN, 2PLN, 3PLN, or logseries.

OBBA bird, and RLS fish surveys) still represent very large spatial extents. Additionally, due to the way the scale gradient was implemented, the spatial relationship between the sections is maintained. This further suggests that multimodality reflects the structure of the communities at these smaller scales, despite the marked decrease in number of species and total abundance as area sampled decreased. While using only AIC_c to select the best-fit model could potentially lead to the prevalence of multimodality being overestimated (Antão et al., 2017), there was consistent and strong support for 2PLN or 3PLN mixtures for these communities. Hence, we are confident the detection of multimodality in our analysis is robust. For these communities, SAD shape was overall conserved across a wide range of areas sampled, highlighting that dramatic shifts in key aspects of community structure are required for the overall SAD shape to change (e.g., Supp and Ernest, 2014). We further note that while we evaluated SAD scaling patterns using a single initial point to establish the scale gradient, in a related analysis we obtained

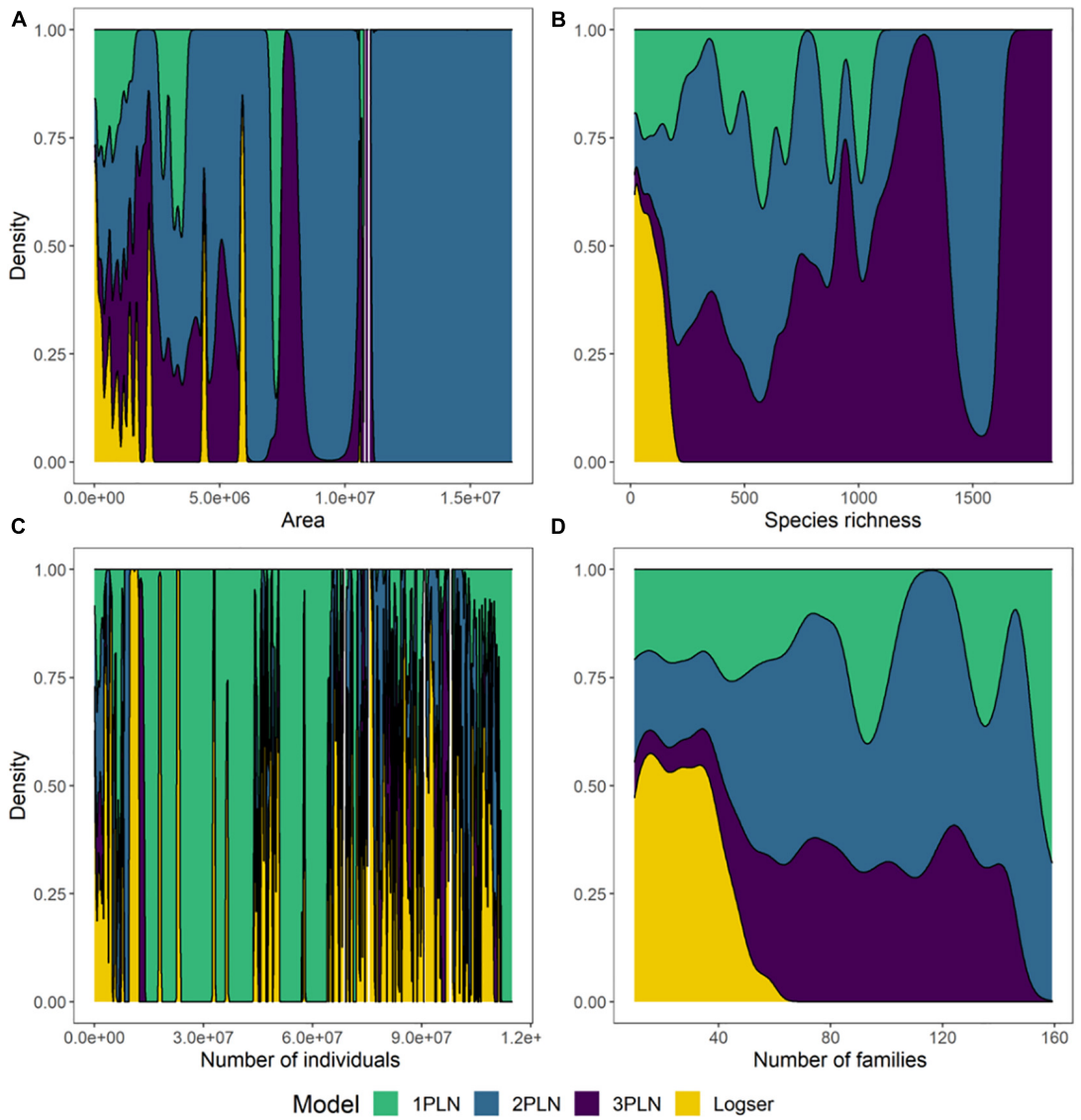


FIGURE 3 | Effect of area sampled **(A)**, species richness **(B)**, total number of individuals **(C)**, and number of families **(D)** on SAD shape (best model selected) aggregated across all the communities (see **Supplementary Figure 1** for plots for each individual community). Plots were built using the function geom_density from the ggplot package in R (Wickham, 2009), which computes smoothed versions of histograms for continuous variables.

consistent beta diversity scaling patterns across ten random initial points (Antão et al., 2019); we are thus confident the initial random point is unlikely to drive our results.

Clear shifts in SAD shape can provide information about relevant ecological and spatial factors affecting community structure at different scales. For the lognormal SADs at the total extent, but for which multimodality was frequently selected for smaller sections, this might be due to haphazard spatial decomposition of the community when splitting the total extent, and/or because of sampling effects. This can occur for instance if the communities become dominated by both rare and very abundant species, thus yielding multiple modes across the SAD (Gray et al., 2005; Green and Plotkin, 2007; see also Šizling et al., 2009b). When the common species are very abundant, 2PLN or 3PLN models are better able to accommodate both the rare and the most abundant species in the distribution, hence being selected despite the increase in number of parameters. Species spatial aggregation patterns can lead to one or a few species becoming extremely (proportionally) abundant for smaller areas. Simultaneously there is also a higher proportion of rare species in smaller samples, and hence a multimodal model provides a better fit than 1PLN. In contrast, logseries systematically failed to simultaneously accommodate rare and very abundant species.

Differences in species aggregation rates have been suggested to be related to the existence of multiple modes in both theoretical and empirical SADs (Green and Plotkin, 2007; Alonso et al., 2008; Dornelas and Connolly, 2008). Our findings suggest that multimodality occurring at smaller scales might be due to the

spatial aggregation of individuals and species, where "hitting or missing" areas where a species is abundant can lead to the appearance of different modes. Hence, a multimodal model provides a better description for the SAD by accommodating both the rarest and the more abundant species, which neither the logseries nor a single PLN are able to do. Conspecific aggregation is one of the fundamental features of ecological communities (McGill, 2010). Our results suggest that species spatial aggregation is likely an important driver across scales [from local (see e.g., Šizling et al., 2009b) to truly continental scales] and taxa, which is consistent with a related analysis focusing on the scaling patterns of beta diversity (Antão et al., 2019). Finally, we also found transitions from logseries to lognormal as scale increased (Preston, 1948; Magurran, 2004). This finding is consistent with studies showing that the rate at which these transitions occur depend on species traits linked to the spatial distribution of individuals (Borda-de-Água et al., 2017; Dantas de Miranda et al., 2019). Overall, we have extended the range of scales usually evaluated in this type of analyses, both highlighting the importance of multimodal SADs, and showing clear patterns of SAD shape variability emerging across scales.

Comparison With Macroecological Theories

Our results clearly contrast with two key macroecological theories. The logseries was never selected as best fit for larger scales (i.e., larger areas, higher diversity or higher number of individuals) and was unable to accurately predict the abundances of both rare and common species. This clearly deviates from neutral models predictions of the logseries as the expected SAD for larger scales, including models with realistic speciation modes, that can produce more flexible metacommunity SADs and reduce the predicted number of singletons (Rosindell et al., 2010, 2011), but are still unable to provide adequate fit for large scale SADs (Fukaya et al., 2020) or accommodate multimodal SADs. Consistent with our results, a study of global rarity patterns in plants also found logseries distributions to inadequately fit global aggregate SAD, being outperformed by lognormal distributions (Enquist et al., 2019); interestingly multimodality was not evaluated. In addition, our findings also show a marked contrast with several studies reporting the success of METE models in characterizing general SAD shape (Harte et al., 2008; White et al., 2012; Xiao et al., 2015). For instance, the logseries was shown to provide a better fit to several empirical SADs with a wide range of "anchor scales" (White et al., 2012). This study included two communities analyzed here, namely the FIA tree and BBS bird datasets, for which multimodality was strongly selected along the scale gradient (although we did not analyze the smallest sample grain level). Yet, in agreement with our results, White et al. (2012) also reported that the logseries tended to overestimate richness for the lowest abundance classes, and the authors suggested that other METE formulations or neutral models could be used as alternatives, since they predict fewer singletons. While we have not directly analyzed either METE or neutral models, our results clearly illustrate that the logseries is not able to simultaneously deal with the rare and the abundant species tails. Hence, given that the logseries has been systematically outperformed, particularly for larger scales [here and in Antão et al. (2017) and Enquist et al. (2019)], it is unlikely that the logseries is an adequate descriptor of SADs across spatial scales. The same inability has been shown for the lognormal distribution, explicitly linked to species turnover and spatial aggregation patterns between the smaller scale SADs being pooled together in the theorethical famework proposed by Šizling et al. (2009a,b). A critical development is then to accommodate the variability in SAD shape with spatial scale, and crucially to include multimodality, given its high prevalence across all scales.

Our results do not invalidate the logseries or the lognormal as "realistic functional forms" for the SAD, as both were selected as best model for several communities (here, as well as in Antão et al., 2017). What our results clearly show is that the logseries is neither adequate as the single SAD distribution, as METE suggests, nor is it more likely to describe SADs at larger scales, contrary to NTB predictions (Hubbell, 2001; Pueyo et al., 2007; Rosindell et al., 2010). A model extending the neutral theory by incorporating size variation and growth dynamics [the size-structured neutral theory model (SSNT)] still assumes a logseries SAD (O'Dwyer et al., 2009). Comparisons of different model formulations for both METE and SSNT showed a better performance for the latter (O'Dwyer et al., 2009; Xiao et al., 2016), with the authors arguing that METE's constraints are not fully capturing relevant biological processes influencing community structure. Furthermore, because derivations for other macroecological patterns depend on a logseries SAD (Harte et al., 2008), testing other SAD distributions to ensure those derivations are robust is warranted. For instance, Šizling et al. (2011) showed that incorporating spatial turnover at different scales affected local species area relationships derived from METE. Moreover, neither neutral nor METE models account for multiple modes in SADs, not to their higher prevalence at larger scales and for more diverse communities. Our results also suggest that the combination of S and N is not sufficient to determine SAD shape across spatial scales (White et al., 2012; Locey and White, 2013; Xiao et al., 2015), specifically as the total abundance across the scale gradient did not exhibit any systematic effect on SAD shape (for the aggregated results). Nonetheless, both area and species richness showed a strong influence on SAD shape, although the two variables are also strongly correlated (Rosenzweig, 1995). One of the advantages of using the METE approach is being able to interpret deviations from the expected distributions solely constrained by richness and abundance as evidence that other ecological features must be important in structuring communities (Harte et al., 2008; White et al., 2012; Xiao et al., 2016). Here, we show this must be the case for several large-scale empirical communties, and suggest that spatial aggregation patterns and ecological heterogeneity as potential such factors.

One potential explanation for the departure of the findings presented here from the two macroecological frameworks considered could be the scale gradient examined. This gradient spanned several orders of magnitude, and included very large areas, even for the smallest levels analyzed. Hence, it is possible that discrepancies between our results and those theoretical

predictions might be at least partially attributable to differences in the spatial scales investigated (as noted above for the FIA and BBS datasets). The same is true for the spatial theory developed by Šizling et al. (2009a,b), which mainly focused on upscaling SADs for larger plots, but which remain much smaller than the smallest level in our analysis. In addition, and while we analyzed extensively and consistently sampled datasets with hundreds to several thousands of samples, it is particularly hard to obtain thorough abundance records at such large extents. Thus, sampling effects could also potentially influence our results, particularly at the largest scales, where sampling effects known to affect SAD shape may be more severe. Nonetheless, our results strongly support the suggestion that both NTB and METE might be more adequate for smaller scales (smaller areas and/or fewer individuals) (McGill, 2010), while the traditional SAD distributions are unable to accommodate SAD shape variability with scale, as we have clearly shown with consistent empirical patterns across taxa, and for both marine and terrestrial habitats.

CONCLUSION

Spatial scale and taxonomic diversity emerged as major drivers of variability in SAD shape. The systematic assessment of SADs at different spatial scales and for different taxa allows us to make stronger inferences about the commonness and rarity of species across scales. Our findings clearly show that neither NTB nor METE formulations are able to accommodate the variability in SADs shape across spatial scales. The interplay of SAD shape at different scales can highlight important mechanisms acting on ecological communities, namely both inter- and intraspecific spatial patterns that lead to different SAD shape as spatial scale changes. Furthermore, a critical development for macroecological theories is to predict or accommodate multimodal SADs, and crucially to incorporate the effect of spatial scale and ecological heterogeneity in determining SAD shape.

DATA AVAILABILITY STATEMENT

Publicly available datasets were analyzed in this study. This data can be found here: All the datasets used in this study can be accessed through either the BioTIME database (on Zenodo—https://doi.org/10.5281/zenodo.1211105; or the BioTIME website—http://biotime.st-andrews.ac.uk/) or the EcoData Retriever, or alternatively from the original sources cited.

AUTHOR CONTRIBUTIONS

LA and MD designed the study. LA assembled the data, performed all the analyses, and wrote the first draft of the manuscript. All authors have discussed the results and contributed extensively to the final manuscript.

FUNDING

LA was supported by the Fundação para a Ciência e Tecnologia, Portugal (POPH/FSE SFRH/BD/90469/2012) and by the Jane and Aatos Erkko Foundation. AM and the BioTIME database were funded by the European Research Council grants AdG BioTIME (250189) and PoC BioCHANGE (727440). MD was funded by a Leverhulme Fellowship and by the John Templeton Foundation (grant #60501 "Putting the Extended Evolutionary Synthesis to the Test"). Open access was funded by Helsinki University Library.

ACKNOWLEDGMENTS

We are grateful to all the scientists, data collectors and their funders for making data publicly available. We thank the University of St. Andrews Bioinformatics Unit (Wellcome Trust ISSF grant [105621/Z/14/Z]). Faye Moyes assisted with data management. The two icons in **Figure 2** are from the Noun Project under CCBY license: *land* by A. Skowalsky, and *wave* by B. Farias.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fevo.2021.626730/full#supplementary-material

REFERENCES

Alonso, D., Ostling, A., and Etienne, R. S. (2008). The implicit assumption of symmetry and the species abundance distribution. *Ecol. Lett.* 11, 93–105. doi: 10.1111/j.1461-0248.2007.01127.x

Antão, L. H., Connolly, S. R., Magurran, A. E., Soares, A., and Dornelas, M. (2017). Prevalence of multimodal species abundance distributions is linked to spatial and taxonomic breadth. *Glob. Ecol. Biogeogr.* 26, 203–215. doi: 10.1111/geb.12532

Antão, L. H., McGill, B., Magurran, A. E., Soares, A. M. V. M., and Dornelas, M. (2019). β-diversity scaling patterns are consistent across metrics and taxa. *Ecography* 42, 1012–1023. doi: 10.1111/ecog.04117

Barabás, G., D'Andrea, R., Rael, R., Meszéna, G., and Ostling, A. (2013). Emergent neutrality or hidden niches? *Oikos* 122, 1565–1572. doi: 10.1111/j.1600-0706.2013.00298.x

Bivand, R., and Rundel, C. (2016). *rgeos: Interface to Geometry Engine - Open Source (GEOS)*. R package version 0.3-19.

Borda-de-Água, L., Borges, P. A. V., Hubbell, S. P., and Pereira, H. M. (2012). Spatial scaling of species abundance distributions. *Ecography* 35, 549–556. doi: 10.1111/j.1600-0587.2011.07128.x

Borda-de-Água, L., Hubbell, S. P., and McAllister, M. (2002). Species-area curves, diversity indices, and species abundance distributions: a multifractal analysis. *Am. Nat.* 159, 138–155. doi: 10.1086/324787

Borda-de-Água, L., Whittaker, R. J., Cardoso, P., Rigal, F., Santos, A. M. C., Amorim, I. R., et al. (2017). Dispersal ability determines the scaling properties of species abundance distributions: a case study using arthropods from the Azores. *Sci. Rep.* 7:3899. doi: 10.1038/s41598-017-04126-5

Brown, S. K. R., Zwanenburg, K., and Branton, R. (2005). *East Coast North America Strategic Assessment Groundfish Atlas - ECNASAP. OBIS Canada*. Dartmouth,

NS: Bedford Institute of Oceanography. Available online at: http://www.iobis.org/ (accessed on 2013).

Bulmer, M. (1974). On fitting the Poisson lognormal distribution to species-abundance data. *Biometrics* 30, 101–110. doi: 10.2307/2529621

Burnham, K. P., and Anderson, D. R. (2002). *Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach*, 2nd Edn. New York, NY: Springer.

Condit, R. S., Ashton, P. S., Baker, P. J., Bunyavejchewin, S., Gunatilleke, S., Gunatilleke, N., et al. (2000). Spatial patterns in the distribution of tropical tree species. *Science* 288, 1414–1418. doi: 10.1126/science.288.5470.1414

Connolly, S. R., Hughes, T. P., Bellwood, D. R., and Karlson, R. H. (2005). Community structure of corals and reef fishes at multiple scales. *Science* 309, 1363–1365. doi: 10.1126/science.1113281

Dantas de Miranda, M., Borda-de-Água, L., Pereira, H. M., and Merckx, T. (2019). Species traits shape the relationship between local and regional species abundance distributions. *Ecosphere* 10:e02750. doi: 10.1002/ecs2.2750

DATRAS (2010a). *Fish trawl survey: ICES North Sea International Bottom Trawl Survey for commercial fish species*. ICES Database of trawl surveys (DATRAS). Copenhagen: International Council for the Exploration of the Sea (ICES). Available online at: http://www.emodnet-biology.eu/data-catalog??module=dataset&dasid=2763 (accessed on 2013).

DATRAS (2010b). *Fish trawl survey: Irish Ground Fish Survey for commercial fish species*. ICES Database of trawl surveys (DATRAS). Copenhagen: International Council for the Exploration of the Sea (ICES). Available online at: www.emodnet-biology.eu/data-catalog?module=dataset&dasid=2762 (accessed on 2013).

Dewdney, A. K. (1998). A general theory of the sampling process with applications to the "veil line.". *Theor. Popul. Biol.* 54, 294–302. doi: 10.1006/tpbi.1997.1370

Dornelas, M., Antão, L. H., Moyes, F., Bates, A. E., Magurran, A. E., Adam, D., et al. (2018). BioTIME: a database of biodiversity time series for the anthropocene. *Glob. Ecol. Biogeogr.* 27, 760–786. doi: 10.1111/geb.12729

Dornelas, M., and Connolly, S. R. (2008). Multiple modes in a coral species abundance distribution. *Ecol. Lett.* 11, 1008–1016. doi: 10.1111/j.1461-0248.2008.01208.x

Dornelas, M., Connolly, S. R., and Hughes, T. P. (2006). Coral reef diversity refutes the neutral theory of biodiversity. *Nature* 440, 80–82. doi: 10.1038/nature04534

Edgar, G. J., and Stuart-Smith, R. D. (2008). *Reef Life Survey (RLS): Invertebrates*. Battery Point, TAS: Institute for Marine and Antarctic Studies (IMAS). Available online at: https://catalogue-rls.imas.utas.edu.au/geonetwork/srv/en/metadata.show?uuid=60978150-1641-11dd-a326-00188b4c0af8b (accessed on 2016).

Edgar, G. J., and Stuart-Smith, R. D. (2014a). *Reef Life Survey (RLS): Global reef fish dataset*. Battery Point, TAS: Institute for Marine and Antarctic Studies (IMAS). Available online at: https://catalogue-rls.imas.utas.edu.au/geonetwork/srv/en/metadata.show?uuid=9c766140-9e72-4bfb-8f04-d51038355c59 (accessed on 2016).

Edgar, G. J., and Stuart-Smith, R. D. (2014b). Systematic global assessment of reef fish communities by the Reef Life Survey program. *Sci. Data* 1:140007. doi: 10.1038/sdata.2014.7

Enquist, B. J., Feng, X., Boyle, B., Maitner, B., Newman, E. A., Jørgensen, P. M., et al. (2019). The commonness of rarity: global and future distribution of rarity across land plants. *Sci. Adv.* 5:eaaz0414. doi: 10.1126/sciadv.aaz0414

Etienne, R. S. (2005). A new sampling formula for neutral biodiversity. *Ecol. Lett.* 8, 253–260. doi: 10.1111/j.1461-0248.2004.00717.x

Etienne, R. S. (2007). A neutral sampling formula for multiple samples and an "exact" test of neutrality. *Ecol. Lett.* 10, 608–618. doi: 10.1111/j.1461-0248.2007.01052.x

Etienne, R. S. (2009). Maximum likelihood estimation of neutral model parameters for multiple samples with different degrees of dispersal limitation. *J. Theor. Biol.* 257, 510–514. doi: 10.1016/j.jtbi.2008.12.016

Fisher, R., Corbet, A., and Williams, C. (1943). The relation between the number of species and the number of individuals in a random sample of an animal population. *J. Anim. Ecol.* 12, 42–58. doi: 10.2307/1411

Fukaya, K., Kusumoto, B., Shiono, T., Fujinuma, J., and Kubota, Y. (2020). Integrating multiple sources of ecological data to unveil macroscale species abundance. *Nat. Commun.* 11:1695. doi: 10.1038/s41467-020-15407-5

Gray, J. S., Bjørgesæter, A., and Ugland, K. I. (2005). The impact of rare species on natural assemblages. *J. Anim. Ecol.* 74, 1131–1139. doi: 10.1111/j.1365-2656.2005.01011.x

Gray, J. S., Bjørgesæter, A., and Ugland, K. I. (2006). On plotting species abundance distributions. *J. Anim. Ecol.* 75, 752–756. doi: 10.1111/j.1365-2656.2006.01095.x

Green, J. L., and Plotkin, J. B. (2007). A statistical theory for sampling species abundances. *Ecol. Lett.* 10, 1037–1045. doi: 10.1111/j.1461-0248.2007.01101.x

Harte, J., and Newman, E. A. (2014). Maximum information entropy: a foundation for ecological theory. *Trends Ecol. Evol.* 29, 384–389. doi: 10.1016/j.tree.2014.04.009

Harte, J., Zillio, T., Conlisk, E., and Smith, A. (2008). Maximum entropy and the state-variable approach to macroecology. *Ecology* 89, 2700–2711. doi: 10.1890/07-1369.1

Hubbell, S. P. (2001). *The Unified Neutral Theory of Biodiversity and Biogeography*. Princeton, NY: Princeton University Press.

Levin, S. (1992). The problem of pattern and scale in ecology: the Robert H. MacArthur award lecture. *Ecology* 73, 1943–1967. doi: 10.2307/1941447

Locey, K. J., and White, E. P. (2013). How species richness and total abundance constrain the distribution of abundance. *Ecol. Lett.* 16, 1177–1185. doi: 10.1111/ele.12154

Magurran, A. E. (2004). *Measuring Biological Diversity*. Oxford: Blackwell Science.

Magurran, A. E., and Henderson, P. A. (2003). Explaining the excess of rare species in natural species abundance distributions. *Nature* 422, 714–716. doi: 10.1038/nature01547

May, R. M. (1975). "Patterns of species abundance and diversity" in *Ecology and Evolution of Communities*, eds M. L. Cody and J. M. Diamond (Cambridge, MA: Belknap Press of Harvard University Press), 81–120.

McGill, B. J. (2003a). A test of the unified neutral theory of biodiversity. *Nature* 422, 881–885. doi: 10.1038/nature01569.1

McGill, B. J. (2003b). Does Mother Nature really prefer rare species or are log-left-skewed SADs a sampling artefact? *Ecol. Lett.* 6, 766–773. doi: 10.1046/j.1461-0248.2003.00491.x

McGill, B. J. (2010). Towards a unification of unified theories of biodiversity. *Ecol. Lett.* 13, 627–642. doi: 10.1111/j.1461-0248.2010.01449.x

McGill, B. J., Etienne, R. S., Gray, J. S., Alonso, D., Anderson, M. J., Benecha, H. K., et al. (2007). Species abundance distributions: moving beyond single prediction theories to integration within an ecological framework. *Ecol. Lett.* 10, 995–1015. doi: 10.1111/j.1461-0248.2007.01094.x

McGlinn, D., and White, E. (2015). *ecoretriever: R Interface to the EcoData Retriever. R package version 0.2.1*. Available online at: https://cran.r-project.org/package=ecoretriever (accessed on 2016).

Morris, B. D., and White, E. P. (2013). The ecodata retriever: improving access to existing ecological data. *PLoS One* 8:e65848. doi: 10.1371/journal.pone.0065848

NatureCounts (2012a). *Maritimes Breeding Bird Atlas (2006-2010): point count data. NatureCounts, a node Avian Knowledge Network. Bird Studies, Canada*. Available online at: http://www.birdscanada.org/birdmon/ (accessed on 2012).

NatureCounts (2012b). *Ontario Breeding Bird Atlas (2001-2005): Point Count Data. NatureCounts, a Node Avian Knowledge Network. Bird Studies, Canada*. Available online at: http://www.birdscanada.org/birdmon/ (accessed on 2012).

O'Dwyer, J. P., Lake, J. K., Ostling, A., Savage, V. M., and Green, J. L. (2009). An integrative framework for stochastic, size-structured community assembly. *Proc. Natl. Acad. Sci. U.S.A.* 106, 6170–6175. doi: 10.1073/pnas.0813041106

Pardieck, K. L., Ziolkowski, D. J. J., Hudson, M.-A. R., and Campbell, K. (2016). *North American Breeding Bird Survey Dataset 1966 - 2015, Version 2015.1*. Laurel, MD: U.S. Geological Survey, Patuxent Wildlife Research Center, doi: 10.5066/F7C53HZN

Pielou, E. C. (1969). *An Introduction to Mathematical Ecology*. New York, NY: Wiley-Interscience.

Pielou, E. C. (1977). *Mathematical Ecology*. New York, NY: John Wiley & Sons.

Preston, F. (1948). The commonness, and rarity, of species. *Ecology* 29, 254–283. doi: 10.2307/1930989

Pueyo, S., He, F., and Zillio, T. (2007). The maximum entropy formalism and the idiosyncratic theory of biodiversity. *Ecol. Lett.* 10, 1017–1028. doi: 10.1111/j.1461-0248.2007.01096.x

R Core Team (2017). *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing.

Rosenzweig, M. L. (1995). *Species Diversity in Space and Time*. Cambridge: Cambridge University Press.

Rosindell, J., and Cornell, S. J. (2013). Universal scaling of species-abundance distributions across multiple scales. *Oikos* 122, 1101–1111. doi: 10.1111/j.1600-0706.2012.20751.x

Rosindell, J., Cornell, S. J., Hubbell, S. P., and Etienne, R. S. (2010). Protracted speciation revitalizes the neutral theory of biodiversity. *Ecol. Lett.* 13, 716–727. doi: 10.1111/j.1461-0248.2010.01463.x

Rosindell, J., Hubbell, S. P., and Etienne, R. S. (2011). The unified neutral theory of biodiversity and biogeography at age ten. *Trends Ecol. Evol.* 26, 340–348. doi: 10.1016/j.tree.2011.03.024

Šizling, A. L., Kunin, W. E., Sizlingová, E., Reif, J., and Storch, D. (2011). Between geometry and biology: the problem of universality of the species-area relationship. *Am. Nat.* 178, 602–611. doi: 10.1086/662176

Šizling, A. L., Storch, D., Reif, J., and Gaston, K. J. (2009a). Invariance in species-abundance distributions. *Theor. Ecol.* 2, 89–103. doi: 10.1007/s12080-008-0031-3

Šizling, A. L., Storch, D., Sizlingova, E., Reif, J., and Gaston, K. J. (2009b). Species abundance distribution results from a spatial analogy of central limit theorem. *Proc. Natl. Acad. Sci. U.S.A.* 106, 6691–6695. doi: 10.1073/pnas.0810096106

Supp, S. R., and Ernest, S. K. M. (2014). Species-level and community-level responses to disturbance: a cross-community analysis. *Ecology* 95, 1717–1723. doi: 10.1890/13-2250.1

USDA Forest Service (2010). *Forest Inventory and Analysis National Core Field Guide (Phase 2 and 3). Version 4.0*. Washington, D.C: USDA Forest Service, Forest Inventory and Analysis.

USFS (2012). *Landbird Monitoring Program (UMT-LBMP)*. US Forest Service. Available online at: http://www.avianknowledge.net/ (accessed on 2012).

Volkov, I., Banavar, J. R., Hubbell, S. P., and Maritan, A. (2003). Neutral theory and relative species abundance in ecology. *Nature* 424, 1035–1037. doi: 10.1038/nature01883

Volkov, I., Banavar, J. R., Hubbell, S. P., and Maritan, A. (2007). Patterns of relative species abundance in rainforests and coral reefs. *Nature* 450, 45–49. doi: 10.1038/nature06197

Wade, E. J. (2011). *Snow Crab Research Trawl Survey Database (Southern Gulf of St. Lawrence, Gulf region, Canada) From 1988 to 2010*. Dartmouth, NS: Bedford Institute of Oceanography Available online at: http://iobis.org/ (accessed on 2012).

White, E. P., Thibault, K. M., and Xiao, X. (2012). Characterizing species abundance distributions across taxa and ecosystems using a simple maximum entropy model. *Ecology* 93, 1772–1778. doi: 10.1890/11-2177.1

Wickham, H. (2009). *ggplot2: Elegant Graphics for Data Analysis*. New York: Springer-Verlag New York.

Wiens, J. A. (1989). Spatial scaling in ecology. *Funct. Ecol.* 3, 385–397. doi: 10.2307/2389612

Woudenberg, S. W., Conkling, B. L., O'Connell, B. M., LaPoint, E. B., Turner, J. A., and Waddell, K. L. (2010). *The Forest Inventory and Analysis Database: Database description and users manual version 4.0 for Phase 2*. General Technical Report No. RMRSGTR-245. Fort Collins, CO: Department of Agriculture, Forest Service, Rocky Mountain Research Station.

Xiao, X., McGlinn, D. J., and White, E. P. (2015). A strong test of the maximum entropy theory of ecology. *Am. Nat.* 185, E70–E80. doi: 10.1086/679576

Xiao, X., O'Dwyer, J. P., and White, E. P. (2016). Comparing process-based and constraint-based approaches for modeling macroecological patterns. *Ecology* 97, 1228–1238. doi: 10.1890/15-0962.1

Zillio, T., and He, F. (2010). Inferring species abundance distribution across spatial scales. *Oikos* 119, 71–80. doi: 10.1111/j.1600-0706.2009.17938.x

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Ecology and Evolution

ORIGINAL RESEARCH
published: 23 April 2021
doi: 10.3389/fevo.2021.603436

Emergent Rarity Properties in Carabid Communities From Chinese Steppes With Different Climatic Conditions

Noelline Tsafack[1,2], Paulo A. V. Borges[2], Yingzhong Xie[1], Xinpu Wang[1] and Simone Fattorini[3]*

[1] School of Agriculture, Ningxia University, Yinchuan, China, [2] cE3c-Centre for Ecology, Evolution and Environmental Changes/Azorean Biodiversity Group, Faculdade de Ciências e Engenharia do Ambiente, Universidade dos Açores, Angra do Heroísmo, Portugal, [3] Department of Life, Health and Environmental Sciences, University of L'Aquila, L'Aquila, Italy

OPEN ACCESS

Edited by:
Bradley S. Case, Auckland University of Technology, New Zealand

Reviewed by:
Raisa Sukhodolskaya, Tatarstan Academy of Sciences, Russia
Julien Pétillon, University of Rennes 1, France
Agnieszka Kosewska, University of Warmia and Mazury in Olsztyn, Poland

***Correspondence:**
Noelline Tsafack
noelline.tsafack@gmail.com

Specialty section:
This article was submitted to Models in Ecology and Evolution, a section of the journal Frontiers in Ecology and Evolution

Received: *06 September 2020*
Accepted: *06 April 2021*
Published: *23 April 2021*

Citation:
Tsafack N, Borges PAV, Xie Y, Wang X and Fattorini S (2021) Emergent Rarity Properties in Carabid Communities From Chinese Steppes With Different Climatic Conditions. Front. Ecol. Evol. 9:603436. doi: 10.3389/fevo.2021.603436

Species abundance distributions (SADs) are increasingly used to investigate how species community structure changes in response to environmental variations. SAD models depict the relative abundance of species recorded in a community and express fundamental aspects of the community structure, namely patterns of commonness and rarity. However, the influence of differences in environmental conditions on SAD characteristics is still poorly understood. In this study we used SAD models of carabid beetles (Coleoptera: Carabidae) in three grassland ecosystems (desert, typical, and meadow steppes) in China. These ecosystems are characterized by different aridity conditions, thus offering an opportunity to investigate how SADs are influenced by differences in environmental conditions (mainly aridity and vegetation cover, and hence productivity). We used various SAD models, including the meta-community zero sum multinomial (mZSM), the lognormal (PLN) and Fisher's logseries (LS), and uni- and multimodal gambin models. Analyses were done at the level of steppe type (coarse scale) and for different sectors within the same steppe (fine scale). We found that the mZSM model provided, in general, the best fit at both analysis scales. Model parameters were influenced by the scale of analysis. Moreover, the LS was the best fit in desert steppe SAD. If abundances are rarefied to the smallest sample, results are similar to those without rarefaction, but differences in models estimates become more evident. Gambin unimodal provided the best fit with the lowest α-value observed in desert steppe and higher values in typical and meadow steppes, with results which were strongly affected by the scale of analysis and the use of rarefaction. Our results indicate that all investigated communities are adequately modeled by two similar distributions, the mZSM and the LS, at both scales of analyses. This indicates (1) that all communities are characterized by a relatively small number of species, most of which are rare, and (2) that the meta-communities at the large scale maintain the basic SAD shape of the local communities. The gambin multimodal models produced exaggerated α-values,

which indicates that they overfit simple communities. Overall, Fisher's α, mZSM θ, and gambin α-values were substantially lower in the desert steppe and higher in the typical and meadow steppes, which implies a decreasing influence of environmental harshness (aridity) from the desert steppe to the typical and meadow steppes.

Keywords: species abundance models, gambin models, Carabidae, ground beetles, Asia, arid environments, grasslands

INTRODUCTION

Biodiversity loss is the most common consequence of the increasing environmental degradation due to anthropogenic changes (Barnosky et al., 2011; Redford et al., 2015; Kehoe et al., 2017). The current biodiversity crisis is not limited to the increasing rates of species extinctions, but involves all levels of biodiversity (Western, 1992; Borges et al., 2019). Local extinctions or even changes in species' relative abundance lead to alterations in community structure and hence in ecosystem functioning (McCann, 2000; Balvanera et al., 2006). Thus, studies that model variations in community structure in response to changes in environmental characteristics may provide important information to predict how biodiversity will be affected by alterations in the balance of rare versus dominant species (Tsafack et al., 2019a; Ibanez et al., 2020).

Although the study of species abundance distributions (SADs) has a long tradition in ecology (Raunkiaer, 1909; Preston, 1948), the recent development of new statistical tools have led to an increasing interest in the use of SAD models to investigate patterns of species commonness and rarity in biotic communities (Matthews and Whittaker, 2014a, 2015; Fattorini et al., 2016; Picanço et al., 2017; Haddad et al., 2019; Matthews et al., 2019; Pennino et al., 2019). Also, SAD models may be used in conservation studies to predict species extinction risk (Kitzes and Harte, 2015) and ecosystem health (related to disturbance) (Dornelas et al., 2009) and, hence, to inform management actions (Kim et al., 2013; Miličić et al., 2017).

Several models have been proposed to describe the SADs (see Matthews and Whittaker, 2014a,b, 2015; Baldridge et al., 2016; Fattorini et al., 2018). One of the most commonly used models is the so-called log-normal (McGill, 2003). This model has been derived as a null form of the distribution resulting from the central limit theorem (May, 1975), and it is classified among the purely statistical models (McGill et al., 2007), but can be the limit of population dynamics (Engen and Lande, 1996), or niche partitioning (Bulmer, 1974; Sugihara, 1980). The lognormal, however, is problematic because it is a continuous distribution, thus allowing fractional abundances, and does not have an associated sampling theory. To address these problems, it has been proposed to use a Poisson sampling of individuals from a standard lognormal distribution (Matthews and Whittaker, 2014a). Thus, the resulting Poisson lognormal distribution describes the abundances of species in a Poisson sample of a community that follows a lognormal SAD. A SAD model frequently used in opposition to the lognormal is the Fisher log-series, which was initially derived as a purely statistical distribution (Fisher et al., 1943) and subsequently interpreted in relation with ecological processes (Hubbell, 2001; Volkov et al., 2003; Pueyo et al., 2007; Harte et al., 2008). The lognormal distribution is generally considered to best describe SADs of undisturbed and species rich communities, whereas SADs of disturbed communities tend to follow the Fisher logseries (May, 1975; Matthews and Whittaker, 2015; Ulrich et al., 2016). For example, Haddad et al. (2019) found that the log-series best fitted spider abundance distributions in riparian woodlands and grasslands which represent disturbed environments in the study area, while the Poisson log-normal best fitted spider abundance distribution in the hillside, the less disturbed environment.

Another popular SAD model is the metacommunity zero-sum multinomial (mZSM) distribution, which describes the SAD of a sample taken from a neutral metacommunity under random drift (Alonso and McKane, 2004). The mZSM follows one of the three assumptions of the neutral theory, that is, the zero-sum assumption [the two others are the neutrality assumption and the point mutation assumption: see Hubbell (2001) and Etienne et al. (2007)]. The zero-sum assumption states that the resources in the environment limit the individual abundance of the community species to a constant total number, because of zero-sum stochastic processes of birth, death and immigration. Therefore, resources are fully saturated at all times and a community following the mZSM model may be considered like a stable community. The neutral theory predicts that species abundance distribution in a local community is the mZSM distribution rather than the log-normal distribution. Compared to the log-normal distribution, the mZSM distribution has a long tail at the end of rare species, and its length depends on the community size and migration from the metacommunity (Zhou and Zhang, 2008).

As the log-series distribution is a limiting case of the mZSM, and the θ parameter of the mZSM tends to Fisher's α as the number of individuals in the sample increases, the two models typically provide similar fits to SADs (Prado et al., 2018).

A recently proposed model that is gaining increasing popularity in the last years is the so-called gambin model (Ugland et al., 2007; Matthews and Whittaker, 2014a; Ibanez et al., 2020). The gambin model has been proposed as a very flexible model able to fit a wide ranges of distributions (Ugland et al., 2007; Matthews et al., 2014). The gambin model is a combination of a gamma distribution and a binomial sampling method. The gambin model is considered to date as the most parsimonious general model (Ugland et al., 2007; Matthews et al., 2019) as it has one parameter α which describes the shape of the species abundance curve. The gambin model may be uni- or multimodal. Multimodal gambin models have multiple α-values (i.e., one α per mode) (Matthews et al., 2019), and their interpretation is more complex. Gambin has been shown to provide good fits to

a wide variety of empirical datasets, with low α-values indicating logseries-distributions (and hence communities where most species are rare) and higher values indicating more lognormal-distributions (and hence communities where most species are dominant) (Matthews and Whittaker, 2014a; Matthews et al., 2019). It has been also suggested that gambin α could be considered as a diversity index based on species equitability within the community (Ugland et al., 2007; Fattorini et al., 2016).

Use of these models to fit SADs of communities experiencing different environmental conditions may help understanding how habitat changes influence community structure (Fattorini et al., 2016; Picanço et al., 2017; Haddad et al., 2019; Matthews et al., 2019; Pennino et al., 2019). In the present study, we used the carabid beetles (Coleoptera Carabidae) of Chinese grassland ecosystems to investigate how SAD shapes are influenced by differences in environmental conditions, mainly aridity and vegetation cover, and hence productivity.

In China, there is a great variety of steppes, which are classified into four major types on the basis of the main differences in their vegetation, fauna, resource uses and management processes: desert, typical, meadow and alpine (Sun, 2005; Kang et al., 2007). Chinese steppes are undergoing degradation due to climate change and land-use intensification (mainly overgrazing with the increasing of livestock) (Lü et al., 2011; Werger and van Staalduinen, 2012; Tsafack et al., 2019a). These anthropic impacts are changing highly productive grasslands, such as the so-called meadow steppes, into arid areas with sparse vegetation, such as the so-called desert steppe. To slow down the degradation, several measures for grassland diagnosis have been suggested and among them are the unveiling of indicator species (Akiyama and Kawamura, 2007). However, there are no studies that use SAD models to investigate how species abundances vary among different types of steppes. Here, we used SAD models to depict the commonness and rarity of carabid species of three steppes characterized by different climatic conditions: the desert steppe (which is the most arid ecosystem), the typical steppe (which has intermediate conditions) and the meadow steppe (which is the less arid ecosystem). These different climatic regimes are associated with different forms of vegetation and hence productivity. The desert steppe, because of the highest aridity and lowest vegetation cover, is the less productive ecosystem, and it has been already shown that aridity and these parameters influence total carabid abundances (Tsafack et al., 2020).

Carabids are a prominent component of the ground dwelling fauna worldwide (Dajoz, 2002) and one of the most frequently used insect groups as model organisms in ecological and conservation studies (Koivula, 2011; Kotze et al., 2011; Roume et al., 2011; Cardarelli and Bogliani, 2014; Duan et al., 2016; Labruyere et al., 2016; Gobbi et al., 2017; Jouveau et al., 2019). Despite the recognized importance of carabids as indicators of environmental conditions, variations of carabid SADs in response to environmental changes remain poorly explored. However, a study on carabids of coniferous boreal forests revealed that SADs vary according to the forest maturity: mature forests are characterized by a distinct pattern with almost no intermediate species between the scarce ones and the very few abundant ones; interestingly, this "gap" is filled in recently cut forests, but emerges again when trees reach the age of 20–30 years, possibly because mature forests are less favorable to carabids (Niemelä, 1993). A study on the changes occurred in carabid community structure of Alpine environments between the years 1980 and 2009, showed that SADs changed from the lognormal model (or other models expressing good equitability) to the geometric series (a SAD model characteristic of simple and highly dominated communities), thus suggesting a regression toward the first stages of the ecological succession, probably because of habitat changes mainly due to global warming (Pizzolotto et al., 2014).

In this study, we predict that the carabid SAD in the desert steppe, which is the ecosystem with the most challenging environmental conditions, should deviate from log-normal model and should be closer to the log-series. By contrast, in the meadow steppe, which represents the less arid grassland, carabid SAD should deviate from log-series models, while SAD in the typical steppe might converge to either that of the desert steppe or that of the meadow steppe.

After assessing which distribution best captures the SAD of each community, we also applied the gambin model, to explore its feasibility in capturing these SADs and the use of the indicator α-value to compare steppes.

Because the landscape within steppes may be not homogeneous, we conducted the analyses also by subdividing steppes into more homogeneous sectors each occupied by a different form of vegetation. Community delimitation is always problematic; thus, it is difficult to say whether the communities at the steppe level are true local communities, or if they are an aggregation of different local communities associated with the various sectors. This may have profound impacts on SAD shapes at a broad scale. If different local communities are merged into a single pooled community, the SAD may show a large number of species with few individuals, which, in fact, belong to different local communities: the resulting SAD might therefore follow, for example, a log-series, whereas the SADs of the different local communities might be better approximated by log-normal distributions (Borda-de-Água et al., 2012). With coarse (steppe level) and fine (sector level) analyses, we used SAD modeling to see how the scale of analysis influences the associated SADs.

The very different values of total abundance recorded in the various steppes or steppe sectors pose the problem of comparing SADs parameters obtained with samples of different size (Stier et al., 2016; Borges et al., 2020). To explore this problem, we applied two different procedures of rarefaction, using as alternative references the desert steppe (which was the less sampled ecosystem) and the smallest sample in each steppe separately (**Table 1**).

MATERIALS AND METHODS

Study Sites and Data Collection

Our study was carried out in a mountain area in the Ningxia Hui Autonomous region (northern China). Three types of

steppes (desert, typical, and meadow steppes) are present in this region. The desert steppe is characterized by a semi-arid continental monsoonal climate typical of desert environments, shows a low and discontinuous vegetation cover, and is the most arid of the three types of steppe considered in this study. The typical steppe is characterized by a continental monsoon climate. This steppe represents an intermediate state between the desert steppe and the meadow steppe. The meadow steppe is characterized by a dense ground level vegetation and by a semi-humid climate. To reflect within-ecosystem variability of the typical and meadow steppes, on the basis on vegetation characteristics, we identified three habitat types within the typical steppe (ts1, ts2, and ts3), and two habitat types within the meadow steppe (called ms1 and ms2). The ts1 and ts2 sectors were located at the top of the mountain, in natural patches of grass vegetation and in fire belts, respectively; the sector ts3 was selected at the bottom of the mountain occupied by crop fields and natural vegetation. The ms1 and ms2 sectors were located at the south-west side and at the bottom of the mountain peak, respectively. Data were gathered from 90 sampling sites distributed as follows: 15 sites in the desert steppe, 45 sites in the typical steppe (15 sites in each sector) and 30 sites in the meadow steppe (again 15 sites in each sector). Sites were selected haphazardly (i.e., without any regular spatial arrangement) and separated by at least 150 m to avoid, or at least reduce, possible autocorrelation.

At each sampling site, five pitfall traps (separated by at least five meters from each other) were installed. Pitfall traps consisted of 7.15-cm diameter plastic cups, sunk in the ground with the cup-lip level with the soil surface, and filled with 60 ml of a mixture of tap water and vinegar (8%), sugar (4%) and 70% alcohol (4%). Sampling was done from May to September 2017. During the sampling period, pitfall traps were placed in the sites once a month in mid-month, and left in the field for ca. 72 h prior to collection. Traps were composed of two buckets, with the smaller inserted into the larger. At each sampling session, the smallest were extracted to collect the trapped beetles and then placed again in the largest, which were left dug into the soil. This ensured that trap position remained exactly the same over the sampling period and disturbance reduced to minimum. Specimens were identified by trained people lead by the expert taxonomist Professor Liang Hongbin. Further details about study area have been published in a companion paper (Tsafack et al., 2019b). Original data are provided in Tsafack (2018) (see section "Data Availability Statement").

TABLE 1 | Schematic description of the various forms of analyses developed in this article.

Model tested	Scale of analysis	Rarefaction
Classical models (Lognormal, Poisson-lognormal, Fisher's log-series, and metacommunity Zero-Sum Multinomial)	Sectors of the same steppe merged	No rarefaction
		Rarefaction with desert as a reference
		Rarefaction with the smallest sample in each steppe as a reference
	Sectors separated	No rarefaction
		Rarefaction with desert as a reference
		Rarefaction with the smallest sample in each steppe as a reference
Gambin models	Sectors of the same steppe merged	No rarefaction
		Rarefaction with desert as a reference
		Rarefaction with the smallest sample in each steppe as a reference
	Sectors separated	No rarefaction
		Rarefaction with desert as a reference
		Rarefaction with the smallest sample in each steppe as a reference

Data Analysis

Comparing Models of Species Abundance Distributions

We modeled carabid species abundance distributions (SADs) with the most commonly used fitting models to identify which distribution best approximated each community. Specifically, we modeled each community using the lognormal (LN), the Poisson-lognormal (PLN; zero-truncated version), the Fisher log-series (LS), and the metacommunity Zero-Sum Multinomial (mZSM) distributions with the function "*fitsad*" in the library "*sads*" (Prado et al., 2018) of the software R version 3.5.

Following current best practices in the study of species abundance distributions (Connolly et al., 2014; Matthews and Whittaker, 2014a), we used chi-square tests to assess deviations of the observed distributions from the expected ones, and likelihood-based model selection to compare competitive models (Burnham and Anderson, 2002). For this, we used Akaike's information criterion corrected for the small sample size (AICc) (Baldridge et al., 2016) and selected the model with the lowest AICc value as the best one, but alternative models with ΔAIC values ≤ 2 were also considered as receiving equal support (Burnham and Anderson, 2002).

We performed the analyses at two spatial scales. A first set of analyses was performed at the steppe type scale (coarse scale analyses). In this case, data from different sectors of the same type of steppe were merged, to highlight between-steppe differences. In a second set of analyses (fine scale analyses), we considered each sector separately, to better highlight possible differences between different sectors of the same type of steppe (within-steppe differences). Aim of these sets of analyses was to assess if different steppes, and sectors within them, followed different types of SADs, and hence to infer the possible underlaying mechanisms.

Fitting Species Abundance Distributions With the Gambin Model

In addition to the aforementioned methods, we fitted our data also using the gambin model with the following two aims: (1) to identify possible multimodal patterns (which cannot be modeled with the aforementioned types of distributions: LN,

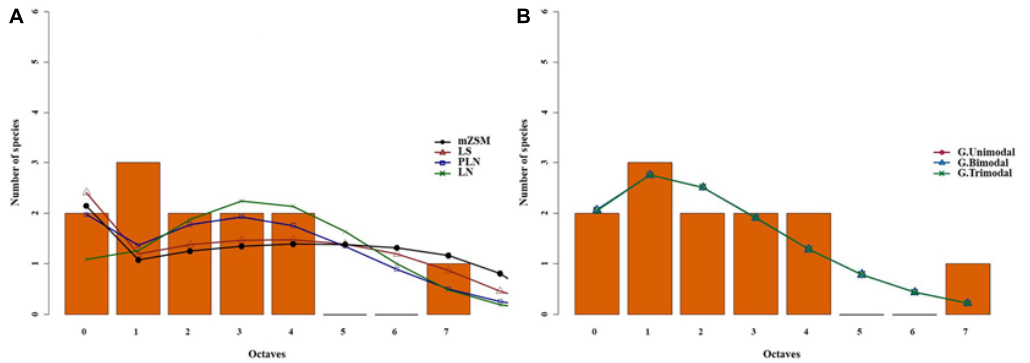


FIGURE 1 | Species abundance distribution (SAD) of carabid beetles in a Chinese desert steppe. SAD was modeled using the lognormal (LN), the Poisson-lognormal (PLN; zero-truncated version), the Fisher log-series (LS), and the metacommunity Zero-Sum Multinomial (mZSM) distributions **(A)** and unimodal, bimodal and trimodal gambin models **(B)**. LN, PLN, LS, and mZSM distributions were fitted to the raw abundance data (i.e., not binned). The data were binned for graphical comparison with gambin models, which require binning.

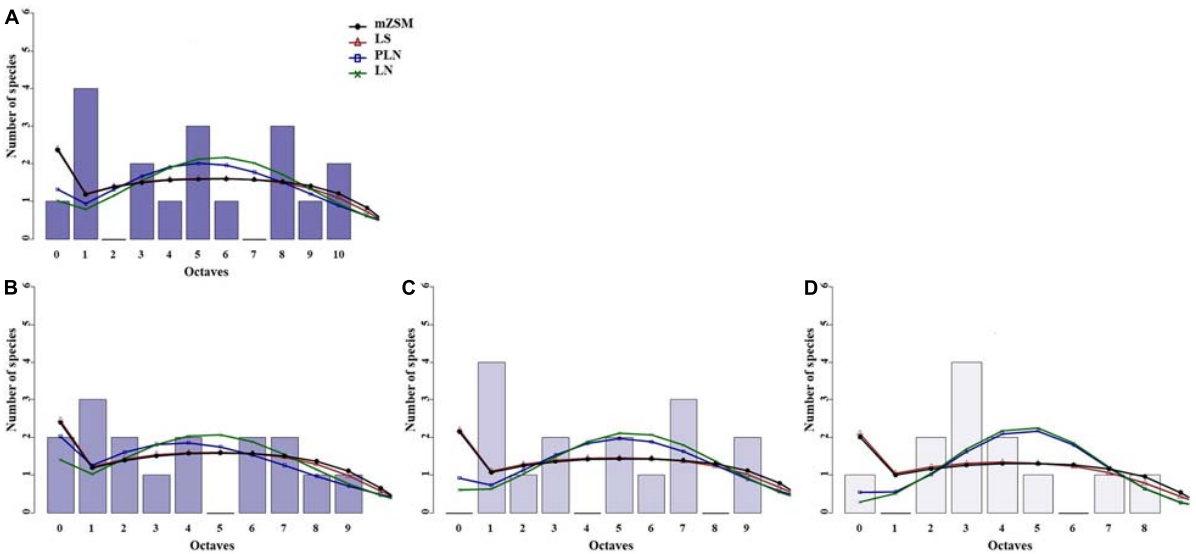


FIGURE 2 | Species abundance distribution (SADs) of carabid beetles in a Chinese typical steppe. SADs were modeled using the lognormal (LN), the Poisson-lognormal (PLN; zero-truncated version), the Fisher log-series (LS), and the metacommunity Zero-Sum Multinomial (mZSM) distributions. Analyses were done for the whole steppe **(A)** and three different sectors separately [**(B)**: ts1 sector, **(C)**: ts2 sector, **(D)**: ts3 sector]. Distributions were fitted to the raw abundance data (i.e., not binned). The data were binned for graphical comparison with gambin models, which require binning.

PLN, LS, and mZSM) and (2) to have a parameter (gambin α) that can be compared across models. Using the model selection procedure described above to identify which type of distribution best describes a certain SAD, we tested whether the various communities followed the same type of distribution, or if different communities followed different models, possibly as a result of different processes. Of course, communities that follow different models cannot be compared in terms of model parameters. With the use of the gambin model, all communities are fitted with the same model (the gambin), which is sufficiently flexible to approximate various shapes. In this case, the objective is not that of finding which type of distribution is best followed by a given community, but to use the same distribution (the gambin) to fit all communities and to see how, given this distribution, its parameter α changes between communities. Thus, the gambin model is not a competitive model to be contrasted with the LN, PLN, LS, and mZSM distributions. Rather, the simultaneous use of the gambin model and the aforementioned distribution models allows a deeper understanding of the studied SADs through the identification of the best fitting distribution (which may vary from a model to another) and the use of the same model (the gambin) whose parameter α can be compared. We used unimodal, bimodal and trimodal gambin models to fit the species abundance distribution of carabid beetles in each type of steppe and in each sector of three steppes separately. To fit gambin models, we used the function "*fit_abundances*" in the "*gambin*" library (Matthews et al., 2014) of the software R version 3.5.

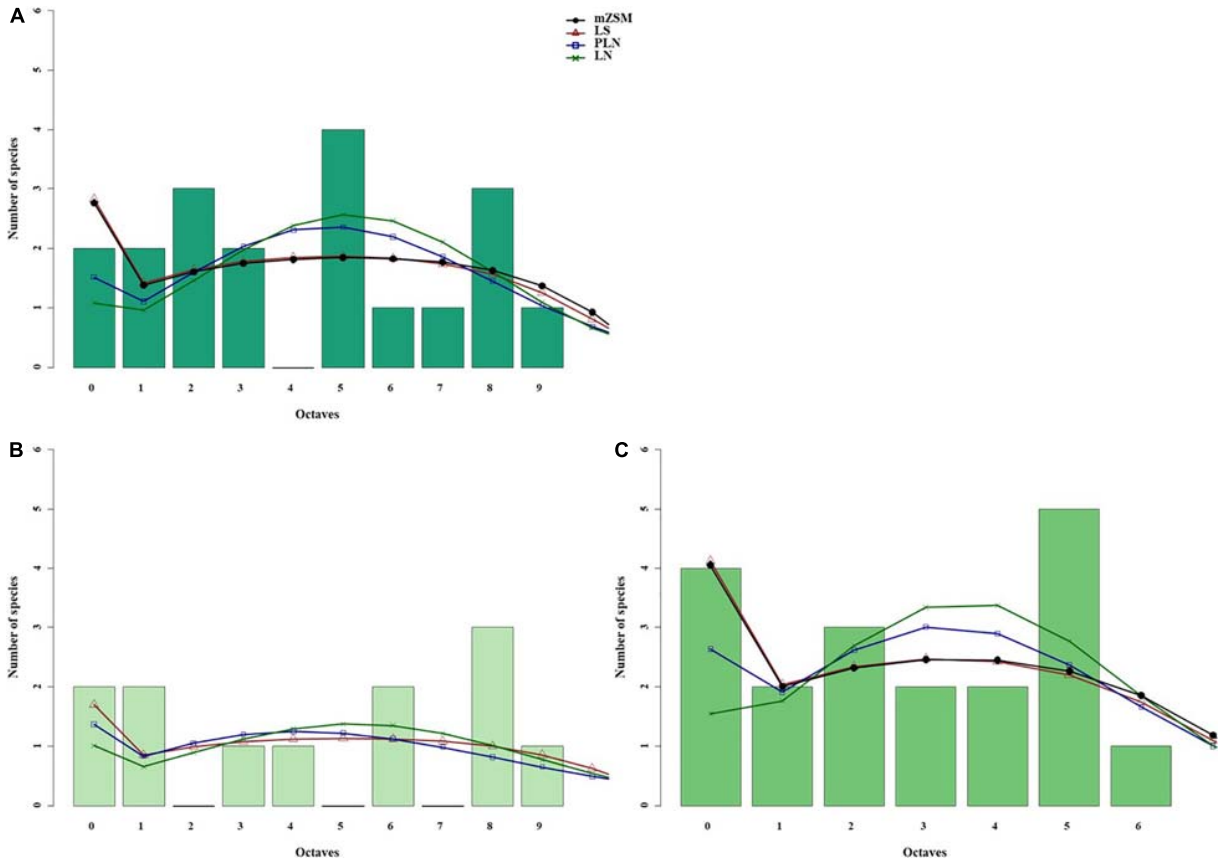


FIGURE 3 | Species abundance distribution (SADs) of carabid beetles in a Chinese meadow steppe. SADs were modeled using the lognormal (LN), the Poisson-lognormal (PLN; zero-truncated version), the Fisher log-series (LS), and the metacommunity Zero-Sum Multinomial (mZSM) distributions. Analyses were done for the whole steppe **(A)** and two different sectors separately [**(B)** ms1 sector, **(C)** ms2 sector]. Distributions were fitted to the raw abundance data (i.e., not binned). The data were binned for graphical comparison with gambin models, which require binning.

Rarefaction

Given that SAD model parameters are sensitive to variations in sample size (Maurer and McGill, 2011) and that we were interested in comparing model parameters across communities that are best fitted by the same models, we used a rarefaction procedure where, for each sample, we subsampled 338 individuals (which corresponded to the number of individuals collected in the desert steppe, the least abundant sample) and fitted the models to this subsample for a coarse-scale (between-steppe) approach.

Using the number of individuals collected in the desert steppe as a reference is appropriate for comparing the three types of steppes with sectors aggregated, because the typical and the meadow steppes have more sectors, and hence were subject to a higher sampling effort (more pitfall traps), than the desert one. However, how conducting rarefaction for a fine-scale (within steppe) approach is less obvious. Using desert sample as a reference has the advantage of allowing comparisons among sectors belonging to different steppes; but to compare different sectors of the same steppe, it might be more appropriate selecting the sector with the lowest total abundance as the reference sample for that given steppe. This might be ecologically sounder, because it uses as a reference the smallest sample of the ecosystem under study, not that of a different ecosystem. For this reason, we conducted analyses using both approaches, i.e., using as a reference sample both the desert steppe sample for all samples, and the smallest sample within each steppe for the various steppes separately.

Rarefaction, however, poses also another problem. If, one hand, rarefaction can be important to obtain comparable model parameters, on the other hand it may change considerably the shape of the distribution. Thus, use of rarefaction in a model selection procedure can be questionable. The aim of our model selection procedure was that of finding the model that best described each community; but, if data are rarefied, the consequent reduction in the number of species due to the exclusion of the rarest ones (because their rarefied abundance becomes zero), might artificially transform a certain original SAD, characteristic of a species rich-community, into a spurious one, characteristic of a species-poor community. When the objective is not that of testing how the parameters of a certain model change across communities, but is that of contrasting models with possibly completely different parameters to find which model best performs in each case, rarefaction might be

not only unnecessary, but also inappropriate. Thus, to explore the effect of rarefaction on the model selection procedures, we conducted the analyses by using both original and rarefied data, and compared the results. A summary of all forms of analyses conducted in this study is given in **Table 1**.

RESULTS

Overall Abundances

Overall, we collected 6,873 individuals belonging to 25 carabid species: 19 species were collected in the meadow steppe, of which 12 in the first sector (ms1) and 19 in the second sector (ms2); 18 species in the typical steppe, of which 16 in the first sector (ts1), 15 in the second sector (ts2), and 12 in the third sector (ts3); and, finally, 12 species in the sole sector of the desert steppe.

Globally, the most abundant species was *Carabus vladimirskyi*, accounting for about 25% of all collected carabids. However, dominant species varied according to the steppe type. In the desert steppe, the most abundant species was *Carabus glyptoterus*, which accounted for about 75% of the carabid individuals sampled in this steppe type. In the typical steppe, the most dominant species were *Carabus glyptoterus* (29%) and *Carabus vladimirskyi* (27%). Namely, *Carabus vladimirskyi* was the most dominant species in the typical steppe sector ts1, where it accounted for 40% of sampled individuals, followed by *Poecilus gebleri* (24%). Similarly, in the typical steppe sector ts2, *Carabus vladimirskyi* accounted for 33% of all sampled individuals, and *Poecilus gebleri* for 31%, respectively. In the typical steppe sector ts3, the most dominant species was *Carabus glyptoterus* (53%), followed by *Poecilus gebleri* (22%). *Carabus vladimirskyi* was the dominant species in the meadow steppe (35%). This species accounted for 42% of total sampled individuals in the meadow steppe sector ms1.

Best Fitting Distributions

Steppe-Level Analysis

The comparative analyses of SADs for the different types of steppes highlighted some differences between ground beetle communities (**Figures 1–3**). For the desert steppe (**Figure 1**), the best model for the carabid SAD was, in terms of AICc value, the LS distribution, although also the mZSM and the PLN gave similarly supported models. Both the typical and the meadow steppe communities were best modeled by the mZSM and the LS distributions. Fisher's α increased in the order: typical steppe < desert steppe < meadow steppe, whereas mZSM θ increased in the order: desert steppe < typical steppe < meadow steppe (**Table 2**).

When samples were rarefied to the number of individuals collected in the desert steppe, both the typical and the meadow steppe communities were best modeled by the mZSM and the LS distributions, and both mZSM θ and Fisher's α increased in the order: desert steppe < typical steppe < meadow steppe, with differences slightly more marked compared to non-rarefied samples (**Table 2**). Using the smallest within-steppe sample as a reference for rarefaction, results remained virtually identical (**Supplementary Table 1**).

Sector-Level Analysis

The three sectors of the typical steppe were best modeled by the mZSM and the LS distributions (**Figure 2**), showing similar values of mZSM θ and Fisher's α. The third sector was best modeled by the PLN distribution regarding the AICc value, but carabid distribution was significantly different from the PLN distribution (**Table 3**). Both sectors of the meadow steppe were best modeled by the mZSM and the LS distributions (**Figure 3**); the second was in fact best modeled only by PLN distribution, but the distribution deviates significantly from this model. In addition, the two meadow sectors had very different values of θ and Fisher's α (**Table 3**).

When samples were rarefied to the numbers of individuals collected in the desert steppe, all sectors were best modeled by the mZSM and the LS distributions, with ts3 showing lower estimates of θ and Fisher's α than ts1 and ts2. All other models received equal support (**Table 3**). Both sectors of the meadow steppe were best modeled by the mZSM and the LS distributions. Values of θ and Fisher's α were virtually identical to those obtained without rarefaction (**Table 3**). Using the smallest within-steppe sample as a reference for rarefaction, results remained virtually identical (**Supplementary Table 2**).

Gambin Models

Steppe-Level Analysis

The gambin unimodal model provided the best fit for the three types of steppe (**Figures 1**, **4**, **5**) regarding the AICc values. However, we found that the observed values deviated significantly

TABLE 2 | Model selection results for the species abundance distributions of carabid beetles in various types of Chinese steppes.

	Without rarefaction		*With rarefaction*	
Model parameter	***Estimate***	***AICc***		
Desert steppe				
LS (Fisher's α)	2.428	92.067		
mZSM (θ)	1.763	92.565		
PLN (μ)	1.302	93.368		
LN (meanlog)	1.937***	94.801		
Typical steppe				
mZSM (θ)	2.580	208.920	2.715	104.850
LS (Fisher's α)	2.411	209.550	2.654	106.247
PLN (μ)	3.051**	214.155	1.681	109.413
LN (meanlog)	3.614***	215.639	2.247	111.253
Meadow steppe				
mZSM (θ)	2.937	201.964	3.071	113.842
LS (Fisher's α)	2.830	202.480	3.068	115.025
PLN (μ)	2.745***	206.486	1.644	117.192
LN (meanlog)	3.217***	207.871	2.194	118.880

Analyses conducted at the steppe-level without and with rarefaction.
*Asterisks indicate significant deviations from the model (**$p < 0.01$, ***$p < 0.001$). Tested models were the lognormal (LN), the Poisson-lognormal (PLN; zero-truncated version), the Fisher log-series (LS), and the metacommunity Zero-Sum Multinomial (mZSM) distributions. AICc, corrected Akaike Information Criterion.*

TABLE 3 | Model selection results for the species abundance distributions of carabid beetles in various types of Chinese steppes.

	Without rarefaction			*With rarefaction*		
Typical steppe	***ts1***	***ts2***	***ts3***	***ts1***	***ts2***	***ts3***
mZSM (θ)	2.576 (158.528)	2.342 (164.737)	2.045 (116.790)	2.552 (97.957)	2.539 (100.408)	2.186 (98.981)
LS (Fisher's α)	2.458 (159.095)	2.213 (165.340)	2.113 (117.168)	2.403 (97.366)	2.494 (101.259)	2.285 (98.696)
PLN (μ)	1.780 (163.870)	3.066*** (167.889)	2.797*** (116.783)	1.111 (102.166)	1.609 (104.537)	2.211 (99.149)
LN (meanlog)	2.874*** (166.362)	3.373*** (168.609)	2.869*** (117.050)	2.223 (103.637)	2.339 (104.245)	2.402 (99.453)
Meadow steppe	***ms1***	***ms2***		***ms1***	***ms2***	
mZSM (θ)	1.868 (130.960)	4.273* (147.846)		1.822 (82.935)	4.265 (139.844)	
LS (Fisher's α)	1.706 (131.675)	4.170 (148.374)		1.653 (83.032)	4.181 (140.408)	
PLN (μ)	2.019 (137.601)	1.612* (153.055)		2.339 (86.640)	1.638 (144.025)	
LN (meanlog)	3.293*** (139.353)	2.111*** (154.418)		2.712 (85.393)	2.046 (146.229)	

Analyses conducted at the sector level without and with rarefaction.
Asterisks indicate significant deviations from the model ($p < 0.05$, ***$p < 0.001$).*
For each sector model parameter estimates are given and, in parentheses, the corrected Akaike Information Criterion (AICc). Tested models were the lognormal (LN), the Poisson-lognormal (PLN; zero-truncated version), the Fisher log-series (LS), and the metacommunity Zero-Sum Multinomial (mZSM) distributions.

from the expected values calculated with the gambin unimodal model. As regards α-values for the unimodal models, the typical and meadow steppes had similar values, much higher than that recorded in the desert steppe. The α-values were consistent between the three modes in desert steppe. In contrast, the α-value for the second mode of bimodal gambin in typical and meadow steppes appeared particularly high (**Table 4**).

Using rarefaction with the desert steppe as reference, the Gambin uni- and bimodal models provided similar fits for both the typical and the meadow steppes, but the α-values for the bimodal models appeared extremely high. As regards α-values for the unimodal models, the typical steppe had a higher value than the meadow steppe (**Table 4**). Using the smallest within-steppe sample as a reference for rarefaction, results remained virtually identical, although α-values for the unimodal model decreased (**Supplementary Table 3**).

Sector-Level Analysis

For the typical steppe sectors (**Figure 4**), the bimodal and trimodal gambin models provided equally supported best fits for the first sector (ts1), but the α-values for the second mode of the bimodal model and for the second and third modes of the trimodal model were extremely high (**Table 5**). For another sector (ts2), the gambin unimodal model provided the best fit regarding the AICc value, but this model showed a significant deviation from observed data. The α-values for the first mode of the bimodal model and for the second and third modes of the trimodal model were extremely high (**Table 5**). Finally, in the third sector (ts3) the gambin unimodal model provided the best fit (**Table 5**).

When the two sectors of the meadow steppe were considered separately, gambin unimodal models provided similar best fit for both sectors. However, *p*-values were significant in both sectors for the unimodal model. In addition, the α-values of the second modes of the bimodal model of both sectors appeared particularly high. The trimodal gambin model seems to be preferable for both sectors, but α-values for the first and third modes in ms1 and for the third mode in ms2 were extremely high. Thus, at the sector level, the gambin models did not provide an acceptable fit of the data (**Table 5** and **Figure 5**).

Using the desert steppe as a reference for rarefaction, the unimodal gambin model provided the best fit for the three sectors of the typical steppe, with similar α-values. Regarding the AICc values, the bimodal model was the best model for the three sectors; however, this bimodal model provided extremely high values in ts2 and relatively high values in ts1 and ts3 (**Table 5**). When the two sectors of the meadow steppe were considered separately, gambin unimodal models provided the best fit for sectors ms1 and ms2 (**Table 5**). Using the smallest within-steppe sample as a reference for rarefaction, for the typical steppe sectors, the unimodal gambin models provided the best fit for the two sectors ts1 and ts2 (**Supplementary Table 4**).

DISCUSSION

We constructed SADs for carabid beetles of three types of Central Asian steppes at different scales, i.e., for different habitats within each steppe type and for each steppe type as a whole. We found that the basic results (i.e., SAD shape) of the analyses conducted at the smaller scale (sectors within steppe types) mirrored those obtained at the larger (steppe type) scale. This finding suggests that the various sectors within the same steppe type are relatively homogeneous, at least for the SADs of carabid communities. Thus, the hypothetical communities that can be delimited on a sector basis actually overlap each to other and the resulting single pooled community arising at the steppe type level maintains the basic SAD shape of the constituting local communities. This is very interesting since in general SAD shapes tend to change with the addition of sub-communities, i.e., with increasing sampling effort (Borda-de-Água et al., 2012).

In all cases, SADs were best modeled by the Fisher log-series (LS) and the metacommunity Zero-Sum Multinomial (mZSM) distributions. This contrasts with our expectation that the best fitting models were different among steppes. In particular, the fact that the LS distribution fitted well also the meadow steppe community was an unexpected result. The LS distribution is

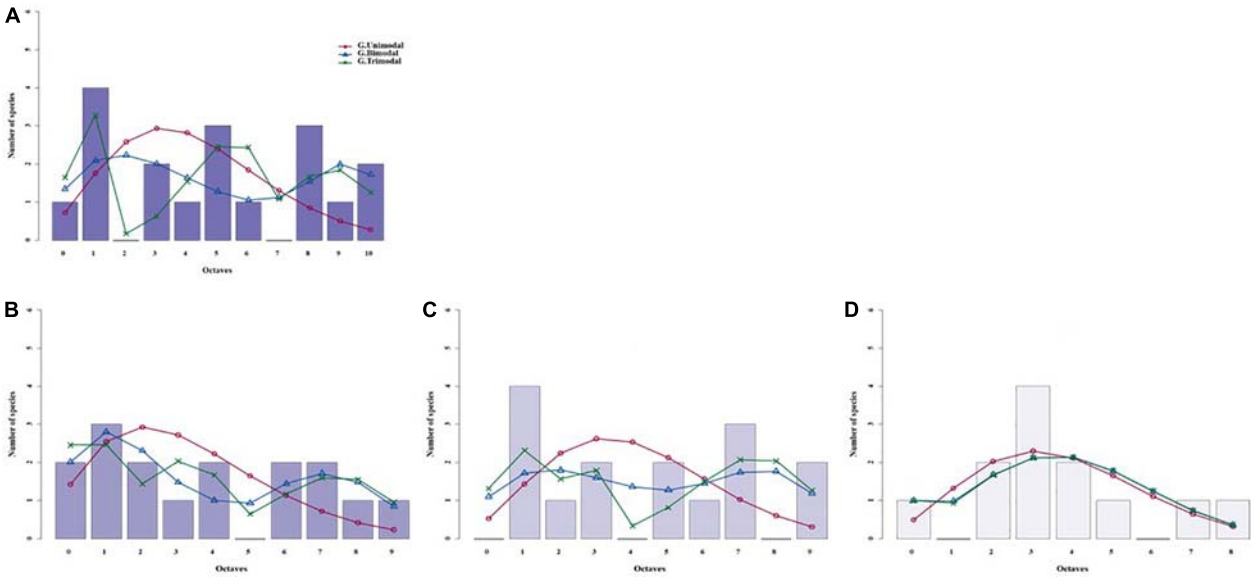


FIGURE 4 | Species abundance distributions (SADs) of carabid beetles in a Chinese typical steppe fitted with unimodal, bimodal and trimodal gambin models. Analyses were done for the whole steppe **(A)** and three different sectors separately [**(B)** ts1 sector, **(C)** ts2 sector, **(D)** ts3 sector].

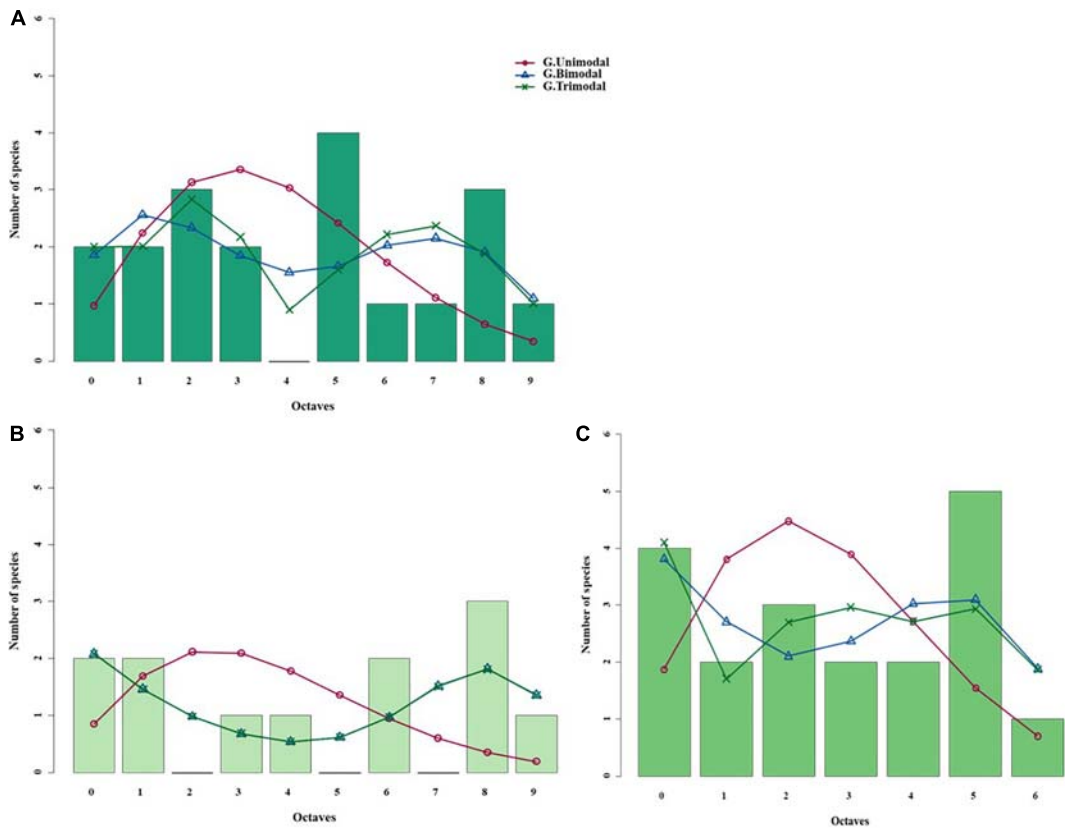


FIGURE 5 | Species abundance distributions (SADs) of carabid beetles in a Chinese meadow steppe fitted with unimodal, bimodal and trimodal gambin models. Analyses were done for the whole steppe **(A)** and two different sectors separately [**(B)** ms1 sector, **(C)** ms2 sector].

TABLE 4 | Values of gambin α and model's corrected Akaike Information Criterion (AICc) for species abundance distributions of carabid beetles in various types of Chinese steppes.

	Without rarefaction				*With rarefaction*			
	α1	**α2**	**α3**	***AICc***				
Desert steppe								
Gambin 1 mode	2.346	–	–	51.664				
Gambin 2 modes	0.643	2.346	–	66.331				
Gambin 3 modes	1.542	2.178	2.380	110.331				
Typical steppe								
Gambin 1 mode**	4.328	–	–	99.619	5.939	–	–	58.677
Gambin 2 modes	2.410	259.829	–	101.112	8091.081	108.341	–	66.915
Gambin 3 modes	26.491	90.910	120.406	114.151	19.536	30.973	120.334	106.290
Meadow steppe								
Gambin 1 mode	4.267	–	–	98.185	4.934	–	–	63.490
Gambin 2 modes	2.651	54.175	–	101.662	23.797	2849.67	–	71.693
Gambin 3 modes	0.858	16.105	38.49039	116.454	9.534	31.437	78.996	98.130

Analyses conducted at the steppe level without and with rarefaction.
*Asterisks indicate significant deviations from the model (**$p < 0.01$).*

TABLE 5 | Values of gambin α and model's corrected Akaike Information Criterion (AICc) for species abundance distributions of carabid beetles in various types of Chinese steppes.

	Without rarefaction				*With rarefaction*			
	α1	**α2**	**α3**	***AICc***	**α1**	**α2**	**α3**	***AICc***
Typical steppe (ts1)								
Gambin 1 mode	3.002	–	–	77.686	5.097	–	–	56.101
Gambin 2 modes	3.511	50.663	–	71.561	53.278	196.029	–	64.782
Gambin 3 modes	5.589	30.528	70.575	71.469	27.324	47.242	169.779	197.674
Typical steppe (ts2)								
Gambin 1 mode	5.216	–	–	78.477	5.973	–	–	55.772
Gambin 2 modes	103.551	2.515	–	84.799	13168.640	2443.329	–	65.283
Gambin 3 modes	14.167	50.683	73.221	106.028	15.301	14.231	93.214	112.813
Typical steppe (ts3)								
Gambin 1 mode	5.586	–	–	56.047	5.348	–	–	50.596
Gambin 2 modes	0.000	7.125	–	70.110	2.194	25.905	–	66.012
Gambin 3 modes	0.581	3.463	7.231	114.038	14.471	18.080	17.608	120.185
Meadow steppe (ms1)								
Gambin 1 mode***	3.499	–	–	70.644	7.437	–	–	45.030
Gambin 2 modes	0.929	132.736	–	73.158	13.022	27.274	–	68.361
Gambin 3 modes	91.872	0.929	150.278	117.153	8.926	13.465	47.152	127.649
Meadow steppe (ms2)								
Gambin 1 mode*	4.527	–	–	84.240	4.191	–	–	79.982
Gambin 2 modes	1.155	44.233	–	86.904	3.400	29.257	–	84.438
Gambin 3 modes	0.021	14.946	41.568	102.521	4.462	12.212	39.270	100.574

Analyses conducted at the sector level without and with rarefaction.
Asterisks indicate significant deviations from the model ($p < 0.05$, ***$p < 0.001$).*

known to fit a wide range of communities, particularly those that have a high frequency of rare species and that are not particularly species rich (Kempton and Taylor, 1974; May, 1975; Magurran and Henderson, 2003; Magurran, 2004; McGill et al., 2007), which is exactly the case of carabids inhabiting arid and semiarid environments. The LS has been criticized as it might result for inadequate sampling (Borda-de-Água et al., 2012; Chen and Shen, 2017). With increasing sampling, the number of collected species also increases, as well as the abundance of species that seem rare at low sampling intensity (Magurran and Henderson, 2003; Green and Plotkin, 2007). As sample size increases, and Preston's "veil line" (Preston, 1948) is pulled back to reveal the mode of the distribution, the log-series turns into log-normal (LN) distribution (Ulrich et al., 2010). However, and at very large sampling efforts (scales) the LS can be recovered again due to the addition of many specialist species or spatially rare species

(Borda-de-Água et al., 2012). We can exclude that our data are affected by these sampling problems, not only because of the intensity of our sampling, but also because carabid communities of Asian arid and semiarid environments are always characterized by a low number of species and a high dominance (see, for example, Li et al., 2016; Liu et al., 2016; Khurelpurev and Pfeiffer, 2017). The good fit provided by the LS distribution also for the carabid community of the meadow steppe indicates that even this ecosystem, which is the less arid of the three investigated here, still hosts a relatively simple and highly dominated community.

The mZSM distribution is strictly related to the LS (the log-series distribution is a limiting case of the mZSM) and so it is not surprising that they provided similar fits. The α parameter of Fisher's LS is a measure of diversity, and reflects the number of singleton species in the community (Kempton and Taylor, 1974; Magurran and Henderson, 2003). The θ parameter of the mZSM tends to Fisher's α as the number of individuals in the sample increases, and this is clearly reflected in our study, with differences between α- and θ-values being very small in the meadow and typical steppes, compared with the desert steppe, from which much fewer individuals were collected.

In our study, Fisher's α increased in the order: typical steppe < desert steppe < meadow steppe, whereas θ increased in the order: desert steppe < typical steppe < meadow steppe. These contrasting results can be a consequence of unequal sampling efforts. As both the overall shape of an empirical SAD (and hence its best fitting model), as well its parameter value(s), are influenced by the overall number of collected individuals (which, in turn, may depend on the sampling effort), it has been proposed of using rarefaction techniques to compare SADs of communities with different total abundances (Stier et al., 2016; Chen and Shen, 2017; Borges et al., 2020). At first glance, this sounds perfectly logical, but one should consider whether different sample sizes are the result of different sampling efforts or reflect true differences in community structure. While in the first case rarefaction is correct, in the second one it can turn a true SAD into a spurious one by unjustifiably wiping out the species with the lowest abundances from a community where they are present. In our case, rarefaction applied to the steppe level appeared justified, because of the different sampling effort (sampling effort in the typical and meadow steppes was two and three times that of the desert steppe). With the use of rarefaction, both Fisher's α-values and θ-values increased in the order: desert steppe < typical steppe < meadow steppe, which indicates an increasing diversity from the most to the least arid type of steppe, which can be in turn related to the increasing vegetation cover, as it seems that communities of ground-dwelling arthropod predators and detritivores are mainly determined by vegetation cover and aboveground plant biomass (Pan et al., 2018; Niu et al., 2020). These results are also consistent with our findings in a previous study (Tsafack et al., 2020), where the desert steppe showed lower Brillouin and Shannon-Wiener diversity indices, while typical and meadow steppes showed equally higher species diversity values; the species richness indices Chao and Margalef showed the same results, increasing in the order: desert steppe < typical steppe < meadow steppe. In a simulation study, Beck and Schwanghart (2010) found that Fisher's α distribution values lose their stability when approaching full completeness sampling, so they recommend to prefer Shannon index or biased-controlled indices to assess biodiversity in full sampled communities. The present study confirms that Fisher's α can be as reliable as Shannon-Wiener index to measure diversity in quite completely sampled sites (as it seems the case in this study) when distributions are correctly rarefied. Regarding the similar increasing values of α- and θ-values, this equivalence have been also described by Hubbell et al. (2008) and Matthews and Whittaker (2014b). In addition, He and Hu (2005) found an analytical relationship between the parameter θ and Simpson's diversity index.

At the sector level, without rarefaction, results were less clear. Fisher's α-values for the typical steppe sectors were similar to that of the desert steppe sector (slightly higher in one case, and slightly lower in two cases), whereas those for the meadow steppe sectors were substantially lower in one case and higher in the other. θ-values for the typical steppe sectors were higher than that of the desert steppe sectors in all three cases, whereas in the meadow steppe one sector had a θ larger than the typical steppe sectors, but the other was lower. Thus, for θ-values, the pattern desert steppe < typical steppe < meadow steppe was confirmed, with the exception of one meadow sector.

These results indicate that, although the basic SAD shape is not influenced by the scale, parameter values are, and α-values were more influenced than θ-values. We also found that, if data are subject to rarefaction, the overall change was a decrease in α- and θ-values. This indicates that the exclusion of the rarest species following rarefaction reduces diversity. As at the sector level sampling effort was homogeneous (same number of traps and same number of days of sampling), rarefaction altered diversity estimates unduly.

The use of gambin model allowed the possibility of investigating how its parameter α (which should not be confused with Fisher's α) varies across carabid communities. Although gambin multimodal models provided an apparently good fit in some cases, their α-values appeared exaggeratedly high, which suggests that the multimodal models actually overfitted the data. The multimodal pattern may indicate that communities are composed by different guilds (Marquet et al., 2004) or groups of species with different dispersal abilities (Borda-de-Água et al., 2017).

At the steppe type level (large scale), the desert steppe was characterized by a gambin α-value much lower than those of the typical and meadow steppes, whereas the typical steppe had a value slightly higher than the meadow steppe. When data were rarefied, the difference between the typical and the meadow steppes became more evident, and gambin α-values followed the same pattern of θ-values for non-rarefied data. At sector level, results were complex and variable. The typical and meadow sectors had gambin α-values always higher than the desert steppe. However, two sectors of the typical steppe had gambin α-values larger than the meadow steppe sectors, but one was lower. Using a rarefaction based on desert steppe sample size, one sector of the meadow steppe had a gambin α-value distinctly higher than those of the typical steppe sectors, and the other lower. Finally, using rarefactions based on within steppe smallest samples, one sector of the meadow steppe had the highest gambin α-value, and the other a value which felt within the range of the typical

steppe sectors. Thus, the gambin α parameter resulted efficient in indicating a strong difference between the most challenging environment (desert steppe) on one hand and the other two steppes on the other hand, but did not provide a clear distinction between the typical and the meadow steppe. On this regard, Fisher's α and mZSM θ values provided a clearer response. However, Fisher's α, mZSM θ, and gambin α values are all substantially lower in the desert steppe and higher in the typical and meadow steppes, which implies a decreasing gradient of environmental harshness from the desert steppe to the typical and the meadow steppes. Possibly *Carabus glyptoterus*, which accounted for about 75% of the carabid individuals sampled in desert steppe, is a generalist species able to cope with severe climatic conditions, but the ecology and biology of this species is still poorly known.

CONCLUSION

In line with previous studies on SADs, the present study contributes to highlight the multiple uses of these models. We compared the SADs of carabid communities of three types of steppes with different aridity conditions (desert, typical and meadow steppes) in China to identify how SADs change between these ecosystems. Contrary to our expectation, all species distributions, including that of the meadow steppe, were best fitted by the LS and the mZSM distributions. However, parameters values (α and θ) were different between the three environments.

Using rarefied data to account for the unequal sampling effort, we found that both Fisher's α- and mZSM θ-values increased in the order: desert steppe < typical steppe < meadow steppe, which indicates an increasing diversity from the most to least arid type of steppe. Regarding gambin models, contrary to our expectations, they were not always efficient in modeling all distributions, although the α-value of the unimodal model clearly indicates a lower diversity of the desert steppe compared to the typical and meadow steppes. In some cases, gambin multimodal models provided good fit, which indicates that communities are composed by different guilds, but we found exaggeratedly high gambin α-values. In conclusion we found that (1) rarefaction emphasized the differences observed between parameters (Fisher's α, mZSM θ, and gambin α) and (2) SAD shapes were not influenced by the scale, but parameter values are, and Fisher's α- and gambin α-values were more influenced than θ-values.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: Tsafack, 2018, "Data_Carabidae in China Grasslands," DOI: 10.7910/DVN/FKBCRQ, Harvard Dataverse, V1.

AUTHOR CONTRIBUTIONS

NT collected the data, conceived and designed the analyses, analyzed the data, prepared the figures and/or tables, and authored and reviewed the drafts of the manuscript. PB conceived and designed the analyses and authored and reviewed the drafts of the manuscript. YX and XW conceived and designed the field experiments, performed the data collection, contributed materials, authored or reviewed the drafts of the manuscript, and approved the final draft. SF conceived and designed the analyses, led manuscript writing, and authored and reviewed the drafts of the manuscript. All authors contributed to content and editing of the manuscript.

FUNDING

The research was funded by the China Postdoctoral Science Foundation (2018M643770) granted to NT; by the first-class discipline of Practaculture Science of Ningxia University (No. NXYLXK2017A01) and the National Natural Science Foundation of China (No. 31660630) granted to XW; by the Fundação para a Ciência e a Tecnologia, project FCT-UIDB/00329/2020-2024 and Fundação Gaspar Fructuoso granted to PB.

ACKNOWLEDGMENTS

We thank the private landowners for allowing us to place the pitfall traps on their lands. We are grateful to Prof. Hongbin Liang (Institute of Zoology, CAS) for identifying the collected specimens and Master students, particularly to Hui Wang of the School of Agriculture Ningxia University, promotion 2017–2018, for their technical support during field and laboratory work. We also thank Thomas J. Matthews, François Rigal, and Pedro Cardoso for their helpful comments on data analysis.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fevo.2021.603436/full#supplementary-material

REFERENCES

Akiyama, T., and Kawamura, K. (2007). Grassland degradation in China: methods of monitoring, management and restoration. *Grassl. Sci.* 53, 1–17. doi: 10.1111/j.1744-697X.2007.00073.x

Alonso, D., and McKane, A. J. (2004). Sampling Hubbell's neutral theory of biodiversity. *Ecol. Lett.* 7, 901–910. doi: 10.1111/j.1461-0248.2004.00640.x

Baldridge, E., Harris, D. J., Xiao, X., and White, E. P. (2016). An extensive comparison of species-abundance distribution models. *PeerJ* 4:e2823. doi: 10.7717/peerj.2823

Balvanera, P., Pfisterer, A. B., Buchmann, N., He, J.-S., Nakashizuka, T., Raffaelli, D., et al. (2006). Quantifying the evidence for biodiversity effects on ecosystem functioning and services. *Ecol. Lett.* 9, 1146–1156. doi: 10.1111/j.1461-0248.2006.00963.x

Barnosky, A. D., Matzke, N., Tomiya, S., Wogan, G. O. U., Swartz, B., Quental, T. B., et al. (2011). Has the Earth's sixth mass extinction already arrived? *Nature* 471, 51–57. doi: 10.1038/nature09678

Beck, J., and Schwanghart, W. (2010). Comparing measures of species diversity from incomplete inventories: an update. *Methods Ecol. Evol.* 1, 38–44. doi: 10.1111/j.2041-210X.2009.00003.x

Borda-de-Água, L., Borges, P. A. V., Hubbell, S. P., and Pereira, H. M. (2012). Spatial scaling of species abundance distributions. *Ecography* 35, 549–556. doi: 10.1111/j.1600-0587.2011.07128.x

Borda-de-Água, L., Whittaker, R. J., Cardoso, P., Rigal, F., Santos, A. M. C., Amorim, I. R., et al. (2017). Dispersal ability determines the scaling properties of species abundance distributions: a case study using arthropods from the Azores. *Sci. Rep.* 7:3899. doi: 10.1038/s41598-017-04126-5

Borges, P. A. V., Gabriel, R., and Fattorini, S. (2019). "Biodiversity erosion: causes and consequences," in *Life on Land*, eds W. Leal Filho, A. M. Azul, L. Brandli, P. G. Özuyar, and T. Wall (Cham: Springer International Publishing), 1–10.

Borges, P. A. V., Rigal, F., Ros-Prieto, A., and Cardoso, P. (2020). Increase of insular exotic arthropod diversity is a fundamental dimension of the current biodiversity crisis. *Insect Conserv. Divers.* 13, 508–518. doi: 10.1111/icad.12431

Bulmer, M. G. (1974). On fitting the poisson lognormal distribution to species-abundance data. *Biometrics* 30, 101–110. doi: 10.2307/2529621

Burnham, K. P., and Anderson, D. R. (2002). *Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach*, 2nd Edn. New York: Springer-Verlag.

Cardarelli, E., and Bogliani, G. (2014). Effects of grass management intensity on ground beetle assemblages in rice field banks. *Agric. Ecosyst. Environ.* 195, 120–126. doi: 10.1016/j.agee.2014.05.004

Chen, Y., and Shen, T.-J. (2017). Rarefaction and extrapolation of species richness using an area-based Fisher's logseries. *Ecol. Evol.* 7, 10066–10078. doi: 10.1002/ece3.3509

Connolly, S. R., MacNeil, M. A., Caley, M. J., Knowlton, N., Cripps, E., Hisano, M., et al. (2014). Commonness and rarity in the marine biosphere. *Proc. Natl. Acad. Sci. U.S.A.* 111:8524. doi: 10.1073/pnas.1406664111

Dajoz, R. (2002). *Les Coléoptères Carabides et Ténébrionidés*. Paris: Lavoisier.

Dornelas, M., Moonen, A. C., Magurran, A. E., and Bàrberi, P. (2009). Species abundance distributions reveal environmental heterogeneity in modified landscapes. *J. Appl. Ecol.* 46, 666–672. doi: 10.1111/j.1365-2664.2009.01640.x

Duan, M., Liu, Y., Yu, Z., Baudry, J., Li, L., Wang, C., et al. (2016). Disentangling effects of abiotic factors and biotic interactions on cross-taxon congruence in species turnover patterns of plants, moths and beetles. *Sci. Rep.* 6:23511.

Engen, S., and Lande, R. (1996). Population dynamic models generating species abundance distributions of the gamma type. *J. Theor. Biol.* 178, 325–331. doi: 10.1006/jtbi.1996.0028

Etienne, R. S., Alonso, D., and McKane, A. J. (2007). The zero-sum assumption in neutral biodiversity theory. *J. Theor. Biol.* 248, 522–536. doi: 10.1016/j.jtbi.2007.06.010

Fattorini, S., Di Lorenzo, T., and Galassi, D. M. P. (2018). Earthquake impacts on microcrustacean communities inhabiting groundwater-fed springs alter species-abundance distribution patterns. *Sci. Rep.* 8:1501. doi: 10.1038/s41598-018-20011-1

Fattorini, S., Rigal, F., Cardoso, P., and Borges, P. A. V. (2016). Using species abundance distribution models and diversity indices for biogeographical analyses. *Acta Oecol.* 70, 21–28. doi: 10.1016/j.actao.2015.11.003

Fisher, R. A., Corbet, A. S., and Williams, C. B. (1943). The relation between the number of species and the number of individuals in a random sample of an animal population. *J. Anim. Ecol.* 12, 42–58. doi: 10.2307/1411

Gobbi, M., Ballarin, F., Brambilla, M., Compostella, C., Isaia, M., Losapio, G., et al. (2017). Life in harsh environments: carabid and spider trait types and functional diversity on a debris-covered glacier and along its foreland. *Ecol. Entomol.* 42, 838–848. doi: 10.1111/een.12456

Green, J. L., and Plotkin, J. B. (2007). A statistical theory for sampling species abundances. *Ecol. Lett.* 10, 1037–1045. doi: 10.1111/j.1461-0248.2007.01101.x

Haddad, C. R., de Jager, L. J. C., and Foord, S. H. (2019). Habitats and cardinal directions are key variables structuring spider leaf litter assemblages under *Searsia lancea*. *Pedobiologia* 73, 10–19. doi: 10.1016/j.pedobi.2019.01.002

Harte, J., Zillio, T., Conlisk, E., and Smith, A. B. (2008). Maximum entropy and the state-variable approach to macroecology. *Ecology* 89, 2700–2711.

He, F., and Hu, X.-S. (2005). Hubbell's fundamental biodiversity parameter and the Simpson diversity index. *Ecol. Lett.* 8, 386–390. doi: 10.1111/j.1461-0248.2005.00729.x

Hubbell, S. (2001). *The Unified Neutral Theory of Biodiversity and Biogeography*. Princeton, NJ: Princeton University Press.

Hubbell, S. P., He, F., Condit, R., Borda-de-Água, L., Kellner, J., and ter Steege, H. (2008). How many tree species are there in the Amazon and how many of them will go extinct? *Proc. Natl. Acad. Sci. U.S.A.* 105(Suppl. 1):11498. doi: 10.1073/pnas.0801915105

Ibanez, T., Keppel, G., Baider, C., Birkinshaw, C., Florens, F. B. V., Laidlaw, M., et al. (2020). Tropical cyclones and island area shape species abundance distributions of local tree communities. *Oikos* 129, 1856–1866. doi: 10.1111/oik.07501

Jouveau, S., Toïgo, M., Giffard, B., Castagneyrol, B., van Halder, I., Vétillard, F., et al. (2019). Carabid activity-density increases with forest vegetation diversity at different spatial scales. *Insect Conserv. Divers.* 13, 36–46. doi: 10.1111/icad.12372

Kang, L., Han, X., Zhang, Z., and Sun, O. J. (2007). Grassland ecosystems in China: review of current knowledge and research advancement. *Philos. Trans. R. Soc. B Biol. Sci.* 362, 997–1008. doi: 10.1098/rstb.2007.2029

Kehoe, L., Romero-Muñoz, A., Polaina, E., Estes, L., Kreft, H., and Kuemmerle, T. (2017). Biodiversity at risk under future cropland expansion and intensification. *Nat. Ecol. Evol.* 1, 1129–1135. doi: 10.1038/s41559-017-0234-3

Kempton, R. A., and Taylor, L. R. (1974). Log-series and log-normal parameters as diversity discriminants for the lepidoptera. *J. Anim. Ecol.* 43, 381–399. doi: 10.2307/3371

Khurelpurev, O., and Pfeiffer, M. (2017). Coleoptera in the Altai Mountains (Mongolia): species richness and community patterns along an ecological gradient. *J. Asia Pacific Biodivers.* 10, 362–370. doi: 10.1016/j.japb.2017.06.007

Kim, D.-H., Cho, W.-S., and Chon, T.-S. (2013). Self-organizing map and species abundance distribution of stream benthic macroinvertebrates in revealing community patterns in different seasons. *Ecol. Inform.* 17, 14–29. doi: 10.1016/j.ecoinf.2013.06.006

Kitzes, J., and Harte, J. (2015). Predicting extinction debt from community patterns. *Ecology* 96, 2127–2136. doi: 10.1890/14-1594.1

Koivula, M. (2011). Useful model organisms, indicators, or both? Ground beetles (Coleoptera, Carabidae) reflecting environmental conditions. *ZooKeys* 100, 287–317. doi: 10.3897/zookeys.100.1533

Kotze, D. J., Brandmayr, P., Casale, A., Dauffy-Richard, E., Dekoninck, W., Koivula, M., et al. (2011). Forty years of carabid beetle research in Europe – from taxonomy, biology, ecology and population studies to bioindication, habitat assessment and conservation. *ZooKeys* 100, 55–148. doi: 10.3897/zookeys.100.1523

Labruyere, S., Bohan, D. A., Biju-Duval, L., Ricci, B., and Petit, S. (2016). Local, neighbor and landscape effects on the abundance of weed seed-eating carabids in arable fields: a nationwide analysis. *Basic Appl. Ecol.* 17, 230–239. doi: 10.1016/j.baae.2015.10.008

Li, F.-R., Liu, J.-L., Sun, T.-S., Ma, L.-F., Liu, L.-L., and Yang, K. (2016). Impact of established shrub shelterbelts around oases on the diversity of ground beetles in arid ecosystems of Northwestern China. *Insect Conserv. Divers.* 9, 135–148. doi: 10.1111/icad.12152

Liu, J.-L., Li, F.-R., Sun, T.-S., Ma, L.-F., Liu, L.-L., and Yang, K. (2016). Interactive effects of vegetation and soil determine the composition and diversity of carabid and tenebrionid functional groups in an arid ecosystem. *J. Arid Environ.* 128, 80–90. doi: 10.1016/j.jaridenv.2016.01.009

Lü, Y., Fu, B., Wei, W., Yu, X., and Sun, R. (2011). Major ecosystems in China: dynamics and challenges for sustainable management. *Environ. Manag.* 48, 13–27. doi: 10.1007/s00267-011-9684-6

Magurran, A. E. (2004). *Measuring Biological Diversity*. Hoboken, NJ: Wiley-Blackwell.

Magurran, A. E., and Henderson, P. A. (2003). Explaining the excess of rare species in natural species abundance distributions. *Nature* 422, 714–716. doi: 10.1038/nature01547

Marquet, P. A., Fernandez, M., Navarrete, S. A., and Valdovinos, C. (2004). "Diversity emerging: toward a deconstruction of biodiversity patterns," in *Frontiers of Biogeography?: New Directions in the Geography of Nature*, eds M. V. Lomolino and L. R. Heaney (Sunderland, MA: Sinauer Associates), 191–209.

Matthews, T. J., and Whittaker, R. J. (2014a). Fitting and comparing competing models of the species abundance distribution: assessment and prospect. *Front. Biogeogr.* 6:2. doi: 10.21425/F5FBG20607

Matthews, T. J., and Whittaker, R. J. (2014b). Neutral theory and the species abundance distribution: recent developments and prospects for unifying niche and neutral perspectives. *Ecol. Evol.* 4, 2263–2277. doi: 10.1002/ece3.1092

Matthews, T. J., and Whittaker, R. J. (2015). REVIEW: on the species abundance distribution in applied ecology and biodiversity management. *J. Appl. Ecol.* 52, 443–454. doi: 10.1111/1365-2664.12380

Matthews, T. J., Borregaard, M. K., Gillespie, C. S., Rigal, F., Ugland, K. I., Krüger, R. F., et al. (2019). Extension of the gambin model to multimodal species abundance distributions. *Methods Ecol. Evol.* 10, 432–437. doi: 10.1111/2041-210X.13122

Matthews, T. J., Borregaard, M. K., Ugland, K. I., Borges, P. A. V., Rigal, F., Cardoso, P., et al. (2014). The gambin model provides a superior fit to species abundance distributions with a single free parameter: evidence, implementation and interpretation. *Ecography* 37, 1002–1011. doi: 10.1111/ecog.00861

Maurer, B. A., and McGill, B. J. (2011). "Measurement of species diversity," in *Biological Diversity: Frontiers in Measurement and Assessment*, eds A. E. Magurran and B. J. McGill (Oxford: Oxford University Press), 55–64.

May, R. (1975). "Patterns of species abundance and diversity," in *Ecology and Evolution of Communities*, eds M. L. D. Cody and J. M. Diamond (Cambridge, MA: Harvard University Press), 81–120.

McCann, K. S. (2000). The diversity–stability debate. *Nature* 405, 228–233. doi: 10.1038/35012234

McGill, B. J. (2003). A test of the unified neutral theory of biodiversity. *Nature* 422, 881–885. doi: 10.1038/nature01583

McGill, B. J., Etienne, R. S., Gray, J. S., Alonso, D., Anderson, M. J., Benecha, H. K., et al. (2007). Species abundance distributions: moving beyond single prediction theories to integration within an ecological framework. *Ecol. Lett.* 10, 995–1015. doi: 10.1111/j.1461-0248.2007.01094.x

Miličić, M., Vujiæ, A., Jurca, T., and Cardoso, P. (2017). Designating conservation priorities for Southeast European hoverflies (Diptera: Syrphidae) based on species distribution models and species vulnerability. *Insect Conserv. Divers.* 10, 354–366. doi: 10.1111/icad.12232

Niemelä, J. (1993). Mystery of the missing species: species-abundance distribution of boreal ground beetles. *Ann. Zool. Fennici* 30, 169–172.

Niu, Y., Ren, G., Lin, G., Di Biase, L., and Fattorini, S. (2020). Fine-scale vegetation characteristics drive insect ensemble structures in a desert ecosystem: the tenebrionid beetles (Coleoptera: Tenebrionidae) inhabiting the Ulan Buh Desert (Inner Mongolia, China). *Insects* 11:410. doi: 10.3390/insects11070410

Pan, C., Feng, Q., Liu, J., Li, Y., Li, Y., and Yu, X. (2018). Community structure of grassland ground-dwelling arthropods along increasing soil salinities. *Environ. Sci. Pollut. Res.* 25, 7479–7486. doi: 10.1007/s11356-017-1011-1

Pennino, M. G., Paradinas, I., Illian, J. B., Muñoz, F., Bellido, J. M., López-Quílez, A., et al. (2019). Accounting for preferential sampling in species distribution models. *Ecol. Evol.* 9, 653–663. doi: 10.1002/ece3.4789

Picanço, A., Rigal, F., Matthews, T. J., Cardoso, P., and Borges, P. A. V. (2017). Impact of land-use change on flower-visiting insect communities on an oceanic island. *Insect Conserv. Divers.* 10, 211–223. doi: 10.1111/icad.12216

Pizzolotto, R., Gobbi, M., and Brandmayr, P. (2014). Changes in ground beetle assemblages above and below the treeline of the Dolomites after almost 30 years (1980/2009). *Ecol. Evol.* 4, 1284–1294. doi: 10.1002/ece3.927

Prado, P., Miranda, M., and Chalom, A. (2018). *sads: Maximum Likelihood Models for Species Abundance Distributions.* [R Package. version 3.5]. Available online at: https://cran.r-project.org/web/packages/sads/index.html

Preston, F. W. (1948). The commonness, and rarity, of species. *Ecology* 29, 254–283. doi: 10.2307/1930989

Pueyo, S., He, F., and Zillio, T. (2007). The maximum entropy formalism and the idiosyncratic theory of biodiversity. *Ecol. Lett.* 10, 1017–1028. doi: 10.1111/j.1461-0248.2007.01096.x

Raunkiaer, C. (1909). Formationsundersogelse og formationsstatistik. *Bot. Tidskr.* 30, 20–132.

Redford, K. H., Huntley, B. J., Roe, D., Hammond, T., Zimsky, M., Lovejoy, T. E., et al. (2015). Mainstreaming biodiversity: conservation for the twenty-first century. *Front. Ecol. Evol.* 3:137. doi: 10.3389/fevo.2015.00137

Roume, A., Ouin, A., Raison, L., and Deconchat, M. (2011). Abundance and species richness of overwintering ground beetles (Coleoptera: Carabidae) are higher in the edge than in the centre of a woodlot. *Eur. J. Entomol.* 108, 615–622.

Stier, A. C., Bolker, B. M., and Osenberg, C. W. (2016). Using rarefaction to isolate the effects of patch size and sampling effort on beta diversity. *Ecosphere* 7:e01612. doi: 10.1002/ecs2.1612

Sugihara, G. (1980). Minimal community structure: an explanation of species abundance patterns. *Am. Nat.* 116, 770–787.

Sun, H.-L. (2005). *Ecosystems of China*. Beijing: Science Press.

Tsafack, N. (2018). *Data_Carabidae in China Grasslands*. doi: 10.7910/DVN/FKBCRQ

Tsafack, N., Di Biase, L., Xie, Y., Wang, X., and Fattorini, S. (2019a). Carabid community stability is enhanced by carabid diversity but reduced by aridity in Chinese steppes. *Acta Oecol.* 99:103450. doi: 10.1016/j.actao.2019.103450

Tsafack, N., Rebaudo, F., Wang, H., Nagy, D. D., Xie, Y., Wang, X., et al. (2019b). Carabid community structure in northern China grassland ecosystems: effects of local habitat on species richness, species composition and functional diversity. *PeerJ* 6:e6197. doi: 10.7717/peerj.6197

Tsafack, N., Xie, Y., Wang, X., and Fattorini, S. (2020). Influence of climate and local habitat characteristics on carabid beetle abundance and diversity in Northern Chinese Steppes. *Insects* 11:19. doi: 10.3390/insects11010019

Ugland, K. I., Lambshead, P. J. D., McGill, B. J., Gray, J. S., O'Dea, N., Ladle, R. J., et al. (2007). Modelling dimensionality in species abundance distributions: description and evaluation of the Gambin model. *Evolutionary Ecol. Res.* 9, 313–324.

Ulrich, W., Ollik, M., and Ugland, K. I. (2010). A meta-analysis of species–abundance distributions. *Oikos* 119, 1149–1155. doi: 10.1111/j.1600-0706.2009.18236.x

Ulrich, W., Soliveres, S., Thomas, A. D., Dougill, A. J., and Maestre, F. T. (2016). Environmental correlates of species rank – abundance distributions in global drylands. *Perspect. Plant Ecol. Evol. Syst.* 20, 56–64. doi: 10.1016/j.ppees.2016.04.004

Volkov, I., Banavar, J. R., Hubbell, S. P., and Maritan, A. (2003). Neutral theory and relative species abundance in ecology. *Nature* 424, 1035–1037. doi: 10.1038/nature01883

Werger, M. J. A., and van Staalduinen, M. A. (2012). *Eurasian Steppes. Ecological Problems and Livelihoods in a Changing World*. Dordrecht: Springer.

Western, D. (1992). The biodiversity crisis: a challenge for biology. *Oikos* 63, 29–38. doi: 10.2307/3545513

Zhou, S., and Zhang, D. (2008). Neutral theory in community ecology. *Front. Biol. China* 3, 1–8. doi: 10.1007/s11515-008-0008-z

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Ecology and Evolution

ORIGINAL RESEARCH
published: 26 May 2021
doi: 10.3389/fevo.2021.633263

Frequency-Dependent Competition Between Strains Imparts Persistence to Perturbations in a Model of *Plasmodium falciparum* Malaria Transmission

Qixin He[1†‡], *Shai Pilosof*[2†‡], *Kathryn E. Tiedje*[3‡], *Karen P. Day*[3‡] *and Mercedes Pascual*[4,5*]

[1] Department of Biological Sciences, Purdue University, West Lafayette, IN, United States, [2] Department of Life Sciences, Ben-Gurion University of the Negev, Beer-Sheva, Israel, [3] Department of Microbiology and Immunology, Bio21 Institute, The University of Melbourne, Melbourne, VIC, Australia, [4] Department of Ecology and Evolution, University of Chicago, Chicago, IL, United States, [5] Santa Fe Institute, Santa Fe, NM, United States

In high-transmission endemic regions, local populations of *Plasmodium falciparum* exhibit vast diversity of the *var* genes encoding its major surface antigen, with each parasite comprising multiple copies from this diverse gene pool. This strategy to evade the immune system through large combinatorial antigenic diversity is common to other hyperdiverse pathogens. It underlies a series of fundamental epidemiological characteristics, including large reservoirs of transmission from high prevalence of asymptomatics and long-lasting infections. Previous theory has shown that negative frequency-dependent selection (NFDS) mediated by the acquisition of specific immunity by hosts structures the diversity of *var* gene repertoires, or strains, in a pattern of limiting similarity that is both non-random and non-neutral. A combination of stochastic agent-based models and network analyses has enabled the development and testing of theory in these complex adaptive systems, where assembly of local parasite diversity occurs under frequency-dependent selection and large pools of variation. We show here the application of these approaches to theory comparing the response of the malaria transmission system to intervention when strain diversity is assembled under (competition-based) selection vs. a form of neutrality, where immunity depends only on the number but not the genetic identity of previous infections. The transmission system is considerably more persistent under NFDS, exhibiting a lower extinction probability despite comparable prevalence during intervention. We explain this pattern on the basis of the structure of strain diversity, in particular the more pronounced fraction of highly dissimilar parasites. For simulations that survive intervention, prevalence under specific immunity is lower than under neutrality, because the recovery of diversity is considerably slower than that of prevalence and decreased *var* gene diversity reduces parasite transmission. A Principal Component Analysis of network features describing parasite similarity reveals that despite lower overall diversity, NFDS is quickly restored after intervention constraining strain structure and maintaining patterns of limiting similarity important to parasite persistence. Given the described enhanced persistence under perturbation, intervention efforts will likely require longer times than the usual practice to

OPEN ACCESS

Edited by:
Luís Borda-de-Água, Universidade do Porto, Portugal

Reviewed by:
Jacob A. Tennessen, Harvard University, United States
Ellen Leffler, The University of Utah, United States
Sonja Lehtinen, ETH Zürich, Switzerland

***Correspondence:**
Mercedes Pascual
pascualmm@uchicago.edu

[†] *These authors have contributed equally to this work*

[‡]ORCID:
Qixin He
orcid.org/0000-0003-1696-8203
Shai Pilosof
orcid.org/0000-0003-0430-5568
Kathryn E. Tiedje
orcid.org/0000-0003-3305-0533
Karen P. Day
orcid.org/0000-0002-6115-6135
Mercedes Pascual
orcid.org/0000-0003-3575-7233

Specialty section:
This article was submitted to Models in Ecology and Evolution, a section of the journal Frontiers in Ecology and Evolution

Received: *25 November 2020*
Accepted: *29 April 2021*
Published: *26 May 2021*

Citation:
He Q, Pilosof S, Tiedje KE, Day KP and Pascual M (2021) Frequency-Dependent Competition Between Strains Imparts Persistence to Perturbations in a Model of Plasmodium falciparum Malaria Transmission. Front. Ecol. Evol. 9:633263. doi: 10.3389/fevo.2021.633263

eliminate *P. falciparum* populations. We discuss implications of our findings and potential analogies for ecological communities with non-neutral assembly processes involving frequency-dependence.

Keywords: strain diversity, stabilizing competition, stochastic assembly, persistence, malaria and antigenic diversity, negative frequency-dependent selection, agent-based model, var genes

1. INTRODUCTION

Under negative frequency dependent selection (NFDS), the relative fitness of a species or a genotype decreases as its abundance and therefore its frequency increases. In the realm of population genetics, NFDS is a form of balancing selection promoting the coexistence of genotypes or of genes, with MHC being a prominent example (Aguilar et al., 2004; Key et al., 2014). In community ecology, NFDS causes species to limit themselves more than others, which stabilizes competition and maintains biodiversity by promoting species coexistence (HilleRisLambers et al., 2012). In some realms, such as microbial ecology, the boundary between these levels of organization is blurred, yet here too, NFDS has been shown to maintain microbial diversity (Cordero and Polz, 2014; Healey et al., 2016). Although the evolutionary time scales at which NFDS operates can be different for different organisms and levels of organization, the mechanisms underlying it can be common. In community ecology, competition for resources and selection for traits that give an edge in resource acquisition can shape the biodiversity of plants and animals (Browne and Karubian, 2016). For example, interaction with specific herbivores leads to negative density-dependent mortality and coexistence of congenerics *Inga spp* in forest ecosystems (Forrister et al., 2019). In infectious diseases, competition for hosts and selection for traits that allow pathogens to evade human immune systems shape the antigenic diversity of pathogens (Gupta and Day, 1994; Grenfell et al., 2004).

Beyond diversity *per se*, NFDS underlies the structure of such diversity in both pathogen populations (Koelle et al., 2006; Volz et al., 2013) and species communities (Scheffer and van Nes, 2006; D'Andrea et al., 2019). There is an unappreciated analogy between the structure of niches in trait space emerging from NFDS in these two lines of research (Pascual, 2020). One infectious disease system in which such strain structure has been investigated is that of the malaria parasite *Plasmodium falciparum*. Strain coexistence in local populations of the malaria parasite results from an on-going assembly process involving the dynamic interplay between ecology (population dynamics of strains, competition for hosts) and evolution (genetic changes via mutation and recombination) (He et al., 2018). The trait of interest concerns here the variation in the major antigen of the blood stage of infection PfEMP1. This protein is exported by the parasite to the surface of the infected blood cells where it becomes the target of the adaptive immune system (Bull et al., 1998). Variations of the protein, or variable surface antigens (VSAs), are encoded by a multigene family known as *var* with about fifty to sixty gene copies per parasite, whose expression is sequential and influences the duration of infection. From this perspective, an individual parasite corresponds to a unique combination of *var* genes—a "repertoire" —encoding for a particular set of VSAs. In high-transmission endemic regions, such as those of sub-Saharan Africa, local diversity of *var* gene types is in the order of tens of thousands. This extremely large genetic pool underlies the vast combinatorial diversity of the repertoires themselves. Both these high levels of diversity result from high recombination rates, which act at the two different levels of organization, mixing repertoires and generating new *var* genes, respectively (Larremore et al., 2013; Zilversmit et al., 2013).

Previous theory and data have shown that frequency-dependent competition for hosts is at play in determining the coexistence of a large number of repertoires, whose population structure is both non-random and non-neutral despite the vast gene pool (Barry et al., 2011; Day et al., 2017; He et al., 2018; Pilosof et al., 2019). Hosts can be viewed as resource patches whose availability depends on their individual history of "consumption" from previous exposure. This is because exposure to a specific VSA leads to the acquisition of specific immunity by individual hosts, which then precludes expression of the corresponding gene and thus shortens infection, reducing the fitness of the parasites that carry this gene. Adaptive immunity therefore implements selection, providing an advantage to strains carrying rare novel *var* genes, and a disadvantage to those composed of common ones. In the language of community ecology, the system experiences stabilizing competition (sensu Chesson, 2000) from trait variation that underlies niche differences. In that of population genetics, such competition is a mechanism of NFDS, which acts as a form of balancing selection and promotes strain (i.e., *var* repertoire) coexistence (He et al., 2018; Pilosof et al., 2019).

We have previously used the malaria system to address the challenge of discerning rules of assembly from complex population patterns, focusing on the large strain variation of *P. falciparum* populations in endemic regions of high transmission. We applied for this purpose a combination of stochastic agent-based models and network analyses of the genetic similarity between repertoires they generate (Artzy-Randrup et al., 2012; He et al., 2018; Pilosof et al., 2019). These studies have shown that networks assembled under NFDS differ in their topology from those assembled in the absence of selection, under neutral processes. Hence, in the malaria system the importance of individual and specific interactions can be detected in features of the macroscopic similarity structure of (strain) diversity. We note that networks, rather than phylogenetic trees, were applied because of the predominant role of recombination in the evolutionary change of the system.

In community ecology, the structure of diversity has been a major theoretical question, because of its hypothesized influence on system "stability" to perturbations (Hutchinson,

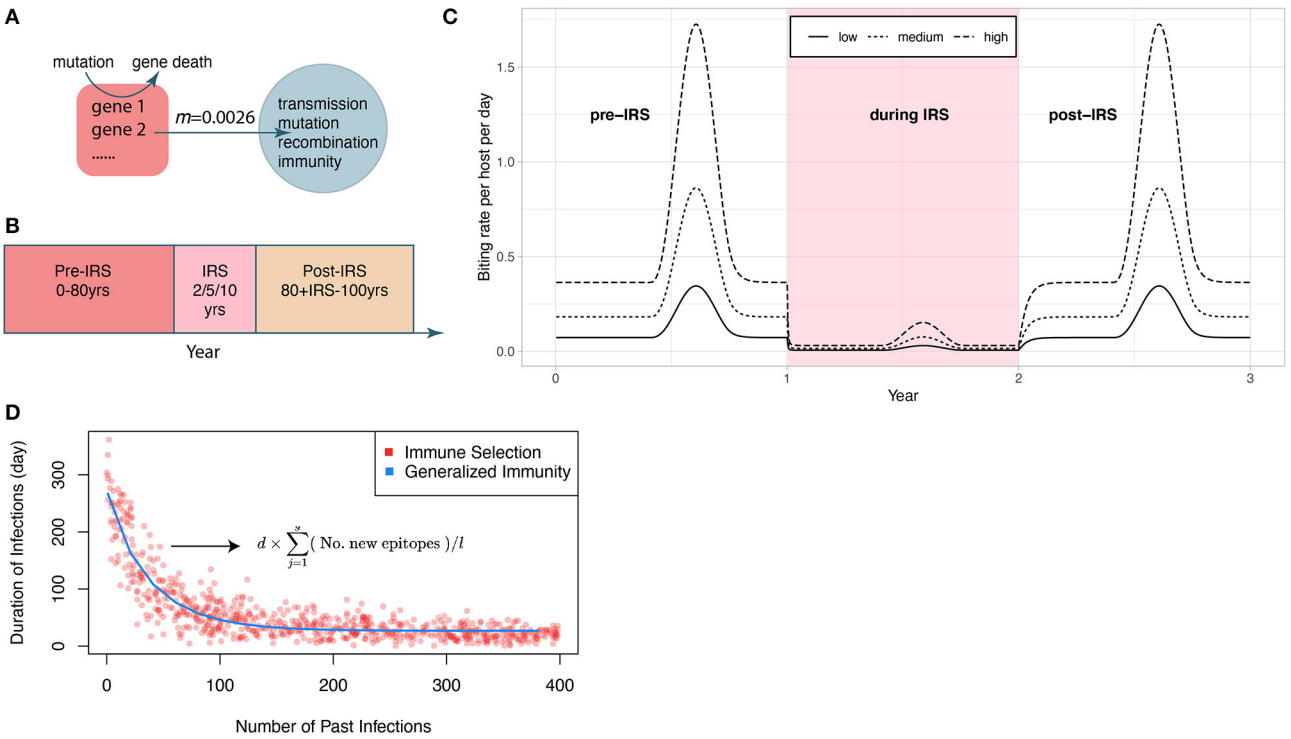


FIGURE 1 | Schematic of the IRS experiment design. **(A)** Initial transmission events and migrant genomes in a local population (blue circle) are sourced from a global pool of genes (red square). Mutation of new genes and the death of existing genes in the global pool occur at the same rate as in the local population (see Methods). The value of the migration rate is inferred from an empirical dataset for a high transmission region in Ghana (see Methods). Individual infections are tracked locally. We also track events of transmission, mutation of genes, recombination within and between genomes, and acquisition and loss of specific immune memory in hosts. **(B)** Each simulation follows three stages (after a burn-in period): a pre-IRS period during which the transmission in the local population reaches a stable state; an IRS period of 2, 5, or 10 years reducing transmission, and a post-IRS period when transmission rates go back to pre-IRS levels. **(C)** Three levels of transmission intensity (biting rates) are explored in the experiments (pre-/post-IRS, low: 44 bites per host per year; medium: 110 bites/h/y; high: 221 bites/h/y). **(D)** The two regimes (with and without NFDS) are ensured to be comparable by specifying the average duration of infection as a function of the number of previous infections in G with a curve fitted to the points generated under S. The expression for the duration of infection under S is given here, and its explanation can be found in the methods (2.2).

1959; Allesina and Tang, 2015). In infectious disease dynamics, a parallel yet little studied line of research regards the response of structured pathogen populations to interventions. Here, we illustrate the combination of stochastic ABMs and network analysis for this purpose, in *P. falciparum* populations under interventions that decrease transmission intensity. We ask whether the malaria system assembled under NFDS is more or less "stable" than a neutral counterpart. We apply press perturbations (Ives and Carpenter, 2007) in a stochastic computational model to represent interventions that target the mosquito vector with indoor residual spraying (IRS), which involves the application of an insecticide to internal walls and ceilings of homes (WHO, 2015). IRS effectively reduces transmission rate and therefore, the growth rate of the parasite population as a whole. We examine two components of stability: (i) persistence—the ability of the parasite's population to withstand an intervention and (ii) recovery: the transient rebound of repertoire abundance or parasite prevalence, and of genetic diversity post-intervention. We consider in particular how quickly repertoire structure is restored. We end by drawing plausible analogies and implications for other ecological systems (Azarian et al., 2019; Forrister et al., 2019).

2. MATERIALS AND METHODS

2.1. Agent-Based Model (ABM) and Model Setup

Malaria transmission and intervention are modeled using an agent-based, discrete-event, continuous-time stochastic system described in detail in He et al. (2018) and Pilosof et al. (2019). Here, we briefly describe the agent-based model (ABM), with an emphasis on the specific implementation of the regional intervention scheme that constitutes the press perturbation.

We model a local parasite population of size N, as well as a global *var* gene pool of size G_p that acts as a proxy for regional parasite diversity (**Supplementary Table 1**). Parasite genomes can migrate from the regional pool to the local population. Each simulation starts with 20 migrant infections to initiate local transmission (**Figure 1A**). Because each parasite genome is a repertoire of 60 *var* genes, migrant genomes are assembled from random sampling of 60 *var* genes from the global pool. Each *var* gene is itself a linear combination of two epitopes—the part of a molecule that acts as an antigen and is recognized by the immune system (Rask et al., 2010; He et al., 2018) (l in **Supplementary Table 1**).

We consider seasonal endemic transmission dynamics in which mosquitoes are not represented as agents in the model, but via biting events. Local transmission events are sampled at the total rate of host population size N_h times the biting rate b, in which a donor and a recipient host are selected randomly. We implemented seasonality as a fluctuation of mosquito bites, which results from density dependence at the egg and larva stages as a function of rainfall typical of Northern Ghana (availability of breeding sites, White et al., 2011). The specific algorithm to obtain the seasonality of the biting rate was described in detail in Pilosof et al. (2019).

The main modification to the model for this work is how the global pool interacts with local transmission. First, instead of remaining static, the global gene pool in this implementation updates its gene composition at the same mutation rate as that of the local population. Specifically, new genes are generated at a rate equal to the product of local parasite population size and the per-allele mutation rate. Once a new gene is generated, the old gene that it mutates from is removed from the global gene pool. Genomes migrate from the global pool to the local population (**Figure 1A**). The number of migrant genomes increases in wet seasons and decreases in dry seasons (see details in section Estimation of Migration Rate). Interventions are assumed to be applied at the regional level so that prevalence of the disease is the same across the region, including the local population (**Figures 1B,C**). Therefore, the proportion of infectious bites from migration is kept the same as that of local transmission.

2.2. Parasite Fitness and Duration of Infection

Our main objective was to compare responses of the parasite population to perturbation in the presence and absence of frequency-dependent competition. Specifically, we aimed to compare an "immune selection scenario," in which immune memory to particular epitopes elicits specific competition between parasite genomes for hosts, to a "neutral scenario" of "generalized immunity" and therefore non-specific competition. In both these scenarios, duration of infection is the relevant trait manifesting the effect of competition for hosts. Parasite fitness is affected by immunity in our model through the duration of infection because shorter duration reduces the probability that a repertoire will be transmitted under a constant transmission rate. In the model with specific immunity, duration of infection depends on the immune history of given *var* epitopes in the host. Hosts gain and lose immunity toward specific epitopes. During infection, expression of the genes composing the given repertoire is sequential and infection ends when the whole repertoire is depleted. The host is considered infectious with the active strain and the expression length d of each gene is controlled by host immunity. When a gene is actively expressed, host immunity "checks" infection history to determine whether any of its epitopes have been seen before. Thus, expression length d shortens as the number of unseen epitopes out of l increases. After the gene is deactivated, the host adds the deactivated gene alleles to its immunity memory. A new gene from the repertoire then becomes immediately active until the end of the repertoire is reached. Therefore, the total duration of infection in a given host with a particular repertoire of g genes is given by

$$\text{Total duration } = d \times \sum_{j=1}^{g} (\text{ No. new epitopes })/l$$

This is the process that confers an advantage to rare new genes and the parasites that carry them, and a disadvantage to common ones—that is, the frequency-dependent selection. The more similar the epitope composition of two repertoires, the stronger their competition for hosts.

In contrast, the model of generalized immunity retains protection conferred by previous infection but does not consider specific memory toward *var* genes. This is a common implicit assumption of most malaria transmission models, including other ABMs and extensions of the well-known compartmental population models of the SIR type (for Susceptible-Infected-Recovered classes), where infection "consumes" susceptible hosts but as a general and shared resource. In our implementation of generalized immunity, the duration of infection depends on the number of previous infections but not on their specific genetic composition. For a meaningful comparison to the immune selection scenario, we parameterized the function of infection duration with number of previous infections, to match the curve generated by the corresponding immune selection scenario (see He et al., 2018) (**Figure 1D**).

We refer hereafter to the scenario of immunity-driven frequency-dependent selection as "S" and to that of generalized immunity as "G." Comparisons between these two scenarios are always made for the same parameter combinations.

2.3. Course of a Simulation and Indoor Residual Spraying (IRS) Intervention

Each simulation follows three stages (**Figure 1B**): (i) a pre-IRS period (0–80 years) during which the local parasite community is assembled and the transmission system reaches a stationary state before the perturbation; (ii) an IRS period of 2, 5, or 10 years during which transmission is decreased; and (iii) a post-IRS period when transmission rates return to pre-IRS levels and the system is allowed to recover (**Figure 1C**). During the IRS interventions, the effectiveness of insecticides in killing adult female mosquitoes is set to be 100% (Wanjala et al., 2015) and the percentage of sprayed household is set to be 90% (Kigozi et al., 2012; West et al., 2014; WHO, 2015; Coleman et al., 2017). Details on the model can be found in White et al. (2011), and our implementation (in Mathematica) at the GitHub repository https://github.com/pascualgroup/Pf_temporal_networks.

2.4. Within-Host Dynamics

The infection and immune history of each host are tracked individually. Upon each biting event, if the donor harbors parasites in the asexual blood stage, then each repertoire has a given probability, c, to be transmitted to the mosquito. Maximum transmissibility is set to $c = 0.5$, because not all bites from

mosquitoes result in successful transmission from the human host. If the host harbors n infections concurrently, each parasite strain experiences a reduced transmissibility equal to c/n. Hosts can harbor up to 10 concurrent infections. Because parasites must go through the sexual reproduction stage within mosquitoes, during meiotic recombination, *var* repertoires picked up by the mosquito at a bite event may recombine with another genome to produce oocysts that mature into sporozoites. Since each oocyst is generated from recombination between two parental genomes, if a mosquito picks up n genomes, each genome has a probability $1/n$ of recombining with itself, producing the same offspring genome, and a probability $1 - 1/n$ of recombining with a different genome, producing recombinants. To generate a recombinant, two sets of genes are pooled from the original genomes, and a random set is sampled. Since *var* genes are distributed across most of the chromosomes, there is a substantial shuffling of genes during meiotic recombination. Although it is possible that some genes are linked and do not recombine, we do not know the exact physical locations of these genes. It is therefore a simplification in our model to take a random sample from two pooled genomes. This sexual recombination process creates variation at the genome level. The total number of *var* repertoires passed onto the receiving host is kept the same as that obtained from biting the infectious donor host.

Once in the blood, each infection progresses with the sequential expression of the *var* genes in random order. During an infection, novelty at the gene level can be generated via mutation or recombination between *var* genes within the same genome. Part of the life cycle of *Plasmodium falciparum* occurs in the liver; we therefore implement a 14 day delay to mimic the transition from the liver stage to the blood stage, where the parasite becomes infectious and expression of the *var* genes is initiated.

Hosts gain and lose immunity toward specific epitopes (l_i), which we implement with a counter to track the boosting of immunity (the number of times a host has been exposed to a given variant) and the loss of immunity. After expression of a given *var* gene, the host gains full protection toward the epitopes in the gene that was previously expressed. However, the immunity loss rate for a specific epitope depends on the number of times the host has been exposed to it. For example, if the host was infected by a parasite that contained two *var* genes, each of which contained an allele encoding epitope x, the counter for this epitope will increase by two. The counter decreases by one at the immunity loss rate of 1/1,000 per day (Collins et al., 1968). When the counter value becomes 0, the host loses protection against the given epitope.

2.5. Estimation of Migration Rate

We use empirical data of *var* genes, hyper-diverse markers that provide a higher resolution for recent migration events, to infer the rate of gene exchange between populations and to obtain a reasonable estimate for migration rates that allows us to implement an open transmission system. We used Jost's D measure of population divergence (Equation 12 in Jost, 2008) to consider highly diverse genetic markers, and to compare *var* gene composition within and between two field sites for which molecular sequences were previously obtained: Soe and Vea/Gowrie in the Bongo District of Ghana at the end of the wet (high transmission) season (Tiedje et al., 2017; He et al., 2018). Using Equation (22) in Jost (2008), we estimated m, the migrant proportion (percentage of migrant bites relative to local transmission events), by dividing D by the mutation/recombination rates of the *var* genes per generation (5.3e-6, Claessens et al., 2014). Thus, the migration rate per day is the product of $m * b$.

2.6. Event Scheduling in the Stochastic Model

The simulation is an ABM implemented with a modified Gillespie algorithm called next-reaction method (Gillespie, 1976), which is a computationally-efficient way to model stochastic dynamics. The algorithm takes a set of events (e.g., infection, biting) and assigns for each event the time of its next occurrence by sampling an exponential distribution with a mean of 1 over its rate, to implement a Poisson process. The next event is then chosen to be the one with the closest event time.

In our simulation, global events include local transmission from biting events, new transmission from migrant *var* repertoires, and birth and death of hosts. In the numerical implementation of the simulation, all possible future events are stored in a single event queue along with their putative times. When an event occurs, it may trigger the addition or removal of future events on the queue, leading to a re-ordering of events, or changes of their rates, and to a recalculation of their putative time. The next-reaction method optimizes the Gillespie first-reaction method (Gillespie, 1976) and allows for faster simulation times, as it targets changes of rate events to a given subset (specified by the structure of the queue).

2.7. Experimental Design

We explored how initial gene pool size, transmission rate and the duration of an IRS intervention influence the system's persistence to, and recovery from, intervention under G and S. Specifically, we set: (i) initial gene pool sizes (G_p) of 1,200, 1,800, 2,400, 3,600, 4,800, 7,200, 9,600, 14,400, and 19,200; (ii) three levels of transmission intensity (biting rates; b): 44, 100, and 221 bites per host per year, corresponding to "low," "medium," and "high" transmission (**Figure 1C**); and (iii) IRS lasting for 2, 5, or 10 years. This design has 81 sets of parameter combinations and we ran 50 replications for each combination. We calculated the probability of extinction, effectively measuring persistence, as the proportion of the 50 replications in which the parasite populations crashed before lifting the IRS.

2.8. Sampling and Gene Similarity Networks

During each simulation, summary statistics including prevalence, multiplicity of infection (MOI; number of genomes within a host), and genetic and allelic diversity, were calculated every 30 days. In addition, 100 infected hosts were randomly sampled to analyze parasite diversity patterns. To evaluate similarity of parasites in the population, pairwise type sharing (PTS) was calculated between all repertoire pairs (regardless of the host in

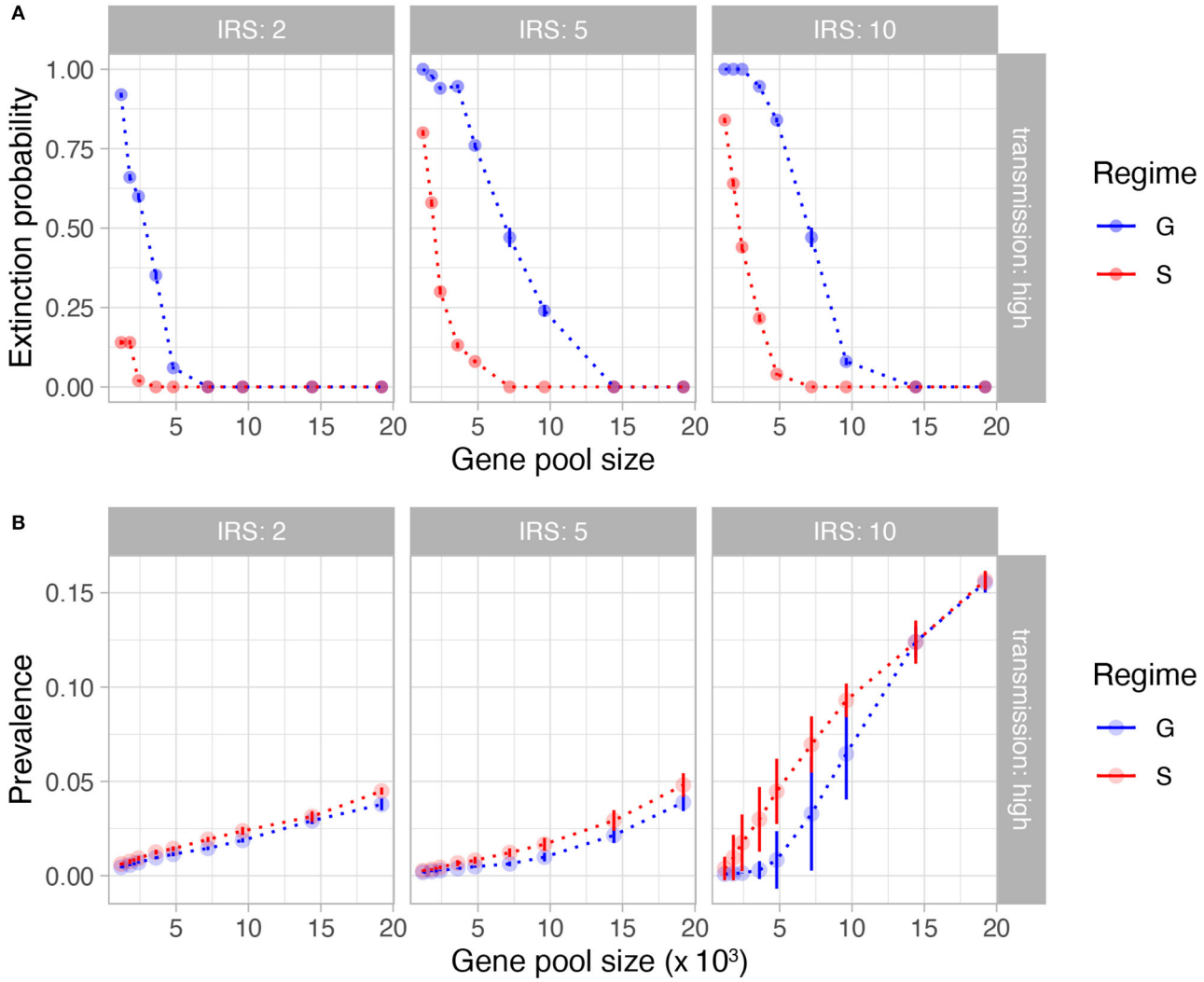


FIGURE 2 | Extinction probability **(A)** and mean prevalence during IRS **(B)** as a function of initial gene pool size and lengths of IRS (columns). Shown here are the results from high transmission intensities. See **Supplementary Figure 1** for results from low and medium transmission intensities. In **(A)**, each point is the proportion of simulation runs for the given parameter combination (out of 50 runs) that crashed before the IRS was lifted. In **(B)**, each point represents the mean value (bars represent standard deviation) of prevalence for those simulation runs that survived the IRS for the given parameter combination. Blue and red colors depict generalized immunity (G, the neutral model) and specific immunity (S, NFDS), respectively. It is clear that despite comparable prevalence, the parasite population is more persistent in the immune selection scenario.

which they are encountered) as $PTS_{ij} = 2n_{ij}/(n_i + n_j)$, where n_i and n_j are the number of unique alleles (corresponding to epitopes) within each repertoire i and j and n_{ij} is the total number of alleles shared between repertoires i and j (Barry et al., 2007). In addition, similarity networks based on the *var* composition were built to investigate changes of parasite genetic population structure through interventions. We calculated genetic similarity of repertoire i to repertoire j as $S_{ij} = (N_i \cap N_j)/N_i$, where N_i and N_j are the number of alleles unique for repertoires i and j, respectively (the genetic similarity of repertoire j to repertoire i was calculated as $S_{ji} = (N_i \cap N_j)/N_j$). Unlike PTS, network edges encode a directional measure that represents the asymmetric competition between repertoires, resulting from different numbers of unique alleles (He et al., 2018).

2.9. Calculation of Network Properties and Principal Component Analysis

We calculated 36 network properties to compare the changes in network structure between the immune selection and neutral scenarios. These properties include metrics of transitivity, degree distributions, component sizes, diameters, reciprocity, and proportion of 3-node graph motifs (see **Supplementary Table 1** in He et al., 2018 for a complete list of properties and definitions). In the network analyses we retained only edges with values within the 95% percentile to focus on the strongest interactions of current competition between strains (He et al., 2018; Pilosof et al., 2019). To minimize the influence of sample size differences across time due to changes in mean MOI, network properties were calculated by resampling 100 repertoires randomly from

the original network. Principal component analysis (PCA) was performed on normalized and centered network properties across time and selection regimes per parameter combination, to inspect overall trends in network structure.

3. RESULTS

3.1. Populations Under Immune Selection Are More Persistent

The interventions implemented in the model reduce the transmission rates by about 90% during IRS. Not surprisingly and especially for the long-lasting interventions, the stochastic simulations are prone to extinction. We found a lower extinction probability under specific immunity (S) than under generalized immunity (G) for the same parameters (**Figure 2A**, **Supplementary Figure 1A**). As expected, longer interventions lead to higher extinction rates, especially under low transmission rates. A higher initial pool size of genetic diversity ensures a lower extinction probability. Pool sizes larger than 10,000 *var* genes, which represent high transmission endemic regimes in sub-Saharan Africa, experience significantly less extinction, even after 10 years of sustained intervention. Interestingly, the difference in extinction probability between S and G does not reflect a trivial effect of overall population size of the parasite: the mean prevalence is comparable during IRS for S and G, only slightly higher for the former, indicating that other factors are responsible for the higher persistence under NFDS (**Figure 2B**, **Supplementary Figure 1B**).

3.2. Persistence Is Enhanced by Repertoire Dissimilarity Under NFDS

To investigate the origin of the difference in persistence as measured by the probability of extinction, we consider the epidemiological parameter that confers the fitness difference to the parasite, namely the duration of infection in the human host. As is typical in epidemiology, the reproductive number of the parasite is proportional to the product of the transmission rate and the duration of infection. In our model, immunity modifies only duration of infection as genes encoding for antigenic variants which have been seen before are not expressed. We note that we set the mean duration of infection with the number of previous exposures under G to exactly match that under S before IRS for a specific parameter combination (**Figure 1D**). Therefore, the mean values of duration of infection under G and S are similar before the IRS by construction, with the values increasing significantly during IRS (**Supplementary Figure 2**). When the longest infections are considered by comparing the top 5% of the distribution of infection duration, we observe higher values under S during IRS obtained in less than a year (**Figure 3A**, **Supplementary Figure 3**). Longer lifetimes of infection at the tail of the distribution provide a buffer against extinction by conferring a higher probability of survival to parasites during the low transmission period imposed by the intervention.

To further understand the extinction patterns, we therefore need to explain the longer durations of infection at the tail of their distribution under NFDS. We turn to the structure of similarity between parasites as reflected by the distribution of pairwise type sharing (PTS) between *var* repertoires (**Figures 3B**, **4**, **Supplementary Figure 4**). **Figure 4** shows the PTS distributions before, during, and after the IRS intervention for the two scenarios. The shapes between the two scenarios differ substantially at low PTS, reflecting clear patterns of high dissimilarity in S. In particular, NFDS generates a monotonically decreasing distribution with the highest frequencies found at the lowest overlap. In contrast, under G the distribution exhibits a mode, with the highest frequency for some intermediate overlap. Due to decreased diversity during IRS, the proportion of repertoire pairs sharing more epitopes increases for both scenarios, and the whole distribution moves toward higher overlap (i.e., higher PTS) (**Figure 4**). However, the change is less pronounced under S, especially for the maintenance of dissimilar repertoires (**Figure 3B**). When the intervention is lifted, the distribution of G shifts significantly toward increased similarity, while the bimodal shape of the PTS distribution in S is maintained, including a considerable fraction of highly dissimilar strains. Since diversity under G does not influence transmission by construction, the change in PTS reflects solely parasite population fluctuations and bottleneck effects during and after IRS. In contrast, the maintenance of low PTS values that are comparable to pre-IRS levels under S is indicative of variant-specific immune selection at work: even under reduced diversity, repertoires are maintained as different from each other as possible, resulting in some longer infections than for a randomly assigned gene composition. A larger fraction of the parasite population with highly dissimilar repertoires generates higher heterogeneity in duration of infection, including longer duration, increasing the persistence during intervention under S.

3.3. Prevalence Recovers to a Lower Level Than Pre-intervention Under NFDS Because of Diversity Loss

We focus next on the recovery from intervention in terms of both antigenic diversity and prevalence. As expected, both antigenic diversity and prevalence decrease with intervention, and longer IRS leads to a higher reduction in antigenic diversity than the shorter 2-year IRS. Post-intervention, diversity settles on a new equilibrium, that is lower than the pre-intervention one (**Figure 5**). Although new genes should be strongly preferred under S (but not G), their generation from mutation, ectopic recombination, and immigration, is likely too slow to rebuild locally to pre-intervention levels given the parameters we use. Moreover, the intervention was implemented to also affect the *var* gene diversity of the regional pool, hindering regeneration of diversity via migration. We touch upon this point further in the Discussion.

All simulations that persist the intervention show a comparable initial rebound of prevalence in the months that immediately follow the control period. The rebound for G is to the same prevalence than that before intervention, as the transmission rate has been restored and the duration of

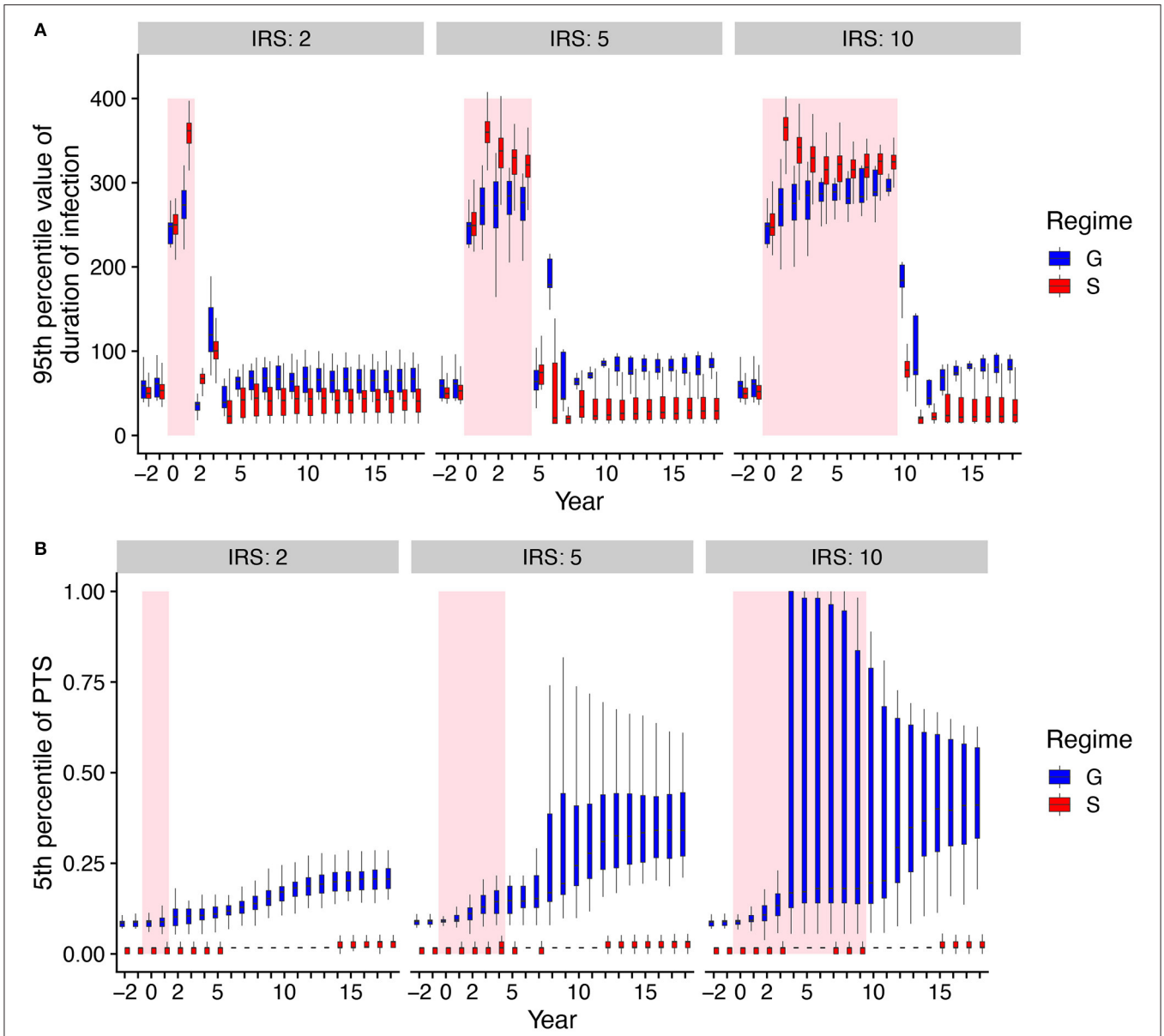


FIGURE 3 | Comparisons of duration of infection and mean PTS over time under high transmission between the two regimes with and without NFDS. Boxplots summarize the distribution among replication runs of the 95th percentile of duration of infection **(A)**, and the distribution of the 5th percentile of PTS between *var* repertoires **(B)**. We therefore focus on the longest periods of infection duration and correspondingly, the most dissimilar parasites. Shaded areas indicate the intervention (IRS) period. Negative years correspond to times prior to IRS.

infection only depends on the exposure rate. In contrast, under S, after an initial overshoot—which is a result of a large human population that is not immune to the parasite due to limited infections during IRS—mean prevalence attains lower values (**Supplementary Figure 5**, **Figure 5B**), because the decreased *var* gene diversity induces a shorter duration of infection. The rebound of diversity under S is therefore a result of immune selection and not a larger parasite population size post-IRS. The recovery of prevalence post-intervention can only be achieved to a lower level than pre-intervention, given that the *var* gene diversity itself cannot be completely rebuilt.

3.4. The Structure of Genetic Diversity Is Restored After Intervention Under NFDS

Until now we have mainly addressed the connection between antigenic diversity and perturbation in relation to similarity (PTS), as one feature of the structure of diversity. We consider next the structure of similarity between *var* repertoires more broadly by applying an analysis of network features (sensu He et al., 2018). In these networks, the nodes are repertoires and the directional links quantify the degree of overlap in *var* gene identity between them. The degree of overlap between two repertoires is a measure of the strength of competition between

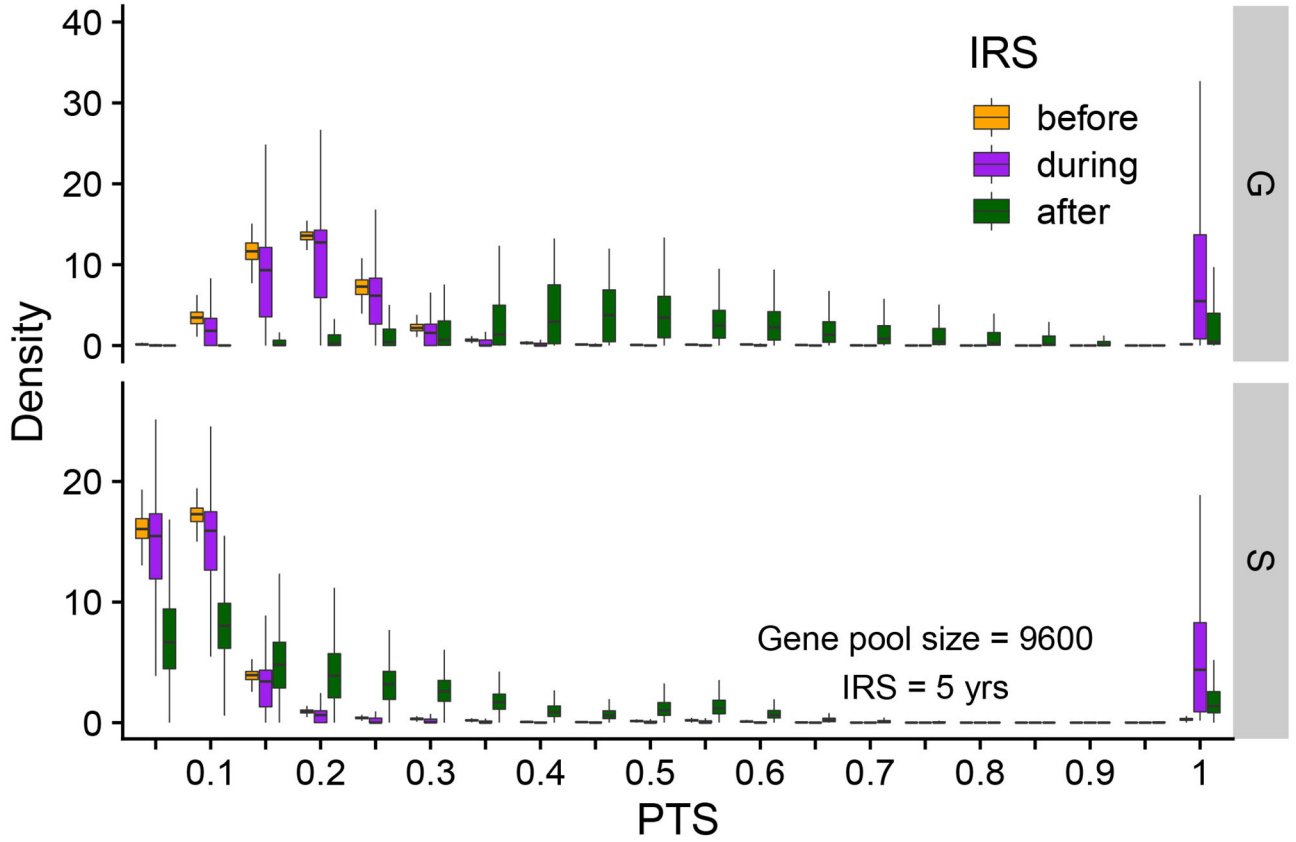


FIGURE 4 | Distribution of pairwise PTS for *var* repertoires before (orange), during (purple), and after (green) intervention under S and G. PTS values are binned into 40 equally-sized bins (of 0.05). Each box shows the variation of PTS values across different runs. Summarized here are runs from a gene pool size of 9,600, high transmission and a 5-year intervention. The PTS distributions differ between G and S. After intervention, the PTS distribution in G drifts away from its initial state, whereas that in S tends to go back to its pre-IRS state.

them for human hosts (i.e., intensity of competition positively correlates with similarity). Therefore, we only retain the links above a given cutoff because it is for these most similar repertoires that we expect the selection against recombinants to be the strongest, and the evidence of frequency-dependent selection to be most apparent (see He et al., 2018). We know from our previous work that we can use ensembles of network features to differentiate between populations of parasites assembled under S and G. Here, we compare the overall similarity structure between these regimes before, during and after intervention, to ask whether and how quickly the effect of NFDS is restored after intervention, as we expect it to be relaxed under the decreased transmission of intervention.

The similarity structure for an ensemble of simulations can be visualized in the 2-D space defined by the two major axes of variation of PCA (Methods). **Figure 6** and **Supplementary Figure 6** show that the networks assembled under S and G do indeed differ and occupy different regions in that space before intervention. Their structure changes and becomes similar for the intervention period as selection/competition is relaxed. Nevertheless, under S the structure is quickly restored after intervention, whereas under G this is not the case, with networks drifting further away from each other and from their initial structure.

The network features contributing the most to the classification of the similarity structure are shown in **Supplementary Figure 7**. These include in particular the 3-node motifs and reciprocity that characterize local structures and divergence between communities within networks that characterize global structures. The properties distinguish the more tree-like local structure of neutrality from that of limiting similarity and more symmetric or balanced diversity of selection (He et al., 2018). The action of frequency-dependent selection on similarity structure is what eventually maintains and restores the distributions of overlap described earlier (**Figure 4**), which explains the longer duration of infection and therefore the higher probability of persistence during intervention.

4. DISCUSSION

Understanding the relationship between structure and persistence in diverse communities is a long-lasting and on-going effort. This relationship likely depends on the processes that generate diversity and assemble its structure. High-transmission

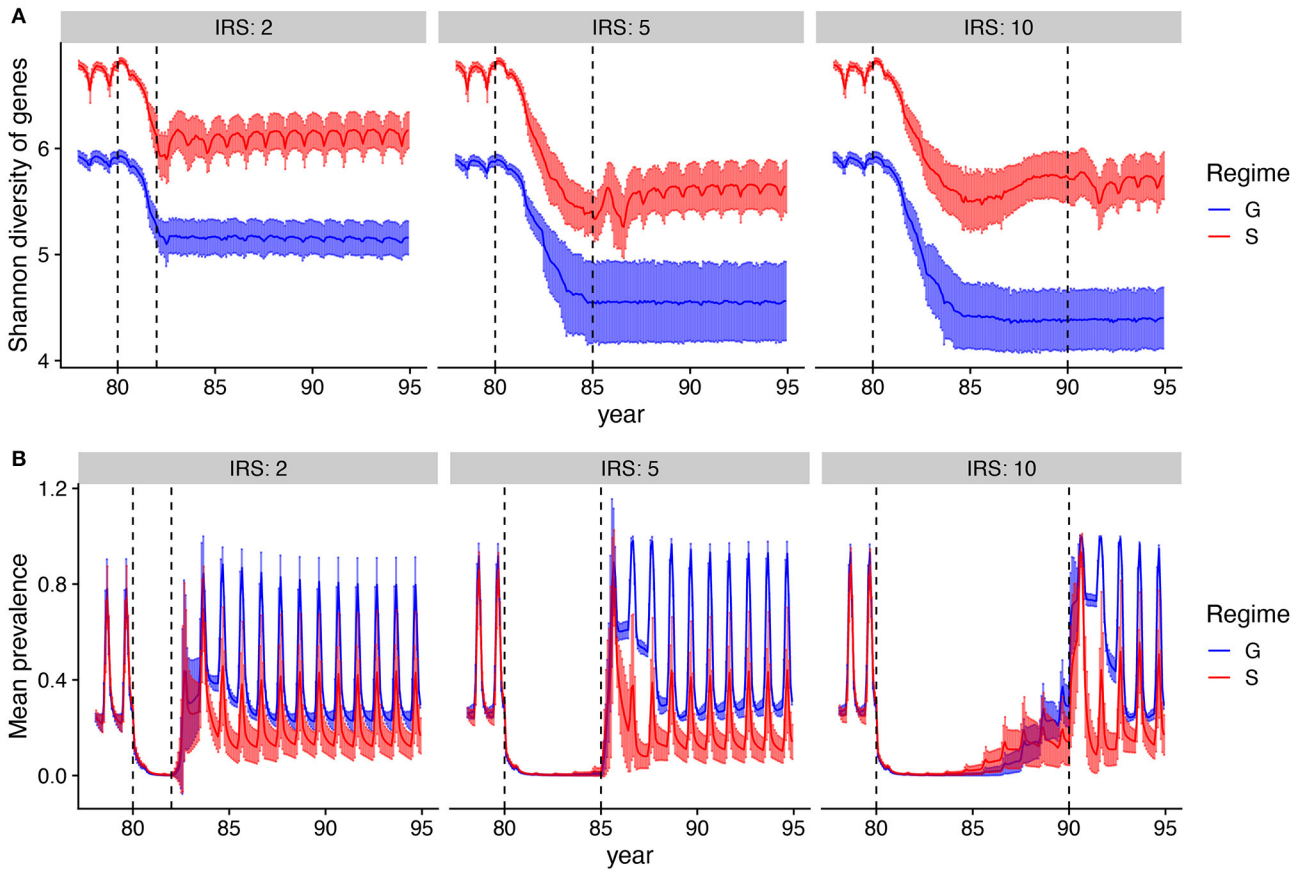


FIGURE 5 | Dynamics of genetic diversity and prevalence during and after IRS. **(A)** Mean Shannon diversity (solid line) and standard deviation (error bars) of *var* gene types are tracked over time. Shaded areas show the extent of the variation across runs. **(B)** Mean changes in prevalence (solid line) and its variation (error bars) over time. After IRS, prevalence returns back to pre-IRS stage under generalized immunity (G), whereas it gravitates to a lower level under specific immunity (S). Parameters: gene pool size equals 9,600 and transmission rate is high. Only runs that persisted after IRS are included.

endemic malaria constitutes a relevant host-pathogen system to investigate this relationship in the context of negative frequency-dependent interactions, which stabilize coexistence and structure diversity via patterns of limiting similarity in antigenic space. Here, we show that NFDS is important not only for coexistence and structuring *P. falciparum* populations, but for the response of these populations to perturbations.

In the malaria system, limiting similarity emerges from the dominant force behind the large antigenic diversity—namely NFDS from competition of parasites for hosts, mediated by adaptive specific immunity. We specifically asked about the connection between the population structure of pathogen genetic diversity and stability to press perturbations that reduce the abundance of the parasite. We have shown that repertoire populations assembled under selection exerted by host immunity are more persistent to intervention than those assembled under a neutral model with a generalized form of immunity. We have linked this persistence to the larger fraction of highly dissimilar repertoires that provides the parasite with a way to infect non-immune hosts. Interestingly, the network analyses of parasite similarity reveal that the effect of NFDS is quickly restored after the intervention is lifted, which indicates that the process acting to maintain patterns of limiting similarity acts strongly on the parasite population. This quick rebound provides a different angle on the stability of the system.

Although the strength of NFDS is quickly restored to maintain low similarity between strains, the *var* gene diversity in our simulations slowly rebounds but does not recover to pre-intervention levels. In part, this is the consequence of a regional intervention where the processes generating new genes are not fast enough to rebuild the original local and regional diversity. The time scale at which new genes can be generated and the selective advantage they represent will be critical parameters determining the speed of diversity recovery, but also whether it can practically rebuild to the original pre-intervention levels. This aspect of the system's recovery will be investigated in future work, in light of the recently introduced concept of a threshold for the accumulation of antigenic diversity, we named the innovation number R_{div}, whose critical value is one (He and Pascual, 2021). Here, we have considered parameter ranges representative of those in nature, although the values of ectopic

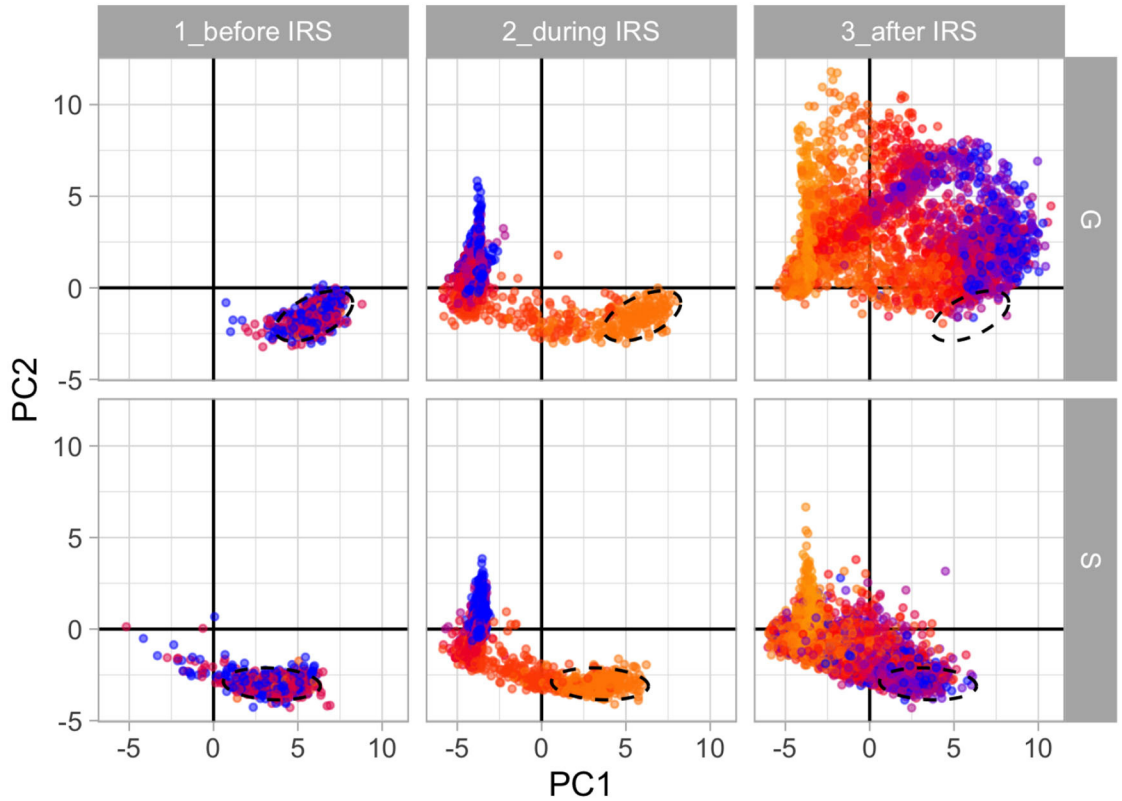


FIGURE 6 | Principal component analysis of 37 network properties of *var* repertoire similarity networks. PCA over time shows how the structure of diversity changes in the 2D space of the PCs for S and G under high transmission. The colors of the networks represent their relative time within the specific time period (i.e., before, during and after IRS) in the simulation. Ellipses show the location of network properties for pre-IRS layers. Parameters: gene pool size equals 9,600 and IRS lasts for 5 years.

recombination rates have been measured only *in vitro* (Claessens et al., 2014).

Our theoretical results suggest that competition for hosts can hamper malaria interventions, by operating to maintain the most dissimilar repertoires (Chen et al., 2011; Ruybal-Pesántez et al., 2017). This is in line with the challenge of malaria elimination that characterizes high transmission regions with *var* gene diversity levels comparable to those considered here, even after intense eradication campaigns. For example, Coleman et al. (2017) showed that after lifting a 7-year long IRS effort, transmission intensity (i.e., entomological inoculation rate, EIR) increased from 30 to 90 infectious bites per person per month in only 2 years, despite a consistent reduction in both EIR and sporozoite rates compared to a nearby control site during the intervention. Thus, transmission persists and prevalence rebounds.

Our model considers a local population embedded within a regional pool that provides a source of genetic variation. This approach is a first step toward developing a more comprehensive theory because in nature malaria is transmitted within and between local human populations, effectively creating a metapopulation of parasites. Therefore, processes that operate on metapopulations, such as dispersal and source-sink dynamics, may influence both the assembly of parasite populations and their stability in addition to local selection. For our purpose, this metapopulation context is particularly relevant in creating a vast pool of genetic variation as documented for endemic regions over larger spatial scales (Day et al., 2017; Tonkin-Hill, 2020) than those of local transmission, and longer temporal scales than those of the intervention we implement here. Addressing metapopulation dynamics explicitly for this highly diverse system would be however computationally extensive. Our approach relies on the initial compromise of a global pool typical of many assembly models in ecology.

While NFDS is a form of balancing selection, we note the difference in the organizational level when referring to genes (such as MHC) or genotypes (combinations of genes). While NFDS can operate at both levels (Takahata, 1990), in this study we have focused on the genotype level. We conjecture that coexistence under NFDS should apply to other ecological systems

where frequency-dependent interactions and resulting selection play an important role in the coexistence of a large number of entities, and possibly concurrently in the establishment of a large pool of underlying genetic/trait diversity (Pascual, 2020). For example, negative density-dependent mortality of offspring in rainforests arises from interactions with natural enemies and mutualists (Janzen, 1970; Mangan et al., 2010; LaManna et al., 2017; Schroeder et al., 2020), and traits associated with such mortality concern phytochemistry (Forrister et al., 2019). In microbial systems, where diversity is typically large, NFDS generates populations composed of multiple coexisting strains and mediates population structure as well as response to vaccination (Corander et al., 2017). Moreover, similarly to malaria, interactions of microbes with their viral predators mediated via the CRISPR adaptive immune system, result in a modular network structure. This structure occurs because viruses with segments of DNA/RNA recognized by the bacterial CRISPR system that are rare, have an advantage over those with segments that are frequent (Pilosof et al., 2020). We speculate that in these and other cases, the resulting community structure should increase stability by reducing overlap between species/genotypes in trait space, allowing for an overall better ability to respond to perturbations. This conjecture is supported by empirical evidence in other pathogen systems for the importance of frequency-dependent interactions in structuring diversity (Corander et al., 2017; Harrow et al., 2021), and by recent results showing predictive effects on strain structure following vaccination (Azarian et al., 2019; McNally et al., 2019). One way to advance relevant theory aiming to explore this hypothesis is via the combination of ABMs and network analyses of the resulting diversity structure. Such theory can guide similar network analyses of complex data sets to address the role of non-neutral processes and associated patterns of similarity based on relevant traits.

DATA AVAILABILITY STATEMENT

The agent-based stochastic simulator of malaria dynamics and the processing scripts to reproduce all the figures are stored and annotated at the GitHub repository: https://github.com/pascualgroup/frontiers_var_processing_code.

AUTHOR CONTRIBUTIONS

SP, QH, and MP conceived the project idea and wrote the manuscript. SP and QH developed the simulations and performed the computations including network analyses. KT and KD contributed to the model formulation, especially the implementation of the intervention scheme and the migration from a regional pool. All authors discussed the results and contributed to the final manuscript.

FUNDING

This study was supported by grants from the Fogarty International Center at the National Institutes of Health (Program on the Ecology and Evolution of Infectious Diseases), Grant number: R01-TW009670 to KD and MP; and the National Institute of Allergy and Infectious Disease, National Institutes of Health (Program on the Ecology and Evolution of Infectious Diseases), Grant number: R01-AI149779 to KD and MP. Salary support for QH was provided by R01-TW009670. Salary support for KT was provided by R01-TW009670 and The University of Melbourne.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fevo.2021.633263/full#supplementary-material

REFERENCES

Aguilar, A., Roemer, G., Debenham, S., Binns, M., Garcelon, D., and Wayne, R. K. (2004). High MHC diversity maintained by balancing selection in an otherwise genetically monomorphic mammal. *Proc. Natl. Acad. Sci. U.S.A.* 101, 3490–3494. doi: 10.1073/pnas.0306582101

Allesina, S., and Tang, S. (2015). The stability-complexity relationship at age 40: a random matrix perspective. *Popul. Ecol.* 57, 63–75. doi: 10.1007/s10144-014-0471-0

Artzy-Randrup, Y., Rorick, M. M., Day, K., Chen, D., Dobson, A. P., and Pascual, M. (2012). Population structuring of multi-copy, antigen-encoding genes in *Plasmodium falciparum*. *Elife* 1:e00093. doi: 10.7554/eLife.00093

Azarian, T., Martinez, P. P., Arnold, B. J., Grant, L. R., Corander, J., Fraser, C., et al. (2019). Predicting evolution using frequency-dependent selection in bacterial populations. *bioRxiv* 420315. doi: 10.1101/420315

Barry, A. E., Leliwa-Sytek, A., Tavul, L., Imrie, H., Migot-Nabias, F., Brown, S. M., et al. (2007). Population genomics of the immune evasion (var) genes of *Plasmodium falciparum*. *PLoS Pathog.* 3:e34. doi: 10.1371/journal.ppat.0030034

Barry, A. E., Trieu, A., Fowkes, F. J. I., Pablo, J., Kalantari-Dehaghi, M., Jasinskas, A., et al. (2011). The stability and complexity of antibody responses to the major surface antigen of *Plasmodium falciparum* are associated with age in a malaria endemic area. *Mol. Cell. Proteomics* 10:M111.008326. doi: 10.1074/mcp.M111.008326

Browne, L., and Karubian, J. (2016). Frequency-dependent selection for rare genotypes promotes genetic diversity of a tropical palm. *Ecol. Lett.* 19, 1439–1447. doi: 10.1111/ele.12692

Bull, P. C., Lowe, B. S., Kortok, M., Molyneux, C. S., Newbold, C. I., and Marsh, K. (1998). Parasite antigens on the infected red cell surface are targets for naturally acquired immunity to malaria. *Nat. Med.* 4, 358–360. doi: 10.1038/nm0398-358

Chen, D. S., Barry, A. E., Leliwa-Sytek, A., Smith, T. A., Peterson, I., Brown, S. M., et al. (2011). A molecular epidemiological study of var gene diversity to characterize the reservoir of *Plasmodium falciparum* in humans in Africa. *PLoS ONE* 6:e16629. doi: 10.1371/journal.pone.0016629

Chesson, P. (2000). Mechanisms of maintenance of species diversity. *Annu. Rev. Ecol. Syst.* 31, 343–366. doi: 10.1146/annurev.ecolsys.31.1.343

Claessens, A., Hamilton, W. L., Kekre, M., Otto, T. D., Faizullabhoy, A., Rayner, J. C., et al. (2014). Generation of antigenic diversity in *Plasmodium falciparum* by structured rearrangement of var genes during mitosis. *PLoS Genet.* 10:e1004812. doi: 10.1371/journal.pgen.1004812

Coleman, S., Dadzie, S. K., Seyoum, A., Yihdego, Y., Mumba, P., Dengela, D., et al. (2017). A reduction in malaria transmission intensity in Northern Ghana after 7 years of indoor residual spraying. *Malaria J.* 16:324. doi: 10.1186/s12936-017-1971-0

Collins, W. E., Warren, M., Skinner, J. C., and Fredericks, H. J. (1968). Studies on the relationship between fluorescent antibody response and the ecology of malaria in Malaysia. *Bull. World Health Organ.* 39, 451–463.

Corander, J., Fraser, C., Gutmann, M. U., Arnold, B., Hanage, W. P., Bentley, S. D., et al. (2017). Frequency-dependent selection in vaccine-associated pneumococcal population dynamics. *Nat. Ecol. Evol.* 1, 1950–1960. doi: 10.1038/s41559-017-0337-x

Cordero, O. X., and Polz, M. F. (2014). Explaining microbial genomic diversity in light of evolutionary ecology. *Nat. Rev. Microbiol.* 12, 263–273. doi: 10.1038/nrmicro3218

D'Andrea, R., Riolo, M., and Ostling, A. M. (2019). Generalizing clusters of similar species as a signature of coexistence under competition. *PLoS Comput. Biol.* 15:e1006688. doi: 10.1371/journal.pcbi.1006688

Day, K. P., Artzy-Randrup, Y., Tiedje, K. E., Rougeron, V., Chen, D. S., Rask, T. S., et al. (2017). Evidence of strain structure in *Plasmodium falciparum* var gene repertoires in children from Gabon, West Africa. *Proc. Natl. Acad. Sci. U.S.A.* 114, E4103–E4111. doi: 10.1073/pnas.1613018114

Forrister, D. L., Endara, M. J., Younkin, G. C., Coley, P. D., and Kursar, T. A. (2019). Herbivores as drivers of negative density dependence in tropical forest saplings. *Science* 363, 1213–1216. doi: 10.1126/science.aau9460

Gillespie, D. T. (1976). A general method for numerically simulating the stochastic time evolution of coupled chemical reactions. *J. Comput. Phys.* 22, 403–434. doi: 10.1016/0021-9991(76)90041-3

Grenfell, B. T., Pybus, O. G., Gog, J. R., Wood, J. L. N., Daly, J. M., Mumford, J. A., et al. (2004). Unifying the epidemiological and evolutionary dynamics of pathogens. *Science* 303, 327–332. doi: 10.1126/science.1090727

Gupta, S., and Day, K. P. (1994). A theoretical framework for the immunoepidemiology of *Plasmodium falciparum* malaria. *Parasite Immunol.* 16, 361–370. doi: 10.1111/j.1365-3024.1994.tb00361.x

Harrow, G. L., Lees, J. A., Hanage, W. P., Lipsitch, M., Corander, J., Colijn, C., et al. (2021). Negative frequency-dependent selection and asymmetrical transformation stabilise multi-strain bacterial population structures. *ISME J.* 15, 1523–1538. doi: 10.1038/s41396-020-00867-w

He, Q., and Pascual, M. (2021). An antigenic diversification threshold for falciparum malaria transmission at high endemicity. *PLoS Comput. Biol.* 17:e1008729. doi: 10.1371/journal.pcbi.1008729

He, Q., Pilosof, S., Tiedje, K. E., Ruybal-Pesántez, S., Artzy-Randrup, Y., Baskerville, E. B., et al. (2018). Networks of genetic similarity reveal non-neutral processes shape strain structure in *Plasmodium falciparum*. *Nat. Commun.* 9:1817. doi: 10.1038/s41467-018-04219-3

Healey, D., Axelrod, K., and Gore, J. (2016). Negative frequency-dependent interactions can underlie phenotypic heterogeneity in a clonal microbial population. *Mol. Syst. Biol.* 12:877. doi: 10.15252/msb.20167033

HilleRisLambers, J., Adler, P. B., Harpole, W. S., Levine, J. M., and Mayfield, M. M. (2012). Rethinking community assembly through the lens of coexistence theory. *Annu. Rev. Ecol. Evol. Syst.* 43, 227–248. doi: 10.1146/annurev-ecolsys-110411-160411

Hutchinson, G. E. (1959). Homage to santa rosalia or why are there so many kinds of animals? *Am. Nat.* 93, 145–159. doi: 10.1086/282070

Ives, A. R., and Carpenter, S. R. (2007). Stability and diversity of ecosystems. *Science* 317, 58–62. doi: 10.1126/science.1133258

Janzen, D. H. (1970). Herbivores and the number of tree species in tropical forests. *Am. Nat.* 104, 501–528. doi: 10.1086/282687

Jost, L. (2008). GST and its relatives do not measure differentiation. *Mol. Ecol.* 17, 4015–4026. doi: 10.1111/j.1365-294X.2008.03887.x

Key, F. M., Teixeira, J. C., de Filippo, C., and Andrés, A. M. (2014). Advantageous diversity maintained by balancing selection in humans. *Curr. Opin. Genet. Dev.* 29, 45–51. doi: 10.1016/j.gde.2014.08.001

Kigozi, R., Baxi, S. M., Gasasira, A., Sserwanga, A., Kakeeto, S., Nasr, S., et al. (2012). Indoor residual spraying of insecticide and malaria morbidity in a high transmission intensity area of Uganda. *PLoS ONE* 7:e42857. doi: 10.1371/journal.pone.0042857

Koelle, K., Cobey, S., Grenfell, B., and Pascual, M. (2006). Epochal evolution shapes the phylodynamics of interpandemic influenza a (H3N2) in humans. *Science* 314, 1898–1903. doi: 10.1126/science.1132745

LaManna, J. A., Mangan, S. A., Alonso, A., Bourg, N. A., Brockelman, W. Y., Bunyavejchewin, S., et al. (2017). Plant diversity increases with the strength of negative density dependence at the global scale. *Science* 356, 1389–1392. doi: 10.1126/science.aam5678

Larremore, D. B., Clauset, A., and Buckee, C. O. (2013). A network approach to analyzing highly recombinant malaria parasite genes. *PLoS Comput. Biol.* 9:e1003268. doi: 10.1371/journal.pcbi.1003268

Mangan, S. A., Schnitzer, S. A., Herre, E. A., Mack, K. M., Valencia, M. C., Sanchez, E. I., et al. (2010). Negative plant-soil feedback predicts tree-species relative abundance in a tropical forest. *Nature* 466, 752–755. doi: 10.1038/nature09273

McNally, A., Kallonen, T., Connor, C., Abudahab, K., Aanensen, D. M., Horner, C., et al. (2019). Diversification of colonization factors in a multidrug-resistant *Escherichia coli* lineage evolving under negative frequency-dependent selection. *MBio* 10:e00644-19. doi: 10.1128/mBio.00644-19

Pascual, M. (2020). "High non-neutral diversity and high-dimensional trait space in pathogen populations and ecological communities," in *Unsolved Problems in Ecology*, eds A. Dobson, D. Tilman, and R. D. Holt (Princeton, NJ: Princeton University Press), 189–200. doi: 10.2307/j.ctvs9fh2n.19

Pilosof, S., Alcalá-Corona, S. A., Wang, T., Kim, T., Maslov, S., Whitaker, R., et al. (2020). The network structure and eco-evolutionary dynamics of CRISPR-induced immune diversification. *Nat. Ecol. Evol.* 4, 1650–1660. doi: 10.1038/s41559-020-01312-z

Pilosof, S., He, Q., Tiedje, K. E., Ruybal-Pesántez, S., Day, K. P., and Pascual, M. (2019). Competition for hosts modulates vast antigenic diversity to generate persistent strain structure in *Plasmodium falciparum*. *PLoS Biol.* 17:e3000336. doi: 10.1371/journal.pbio.3000336

Rask, T. S., Hansen, D. A., Theander, T. G., Pedersen, A. G., and Lavstsen, T. (2010). *Plasmodium falciparum* erythrocyte membrane protein 1 diversity in seven genomes – divide and conquer. *PLoS Comput. Biol.* 6:e1000933. doi: 10.1371/journal.pcbi.1000933

Ruybal-Pesántez, S., Tiedje, K. E., Tonkin-Hill, G., Rask, T. S., Kamya, M. R., Greenhouse, B., et al. (2017). Population genomics of virulence genes of *Plasmodium falciparum* in clinical isolates from Uganda. *Sci. Rep.* 7:11810. doi: 10.1038/s41598-017-11814-9

Scheffer, M., and van Nes, E. H. (2006). Self-organized similarity, the evolutionary emergence of groups of similar species. *Proc. Natl. Acad. Sci. U.S.A.* 103, 6230–6235. doi: 10.1073/pnas.0508024103

Schroeder, J. W., Dobson, A., Mangan, S. A., Petticord, D. F., and Herre, E. A. (2020). Mutualist and pathogen traits interact to affect plant community structure in a spatially explicit model. *Nat. Commun.* 11:2204. doi: 10.1038/s41467-020-16047-5

Takahata, N. (1990). A simple genealogical structure of strongly balanced allelic lines and trans-species evolution of polymorphism. *Proc. Natl. Acad. Sci. U.S.A.* 87, 2419–2423. doi: 10.1073/pnas.87.7.2419

Tiedje, K. E., Oduro, A. R., Agongo, G., Anyorigiya, T., Azongo, D., Awine, T., et al. (2017). Seasonal variation in the epidemiology of asymptomatic *Plasmodium falciparum* infections across two catchment areas in Bongo District, Ghana. *Am. J. Trop. Med. Hyg.* 97, 199–212. doi: 10.4269/ajtmh.16-0959

Tonkin-Hill, G. (2020). Evolutionary analyses of the major variant surface antigen-encoding genes reveal population structure of *Plasmodium falciparum* within and between continents. *PLoS Genet.* 17:e1009269. doi: 10.1371/journal.pgen.1009269

Volz, E. M., Koelle, K., and Bedford, T. (2013). Viral phylodynamics. *PLoS Comput. Biol.* 9:e1002947. doi: 10.1371/journal.pcbi.1002947

Wanjala, C. L., Mbugi, J. P., Ototo, E., Gesuge, M., Afrane, Y. A., Atieli, H. E., et al. (2015). Pyrethroid and DDT resistance and organophosphate susceptibility among *Anopheles* spp. mosquitoes, Western Kenya. *Emerg. Infect. Dis.* 21, 2178–2181. doi: 10.3201/eid2112.150814

West, P. A., Protopopoff, N., Wright, A., Kivaju, Z., Tigererwa, R., Mosha, F. W., et al. (2014). Indoor residual spraying in combination with insecticide-treated nets compared to insecticide-treated nets alone for protection against malaria: a cluster randomised trial in Tanzania. *PLoS Med.* 11:e1001630. doi: 10.1371/journal.pmed.1001630

White, M. T., Griffin, J. T., Churcher, T. S., Ferguson, N. M., Basáñez, M. G., and Ghani, A. C. (2011). Modelling the impact of vector control interventions on *Anopheles gambiae* population dynamics. *Parasit. Vectors* 4:153. doi: 10.1186/1756-3305-4-153

WHO (2015). *Indoor Residual Spraying: An Operational Manual for Indoor Residual Spraying (IRS) for Malaria Transmission Control and Elimination, 2nd Edn*. Geneva: World Health Organization.

Zilversmit, M. M., Chase, E. K., Chen, D. S., Awadalla, P., Day, K. P., and McVean, G. (2013). Hypervariable antigen genes in malaria have ancient roots. *BMC Evol. Biol.* 13:110. doi: 10.1186/1471-2148-13-110

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Ecology and Evolution

ORIGINAL RESEARCH
published: 24 June 2021
doi: 10.3389/fevo.2021.691792

Relating the Strength of Density Dependence and the Spatial Distribution of Individuals

Micah Brush[1]* *and John Harte*[2]

[1] Department of Physics, University of California, Berkeley, Berkeley, CA, United States, [2] Energy and Resources Group, University of California, Berkeley, Berkeley, CA, United States

Spatial patterns in ecology contain useful information about underlying mechanisms and processes. Although there are many summary statistics used to quantify these spatial patterns, there are far fewer models that directly link explicit ecological mechanisms to observed patterns easily derived from available data. We present a model of intraspecific spatial aggregation that quantitatively relates static spatial patterning to negative density dependence. Individuals are placed according to the colonization rule consistent with the Maximum Entropy Theory of Ecology (METE), and die with probability proportional to their abundance raised to a power α, a parameter indicating the degree of density dependence. This model can therefore be interpreted as a hybridization of MaxEnt and mechanism. Our model shows quantitatively and generally that increasing density dependence randomizes spatial patterning. $\alpha = 1$ recovers the strongly aggregated METE distribution that is consistent with many ecosystems empirically, and as $\alpha \to 2$ our prediction approaches the binomial distribution consistent with random placement. For $1 < \alpha < 2$, our model predicts more aggregation than random placement but less than METE. We additionally relate our mechanistic parameter α to the statistical aggregation parameter k in the negative binomial distribution, giving it an ecological interpretation in the context of density dependence. We use our model to analyze two contrasting datasets, a 50 ha tropical forest and a 64 m^2 serpentine grassland plot. For each dataset, we infer α for individual species as well as a community α parameter. We find that α is generally larger in the tightly packed forest than the sparse grassland, and the degree of density dependence increases at smaller scales. These results are consistent with current understanding in both ecosystems, and we infer this underlying density dependence using only empirical spatial patterns. Our model can easily be applied to other datasets where spatially explicit data are available.

Keywords: aggregation, community assembly, density dependence, macroecology, METE, scale, spatial ecology, theoretical ecology

Edited by:
György Barabás,
Linköping University, Sweden

Reviewed by:
Daniel McGlinn,
College of Charleston, United States
Matthieu Barbier,
Institut Natura e Teoria en Pirenèus, France

***Correspondence:**
Micah Brush
micbru@berkeley.edu

Specialty section:
This article was submitted to Models in Ecology and Evolution, a section of the journal Frontiers in Ecology and Evolution

Received: *07 April 2021*
Accepted: *25 May 2021*
Published: *24 June 2021*

Citation:
Brush M and Harte J (2021) Relating the Strength of Density Dependence and the Spatial Distribution of Individuals. Front. Ecol. Evol. 9:691792. doi: 10.3389/fevo.2021.691792

1. INTRODUCTION

Spatial patterns in ecology have been studied extensively (e.g., Wiegand and Moloney, 2013; Diggle, 2014), and contain useful information about what processes shape ecosystems (Law et al., 2009; Brown et al., 2011; Münkemüller et al., 2020). Quantitative understanding of these patterns can therefore be used to infer the importance of various mechanisms, and illuminate underlying

processes (Levin, 1992; Rosenzweig, 1995; Brown et al., 2016). Additionally, models of spatial patterns allow us to better model and predict ecosystem response to natural and anthropogenic disturbances (Thomas et al., 2004; Newman et al., 2020), are critical in understanding the well studied species-area relationship (Arrhenius, 1921; Plotkin et al., 2000; Drakare et al., 2006; Harte and Kitzes, 2015), and have applications in reserve designs and conservation (Kitzes and Shirley, 2016).

A common approach to quantifying these patterns is the use of various summary statistics (Wiegand et al., 2013), which have been shown to be able to distinguish different ecological mechanisms (Brown et al., 2016). Here we take a slightly different approach and directly model the impact of an important mechanism in population dynamics: intraspecific negative density dependence. We focus on the effects of this ecological mechanism on spatial patterning.

More specifically, we consider the effects of intraspecific negative density dependence on the spatially explicit species-level abundance distribution. This distribution, $\Pi(n|A, A_0, n_0)$, is defined as the probability that if a species has n_0 individuals in a plot of area A_0, then it has n individuals in a randomly selected subplot of area A. In this analysis, we will focus on this distribution in bisected plots where $A = A_0/2$. Studying bisections is well-motivated theoretically as it often leads to simpler expressions which can be easily compared across models. Here it keeps our model analytically tractable and facilitates comparison to empirical data. We note limitations to this approach in the section 4.

One prediction of the function $\Pi(n|n_0)$ comes from the Maximum Entropy Theory of Ecology (METE), which successfully and simultaneously predicts many macroecological patterns (Harte, 2011; Harte and Newman, 2014) across a wide range of spatial scales, taxa, and habitats (White et al., 2012; Xiao et al., 2015). METE predicts very strong spatial aggregation, which is consistent with many observed ecosystems, and obtains the functional form of $\Pi(n)$ by maximizing entropy while constraining the mean number of individuals in a subplot. However, the same functional form can be obtained using a colonization rule, which is the approach we will use in our model.

Colonization rules assign spatial locations to new individuals based on the location of existing individuals. Chapter 4.1.2 in Harte (2011) shows that using the Laplace rule of succession as a colonization rule results in the same geometric distribution for $\Pi(n)$ that METE predicts. Because METE agrees well with empirical data in many cases, we will use this colonization rule in our model. Occasionally, however, we see that the empirical degree of aggregation is less than the METE prediction (Conlisk et al., 2012; McGlinn et al., 2015). To study this, Conlisk et al. (2007) added an extra parameter to the relevant colonization rule that allows $\Pi(n)$ to vary, but it has no mechanistic interpretation and is used only as a free fit parameter.

We derive a new model that uses the colonization rule consistent with METE and adds a density dependent death rule. This means our model can be viewed as a density dependent extension of METE, and in that sense hybridizes MaxEnt and mechanism. Our model introduces a parameter α which quantifies the degree of intraspecific negative density dependence. This parameter can be fit to empirical spatial data to predict the strength of underlying density dependence. However, as with all models inferring process from pattern, there are many underlying mechanisms that lead to similar spatial patterns (Vellend, 2016; Leibold and Chase, 2018), and we cannot definitively attribute any pattern to a single process.

More generally, our model predicts a more random spatial arrangement with stronger negative density dependence and more spatial aggregation with weaker density dependence. While empirically there is an apparent qualitative relationship between species density and aggregation (Condit et al., 2000; Bagchi et al., 2011; Comita et al., 2014), our aim here is to establish a general quantitative statement relating density dependence and spatial aggregation.

2. MATERIALS AND METHODS

In this section, we review the Maximum Entropy Theory of Ecology (METE) and its prediction for the species-level abundance distribution, $\Pi(n)$. We then contrast this prediction of strong aggregation to the well-known random placement model (Coleman, 1981), which predicts no spatial aggregation. Given that most species are aggregated (He and Gaston, 2000; Kitzes, 2019), but not all are as aggregated as predicted by METE (Conlisk et al., 2012), the aggregation of most species should fall somewhere between these two predictions for $\Pi(n)$.

We then introduce a density dependent death rule to combine with the colonization rule consistent with METE, and derive the resulting $\Pi(n)$ distribution. This derivation assumes a steady state between deaths and new individuals in a single species, but our results should hold if this assumption is relaxed (see section 4).

Finally, we discuss the techniques used to compare our predicted distribution to data, and describe the datasets used in our analysis.

Relevant code for the resulting $\Pi(n)$ distribution and data analysis is available at https://github.com/micbru/density_dependence_public.

2.1. The Maximum Entropy Theory of Ecology (METE)

In METE, the $\Pi(n)$ function is given by maximizing the information entropy of the $\Pi(n)$ distribution given the following constraint (Harte et al., 2008; Harte, 2011, Chapter 7.4):

$$\sum_{n=0}^{n_0} n\Pi(n|A, A_0, n_0) = \frac{n_0 A}{A_0}. \tag{1}$$

This leads to the following distribution

$$\Pi(n|A, A_0, n_0) = \frac{e^{-\lambda_\Pi n}}{Z_\Pi} \tag{2}$$

where Z_Π is a normalization factor, and λ_Π is the Lagrange multiplier determined by the constraint condition.

In the case of a bisection, $A = A_0/2$ and the Π function simplifies to

$$\Pi(n|n_0) = \frac{1}{n_0 + 1}, \tag{3}$$

which is independent of n. This means that given n_0 individuals, any arrangement of them on the two sides of a bisected plot or quadrat is just as likely as any other. In other words, this is equivalent to equal probability for each unique spatial arrangement of unlabeled individuals (Haegeman et al., 2010).

Ecologically this prediction translates to very strong spatial aggregation, as individuals are equally as likely to all be on one side of the bisection as to be evenly divided on each half. This is in agreement with many datasets (Harte et al., 2008; Harte, 2011, Chapter 8.3) but fails in others, where the theory over-predicts aggregation (Conlisk et al., 2007; McGlinn et al., 2015). This empirical agreement is why we choose the METE distribution as our starting point.

The prediction from METE is equivalent to the distribution obtained from using the Laplace rule of succession as a colonization rule (Harte, 2011, Chapter 4.1.2). This rule states that in a colonization process, the probability of placing an individual on one side of the bisected area is roughly proportional to the fraction of individuals already there. This "rich get richer" effect results in strong spatial aggregation. The probability for placing an individual on the left half of a bisected plot with n_L individuals on the left and n_R individuals on the right is

$$p_L = \frac{n_L + 1}{n_L + n_R + 2}.$$

To make our notation consistent with that above, let the number on the left be n and the total number to be n_0. The probabilities of a new individual arriving on the left or on the right are then:

$$\begin{aligned} p_L(n|n_0) &= \frac{n+1}{n_0+2}, \\ p_R(n|n_0) &= \frac{n_0 - n + 1}{n_0 + 2}. \end{aligned} \tag{4}$$

If we place n_0 individuals using this rule, the resulting probability distribution is given by Equation (3).

2.2. Random Placement

Another model for spatial ecology, perhaps the simplest, is the random placement model (Coleman, 1981). Instead of the placement rules in Equation (4), each individual is placed randomly. In a bisected plot this means each individual has a 50 percent chance of being placed on either side, $p_L = p_R = 0.5$. Placing n_0 individuals this way gives the binomial distribution

$$\Pi_{\mathrm{RP}}(n|n_0) = \binom{n_0}{n}\left(\frac{1}{2}\right)^{n_0} \tag{5}$$

which, if n_0 is large, means we are very likely to have roughly half the individuals on each side. This is equivalent to having no spatial aggregation; there is no preference for any new individual to stay close to any previous individual as each placement is a random coin flip.

2.3. Deriving the $\Pi(n)$ Distribution With a Density Dependent Death Rule

We now introduce an intraspecific density dependent death rule in addition to the METE colonization rule in Equation (4). To allow for general density dependence, we set the death rate proportional to n^α. The parameter α determines the strength of the density dependence, and can be inferred from the data. Density dependence may result from resource limitation, or some other mechanism (e.g., the Janzen-Connell effect, Janzen, 1970; Connell, 1971).

In the case of a bisected plot, each death must be on the left or right. Thus, given that we have one death in a species, the probabilities that the death is on the left, $p_{D,L}$, or on the right, $p_{D,R}$, are

$$\begin{aligned} p_{D,L}(n|n_0) &= \frac{n^\alpha}{n^\alpha + (n_0 - n)^\alpha} \\ p_{D,R}(n|n_0) &= \frac{(n_0 - n)^\alpha}{n^\alpha + (n_0 - n)^\alpha}. \end{aligned} \tag{6}$$

Now that we have the colonization and death rules (Equations 4 and 6), we can derive the general $\Pi_\alpha(n|n_0)$ for bisections. We will assume the population size of the species is constant and step the model forward over time, where at each step in the model we will have one death followed by the placement of one new individual within a species. Each placement can be interpreted ecologically as a birth or as the immigration of an individual from the same species. We can then solve for the resulting steady state distribution where we reach an equilibrium in the spatial pattern.

There are several approaches for deriving the steady state solution for such a system. Here, we equate the rates leaving and entering any individual state $\Pi_\alpha(n|n_0)$. We take the probability that we start with n individuals on the left, one on the right dies and then one is placed on the left resulting in $n + 1$ individuals on the left, and equate that to the probability that we have $n + 1$ individuals on the left, one on the left dies and then one is placed on the right resulting in n individuals on the left. We could have equivalently done the same thing with n and $n-1$. Equating these rates using the probabilities in Equations (4) and (6) leads to a recursion relation. Solving it gives a general stationary solution for Π_α with a given n_0 and α:

$$\Pi_\alpha(n|n_0) = \frac{n^\alpha + (n_0 - n)^\alpha}{C(n_0, \alpha) n_0^\alpha}\binom{n_0}{n}^{\alpha - 1} \tag{7}$$

where $C(n_0, \alpha)$ is the overall normalization that does not have a closed analytic form. In the case that $n_0(\alpha - 1)$ is large, an approximate form for the normalization is

$$C = \frac{2^{n_0(\alpha-1)}\pi n_0}{\sqrt{\alpha - 1}}\left(\frac{1}{2\pi n_0}\right)^{\alpha/2}. \tag{8}$$

See **Supplementary Material 1** for the details of this derivation.

If $\alpha = 2$, we can solve for the normalization explicitly to get

$$\Pi_{\alpha=2}(n|n_0) = \frac{n^2 + (n_0 - n)^2}{2^{n_0 - 1} n_0(n_0 + 1)}\binom{n_0}{n}, \tag{9}$$

and if $\alpha = 1$, we recover the METE prediction $\Pi_{\alpha=1}(n|n_0) = \frac{1}{n_0+1}$.

2.4. Comparing to Data

Inferring the degree of intraspecific density dependence in empirical data requires obtaining a value of α consistent with the data. Bisection predictions can be compared to data by rank ordering the fraction of individuals present in one half of the plot for each species (e.g., Harte, 2011). This method, however, ignores species abundance and does not account for the likelihood of individual data points. This can lead to incorrect conclusions about which model is preferred (see **Supplementary Material 2**).

We instead find the maximum likelihood α given the data, where we minimize the sum of the negative logs of the probabilities given data points n_i and $n_{0,i}$, where i labels each quadrat for a given species. Inferring α using this method gives us the values that are the most consistent with the data, even if they may not look like they agree with the rank ordered fractions (**Supplementary Figure 1** and **Supplementary Table 1**).

The statistical error in estimating α this way goes as $1/\sqrt{p}$ where p is the sample size (see **Supplementary Material 3**). We can also get some idea of error from the maximum likelihood estimate itself by considering the width of the likelihood distribution, however for **Figures 3**, **4**, we do not include these error bars as they are smaller than the data points.

For determining α for individual species, we will require multiple bisections and the sample size p will be roughly the number of cell pairings, $p \approx 2^{b-1}$, where b is the number of bisections. There will be fewer data points in practice as some cell pairings will be empty.

We can also define a community α, assuming each species follows the same death rule with identical α. In this case, we will have a larger sample size. For a single bisection, we will have a sample size p equal to the number of species, $p = S_0$. For multiple bisections where we consider the species on aggregate, the sample size will be roughly equal to the number of species multiplied by the number of cells, $p \approx S_0 2^{b-1}$. Again, the equality is not exact as not all cell pairings beyond the first bisection will have all of the species present at the single bisection level.

In both the case of the species-level and community-level α, we will bisect single plots more than once (into quadrants, then into 8 cells, etc.) when comparing the model to data. In our analysis, we begin by bisecting the plot in half in one direction, then bisecting each of the resulting plots in the opposite direction. We alternate this bisection pattern until we have 2^b cells. We can then combine adjacent cells (either left/right or up/down) as if they were single plots with abundance $n_{0,i}$, where i will index the plots and range from 1 to 2^{b-1}. We then choose the abundance on one half to be n_i. This method gives us 2^{b-1} points.

Additionally, in this analysis we will only consider species that could have at least one individual per bisection ($n_0 > 2^{b-1}$). The smallest scale we consider in our datasets is $b = 8$, so when we bisect the plot more than once we will only consider species with more than 128 individuals. This restriction ensures that we do not have too many plots with only a few individuals present. If n_0 is very small, $\Pi_\alpha(n)$ is not particularly sensitive to α and it becomes very difficult to reliably infer α from the data. For $n_0 \leq 2$, $\Pi_\alpha(n)$ does not depend on α.

2.5. Data Used

We will compare our results to two contrasting datasets. First, we will use data from a sparse Californian serpentine grassland site (Green et al., 2003, 2019) at the McLaughlin University of California Natural Reserve censused in 1998. This is a 64 m^2 plot divided into 256 cells with 24 species and 37 182 individuals. There are 10 species with abundance greater than 128 individuals that constitute 36 783 individuals.

Second, we will use data from Barro Colorado Island (BCI) in Panama (Condit, 1998; Hubbell et al., 1999, 2005; Condit et al., 2019), a 50 ha plot in a moist tropical forest. We will work with the 2005 census and consider plants with a diameter at breast height (dbh) greater than 10 cm. This dataset has 229 species and 20,852 individuals, and 40 species with abundance greater than 128 individuals that constitute 15,960 individuals.

3. RESULTS

3.1. Comparison to METE and Random Placement

Figure 1 compares the bisection predictions for $\Pi(n)$ from METE, random placement, and our density dependent model for various α, at $n_0 = 10$ and 50. In general, our model predicts that increasing negative density dependence (larger α) leads to more random spatial patterning, and less density dependence (smaller α) leads to stronger aggregation.

We can relate our distribution directly to both the METE and random placement distributions for different values of α. $\alpha = 1$ corresponds exactly to the METE solution, which makes sense given that the placement and death rules are both linear in n. As $\alpha \to 2$, our distribution approaches the random placement prediction if n_0 is large enough (**Supplementary Material 4** shows this result analytically). For $1 < \alpha < 2$, we vary continuously between METE and random placement. We can make the distribution even more spatially aggregated than METE with $\alpha < 1$ and even less than random placement (overdispersed) with $\alpha > 2$.

We can also relate this distribution to the commonly used conditional negative binomial distribution (Bliss and Fisher, 1953; He and Gaston, 2000, 2003; Green and Plotkin, 2007) in the limit of large n_0, assuming that matching the peak of the distributions is a good approximation for the entire distribution. In that limit, the aggregation parameter k is approximately related to the density dependent parameter α by

$$k \approx \frac{n_0}{2}\left(\frac{\alpha - 1}{2 - \alpha}\right) + 1. \tag{10}$$

Note that this approximation holds for $1 \leq \alpha \leq 2$, which should be the ecologically relevant range for most species as most species will be more aggregated than random placement, and less aggregated than METE. This also allows the aggregation parameter k to be interpreted mechanistically as the degree of density dependence, in that higher k corresponds to higher α and

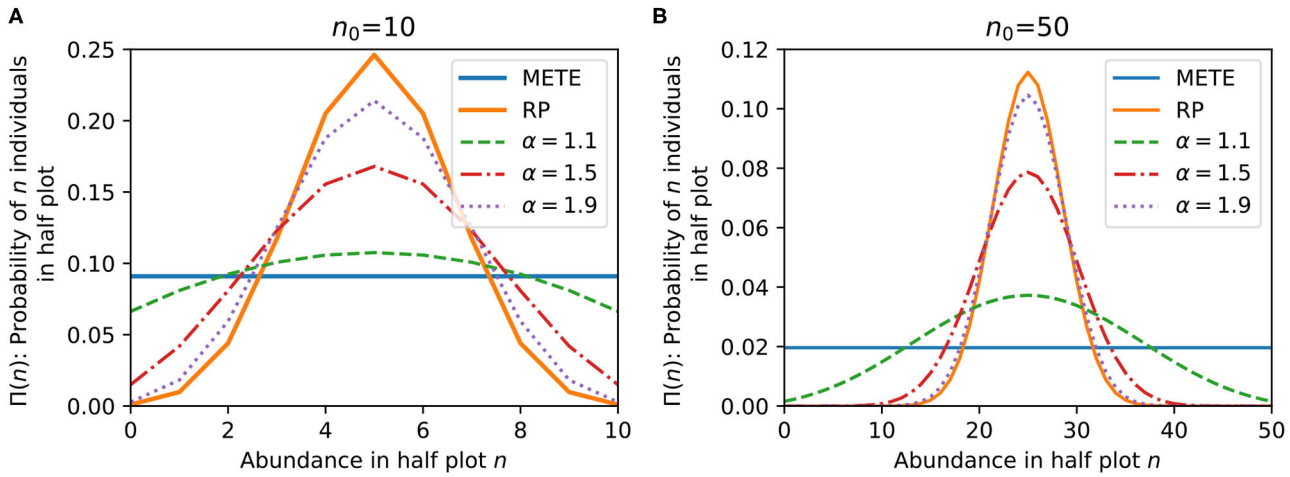


FIGURE 1 | Comparison of the bisection probability distributions $\Pi(n)$ from METE, random placement (RP), and our density dependent model with varying α at **(A)** $n_0 = 10$ and **(B)** $n_0 = 50$. At $\alpha = 1$, our model corresponds exactly to METE. At larger n_0, $\alpha \to 2$ approaches the random placement distribution. Our model varies continuously between METE and random placement for $1 < \alpha < 2$.

greater density dependence. See **Supplementary Material 5** for the derivation.

3.2. Individual Species

Since the Π function is defined on the species level, we can consider each species separately and find the maximum likelihood α for each. To do this we have to go beyond the first bisection to get multiple data points for the same species at smaller scales.

For the serpentine data, we exclude *Eriogonum nudum* from the following figures as an outlier (see section 4). This leaves 9 species with abundance greater than 128 individuals.

Figure 2 shows the distribution of α values among the species at each scale, for both the serpentine and BCI data. The median α increases at smaller scales for both datasets, and is higher overall at the BCI dataset, even though the absolute scale is much larger. The spread in α is quite large, but this variation is expected considering the small number of data points, especially for rarer species. Most species have an α between 1 and 2, which is somewhere between the aggregation predicted by METE and random placement.

3.3. Community α

We can instead treat α as a community parameter, using each species as a single data point to recover a community α. **Figure 3** shows the direct comparison between our model prediction and the serpentine and BCI datasets at the single bisection level. Each data point is the observed fraction of individuals in one half of the plot vs. the species abundance. The curves in this figure show the 95% contour intervals for the $\Pi(n|n_0)$ distributions predicted by METE, random placement, and our density dependent model with the maximum likelihood α value. We can see that with increasing n_0, the random placement model narrows quickly to having most of its probability weight around 0.5, whereas the METE contours are very wide.

At the single bisection level, the maximum likelihood result for the serpentine dataset is nearly indistinguishable from $\alpha = 1$, so the confidence interval curves on the plot for METE and the density dependent model overlap for most n_0. For the BCI data, the maximum likelihood value is $\alpha = 1.12$, slightly larger than 1. In this case, where $1 < \alpha < 2$, we see the width of the predicted distribution is between METE and random placement. The likelihoods for each of the models are shown in **Table 1**.

3.4. Scale Dependence in Community α

Going beyond the first bisection allows us to see how α varies depending on the scale of our plot. **Figure 4** shows how α scales with fractional area for both the serpentine and BCI plots. Density dependence increases at smaller scales in both datasets. The trend in community α across scales is similar to the median α in the single species analysis, though the median α is in general slightly larger than the community α. Note that here we restrict our analysis to species with $n_0 > 128$ for all scales so that we are including the same species across scales.

4. DISCUSSION

Our model establishes a quantitative relationship between the spatially explicit distribution $\Pi(n)$ and the parameter α, which measures the strength of negative density dependence. This can be seen in **Figure 1**, where the $\Pi_\alpha(n)$ distribution flattens with smaller α, indicating greater aggregation, and broadens as α increases. Importantly, the parameter α has a direct interpretation as quantifying the strength of negative density dependence. Further, our relationship in Equation (10) allows us to interpret the parameter k in the negative binomial distribution in the same intuitive way.

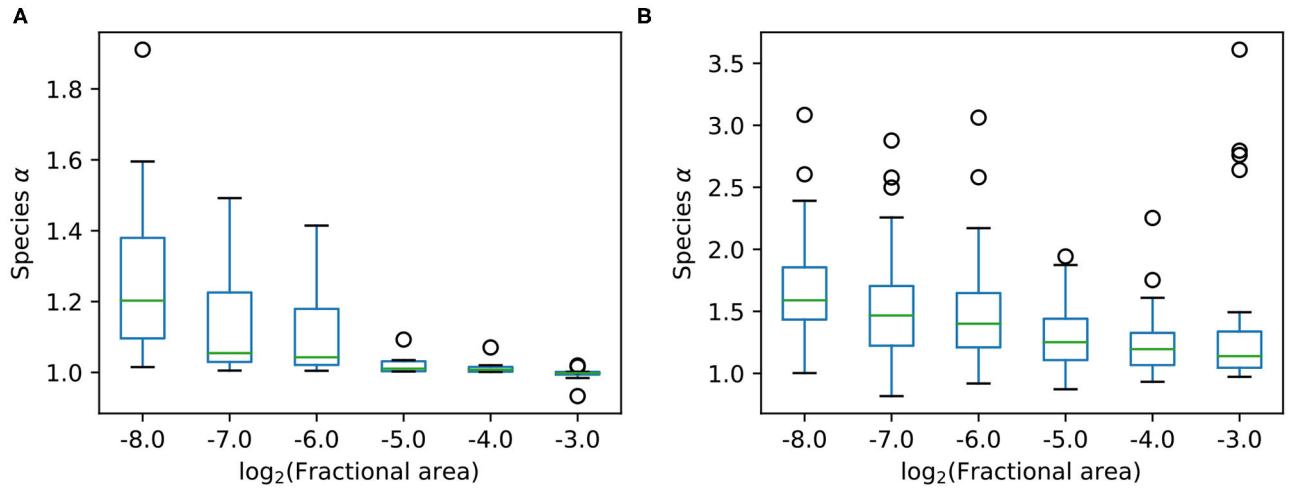


FIGURE 2 | Boxplots for α among the species at different scales at both sites, where panel **(A)** shows 10 species from the Serpentine dataset, and panel **(B)** shows 40 species from the BCI dataset. In both cases at smaller scales α is larger, and we see a relatively large spread in α across species at the same scale. The boxplots show boxes from quartile 1 (Q1) to quartile 3 (Q3) with a line at the median. The whiskers extend to 1.5×(Q3-Q1). The remaining points are plotted as individual circles.

4.1. Comparing Species and Community α

In our analysis, we consider α both as a separate parameter for each species (as in **Figure 2**), and as a community parameter where each species has the same α (as in **Figures 3**, **4**). The community α is harder to interpret ecologically, but we include it to allow for comparisons with models with community level aggregation parameters (e.g., Volkov et al., 2005; Conlisk et al., 2007). To analyze and compare the accuracy of the species-level α and the community α, we considered the Akaike Information Criterion (AIC) in both cases across scales (**Table 2**). This was calculated for single species as the negative log-likelihood summed over each species with the number of parameters equal to the number of species, whereas for the community α there was only a single α parameter. For both serpentine and BCI at all scales considered, we find that the AIC is lower with species-level α compared to a single community α, despite the inclusion of 9 more parameters in the case of the serpentine data and 228 more parameters in the case of the BCI data. We therefore conclude that a separate α for each species describes the data better than a single community α.

4.2. Comparing Serpentine and BCI

We use our model to directly compare our results between our two contrasting datasets, serpentine and BCI. Because the serpentine site was very sparse, whereas the BCI forest is tightly packed, we expect higher α values and greater density dependence at BCI than at the serpentine site. This is consistent with our inferred values of α at both the individual species level and at the community level.

Another difference between the serpentine and BCI sites is how well other macroecological distributions agree with METE. METE well describes other patterns at the serpentine site, and does less well at explaining the BCI data. Given that $\alpha = 1$ corresponds to the METE prediction for $\Pi(n)$, we might expect that ecosystems well-described by other METE predictions will have $\alpha \approx 1$, as these systems will generally be consistent with METE. This is consistent with our analysis here as the median and community αs for the serpentine data are approximately 1 at the largest scale, whereas at BCI the median and community αs are larger than 1.

Because METE predictions seem to hold for relatively static and undisturbed ecosystems (Newman et al., 2020), this suggests interpreting an increase in density dependence away from METE ($\alpha > 1$) as a kind of ecological disturbance. A biological example of strong density dependent mortality as a result of disturbance could be the self-thinning of trees in forest recovery from wildfire, such as bishop pines in coastal California (Harvey et al., 2014). This interpretation is in line with the recently proposed DynaMETE theory (Harte et al., 2021), which models specific mechanistic disturbances away from METE to predict macroecological patterns.

4.3. Scaling

Our scaling results in both **Figures 2**, **4** make ecological sense. We expect that at smaller scales, the density dependence would be larger as individuals compete more for resources at that scale. At large scales, we expect α to be close to 1 as the individuals do not compete over large distances. This means that the spatial distributions look more aggregated on large scales than on small scales as the individuals within species broadly group together, but repel each other at small scales. We see this trend at the individual level in **Figure 2** as the medians increase at smaller scales, and for the community α in **Figure 4**.

Our repeated bisection analysis also indicates at which scale density dependence becomes important. This will appear as a shoulder in the data where α moves away from ≈ 1. We could do this for individual species by tracing α and looking for a shoulder in **Figure 2**, but here we will look at the community results in

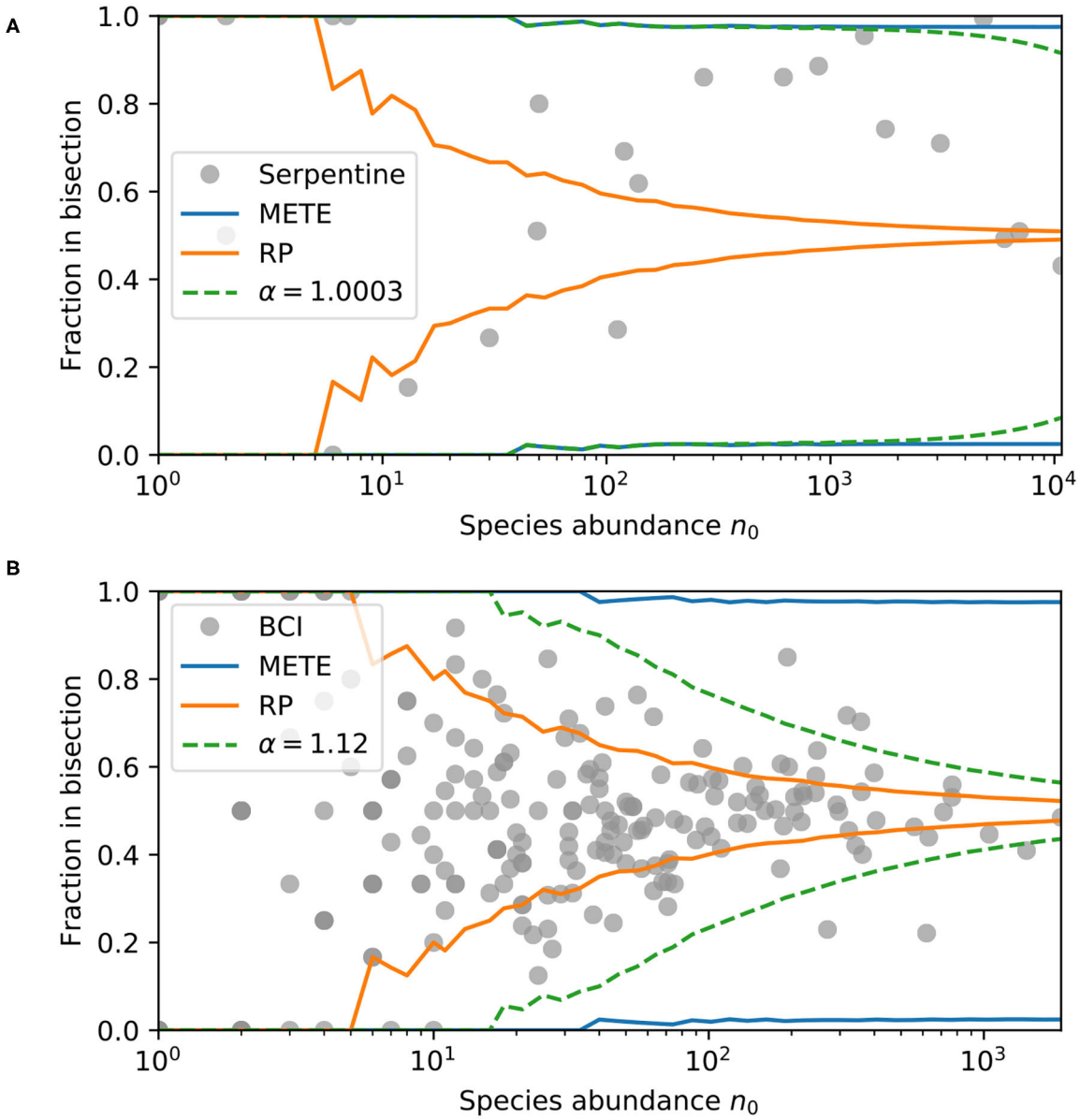


FIGURE 3 | Ninety-five percent contour intervals for the predicted bisection probability distributions $\Pi(n|n_0)$ from METE, random placement, and our density dependent model with maximum likelihood community α, and bisection data for each species in **(A)** the serpentine dataset, and **(B)** the BCI dataset. The data is plotted for each species, where the y-axis is the fraction in one half plot and the x-axis is the total species abundance in that plot. The contours are calculated at each n_0. For our density dependent model with a community α, $\alpha = 1.0003$ maximizes the log-likelihood for the serpentine dataset, and $\alpha = 1.12$ maximizes log-likelihood for BCI dataset.

order to compare to Conlisk et al. (2007) and Volkov et al. (2005). Looking at **Figure 4**, we find that the shoulder in absolute scale corresponds to < 0.5 m^2 for the serpentine plot and < 1.6 ha for the BCI dataset. This again makes sense given that the serpentine grassland is much more sparse than the BCI forest.

We first compare to Conlisk et al. (2007), who introduce a fit parameter ϕ that modifies the colonization rule Equation (4) and allows the Π distribution to vary continuously between random placement and METE. They compare their estimated community ϕ parameter to both the serpentine and BCI data in their Figure 6. For the serpentine data, they find that at scales larger than around 0.5 m^2 (the 8th bisection), ϕ approaches 0.5, which corresponds to the METE prediction. At scales around 0.5 m^2 or smaller, $\phi \approx 0.25$, where $\phi = 0$ corresponds to random placement. This is consistent with our scaling results in **Figure 4**. For BCI, they find that at all but the largest scales $\phi \approx 0.25$. Our result that α is larger at BCI than at the serpentine site across scales, which corresponds to less spatial aggregation, is not consistent

TABLE 1 | Log-likelihood values for the three different models, with α as a community parameter.

Serpentine		BCI	
Model	**Log-likelihood**	**Model**	**Log-likelihood**
METE	−114.8	METE	−729
RP	−5188.6	RP	−963
$\alpha = 1.0003$	−114.6	$\alpha = 1.12$	−660

We can compare our model to METE using the deviance in log-likelihood and obtain a p-value. The deviance is defined as twice the difference in log-likelihood. For the serpentine dataset, the deviance is 0.6 which corresponds to a p-value of 0.45. For the BCI dataset, the deviance is 138 which corresponds to a p-value of $< 10^{-30}$.

with their findings. We believe this is due to a difference in how the data are analyzed.

In Conlisk et al. (2007), the species abundance n_0 is measured at the scale of the full plot, and the bisection prediction is recursively iterated to smaller scales (see their Theorem 2). Here, we treat each bisection at smaller scales separately. For example, after dividing the plot into 128 quadrants (8 bisections), we look at the species abundance in each individual quadrat without considering n_0 at the scale of the entire plot. In principle, we could conduct our analysis in the same way and anchor at the largest scale, though this would be difficult analytically, and our approach makes use of the empirical data available at each scale rather than only at the largest scale. Further, the method in Conlisk et al. (2007) depends implicitly on the chosen size of the overall study plot. This is not true in our analysis as a bisection studied at any scale does not depend on information at any other scale. In practice, this means that in our analysis there is no difference between studying species in a 1 m^2 subplot embedded in a 100 m^2 plot vs. studying the same 1 m^2 plot independently. This difference in how n_0 is treated across scales could lead to different predictions for α (or ϕ).

We can further compare our results to Conlisk et al. (2007) by relating our α to ϕ, using their relationship between ϕ and k and our Equation (10). This relationship depends on n_0, which may affect comparisons between these parameters across scales. Finally, an additional difference between our analyses is our different cutoff of $n_0 > 128$, and for BCI, dbh > 100 mm, however this does not explain all of the difference between our results. **Supplementary Material 6** derives an approximate relationship between α and ϕ, **Supplementary Figure 4** uses that relationship to transform our **Figure 4** to a relationship in ϕ, and **Supplementary Figure 5** shows how our result changes if we remove our abundance threshold. A takeaway from this comparison is that these scaling results depend at least in part on the choice of model and the data analysis methods.

Volkov et al. (2005) showed that intraspecfic and symmetric density dependence can explain different shapes for the species-abundance distribution. Their added parameter c is interpreted as a measure of the strength of symmetric density dependence, where $c = 0$ corresponds to no density dependence. This parameter is therefore similar to our community α in that all species have the same degree of negative density dependence. They then show at what density these effects become important in their Figure 3. For BCI, they find $c = 1.80$, and the density dependent effects are visible for species with $n > 27$. To convert this to area, we need to look at scales of the total area divided by the abundance where density dependent effects become visible. Thus, this corresponds to density dependence entering at scales smaller than a fractional area of $1/27 = 1/2^{4.75}$, which is close to the same scale where we see α increase away from 1 in **Figure 4**.

Across these results, we interpret increasing α at small scale as an increase in density dependence. However, at smaller spatial scales where there are fewer individuals it becomes more difficult to distinguish between different patterns of aggregation. In particular, when $n_0 A/A_0 << 1$, it is difficult to determine if the empirical pattern is due to noise or a specific clustering process (Harte, 2011, p. 63). This sampling effect should be small here, as even at the smallest scale the median $n_0 A/A_0$ is greater than 1 for both datasets (**Supplementary Material 7**).

4.4. Trends for Individual Species at BCI

At the individual species level at BCI, we find overall that most species at all scales are more aggregated than random ($\alpha < 2$ in **Figure 2**). This is consistent with results from Condit et al. (2000). We also find that species tend to be more aggregated at large scales than at small scales (median $\alpha > 1$ at small scales and $\alpha \approx 1$ at large scales in **Figure 2**), which makes sense as we expect some species to only be present in certain areas of the plot.

More broadly, we might expect to find trends in inferred density dependence with species abundance or size. More abundant species may be competing more for the same resources, or larger species may compete over larger distances. For example, Condit et al. (2000) find that both rarer species and smaller individuals tend to be more aggregated, however at a much smaller scale (within a 10 m radius). We looked for trends in abundance, mean dbh, and total energy for each species at BCI with $n_0 > 128$ across all scales considered (as in **Figure 2**).

In terms of abundance (**Supplementary Figure 7** and **Supplementary Table 2**), we do not find any species with high α and high abundance (no highly density dependent high abundance species), and we find that the variance in α decreases with abundance. We also find that at all scales except the two smallest, α decreases slightly with increasing log of abundance. Thus, we find that at larger scales, more abundant species are slightly more aggregated than less abundant species.

We find no evidence of a trend with species' mean dbh (**Supplementary Figure 8** and **Supplementary Table 3**), though it is possible this trend is obscured by variance in individual size within a species, or that the range of mean dbh we considered (about $100 - 500$ mm) is too small to see its effect. Finally, we looked for an overall energy effect. Considering that the most abundant species tend to be smaller, it may be that density dependence depends on the total metabolic rate of a species. Plotting this relationship (**Supplementary Figure 9** and **Supplementary Table 4**) again does not reveal a significant scaling relationship at all scales except one ($\log_2(A/A_0) = -6$).

A plausible mechanism for the observed density dependence at BCI is the Janzen-Connell effect (Janzen, 1970; Connell, 1971),

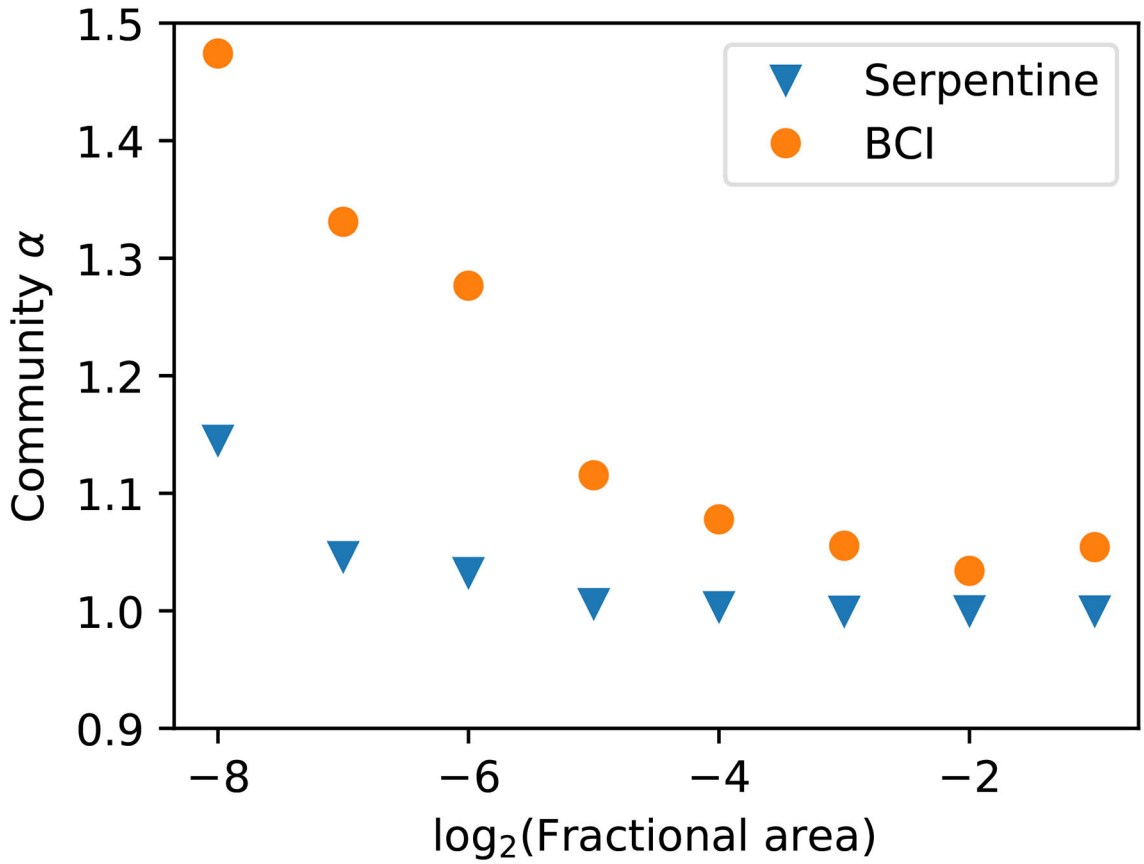


FIGURE 4 | Community α scaling with area for species with abundance $n_0 > 128$. The density dependence again increases at smaller scales and the trend is similar to the single species analysis. The serpentine dataset has 36,783 individuals and the BCI dataset has 15,960 individuals.

TABLE 2 | Comparison of the Akaike Information Criterion (AIC) for α defined at the individual species level and at the community level in both the serpentine and BCI data and across scales.

	Scale (A/A_0)	2^{-8}	2^{-7}	2^{-6}	2^{-5}	2^{-4}	2^{-3}
Serpentine	Species α, AIC	474	769	1,294	1,931	3,208	4,881
	Community α, AIC	485	777	1,321	2,079	3,420	5,182
BCI	Species α, AIC	10,133	7,541	5,079	3,409	2,109	1,271
	Community α, AIC	10,207	7,621	5,148	3,522	2,171	1,307

At the individuals species level, the number of parameters is equal to the number of species, whereas at the community level there is only a single parameter. The AIC is lower at the species level in all cases.

whereby areas near parent trees are inhospitable for offspring, resulting in density dependence. Various studies (Harms et al., 2000; Carson et al., 2008; Comita et al., 2014) have observed this effect at BCI, which is consistent with our result that $\alpha > 1$ for most species at smaller scales there.

See **Supplementary Material 8** for more information on these trends.

4.5. Notable Species

For individual species at the BCI dataset, *Gustavia superba* stood out with an average α of 1.001 across scales. This species is largely limited to 2 hectares of young secondary forest along the edge of the plot, (J. Wright, personal communication, 2019) making it look especially aggregated and resulting in a maximum likelihood α close to 1.

In the serpentine dataset, we excluded *Eriogonum nudum* as an outlier for part of our analysis. The maximum likelihood α was > 6 at the smallest scale and the maximum itself was very shallow. This species has a large canopy compared to the other grassland plants, and tends to be found far from other individuals. It makes sense that it would be overdispersed with $\alpha > 2$.

4.6. Implications for the Species-Area Relationship

In METE, the spatial distribution is used together with the species-abundance distribution to predict the species-area relationship (Harte, 2011, Chapter 7.5), and to upscale predictions of biodiversity (Harte and Kitzes, 2015). These predictions should hold in ecosystems like the serpentine grassland analyzed here, as the observed species aggregation agrees with the METE prediction. However, different levels of aggregation will impact the species-area relationship. The impact of aggregation is discussed in Wilber et al. (2015). They find that increasing randomization decreases the predicted slope of the species-area relationship at the same scale, and therefore upscaling METE will overpredict species richness. In addition, they analyze the effect of variation in aggregation among species, which slightly decreases the slope at small scales and increases the slope at larger scales. This results in a species-area relationship that more closely resembles a power law. They also consider the effect of decreasing aggregation across scale, which results in a species-area relationship that no longer displays scale collapse. We observe both of these effects here.

4.7. Limitations and Assumptions

As with all models inferring process from pattern, we can never be sure the pattern we observe can be completely attributed to the process we model. There are many different underlying processes that can lead to aggregation, including environmental filtering and dispersal limitation (Vellend, 2016; Leibold and Chase, 2018), and it is not possible for any one model to include every effect. Our empirical results here are consistent with our interpretation of α as a parameter that relates to the strength of intraspecific negative density dependence, however there are certainly other important mechanisms in these datasets. Regardless of our ability to infer process from pattern, our theoretical result that increasing density dependence increases spatial randomization holds.

Our model is also limited in that it only considers bisections, and it would be useful to extend it to be more general. There are many spatial arrangements that can not be accurately captured by dividing plots into bisection, and in general a single functional summary statistic does not completely describe the observed spatial pattern (Wiegand et al., 2013). For example, if we divide our plot into an m by m grid, and have one individual per cell, we would see exactly 0.5 as the fraction for each bisection. This result would be consistent with random placement with a large number of individuals, which does not well describe this exceptionally uniform arrangement. There could also be different degrees of spatial aggregation within a cell that we will not accurately capture with a bisection. Despite these limitations, bisections are useful for understanding commonly observed macroscopic spatial patterns.

A conceptually simple extension to our model is to divide plots into quadrisections rather than bisections. The colonization and death rules then have three unknowns rather than one (the number of individuals in each quadrant, where the fourth is determined by constraining the sum to be n_0). This makes it hard to solve analytically, however we can simulate the birth-death process until it reaches steady state. We find no significant difference in our simulation compared to our prediction from two bisections, and find that a community $\alpha = 1.12$ is still consistent with the BCI data.

Because we consider the steady state solution in our model, we are assuming that the density dependence time scale is longer than the time scale of individual births or deaths. That is, α must not change too rapidly in time. This assumption is justified for many systems roughly in steady state with their environment and not undergoing rapid change (Newman et al., 2020).

Solving for the steady state solution also assumes that births and deaths are in balance. We assume here that there is a single death followed by a placement, however simulating two deaths followed by two placements gives a probability distribution consistent with our analytic prediction. We expect our result to hold with other numbers of deaths and placements. Assuming that births and deaths are in balance also implicitly assumes some amount of negative density dependence, and here α provides a quantitative measure of the degree of density dependence.

Another assumption in our model is the choice of colonization rule itself, though if we had chosen a different colonization rule many of our conclusions would remain the same. We use the colonization rule consistent with METE because of its good empirical agreement (Harte, 2011, Chapter 8.3). This allows us to interpret the $\alpha = 1$ case as consistent with METE. This is useful as METE can be thought of as a null theory that holds in ecosystems that are undisturbed and relatively static (White et al., 2012; Xiao et al., 2015; Harte et al., 2017; Newman et al., 2020), and $\alpha \neq 1$ can be thought of as a density dependent correction, away from the MaxEnt distribution. In this sense, this model hybridizes MaxEnt and mechanism.

Instead, as an example, we could have chosen the colonization rule resulting in the random placement distribution. For a bisection this rule is just $p_L = p_R = 0.5$. In this case, $\alpha = 1$ would recover the binomial distribution, which we know does not well describe most spatial data (Condit et al., 2000; He and Gaston, 2000), and so we cannot interpret $\alpha \neq 1$ as a density dependent correction. As another example, if we had chosen the more general colonization rule in (Conlisk et al., 2007) we would have two parameters to tune, making it difficulty to differentiate between colonization and death. In ecosystems where we suspect a different colonization rule may be in play, we could modify our theory appropriately. In any of these cases, our general results would remain largely unchanged.

4.8. Future Work

One advantage of the bisection approach is that it can make predictions about inter-quadrat correlations. McGlinn et al. (2015) examined these correlations and compared empirical distance-decay relationships with the spatial predictions of METE ($\alpha = 1$ in this model). They found that the predicted distance-decay was much stronger than observed. We would expect the predicted distance-decay relationship to be weaker with $\alpha > 1$ in our model. Conlisk et al. (2007) note that $\phi > 0$ in their model produces more realistic looking distance-decay than random placement. Together, this means that with $1 < \alpha < 2$

our model should predict a more realistic shape for the distance-decay relationship compared to random placement, but with less steep of a slope than predicted by METE. However, Conlisk et al. (2007) also note that the analysis of these inter-quadrat correlations makes use of distance between cell pairs rather than physical distance, which limits the analysis [though note Ostling et al. (2004) provides a set of user rules to reduce this effect]. This issue is also present in our model. Future comparisons to empirical distance-decay relationships could provide another method of estimating α and testing this framework.

Another advantage of our approach is that it only requires static spatial data. However, analyzing a single dataset over time could provide an interesting test of our interpretation of α as a measure of density dependence. This would be particularly appropriate with data where strong density dependent mortality is known to occur, for example a self-thinning forest recovering from wildfire (Harvey et al., 2014; Newman et al., 2020).

Finally, while our analysis here compares two contrasting datasets, future work could analyze more ecosystems to look for effects of habitat type, species richness, or average density.

5. CONCLUSION

Our model robustly predicts that increased intraspecific negative density dependence leads to more random spatial patterning, and establishes a quantitative relationship between the degree of density dependence described by the parameter α and spatial patterning described by the metric $\Pi(n)$. We predict that this result is general across ecosystems and taxonomic groups. We find that at all but the smallest scales, the serpentine grassland site is consistent with the absence of a density dependent correction and has the strong spatial aggregation predicted by METE. This is true for both the median individual species and at the community level. At the tropical forest site, our results indicate that negative density dependence is important: the median species α and the community α are both greater than 1 at even the largest scales. Both ecosystems show scaling of α consistent with its interpretation as the strength of negative density dependence. Median species α and community α are larger at smaller scales, and increase away from 1 at scales consistent with other analyses. Overall, our analysis of α is consistent with the interpretation of density dependence at both sites. Because this model uses only static spatial patterning, it can be applied in any ecosystem with spatially explicit data.

DATA AVAILABILITY STATEMENT

The serpentine data are available from the Dryad Digital Repository at https://doi.org/10.6078/D1MQ2V. The BCI data can be found at https://forestgeo.si.edu/explore-data and are available from the Dryad Digital Repository at https://doi.org/10.15146/5xcp-0d46. Relevant code is available at https://github.com/micbru/density_dependence_public. The original contributions presented in the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding author/s.

AUTHOR CONTRIBUTIONS

MB conducted the formal analysis and led the writing of the manuscript. Both authors were involved in conceptualization.

FUNDING

This work was supported by the National Science Foundation under Grant no. DEB-1751380. MB acknowledged the support of the Natural Sciences and Engineering Research Council of Canada (NSERC) [PGSD2-517114-2018]. Publication made possible in part by support from the Berkeley Research Impact Initiative (BRII) sponsored by the UC Berkeley Library.

ACKNOWLEDGMENTS

JH thanks the Santa Fe Institute and the Rocky Mountain Biological Laboratory for their hospitality. We thank Kaito Umemura for valuable discussion and feedback, and Egbert Leigh and Joseph Wright for their help with the BCI dataset. We thank Jessica Green for the serpentine data. The BCI forest dynamics research project was founded by S. P. Hubbell and R. B. Foster and is now managed by R. Condit, S. Lao, and R. Perez under the Center for Tropical Forest Science and the Smithsonian Tropical Research in Panama. Numerous organizations have provided funding, principally the U.S. National Science Foundation, and hundreds of field workers have contributed.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fevo.2021.691792/full#supplementary-material

REFERENCES

Arrhenius, O. (1921). Species and area. *J. Ecol.* 9, 95–99. doi: 10.2307/2255763

Bagchi, R., Henrys, P. A., Brown, P. E., Burslem, D. F. R. P., Diggle, P. J., Gunatilleke, C. V. S., et al. (2011). Spatial patterns reveal negative density dependence and habitat associations in tropical trees. *Ecology* 92, 1723–1729. doi: 10.1890/11-0335.1

Bliss, C. I., and Fisher, R. A. (1953). Fitting the negative binomial distribution to biological data. *Biometrics* 9, 176–200. doi: 10.2307/3001850

Brown, C., Illian, J. B., and Burslem, D. F. R. P. (2016). Success of spatial statistics in determining underlying process in simulated plant communities. *J. Ecol.* 104, 160–172. doi: 10.1111/1365-2745.12493

Brown, C., Law, R., Illian, J. B., and Burslem, D. F. R. P. (2011). Linking ecological processes with spatial and non-spatial patterns in plant communities. *J. Ecol.* 99, 1402–1414. doi: 10.1111/j.1365-2745.2011.01877.x

Carson, W. P., Anderson, J., Leigh, E., and Schnitzer, S. A. (2008). "Challenges associated with testing and falsifying the Janzen-Connell hypothesis: a review and critique," in *Tropical Forest Community Ecology*, eds W. P. Carson and S. A. Schnitzer (Oxford: Wiley-Blackwell), 210–241.

Coleman, B. D. (1981). On random placement and species-area relations. *Math. Biosci.* 54, 191–215. doi: 10.1016/0025-5564(81)90086-9

Comita, L. S., Queenborough, S. A., Murphy, S. J., Eck, J. L., Xu, K., Krishnadas, M., et al. (2014). Testing predictions of the Janzen-Connell hypothesis: a meta-analysis of experimental evidence for distance- and density-dependent seed and seedling survival. *J. Ecol.* 102, 845–856. doi: 10.1111/1365-2745.12232

Condit, R. (1998). *Tropical Forest Census Plots: Methods and Results from Barro Colorado Island, Panama and a Comparison with Other Plots*. Environmental Intelligence Unit. Springer-Verlag, Berlin; Heidelberg.

Condit, R., Ashton, P. S., Baker, P., Bunyavejchewin, S., Gunatilleke, S., Gunatilleke, N., et al. (2000). Spatial patterns in the distribution of tropical tree species. *Science* 288, 1414–1418. doi: 10.1126/science.288.5470.1414

Condit, R., Perez, R., Aguilar, S., Lao, S., Foster, R., and Hubbell, S. (2019). Complete data from the Barro Colorado 50-ha plot: 423617 trees, 35 years. *Dryad Digital Repository*. doi: 10.15146/5xcp-0d46

Conlisk, E., Bloxham, M., Conlisk, J., Enquist, B., and Harte, J. (2007). A new class of models of spatial distribution. *Ecol. Monogr*. 77, 269–284. doi: 10.1890/06-0122

Conlisk, J., Conlisk, E., Kassim, A. R., Billick, I., and Harte, J. (2012). The shape of a species' spatial abundance distribution. *Glob. Ecol. Biogeogr*. 21, 1167–1178. doi: 10.1111/j.1466-8238.2011.00755.x

Connell, J. (1971). "On the role of natural enemies in preventing competitive exclusion in some marine animals and in rain forest trees," in *Dynamics of populations: Proceedings of the Advanced Study Institute on Dynamics of Numbers in Populations*, P. J. D. Boer and G. R. Gradwell (Wageningen), 298–312.

Diggle, P. (2014). *Statistical Analysis of Spatial and Spatio-Temporal Point Patterns, 3rd Edn*. Boca Raton, FL: CRC Press, Taylor & Francis Group.

Drakare, S., Lennon, J. J., and Hillebrand, H. (2006). The imprint of the geographical, evolutionary and ecological context on species-area relationships. *Ecol. Lett*. 9, 215–227. doi: 10.1111/j.1461-0248.2005.00848.x

Green, J., Harte, J., and Ostling, A. (2019). Data from: species richness, endemism, and abundance patterns: tests of two fractal models in a serpentine grassland. *Dryad Digital Repository*. doi: 10.6078/D1MQ2V

Green, J. L., Harte, J., and Ostling, A. (2003). Species richness, endemism, and abundance patterns: tests of two fractal models in a serpentine grassland. *Ecol. Lett*. 6, 919–928. doi: 10.1046/j.1461-0248.2003.00519.x

Green, J. L., and Plotkin, J. B. (2007). A statistical theory for sampling species abundances. *Ecol. Lett*. 10, 1037–1045. doi: 10.1111/j.1461-0248.2007.01101.x

Haegeman, B., Etienne, R., Rossberg, A. E. A. G., and McPeek, E. M. A. (2010). Entropy maximization and the spatial distribution of species. *Am. Nat*. 175, E74–E90. doi: 10.1086/650718

Harms, K. E., Wright, S. J., Calderon, O., Hernandez, A., and Herre, E. A. (2000). Pervasive density-dependent recruitment enhances seedling diversity in a tropical forest. *Nature* 404, 493–495. doi: 10.1038/35006630

Harte, J. (2011). *Maximum Entropy and Ecology: A Theory of Abundance, Distribution, and Energetics*. Oxford: Oxford University Press. doi: 10.1093/acprof:oso/9780199593415.001.0001

Harte, J., and Kitzes, J. (2015). Inferring regional-scale species diversity from small-plot censuses. *PLoS ONE* 10:e0117527. doi: 10.1371/journal.pone.0117527

Harte, J., and Newman, E. A. (2014). Maximum information entropy: a foundation for ecological theory. *Trends Ecol. Evol*. 29, 384–389. doi: 10.1016/j.tree.2014.04.009

Harte, J., Newman, E. A., and Rominger, A. J. (2017). Metabolic partitioning across individuals in ecological communities. *Glob. Ecol. Biogeogr*. 1, 993–997. doi: 10.1111/geb.12621

Harte, J., Umemura, K., and Brush, M. (2021). DynaMETE: a hybrid MaxEnt-plus-mechanism theory of dynamic macroecology. *Ecol. Lett*. 24, 935–949. doi: 10.1111/ele.13714

Harte, J., Zillio, T., Conlisk, E., and Smith, A. B. (2008). Maximum entropy and the state-variable approach to macroecology. *Ecology* 89, 2700–2711. doi: 10.1890/07-1369.1

Harvey, B. J., Holzman, B. A., and Forrestel, A. B. (2014). Forest resilience following severe wildfire in a semi-urban national park. *Fremontia* 42, 14–18. Available online at: https://cnps.org/wp-content/uploads/2018/03/FremontiaV42.3.pdf

He, F., and Gaston, K. (2003). Occupancy, spatial variance, and the abundance of species. *Am. Nat*. 162, 366–375. doi: 10.1086/377190

He, F., and Gaston, K. J. (2000). Estimating species abundance from occurrence. *Am. Nat*. 156, 553–559. doi: 10.1086/303403

Hubbell, S., Condit, R., and Foster, R. (2005). *Forest Census Plot on Barro Colorado Island*. Available online at: http://ctfs.si.edu/webatlas/datasets/bci/

Hubbell, S. P., Foster, R. B., O'Brien, S. T., Harms, K. E., Condit, R., Wechsler, B., et al. (1999). Light-gap disturbances, recruitment limitation, and tree diversity in a neotropical forest. *Science* 283, 554–557. doi: 10.1126/science.283.5401.554

Janzen, D. H. (1970). Herbivores and the number of tree species in tropical forests. *Am. Nat*. 104, 501–528. doi: 10.1086/282687

Kitzes, J. (2019). Evidence for power-law scaling in species aggregation. *Ecography* 42, 1224–1225. doi: 10.1111/ecog.04159

Kitzes, J., and Shirley, R. (2016). Estimating biodiversity impacts without field surveys: a case study in northern Borneo. *Ambio* 45, 110–119. doi: 10.1007/s13280-015-0683-3

Law, R., Illian, J., Burslem, D. F. R. P., Gratzer, G., Gunatilleke, C. V. S., Gunatilleke, I., et al. (2009). Ecological information from spatial patterns of plants: insights from point process theory. *J. Ecol*. 97, 616–628. doi: 10.1111/j.1365-2745.2009.01510.x

Leibold, M. A., and Chase, J. M. (2018). *Metacommunity Ecology*. Princeton, NJ: Princeton University Press.

Levin, S. A. (1992). The problem of pattern and scale in ecology: the Robert H. MacArthur Award Lecture. *Ecology* 73, 1943–1967. doi: 10.2307/1941447

McGlinn, D. J., Xiao, X., Kitzes, J., and White, E. P. (2015). Exploring the spatially explicit predictions of the maximum entropy theory of ecology. *Glob. Ecol. Biogeogr*. 24, 675–684. doi: 10.1111/geb.12295

Münkemüller, T., Gallien, L., Pollock, L. J., Barros, C., Carboni, M., Chalmandrier, L., et al. (2020). Dos and don'ts when inferring assembly rules from diversity patterns. *Glob. Ecol. Biogeogr*. 29, 1212–1229. doi: 10.1111/geb.13098

Newman, E. A., Wilber, M. Q., Kopper, K. E., Moritz, M. A., Falk, D. A., McKenzie, D., et al. (2020). Disturbance macroecology: a comparative study of community structure metrics in a high-severity disturbance regime. *Ecosphere* 11:e03022. doi: 10.1002/ecs2.3022

Ostling, A., Harte, J., Green, J., and Kinzig, A. (2004). Self-similarity, the power law form of the species-area relationship, and a probability rule: a reply to maddux. *Am. Nat*. 163, 627–633. doi: 10.1086/382663

Plotkin, J. B., Potts, M. D., Leslie, N., Manokaran, N., Lafrankie, J., and Ashton, P. S. (2000). Species-area curves, spatial aggregation, and habitat specialization in tropical forests. *J. Theoret. Biol*. 207, 81–99. doi: 10.1006/jtbi.2000.2158

Rosenzweig, M. L. (1995). *Species Diversity in Space and Time*. Cambridge: Cambridge University Press. doi: 10.1017/CBO9780511623387

Thomas, C. D., Cameron, A., Green, R. E., Bakkenes, M., Beaumont, L. J., Collingham, Y. C., et al. (2004). Extinction risk from climate change. *Nature* 427, 145–148. doi: 10.1038/nature02121

Vellend, M. (2016). *The Theory of Ecological Communities*. Princeton, NJ: Princeton University Press.

Volkov, I., Banavar, J. R., He, F., Hubbell, S. P., and Maritan, A. (2005). Density dependence explains tree species abundance and diversity in tropical forests. *Nature* 438, 658–661. doi: 10.1038/nature04030

White, E. P., Thibault, K. M., and Xiao, X. (2012). Characterizing species abundance distributions across taxa and ecosystems using a simple maximum entropy model. *Ecology* 93, 1772–1778. doi: 10.1890/11-2177.1

Wiegand, T., He, F., and Hubbell, S. P. (2013). A systematic comparison of summary characteristics for quantifying point patterns in ecology. *Ecography* 36, 92–103. doi: 10.1111/j.1600-0587.2012.07361.x

Wiegand, T., and Moloney, K. A. (2013). *Handbook of Spatial Point-Pattern Analysis in Ecology*. Boca Raton, FL: CRC Press, Taylor & Francis Group. doi: 10.1201/b16195

Wilber, M. Q., Kitzes, J., and Harte, J. (2015). Scale collapse and the emergence of the power law species-area relationship. *Glob. Ecol. Biogeogr*. 24, 883–895. doi: 10.1111/geb.12309

Xiao, X., Mcglinn, D. J., and White, E. P. (2015). A strong test of the maximum entropy theory of ecology. *Am. Nat*. 185, E70–E80. doi: 10.1086/679576

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Ecology and Evolution

ORIGINAL RESEARCH
published: 02 August 2021
doi: 10.3389/fevo.2021.644328

Biogeographic Drivers of Evolutionary Radiations

*Ran Tao[1], Lawren Sack[2] and James Rosindell[1]**

[1] Department of Life Sciences, Imperial College London, Ascot, United Kingdom, [2] Department of Ecology and Evolutionary Biology, University of California, Los Angeles, Los Angeles, CA, United States

Some lineages radiate spectacularly when colonizing a region, but others do not. Large radiations are often attributed to species' adaptation into niches, or to other drivers, such as biogeography including dispersal ability and spatial structure of the landscape. Here we aim to disentangle the factors determining radiation size, by modeling simplified scenarios without the complexity of explicit niches. We build a spatially structured neutral model free from niches and incorporating a form of protracted speciation that accounts for gene flow between populations. We find that a wide range of radiation sizes are possible in this model depending on the combination of geographic isolation and species' dispersal ability. At extremely low rates of dispersal between patches, each patch maintains its own endemic species. Intermediate dispersal rates foster larger radiations as they allow occasional movement between patches whilst sufficiently restricting gene flow to support further speciation in allopatry. As dispersal rates increase further, a critical point is reached at which demographically identical lineages may vary greatly in radiation size due to rare and stochastic dispersal events. At the critical point in dispersal frequency, some lineages remain a single species for a comparatively long time, whilst others with identical characteristics produce the largest radiations of all via a new mechanism for rapid radiation that we term a 'radiation cascade'. Given a single species covering many patches connected with gene flow, a radiation cascade is triggered when stochastic dispersal is unusually low for a period, leading to an initial speciation event. This speciation means there are fewer individuals per species and thus further reduced gene flow between conspecifics. Reduced gene flow in turn makes it easier for further speciation to occur. During a radiation cascade, dispersal of individuals between patches continues at the same rate as before, but due to the increasing diversity it primarily introduces novel species that will later speciate, rather than adding to gene flow of existing species. Once a radiation cascade begins, it continues rapidly until it is arrested by a new equilibrium between speciation and extinction. We speculate that such radiation cascades may occur more generally and are not only present in neutral models. This process may help to explain rapid radiation, and the extreme radiation sizes of certain lineages with dispersing ancestors. Whilst niches no doubt play a role in community assembly, our findings lead us to question whether diversification and adaptation into niches is sometimes an effect of speciation and rapid radiation, rather than its cause.

Keywords: neutral theory, radiation, adaptive radiation, rapid burst, island biogeography, niche, dispersal limitation, gene flow

OPEN ACCESS

Edited by:
John Maxwell Halley, University of Ioannina, Greece

Reviewed by:
Kostas Kougioumoutzis, National and Kapodistrian University of Athens, Greece
Paulo A. V. Borges, University of the Azores, Portugal

****Correspondence:***
James Rosindell
j.rosindell@imperial.ac.uk

Specialty section:
This article was submitted to Models in Ecology and Evolution, a section of the journal Frontiers in Ecology and Evolution

Received: *20 December 2020*
Accepted: *24 June 2021*
Published: *02 August 2021*

Citation:
Tao R, Sack L and Rosindell J (2021) Biogeographic Drivers of Evolutionary Radiations. Front. Ecol. Evol. 9:644328. doi: 10.3389/fevo.2021.644328

INTRODUCTION

Radiations are an evolutionary phenomenon in which one species proliferates into many (Erwin, 1992; Stroud and Losos, 2016). Radiations are often studied on island and archipelago settings as their spatial structure provides prominent opportunities for allopatric speciation, with celebrated examples including Darwin's finches of the Galapagos (Grant and Grant, 2011, 2014), anolis lizards of the Caribbean (Losos, 2011), lobeliads and silverswords of Hawaii (Robichaux et al., 1990; Givnish et al., 2009) and others (Lovette et al., 2002; Kocher, 2004). Since Darwin's early work, the mechanism behind radiation has captured the interest of a wide range of ecologists and evolutionary biologists. Most studies have focused on adaptive radiations (Harvey and Rambaut, 2000; Yoder et al., 2010; Wellborn and Langerhans, 2015), in which the radiation was accompanied by diversification of phenotypes enabling the species to fill different niches (Schluter, 2000). Geographic radiations, in which spatially structured landscapes are the main driver of species proliferation, have been relatively overlooked (Ibanez et al., 2018). A combination of adaptation and geography may explain the overall pattern of species proliferation in most cases (Rundell and Price, 2009; Simões et al., 2016).

Adaptive radiation is explained by ecological opportunity (Simpson, 1953; Losos, 2010; Losos et al., 2010), but some phenomena cannot be explained in this way. For example, in the Galapagos, Darwin's finches proliferated into fourteen species (Lamichhaney et al., 2015; Grant, 2017), while mockingbirds only four species (Arbogast et al., 2006) and most other bird lineages remained as single taxa. For the Hawaiian flora, a few lineages have large species richness, while many colonist lineages did not radiate at all (**Figure 1**). What explains such disparity in radiation size? If ecological opportunity is the important factor, what prevented such opportunities from being filled more evenly by the various immigrating lineages? The answer might be partly to do with the order or timing of immigration (Fukami, 2015; De Meester et al., 2016) and/or the influences of geography and stochasticity on dispersal. Indeed, propagule size, dispersal mechanism and breeding system can interact with geography and the extent of natural barriers between islands and habitat patches to influence radiation size in plants (Price and Wagner, 2018).

Recent research has revealed that, in many cases, geographic and ecological factors both influence radiation size (Bennett and O'Grady, 2013; Yu et al., 2014; Simões et al., 2016; Schenk and Steppan, 2018). Some have proposed that a diverse radiated lineage might undergo both adaptive and geographic radiation (Simões et al., 2016). However, such hypotheses are hard to test. We suggest that stochasticity and geography should not simply be a 'backup' explanation invoked only if an adaptive story cannot be found, as they are likely to work alongside adaptation in a much greater proportion of cases. One approach to disentangle the role of adaptation is by modeling radiation in the absence of niches or adaptation, to explore the effects of stochasticity, dispersal and geography on radiation size. Neutral theory (Hubbell, 2001) provides a natural direction for building such a model because it does not include specific niches and instead emphasizes biogeographic factors such as dispersal and landscape spatial structure (Rosindell et al., 2011).

We investigated the quantitative influence of biogeographic drivers on radiation size with a mechanistic model based on neutral theory (Hubbell, 2001). Our model integrates both spatial structure and a protracted speciation process that explicitly accounts for gene flow between multiple habitat patches (Rosindell et al., 2010; Rosindell and Phillimore, 2011; Gascuel et al., 2016). We found that biogeography alone can drive a very large variation in radiation size. We also find a potential new mechanism for rapid radiation: a 'radiation cascade' in which initial speciation of a widespread species reduces gene flow between conspecifics and as a result leads to yet more speciation in a reinforcing cycle.

MATERIALS AND METHODS

We built a neutral model with spatial structure to explicitly simulate the proliferation of one or more radiating lineages within a network of connected habitat patches, which could represent islands. We conducted experiments simulating different sets of parameters to identify the relationship between radiation size and biogeographic factors.

Model Description

Our neutral model is based on birth-death cycles, where the total number of individuals is constant (zero-sum). Unlike the classic neutral model of Hubbell (2001), which contains one local community and a single metacommunity, our model uses a spatially structured community network containing multiple patches linked together (Economo and Keitt, 2008; Gascuel et al., 2016) as well as a remote metacommunity (Hubbell, 2001). The community network can receive external immigrants from the metacommunity. Within the network, local patches represent local communities that may be connected to a greater or lesser extent. There are internal dispersal events between local patches. An isolated archipelago of islands is one example of a system that can naturally be represented in this way. The assumption of neutral theory, and our model, is that an individual's chances of birth, death and movement are independent of its species identity. This neutrality assumption aids understanding of the ecological processes other than niches and selection, with the constraint that species are from a single trophic level and guild (Hubbell, 2001).

In our simulation, one individual dies at each time step, leaving a gap that is filled by the new-born offspring of either a local parent, or an immigrant from another local patch or an external immigrant from the metacommunity (**Figure 2**). The network of dispersal opportunities between every pair of local patches is represented by a probability-matrix P. In this matrix, $P_{i,j}$ gives the probability for the gap in patch i to be filled by an individual from patch j. $P_{i,i}$ gives the probability of a deceased individual in patch i being replaced by the offspring of a local individual. In practice we simulated a simplified scenario in which all local patches were of the same size J, in terms of total number

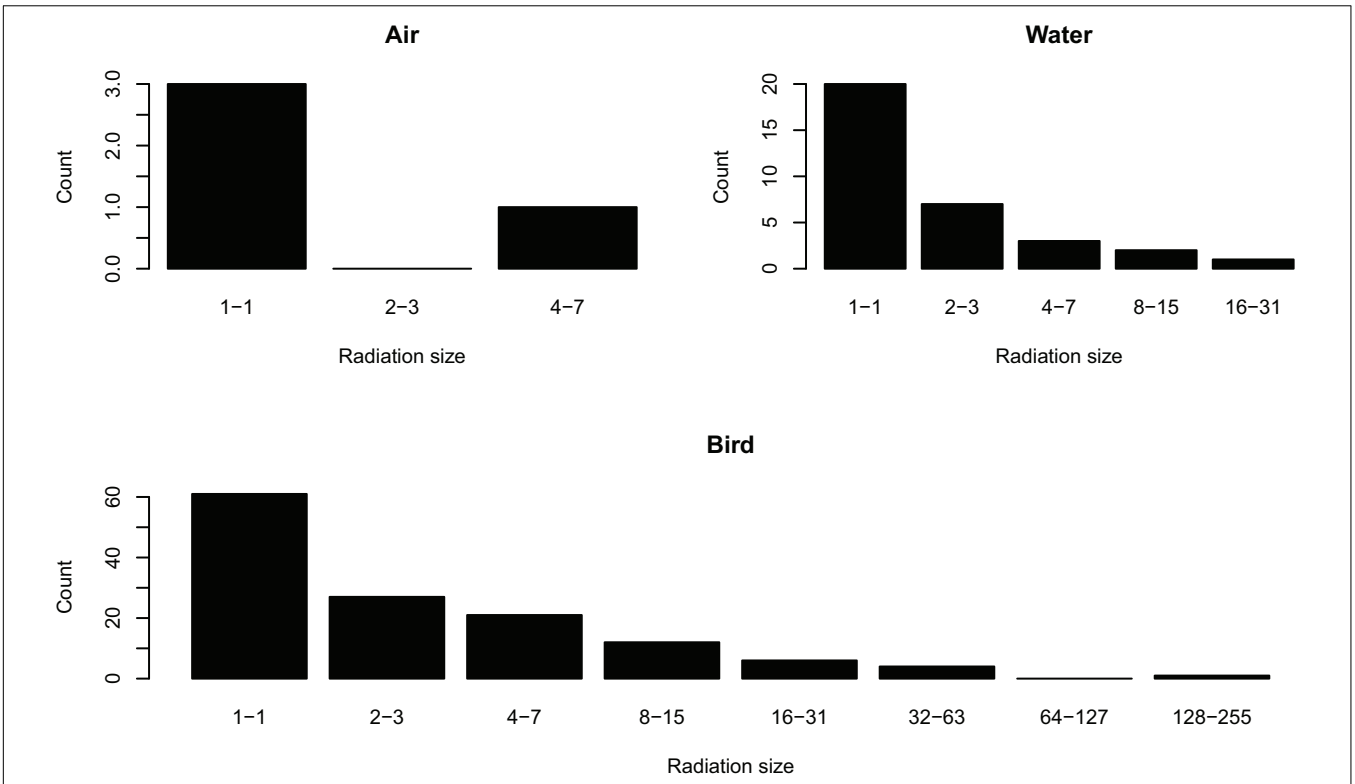


FIGURE 1 | Radiation sizes of Hawaiian angiosperm lineages, separated by dispersal vector. The data are from Price and Wagner (2018). Lineages dispersed by air never reached radiation sizes beyond the 4–7 category. Those dispersing over water were able to reach radiation sizes in the 16–32 category. All the lineages with the greatest radiation sizes of 32 or more were dispersed by birds.

of individuals, and, where present, any links between patches were of the same strength given by m. As a result $P_{i,j}$, where $i \neq j$, is either 0 or m and $P_{i,i}$ is set so that $\sum_{\mathrm{j}=1}^{\mathrm{N}} P_{\mathrm{i.j}} = 1 - m$. Here, N is the total number of local patches in the system and M is the probability of dispersal from the external metacommunity. To simplify the simulation further and improve tractability, species abundances within the metacommunity resource pool were not modeled explicitly; any individual arriving from the metacommunity was considered a different species, which would be reasonable for a large, diverse, and distant metacommunity.

Speciation Process

We modeled speciation as a protracted process (Rosindell et al., 2010) that takes time to complete. Each possible new species has a countdown 'clock' (measured in generations) until its speciation process has completed. Given that generations are overlapping in our model, we defined one generation as the amount of time required for half of the individuals in the community to have died and been replaced with offspring of other individuals (Rosindell et al., 2010). Following the protracted speciation model with gene flow of Rosindell and Phillimore (2011), any new colonization of a local patch will start the countdown to a potential speciation event that would result in the newly colonized local patch becoming a new species (**Figure 2**). The shortest possible duration of speciation is given by a parameter (τ), the number of generations taken for allopatric speciation in the absence of any gene flow between local patches. Any immigration event of conspecific individuals between different local patches increases the amount of time until speciation will complete for that species by a predefined parameter G giving the gene flow effect. This increases the amount of time until the populations within the local patches concerned can be considered as two different species (**Figure 2**). Here, we expand the two-community (island mainland) model of Rosindell and Phillimore (2011) to allow arbitrary networks of local patches. In cases where there are more than two patches, every possible pair of patches has its own independent speciation clock for each species, keeping track of gene flow events between those two patches. Individuals in a local patch (or collection of local patches) speciate in allopatry from their former conspecifics in other local patches when they can be split into two groups so that the speciation clock for every local patch in one group has counted down to zero (or less) in terms of connection to every local patch in the other group (**Figure 3**). If the speciation clock for two local patches has counted down to zero but they are still linked through other local patches with speciation clocks that have not yet counted down to zero, speciation does not yet complete and gene flow between disconnected links may resume if new dispersal events take place. We simulated our model forwards in time rather than using a backward in time coalescence approach (Rosindell et al., 2008;

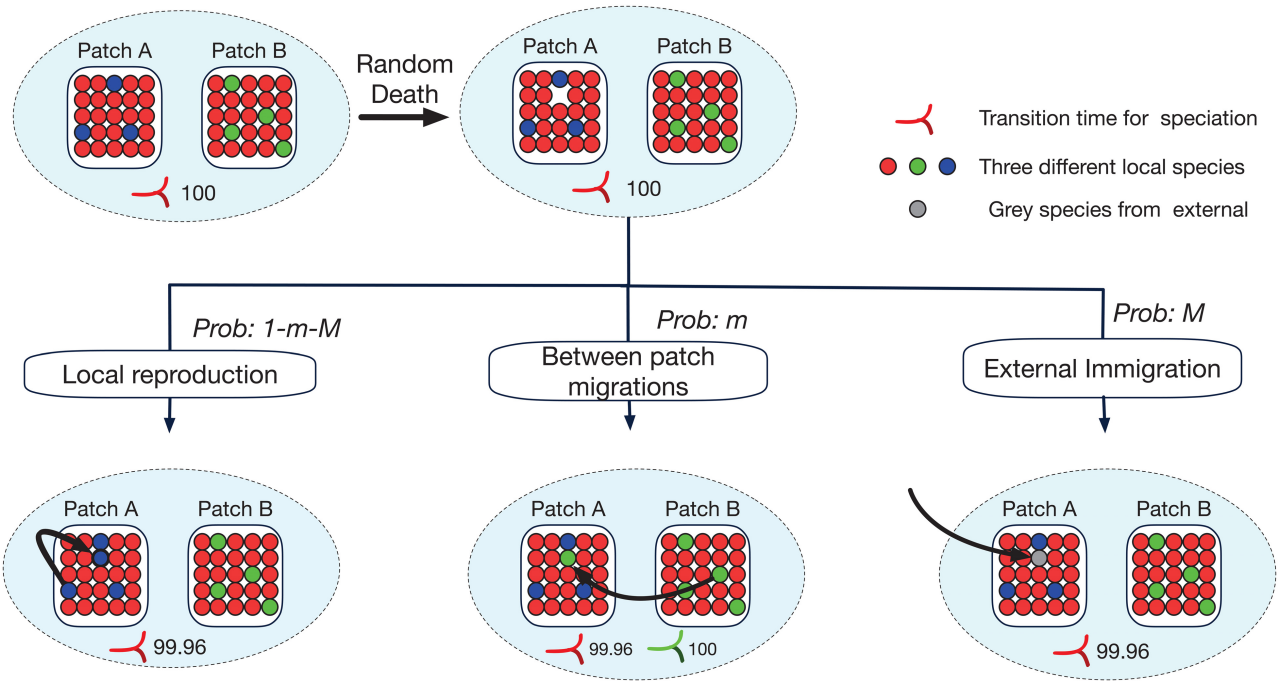


FIGURE 2 | Schematic representation of the network structure used in our study. In this example, we show three initial species (red, blue, and green) in the community network of two patches (A and B). There are 50 individuals in total. The time step shown represents a birth-death cycle: that is 2/50 = 0.04 of a generation. The minimum duration of speciation (τ) is set at 100 generations meaning that in the absence of gene flow speciation will complete in 100 generations. The gene flow effect (G) is 5 in this example meaning that every individual gene flow event causes a 5-generation delay to the conclusion of the speciation process. The numbers at the bottom of each panel are the speciation 'clocks' for each species showing how many generations without gene flow are required to completely sever that species into two. Only when a species has conspecific populations in different patches is a speciation 'clock' required. Any 'clock' is initialized with τ generations (here 100), and will then count down over time but increase by G whenever a dispersal event of a conspecific occurs (gene flow).

Thompson et al., 2020). This is because a model with multiple habitat patches and protracted speciation with gene flow can only be approximated with coalescence (Rosindell and Phillimore, 2011; Gascuel et al., 2016).

Analytical and Scaling Results

We expect that a critical point in the system will be reached when m, the rate of dispersal between any pair of local patches, passes $m_{\text{crit}} = \frac{1}{G \times J}$. Here, m_{crit} represents the dispersal rate where, in a community with only one species on all islands, gene flow events add to the speciation clock at precisely the same rate that time erodes it. G is the gene flow effect and J represents the population size of each local patch. Gene flow between a pair of patches can come from dispersal in either direction, so there is a pool of $2 \times J$ individuals who could disperse with probability m_{crit} per reproduction event. However, given the way a generation is defined above, only half of those of $2 \times J$ individuals will reproduce per generation. The total effect of gene flow per generation will therefore be given by $m_{\text{crit}} \times G \times J$ and setting this to one gives our solution for m_{crit}. We will compare the critical point m_{crit} against our simulated results for radiation size across a wide range of values for dispersal rate m.

We also wish to test the idea that increasing the number of links in a network of local patches has a similar effect on radiation size as increasing the strength (amount of immigration) of each link. To do this we plot of graph against the compound parameter (immigration rate × number of links) for a range of different spatial network structures to observe if the structure itself has any effect beyond its raw number of links.

Experimental Design

We designed a series of experiments to investigate different biogeographic drivers of radiation patterns in a network of local patches. We investigated three main spatial structures for the network of dispersal between local patches: 'line' structure, 'fully-connected' structure, and 'random' structure. The 'Line' structure represents a chain of local patches connected only to their direct neighbors. The 'Fully-connected' structure represents a network where every local patch is connected to every other local patch with equal strength. The 'Random' structure is a network of local patches with random connections between them starting with a line structure (the minimum number of links to connect the network) and choosing a given number of additional links to add at random, up to a maximum corresponding to a fully connected network. See **Table 1** for a complete list of all model parameters.

To identify the relationship between habitat size and radiation, we varied the number of local patches as well as the number of individuals in each local patch. We tested a wide range of internal dispersal rates and rates of dispersal from the metacommunity. Our simulations were conducted in groups,

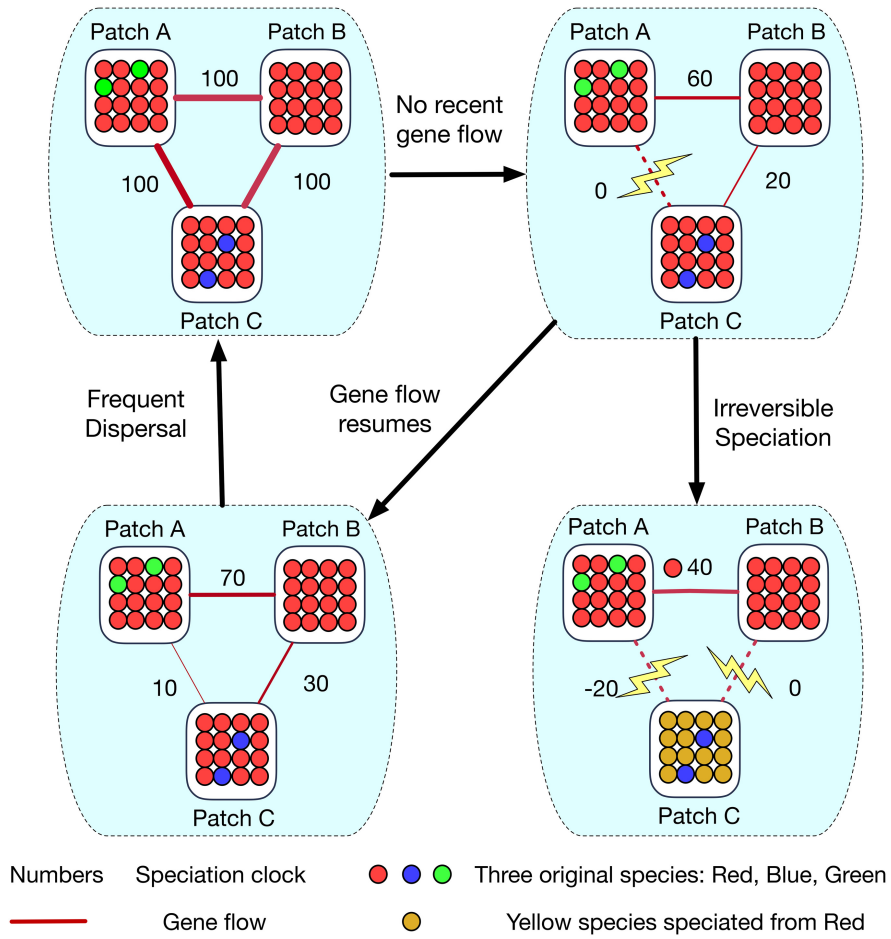


FIGURE 3 | A schematic diagram showing how allopatric speciation (cladogenesis) happens in a community network of patches. Assume there are 3 patches (A, B, and C) fully connected to each other. In this example, the red species has spread to all 3 patches, so gene flow occurs between all these patches and links them as a gene flow network. The gene flow strength is shown by the width of the red line. The 'clock' for speciation is initially equal to $\tau = 100$. When the conspecific link between red species population in patch A and C has been eroded by lack of gene flow, these populations remain conspecific because they are connected via patch B: red individuals in patch A and B are conspecific and the same is true for red individuals in patch C and B - so patch A and C must also remain conspecific. Irreversible speciation finally occurs when the conspecific link between patch C and B is cut, and the conspecific link between patch A and C remains unconnected so that the red population in patch C can now speciate from its relatives in patches A and B to from a new yellow species by allopatric speciation. In the scenario where gene flow resumes, the conspecific link between patch A and C has been revived, so the red species in all three local patches is back to being a complete gene flow network.

each with a distinct purpose (**Table 2**). After test simulations, the first groups (1–3) used a single immigration event from the metacommunity, a simplification to focus on the effects of dispersal between patches, network structure and gene flow effect. This is equivalent to assuming $M \rightarrow 0$ where M represents the rate of immigration from the metacommunity and not to be confused with m representing the rate of immigration between connected local community patches. Simulation group 1 investigated how radiation size varied with changes of all simulation parameters except for M. Simulation group 2 then focused on parameters values where the radiation cascade was observed and replicated simulations of those one hundred times with different random seeds in order to obtain a more complete picture of the radiation cascades. Simulation group 3 varied the number of links within the network as well as the internal dispersal strength to capture their co-effect on radiation size. Finally, group 4 simulations incorporated external immigration from the metacommunity ($M > 0$) to observe the radiation patterns resulting from multiple immigration events.

To ensure simulation groups 1–3 reached equilibrium, we discarded the first half of all simulated data as a burn in. To show this was sufficient we also produced a second set of results that retained only the last quarter of the simulation. The simulations in group 4, where $M > 0$ detected equilibrium as the time when the island was populated entirely with lineages that immigrated onto the island during the simulation. Thus, at equilibrium, the individuals that formed the initial state on the island had all died out leaving no descendants and no meaningful way to influence the outcome (Rosindell and Harmon, 2013).

TABLE 1 | A list of all parameters used in our simulations and their meaning.

Parameter	Explanation
N	Number of local patches in the network. We used 4 or 8.
S	The spatial network structure of dispersal between local patches used to build the dispersal matrix *P*. We investigated 'Line' - a chain of local patches with *N*-1 links, 'Star' - one patch at center connects with all others, with *N*-1 links, 'Fully-connected' - every patch connects to every other, with *0.5N(N-1)* links and 'Random' with no fixed structure and number of links in the range between *N-1* and *0.5N(N-1)*.
J	Number of individuals in each local patch. We used 500 or 1000.
m	Internal dispersal rate within network of local patches. We used a range between 10^{-7} and 10^{-2}.
M	External dispersal rate from metacommunity. We used a range between 10^{-7} and 10^{-2}.
G	Gene flow effect. We used 0.1 or 5.
τ	Minimum duration of protracted speciation in the absence of gene flow. We used 100 generations.

TABLE 2 | Parameters used for all groups of simulations.

Simulation group	Parameter range					
	S	***N***	***J***	***m***	***M***	***G***
Test	F	4	1000	$10^{-3.5}$	$10^{-4.0}$, $10^{-4.5}$	0.1
1	F, L	4, 8	500, 1000	Range $10^{-7} \leq m \leq 10^{-2}$	M→0	0.1,5
2	F, L	4	1000	$10^{-2.0}$	M→0	0.1
3	F, L, St, R	8	1000	Range $10^{-7} \leq m \leq 10^{-2}$	M→0	0.1
4	F	4	1000	$10^{-3.5}$	Range $10^{-7} \leq M \leq 10^{-2}$	0.1

*Refer also to **Table 1** for parameter meanings. There are four kinds of structure (S): Fully-connected (F), Line (L), Star (St), and Random (R). In group 3, we studied how spatial structure affected radiation size through the different dispersal links. For the line and star structure, there were 7 dispersal links were 7; for the fully-connected structure, there were 28 dispersal links; we set 5 different random structures with 10, 13, 17, 21, and 25 dispersal links. In terms of the value of internal dispersal rate (m) 'Range' refers to (10^{-2}, $10^{-2.2}$,…,$10^{-4.8}$, 10^{-5}, 10^{-6}, 10^{-7}), and the equivalent 'Range' of M refers to ($10^{-3.5}$, $10^{-3.6}$,…,$10^{-5.3}$, $10^{-5.4}$). Each experiment had 10 repeated simulations for every set of parameters, except for group 2 which had 100 repeats. In groups 1–3, we assumed that only one lineage of the guild under consideration occupies the network of patches, and after the first colonization, there will be no further immigration, this corresponds to the limit of M tends to 0. τ was constant at 100 throughout.*

RESULTS

In the test simulations, the radiation size produced by our model fluctuated over time and included a mixture of both single patch endemics and multi-patch species (**Figure 4**). It is normal in this model to observe extinction of entire lineages and recovery of nearly extinct lineages to radiate again (**Figure 4**). For the simulations in groups 1–3 with only one immigrating lineage, high immigration between local patches resulted in only one species across all simulations. In contrast, if immigration between local patches was extremely low each local patch would hold its own locally endemic species, so radiation size equals the number of local patches. The largest radiation size occurred when between-patch dispersal was intermediate (**Figures 5**, **6**). These results were broadly robust to the length of burn in period (see **Supplementary Material**). The only exception was at the critical point where a limited number of additional lineages underwent radiation cascades after the longer burn in (see **Supplementary Material**).

At a critical point where the immigration between local patches was high, but not yet high enough to homogenize the patches and result in a single species, the radiation size was either 1 or very high (e.g., over 40, **Figure 5**). At this critical point in immigration between patches, lineages could radiate dramatically at any time, but after this would be stable in a species rich configuration (**Figure 7**). The critical point was comparable to the one predicted from our proposed formula of $m = \frac{1}{G \times J}$ (**Figure 6**).

We found that fragmentation and community size had positive effects on radiation size. A network with eight patches showed approximately twice the radiation size of a network with four patches. The Radiation size also increased in response to larger local patches without changing the number of patches. When gene flow effect was high (5), the fully-connected network showed a larger radiation than the line-structured one. When the gene flow effect was low (0.1) and immigration rate between local patches was high, the line-structured network had the larger radiations (**Figure 6**). We found that when gene flow effect was low (0.1) and internal immigration rate was also relatively low (how low depends on the number of links), the radiation size was not regulated by internal dispersal independently but rather by the combined effect of internal dispersal rate multiplied by number of dispersal links between local patches (**Figure 8**). These results were robust to the length of burn in period (see **Supplementary Material**).

For the simulations with a metacommunity, the radiation size for multiple immigrants varied from no radiation to a radiation of size 24 with our parameters. There was a relatively narrow range between lower quartile and upper quartile, but a much larger range of extremes was possible, especially at lower rates of external immigration (**Figure 9**).

DISCUSSION

In our model, dispersal within the community of patches was a key determinant of radiation size. Internal dispersal could impede radiation by maintaining gene flow or promote radiation by providing opportunities for small populations to jump across barriers and later go on to speciate. An intermediate rate of internal dispersal therefore led to the largest radiations, consistent with the earlier findings from a simpler neutral model (Rosindell and Phillimore, 2011) as well as the intermediate dispersal model (Claramunt et al., 2012; Agnarsson et al., 2014), which predicts that taxa with intermediate dispersal ability diversify more readily. Our analyses extended these previous results with a fully mechanistic model that explicitly allows for multiple islands or patches in the community. We found

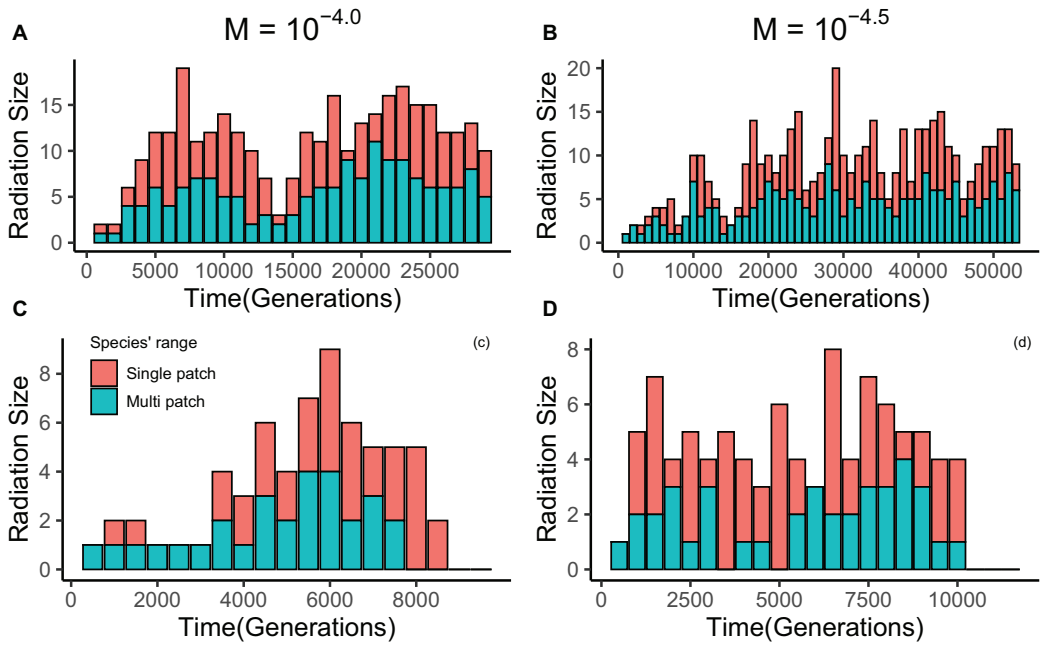


FIGURE 4 | Examples of lineage radiation size through time for 4 well-radiated lineages from two simulations. From each simulation, we chose one lineage that still exists at the end of the simulation and one lineage that has gone extinct. The lineages in panels **(A,C)** are from one simulation, the lineages in panels **(B,D)** from another simulation. Data in panels **(A,B)** were collected every 1000 generations, while the interval in panels **(C,D)** was 500 generations. The external immigration rate in panels **(A,C)** is 10^{-4}, and in panels **(B,D)** is $10^{-4.5}$. The features of the community simulated were: 4 local patches, each with 1000 individuals; fully-connected spatial structure; internal dispersal rate is $10^{-3.5}$; gene flow effect (G) is 0.1. The red portion of the stacked bar represents the species that are only endemic to a single local patch; the portion represents endemic species that have spread onto multiple patches. This figure illustrates that, for our model, fluctuations in radiation size are normal, as is extinction of entire lineages and recovery of nearly extinct lineages to radiate again.

that under these conditions the largest radiations occur at intermediate levels of dispersal. We also found, however, that the largest radiations arise due to conditions at singular critical points, at which lineages may radiate dramatically, or remain as one species for a relatively long period of time. The radiation size of a single lineage naturally fluctuates over time and may either go extinct or recover from a nearly extinct state to radiate again (**Figure 4**). These fluctuations appear reminiscent of the taxon cycle (Ricklefs and Bermingham, 2002), though they are driven by different mechanisms.

In larger and more complex communities the connectivity between any two local communities becomes a complex property determined by multiple factors, including both direct dispersal between the communities and indirect dispersal between them using other communities as stepping-stones. The largest radiations occurred when internal dispersal rates were at a critical point. In systems at this critical point, a period of slightly fewer than normal stochastic dispersal events to a patch can result in a very large radiation, but otherwise the same system may simply remain stable for a long while with no radiation at all. When radiation does occur in such cases, it is very rapid and forms a feedback loop that we call a 'radiation cascade'. The mechanism behind a radiation cascade is that once the first allopatric speciation has occurred, population sizes of the resulting daughter species are smaller. Gene flow between conspecific populations on different patches is the product of number of individuals and per individual dispersal rate. Consequently, whilst dispersal rate has not changed, the reduction in number of individuals leads to less gene flow for each species. A decrease in gene flow in turn promotes further speciation and a positive feedback cycle is formed (**Figure 10**). The feedback is finally broken when species decline to abundances low enough to promote extinction and thus speciation is balanced by extinction with a new dynamic equilibrium (MacArthur and Wilson, 1967; Hubbell, 2001).

We propose that radiation cascades might explain the paradox of how speciation can be sustained at a rapid rate in early bursts of speciation (Martin and Richards, 2019). Radiation cascades may only happen when inter patch dispersal is close to a critical point and also depend on stochastic outcomes; this might explain why 'early bursts' are relatively rare in nature (Harmon et al., 2010).

Our work suggests that radiation cascades may be added to the list of many previously proposed mechanisms for rapid radiation, including diversity begets diversity, sexual signal complexity, the transporter hypothesis, ecological opportunity, fitness landscape connectivity, and plasticity first (Martin and Richards, 2019). These previously described mechanisms may not fully describe radiation processes in nature. According to the diversity begets diversity hypothesis, speciation opportunities arise directly from the number of species themselves, such that the speciation process becomes positive feedback cycle. However,

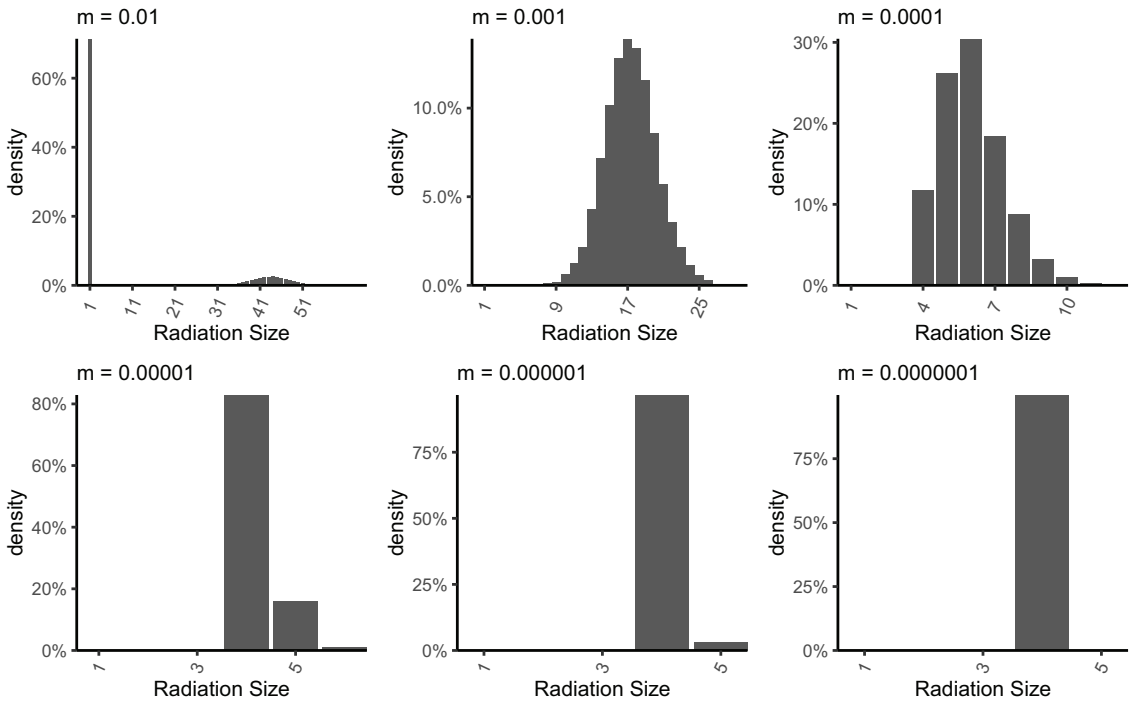


FIGURE 5 | Radiation size distributions across repeat simulations for a single immigrating lineage under different rates of internal dispersal (m). The numbers given above each panel indicate the internal dispersal rate (m). The habitat network simulated was a line structure with 4 patches, each having 1000 individuals. Gene flow effect (G) was 0.1.

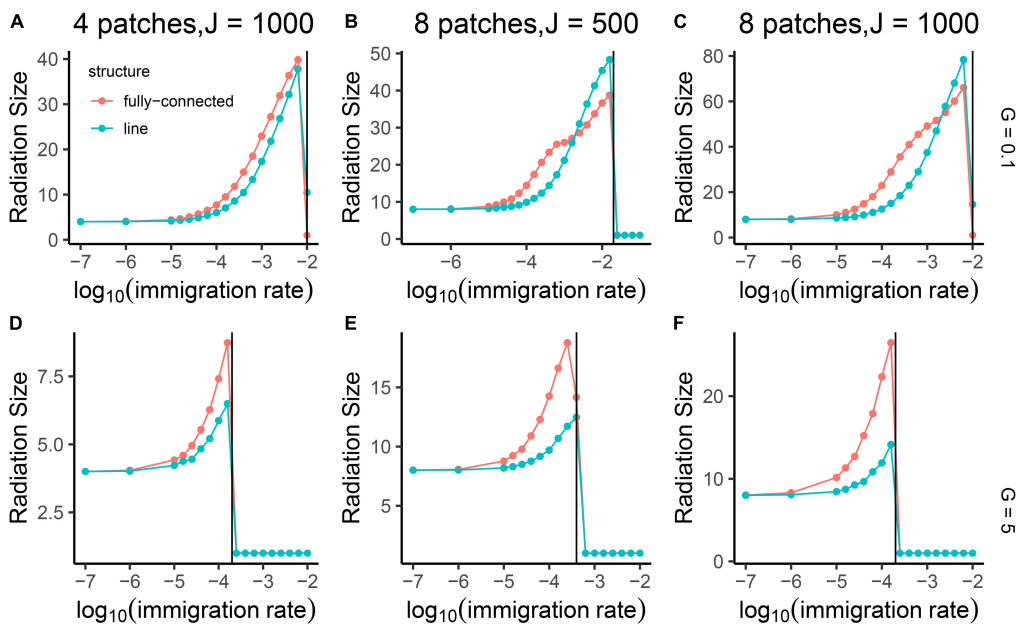


FIGURE 6 | Mean radiation size of the single lineage in the community network, shown as a function of internal dispersal rate (m) on log 10 scale. The three columns of panels correspond to three different sizes of community network: 4 patches each with 1000 individuals **(A,D)**, 8 patches each with 500 individuals **(B,E)**, and 8 patches each with 1000 individuals **(C,F)**. The gene flow effect (G) was set to 5 and 0.1 shown as two rows of panels. Blue lines represent results from a line-structured network of local patches, and red lines represent results from a fully-connected structure. The vertical black lines give the predicted value of the critical point m_{crit} for each panel based on the formula $m_{crit} = \frac{1}{G \times J}$ that was described in the Material and Methods section.

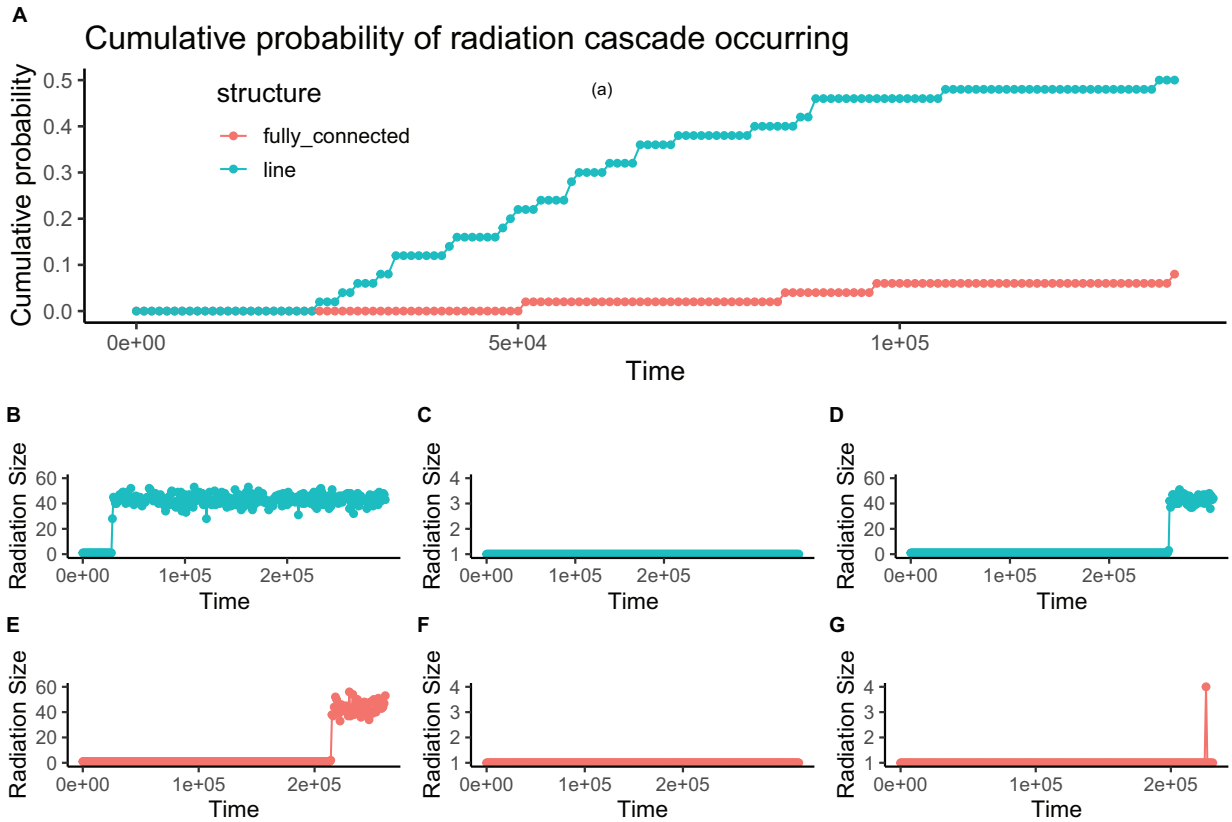


FIGURE 7 | Example time series of radiation size (panels **B–G**) showing radiation cascades occurring in some simulation runs, but not in others with the same parameters. Two extreme scenarios are possible: the radiation can be very large or non-existent. After a radiation cascade has occurred, it cannot easily be reversed. The cumulative probability for a lineage to undergo a radiation cascade increases but shows that in some cases the cascade may not occur for a long time (panel **A**). Blue lines represent results from a line-structured network of local patches and red lines represent results from a fully connected structure of local patches. The figure was produced with gene flow effect (G) equal to 0.1 and community network containing 4 patches each with 1000 individuals, within the network, the internal dispersal rate (m) was 0.01.

this mechanism would not explain how initial diversity originated in the first place, or why diversification would ultimately cease rather than expanding indefinitely at ever increasing rates. The sexual signal complexity mechanism, in which sexual signaling promotes speciation, may also explain the ease of rapid radiation, but not the extent to which a lineage will ultimately radiate before speciation slows. The remaining mechanisms provide different ways in which a lineage may be prone to future radiation due to its genes (transporter hypothesis), accessible niches (ecological opportunity and fitness landscape connectivity) or phenotypes (plasticity first). While none of these factors were added in our model, we propose that the radiation cascade mechanism is another likely candidate for the mechanism behind some rapid bursts of radiation that is independent of species numbers or biological features invoked by the previously-proposed mechanisms. What is perhaps unique to the radiation cascade mechanism is that in some cases, with the right dispersal and gene flow effect parameters, a single species and a large radiation form quasi-stable states. In this case intermediate radiation sizes are inherently unstable and thus very quickly result in further rapid speciation, or occasionally in extinction and return to the single species state.

The role of the connectivity among communities in enabling radiations is clarified by a scaling collapse (**Figure 8**) identified in our study, that is a relationship between model parameters and results that simplifies the model and increases our understanding of the system in particular cases. Specifically, a local community network with more links can be equivalent to one with stronger links, provided that the overall connectivity between local communities is still relatively weak. The intuitive rationale for this invariance to the precise network structure is that when dispersal is rare, every dispersal event results in a new species if it survives (ongoing gene flow is negligible). Consequently, the total number of dispersal events is all that matters for model dynamics as this gives the potential number of new species later on. The total number of dispersal events is given by the product of number of connections and rate of movement along each connection; this can be increased by having more connections, or by strengthening existing connections.

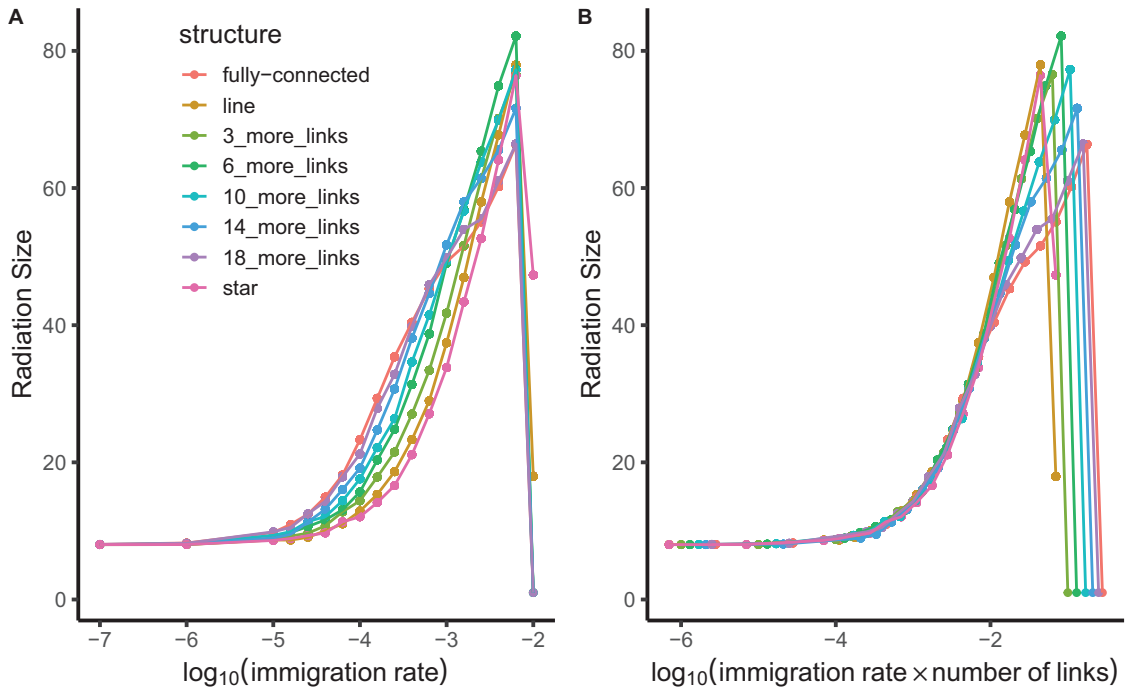


FIGURE 8 | Mean radiation size of a single lineage in a community network as a function of internal immigration rate (m) on log 10 scale. Results show a broad range of connection networks with different numbers of links in different colors. Panel **(A)** shows the raw data. Panel **(B)** shows a scaling collapse based on the same data; here, the horizontal axis shows immigration rate multiplied by number of links in the network. The leftmost part of the graph for relatively low immigration rates thus has a radiation size predicted by immigration rate × number of links rather than depending independently on those two parameters. The figure was produced with gene flow effect (G) equal to 0.1 and community network containing 8 patches each with 1000 individuals.

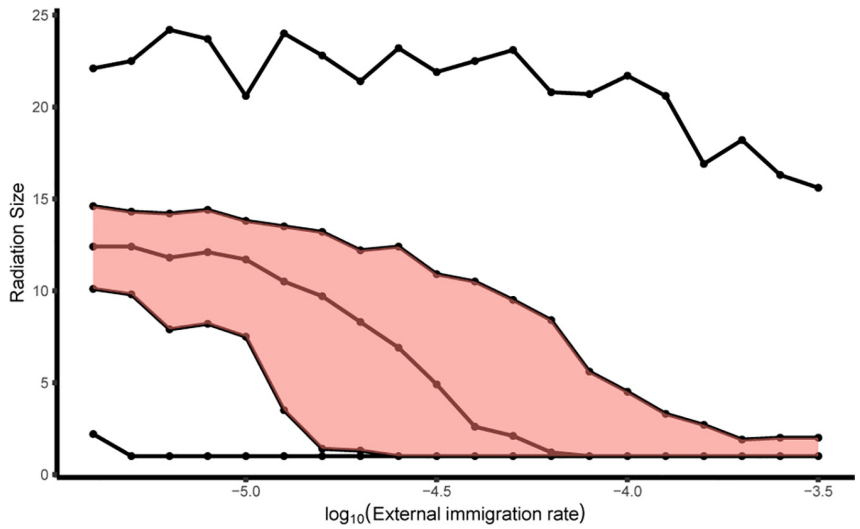


FIGURE 9 | The range of radiation sizes for all lineages in a community network as a function of the external immigration rate (M) on log 10 scale. The five lines correspond to: minimum, lower quartile, median, upper quartile, and maximum, of the distribution of radiation sizes. The range between upper quartile and lower quartile is shaded in red. The gene flow effect (G) was 0.1. Internal dispersal rate (m) was $10^{-3.5}$ and the community network was simulated with 4 patches, each having 1000 individuals.

Our models show that more isolated community networks with fewer external immigrants will naturally have larger radiation sizes. The reason is likely that external immigrants compete neutrally with pre-existing lineages and thus restrict the extent to which the pre-existing lineages can increase their range and abundance. This in turn reduces the amount of

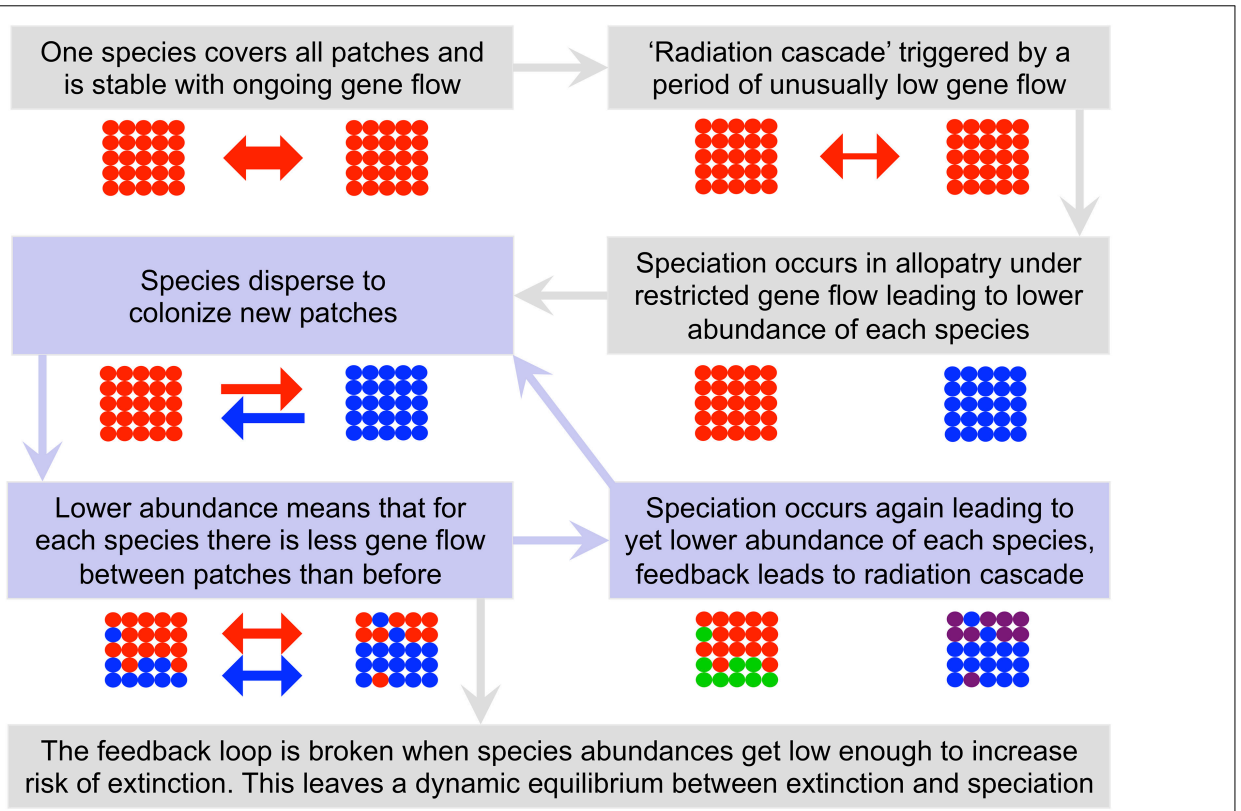


FIGURE 10 | A cartoon explaining the radiation cascade mechanism. In this example the community consists of two local habitat patches each with 25 individuals. Colored circles represent individual organisms with color representing species. The arrow(s) between the two local habitat patches indicate the strength and direction of dispersal (and thus gene flow) for each species individually. The gene flow arrows are colored to match the color of the species to which they refer. Shaded boxes and lightly shaded arrows between them indicate the stages of a radiation cascade, with the focal feedback loop that leads to the cascade being shown in a light blue.

diversification taking place in pre-existing lineages. Colonizers that do arrive on more isolated communities, with fewer immigrating competitors, have more resources to proliferate and can therefore attain larger radiation sizes.

We have framed our work in terms of local community networks and a metacommunity. Yet, the concepts would also apply an archipelago of isolated islands and a very distant mainland source pool. These processes would also occur for patches where speciation occurs in allopatry at a sub-island scale. For example, many islands are volcanic, providing isolated topographical patches throughout, e.g., at different altitudes or cut off by topographic complexity. Indeed, Losos and Parent (2009) proposed that larger areas hold larger habitat diversity leading to more radiation. Our model would be consistent with this view if the larger habitats and increased habitat diversity manifest as a more complex network of local communities. However, our model also shows that habitat diversity in itself is not a prerequisite for large radiation sizes, which could also be attributed to dispersal and isolation across a single habitat type as has been observed in diversification of micro-snail species complexes (Hendriks et al., 2019).

The radiation cascades we report here are shown only under idealized simulated conditions. It could be that in the real world, natural amounts of gene flow between pairs of patches are so variable that not many connections are at a critical point at the same time, preventing such an extreme cascade from occurring. It could be that natural variation over time of factors that affect gene flow will more likely drive the radiation forward once it has started. In real world island systems, the dispersal ability of immigrating lineages may be reduced through evolutionary processes to avoid loss of dispersers (or propagules) into the ocean (Gillespie and Baldwin, 2010). Whilst this was not explicit in our model, an evolutionary reduction in dispersal could provide another way in which a radiation cascade could suddenly be triggered: as a deterministic outcome of evolutionary change reducing dispersal rate, rather than as a rarely seen paucity of stochastic dispersal events over a sufficiently long period of time. Despite the caveats that apply to our precise model formulation, we suggest that the primary mechanisms we describe as driving radiation cascades may apply in many real systems. The fundamental conditions for a radiation cascade are simply that (i) reduced gene flow promotes speciation and (ii) speciation reduces abundance and thus reduces gene flow between conspecific populations within each new species. Together these conditions create a feedback mechanism and explain part of the reason why radiations in geographically

isolated structures (such as archipelagos) are so rapid for some species and non-existent for other species.

It seems logical, as we observe radiations with myriad functional forms adapted to different niches, to assume that the niches and adaptive processes were the only, or at least the primary driving force behind the radiation. However, our results suggest the possibility that, in some cases, the intertwined factors of stochastic dispersal events and spatial structure are the primary drivers for speciation and radiation. The resulting species that survive of course become adapted to their local niches, but as an effect of their isolation and speciation, rather than the cause of their isolation and speciation. In a radiation of Saxifragales, a major clade of temperate angiosperms, the rates of niche and phenotype evolution indeed lagged behind speciation (diversification in macroevolutionary terms) (Folk et al., 2019).

There is further support for the idea of niche adaptation as an effect rather than a cause in the case of generalist species. Suppose that a founding lineage displaying rapid radiation into niches had a high dispersal ability, and/or arrived in conditions enabling easy dispersal between the different habitat types. Gene flow may thereby continue between the different habitats, preventing radiation into locally endemic species in different niches, and instead creating the conditions for a single more generalist and widespread species to evolve that can tolerate a wide range of environmental conditions. This idea is consistent with recent observations among oceanic island plants that lineages undergoing spectacular radiations yield species that are rare and with narrow niches: "evolutionary winners are ecological losers" (Fernández-Palacios et al., 2021). Thus, we propose that radiation—or lack thereof—may arise more from dispersal limitation or genetic factors rather than the existence of the niches themselves.

Of course, in practice both perspectives will likely have elements of truth, with niches being both a cause and an effect of speciation and radiation to some extent. Ecologically distinct environments will hasten the speciation process, but with the process still requiring some degree of dispersal limitation and thus still being also subject to stochastic factors that determine radiations – including, possibly, radiation cascades.

There is much of scope for further work in this field. For example, one could experiment with different implementations of the speciation clock, which may be sensitive to population size or restrict the maximum time the speciation clock can reach, even given unlimited gene flow. More complex network structures, including structures with unequal connections between pairs of patches, would enable a more realistic view closer to real networks of local community patches (or islands). This is important as it might soften the sharp binary edge of a radiation cascade either occurring or not occurring. A more detailed study of radiation size, as a time series under different conditions, may shed further light on the new radiation cascade mechanism and when it arises. There is potential to add explicit niches, selection and adaptation into the model. It has been pointed out before that species commonness and rarity (that is, species abundance, measured in terms of numbers of individuals) could provide additional insight into community processes on islands (Rosindell and Phillimore, 2011). Indeed, neutral models of the kind we study here have often been both criticized (Enquist et al., 2002) and praised based on their predictions of commonness and rarity (Hubbell, 2001). In the context of determining radiation size, the combination of phylogenetic information and species abundances could prove to be a powerful approach to diagnosing the true underlying processes (Overcast et al., 2020).

Our study provided a mechanistic model to help understand the effects of biogeography and stochasticity on radiation size in the idealized scenario where there are no effects of adaptation. Our main finding suggests possible new drivers of radiation size in connected groups of isolated communities, including radiation cascades. This leads us to tentatively suggest that niche differentiation can be an effect rather than a cause of rapid radiations. We hope that our work will promote development of further mechanistic models of radiations and further thinking about the possible causes for rapid radiation, toward resolving why some lineages radiate so spectacularly when others do not.

DATA AVAILABILITY STATEMENT

Publicly available datasets were analyzed in this study and are available here: https://onlinelibrary.wiley.com/doi/full/10.1111/jse.12465.

AUTHOR CONTRIBUTIONS

JR proposed the model and study after early discussion with LS. RT wrote the code, did all the analyses, and produced all the graphs, with regular supervisory input from JR throughout. RT and JR produced the conceptual figures and wrote the manuscript. LS provided comments on manuscript drafts. All authors contributed to the article and approved the submitted version.

FUNDING

JR was funded by fellowships from the Natural Environment Research Council (NERC) (NE/I021179/1 and NE/L011611/1). This study is an output of the Georgina Mace centre for the Living Planet at Imperial College London. All simulations were run on high-throughput computing systems run by research computing services at Imperial College London.

ACKNOWLEDGMENTS

We thank Tim Barraclough, Rampal Etienne, Luke Harmon, and Ally Phillimore for insightful conversations about the topic, some of which occurred a long time in the past but were not forgotten.

We thank Ryan Chisholm for valuable discussions and for his comments on a draft manuscript. We also thank two referees and the editors for their comments and suggestions. RT thanks Hongye Wang for her support as well as for valuable suggestions about the figure presentation from a reader's perspective.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fevo.2021.644328/full#supplementary-material

REFERENCES

Agnarsson, I., Cheng, R.-C., and Kuntner, M. (2014). A multi-clade test supports the intermediate dispersal model of biogeography. *PLoS One* 9:e86780. doi: 10.1371/journal.pone.0086780

Arbogast, B. S., Drovetski, S. V., Curry, R. L., Boag, P. T., Seutin, G., Grant, P. R., et al. (2006). The origin and diversification of Galapagos mockingbirds. *Evolut. Int. J. Organic Evolut.* 60, 370–382. doi: 10.1554/03-749.1

Bennett, G. M., and O'Grady, P. M. (2013). Historical biogeography and ecological opportunity in the adaptive radiation of native Hawaiian leafhoppers (Cicadellidae: Nesophrosyne). *J. Biogeogr.* 40, 1512–1523. doi: 10.1111/jbi.12099

Claramunt, S., Derryberry, E. P., Remsen, J. V. Jr., and Brumfield, R. T. (2012). High dispersal ability inhibits speciation in a continental radiation of passerine birds. *Proc. Biol. Sci. R. Soc.* 279, 1567–1574. doi: 10.1098/rspb.2011.1922

De Meester, L., Vanoverbeke, J., Kilsdonk, L. J., and Urban, M. C. (2016). Evolving perspectives on monopolization and priority effects. *Trends Ecol. Evolut.* 31, 136–146. doi: 10.1016/j.tree.2015.12.009

Economo, E. P., and Keitt, T. H. (2008). Species diversity in neutral metacommunities: a network approach. *Ecol. Lett.* 11, 52–62.

Enquist, B. J., Sanderson, J., and Weiser, M. D. (2002). Modeling macroscopic patterns in ecology. *Science* 295, 1835–1838.

Erwin, D. H. (1992). A preliminary classification of evolutionary radiations. *Hist. Biol.* 6, 133–147. doi: 10.1080/10292389209380423

Fernández-Palacios, J. M., Otto, R., Borregaard, M. K., Kreft, H., Price, J. P., Steinbauer, M. J., et al. (2021). Evolutionary winners are ecological losers among oceanic island plants. *J. Biogeogr.* doi: 10.1111/jbi.14143 [Epub ahead of print].

Folk, R. A., Stubbs, R. L., Mort, M. E., Cellinese, N., Allen, J. M., Soltis, P. S., et al. (2019). Rates of niche and phenotype evolution lag behind diversification in a temperate radiation. *Proc. Natl. Acad. Sci. U S A.* 116, 10874–10882. doi: 10.1073/pnas.1817999116

Fukami, T. (2015). Historical contingency in community assembly: integrating niches, species pools, and priority effects. *Annu. Rev. Ecol. Evolut. Systemat.* 46, 1–23. doi: 10.1146/annurev-ecolsys-110411-160340

Gascuel, F., Laroche, F., Bonnet-Lebrun, A.-S., and Rodrigues, A. S. L. (2016). The effects of archipelago spatial structure on island diversity and endemism: predictions from a spatially-structured neutral model. *Evol. Int. J. Organic Evolut.* 70, 2657–2666. doi: 10.1111/evo.13067

Gillespie, R. G., and Baldwin, B. G. (2010). "Island biogeography of remote archipelagoes: interplay between ecological and evolutionary processes," in *The theory of island biogeography at 40: impacts and prospects*, eds J. B. Losos and R. E. Ricklefs (Princeton, NJ: Princeton University Press), 358–378. doi: 10.1515/9781400831920.358

Givnish, T. J., Millam, K. C., Mast, A. R., Paterson, T. B., Theim, T. J., Hipp, A. L., et al. (2009). Origin, adaptive radiation and diversification of the Hawaiian lobeliads (Asterales: Campanulaceae). *Proc. R. Soc. B Biol. Sci.* 276, 407–416. doi: 10.1098/rspb.2008.1204

Grant, P. R. (2017). *Ecology and Evolution of Darwin's Finches (Princeton Science Library Edition): Princeton Science Library Edition*. Princeton, NJ: Princeton University Press.

Grant, P. R., and Grant, B. R. (2011). *How and Why Species Multiply: The Radiation of Darwin's Finches*. Princeton, NJ: Princeton University Press.

Grant, P. R., and Grant, B. R. (2014). *40 Years of Evolution: Darwin's Finches on Daphne Major Island*. Princeton, NJ: Princeton University Press.

Harmon, L. J., Losos, J. B., Jonathan Davies, T., Gillespie, R. G., Gittleman, J. L., Bryan Jennings, W., et al. (2010). Early bursts of body size and shape evolution are rare in comparative data. *Evolut. Int. J. Organic Evolut.* 64, 2385–2396.

Harvey, P. H., and Rambaut, A. (2000). Comparative analyses for adaptive radiations. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* 355, 1599–1605. doi: 10.1098/rstb.2000.0721

Hendriks, K. P., Alciatore, G., Schilthuizen, M., and Etienne, R. S. (2019). Phylogeography of Bornean land snails suggests long-distance dispersal as a cause of endemism. *J. Biogeogr.* 46, 932–944. doi: 10.1111/jbi.13546

Hubbell, S. P. (2001). *The Unified Neutral Theory of Biodiversity and Biogeography (MPB-32)*. Princeton, NJ: Princeton University Press.

Ibanez, T., Keppel, G., Baider, C., Birkinshaw, C., Culmsee, H., Cordell, S., et al. (2018). Regional forcing explains local species diversity and turnover on tropical islands. *Glob. Ecol. Biogeogr.* 27, 474–486. doi: 10.1111/geb.12712

Kocher, T. D. (2004). Adaptive evolution and explosive speciation: the cichlid fish model. *Nat. Rev. Genet.* 5, 288–298. doi: 10.1038/nrg1316

Lamichhaney, S., Berglund, J., Almén, M. S., Maqbool, K., Grabherr, M., Martinez-Barrio, A., et al. (2015). Evolution of Darwin's finches and their beaks revealed by genome sequencing. *Nature* 518, 371–375. doi: 10.1038/nature14181

Losos, J. B. (2010). Adaptive radiation, ecological opportunity, and evolutionary determinism. American Society of Naturalists E. O. Wilson award address. *Am. Natural.* 175, 623–639. doi: 10.1086/652433

Losos, J. B. (2011). *Lizards in an evolutionary tree: ecology and adaptive radiation of anoles*, Vol. 10. California: Univ of California Press.

Losos, J. B., and Parent, C. E. (2009). *The speciation-area relationship. The Theory of Island Biogeography Revisited*. Princeton, NJ: Princeton University Press, 415–438.

Losos, J. B., Mahler, D. L., and Others. (2010). Adaptive radiation: the interaction of ecological opportunity, adaptation, and speciation. *Evolut. Darwin First* 150, 381–420.

Lovette, I. J., Bermingham, E., and Ricklefs, R. E. (2002). Clade-specific morphological diversification and adaptive radiation in Hawaiian songbirds. *Proc. Biol. Sci. R. Soc.* 269, 37–42. doi: 10.1098/rspb.2001.1789

MacArthur, R. H., and Wilson, E. O. (1967). *The Theory of Island Biogeography*. Princeton, NJ: Princeton University Press.

Martin, C. H., and Richards, E. J. (2019). The paradox behind the pattern of rapid adaptive radiation: how can the speciation process sustain itself through an early burst? *Annu. Rev. Ecol. Evolut. Systemat.* 50, 569–593. doi: 10.1146/annurev-ecolsys-110617-062443

Overcast, I., Ruffley, M., Rosindell, J., Harmon, L., Borges, P. A., Emerson, B. C., et al. (2020). A unified model of species abundance, genetic diversity, and functional diversity reveals the mechanisms structuring ecological communities. *BioRxiv*. [Preprint].

Price, J. P., and Wagner, W. L. (2018). Origins of the Hawaiian flora: Phylogenies and biogeography reveal patterns of long—distance dispersal. *J. Systemat. Evolut.* 56, 600–620. doi: 10.1111/jse.12465

Ricklefs, R. E., and Bermingham, E. (2002). The concept of the taxon cycle in biogeography. *Glob. Ecol. Biogeogr. J. Macroecol.* 11, 353–361. doi: 10.1046/j.1466-822x.2002.00300.x

Robichaux, R. H., Carr, G. D., Liebman, M., and Pearcy, R. W. (1990). Adaptive Radiation of the Hawaiian Silversword Alliance (Compositae- Madiinae): Ecological, Morphological, and Physiological Diversity. *Ann. Missouri Botanic. Garden.* 77, 64–72. doi: 10.2307/2399626

Rosindell, J., and Harmon, L. J. (2013). A unified model of species immigration, extinction and abundance on islands. *J. Biogeogr.* 40, 1107–1118. doi: 10.1111/jbi.12064

Rosindell, J., and Phillimore, A. B. (2011). A unified model of island biogeography sheds light on the zone of radiation. *Ecol. Lett.* 14, 552–560. doi: 10.1111/j.1461-0248.2011.01617.x

Rosindell, J., Cornell, S. J., Hubbell, S. P., and Etienne, R. S. (2010). Protracted speciation revitalizes the neutral theory of biodiversity. *Ecol. Lett.* 13, 716–727. doi: 10.1111/j.1461-0248.2010.01463.x

Rosindell, J., Hubbell, S. P., and Etienne, R. S. (2011). The unified neutral theory of biodiversity and biogeography at age ten. *Trends Ecol. Evol.* 26, 340–348. doi: 10.1016/j.tree.2011.03.024

Rosindell, J., Wong, Y., and Etienne, R. S. (2008). A coalescence approach to spatial neutral ecology. *Ecol. Inform.* 3, 259–271. doi: 10.1016/j.ecoinf.2008.05.001

Rundell, R. J., and Price, T. D. (2009). Adaptive radiation, nonadaptive radiation, ecological speciation and nonecological speciation. *Trends Ecol. Evolut.* 24, 394–399. doi: 10.1016/j.tree.2009.02.007

Schenk, J. J., and Steppan, S. J. (2018). The Role of Geography in Adaptive Radiation. *Am. Natural.* 192, 415–431. doi: 10.1086/699221

Schluter, D. (2000). *The ecology of adaptive radiation*. Oxford: OUP Oxford.

Simões, M., Breitkreuz, L., Alvarado, M., Baca, S., Cooper, J. C., Heins, L., et al. (2016). The Evolving Theory of Evolutionary Radiations. *Trends Ecol. Evolut.* 31, 27–34. doi: 10.1016/j.tree.2015.10.007

Simpson, G. G. (1953). *The Major Features of Evolution*. Columbia: Columbia University Press.

Stroud, J. T., and Losos, J. B. (2016). Ecological opportunity and adaptive radiation. *Annu. Rev. Ecol. Evolut. Systemat.* 47, 507–532. doi: 10.1146/annurev-ecolsys-121415-032254

Thompson, S. E., Chisholm, R. A., and Rosindell, J. (2020). pycoalescence and rcoalescence: packages for simulating spatially explicit neutral models of biodiversity. *Methods Ecol. Evol.* 11, 1237–1246. doi: 10.1111/2041-210X.13451

Wellborn, G. A., and Langerhans, R. B. (2015). Ecological opportunity and the adaptive diversification of lineages. *Ecol. Evolut.* 5, 176–195. doi: 10.1002/ece3.1347

Yoder, J. B., Clancey, E., Des Roches, S., Eastman, J. M., Gentry, L., Godsoe, W., et al. (2010). Ecological opportunity and the origin of adaptive radiations. *J. Evolut. Biol.* 23, 1581–1596. doi: 10.1111/j.1420-9101.2010.02029.x

Yu, X. Q., Maki, M., Drew, B. T., Paton, A. J., Li, H. W., Zhao, J. L., et al. (2014). Phylogeny and historical biogeography of Isodon (Lamiaceae): rapid radiation in south-west China and Miocene overland dispersal into Africa. *Mol. Phylogenet. Evolut.* 77, 183–194. doi: 10.1016/j.ympev.2014.04.017

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Ecology and Evolution

ORIGINAL RESEARCH
published: 29 October 2021
doi: 10.3389/fevo.2021.686403

The Dynamic Hypercube as a Niche Community Model

John M. Halley[1]* *and Stuart L. Pimm*[2]

[1] *Department of Biological Applications & Technology, University of Ioannina, Ioannina, Greece,* [2] *Nicholas School of the Environment, Duke University, Durham, NC, United States*

Different models of community dynamics, such as the MacArthur–Wilson theory of island biogeography and Hubbell's neutral theory, have given us useful insights into the workings of ecological communities. Here, we develop the niche-hypervolume concept of the community into a powerful model of community dynamics. We describe the community's size through the volume of the hypercube and the dynamics of the populations in it through the fluctuations of the axes of the niche hypercube on different timescales. While the community's size remains constant, the relative volumes of the niches within it change continuously, thus allowing the populations of different species to rise and fall in a zero-sum fashion. This dynamic hypercube model reproduces several key patterns in communities: lognormal species abundance distributions, 1/*f*-noise population abundance, multiscale patterns of extinction debt and logarithmic species-time curves. It also provides a powerful framework to explore significant ideas in ecology, such as the drift of ecological communities into evolutionary time.

Keywords: dynamic population model, extinction, biodiversity, community model, 1/*f* noise, hypervolume, Hutchinsonian niche

OPEN ACCESS

Edited by:
György Barabás,
Linköping University, Sweden

Reviewed by:
Thomas Koffel,
Michigan State University, United States
Daniel Maynard,
ETH Zürich, Switzerland

***Correspondence:**
John M. Halley
jhalley@uoi.gr

Specialty section:
This article was submitted to Models in Ecology and Evolution, a section of the journal Frontiers in Ecology and Evolution

Received: *26 March 2021*
Accepted: *22 September 2021*
Published: *29 October 2021*

Citation:
Halley JM and Pimm SL (2021) The Dynamic Hypercube as a Niche Community Model. Front. Ecol. Evol. 9:686403. doi: 10.3389/fevo.2021.686403

INTRODUCTION

Empirical studies of ecological communities often call for a greater emphasis on time and on longer time intervals (Ripa and Lundberg, 2000; Hastings, 2004; Magurran, 2007). This call is urgent as we live in an age of major changes in biodiversity across the Earth via landscape transformation or migration (Dornelas et al., 2014). Additionally, in the last 50 years, there has been a reassessment of the dominant paradigm of community organization. The "balance of nature" paradigm has receded somewhat (Cuddington, 2001; Ergazaki and Ampatzidis, 2012; Simberloff, 2014) and we now have a more dynamic conception of persistence in natural ecosystems (Pimm, 1991). Variability exists on all timescales both within the ecological community itself (Pimm and Redfearn, 1988; Halley, 1996) and in the environment to which it responds (Wunsch, 2003; Franzke et al., 2020). Thus, we expect ecological equilibria to be more dynamic, more provisional and more diffuse (Halley and Inchausti, 2004). The simplifications of assuming an equilibrium will remain useful and compelling, such as for flux calculations and for the species-area relationship. That said, the greater attention to time calls for model development to help insights specifically into dynamic effects in communities.

The basis of most modeling in community ecology is the Lotka–Volterra model. This pair of equations with four interaction terms can describe relationships of competition, predation, herbivory and mutualism. While the Lotka–Volterra system is phenomenological, it can also be related to more mechanistic resource models (Schoener, 1973). A natural development for dynamic community modeling is to generalize this system to S species described by a set of first-order non-linear coupled differential equations. This generalized Lotka–Volterra system has had enormous

impact on community ecology, including classical treatments such as that of May (May, 1973) and continues to be the subject of novel approaches (Forte and Vrscay, 1996; Bertuzzo et al., 2011; Fisher and Mehta, 2014; Kessler and Shnerb, 2015). The growth in computing power has circumvented many of the obvious technical difficulties of solving a system of this size. A more persistent challenge is how to populate the S^2 interaction terms with plausible parameter values. Assigning interaction strengths in an ecological community of hundreds of species remains a formidable problem and drives the development of various other community models, though these could arguably be seen as special cases of the generalized Lotka–Volterra system (Kessler and Shnerb, 2015).

One of the most widely applied community theories is that of island biogeography (MacArthur and Wilson, 1967). It ignores interactions entirely, so extinction and colonization processes drive the system's biodiversity. The simplicity of this theory (MacArthur and Wilson, 1967) allowed many developments using metapopulation-type communities (Tilman, 1994; Matter et al., 2002). Another development for biodiversity ecology was Hubbell's unified neutral theory of biodiversity (NTB). Hubbell's model made the remarkable assumption of following Kimura's model of genetic neutral evolution (Kimura, 1955), considering species to be ecologically indistinguishable from one another and then assuming that all stochasticity is demographic (Hubbell, 2001). This model violated obvious ecological realities. Nonetheless, it explained a wide variety of ecological patterns and processes such as near-lognormal species abundance distributions (McGill, 2003), Fisher's α-parameter (Chave, 2004), and spatial clumping (Chave and Leigh, 2002). The application of neutral theory to extinction debt (Halley and Iwasa, 2011) that specifically used its dynamics was probably its first potential use in conservation. The NTB has since stimulated much theoretical work in community ecology — including this special issue. Nevertheless, limitations in the NTB (Clark and Mclachlan, 2003; Ricklefs, 2006; Leigh, 2007) show that it cannot be a useful tool for all phenomena. Its most obvious violation of ecological reality is it ignores niches. Various attempts have been made to merge these. Proposals combining niche and neutral models explore community dynamics (Adler et al., 2007; Zillio and Condit, 2007; Chisholm and Pacala, 2010). Since stochasticity is crucial to the operation of the NTB, the need to involve environmental stochasticity has been noted as particularly important (Kalyuzhny et al., 2015; Engen et al., 2017). Species-area relations have also been a problem. The neutral model found it difficult to reproduce observed spatial patterns such as the Arrhenius species-area relationship (Leigh, 2007; Halley et al., 2014).

Hutchinson's niche-hypervolume concept (Hutchinson, 1957) has been an attractive and popular approach because of its intuitive simplicity and because it is readily visualized (e.g., **Figures 1**, **2** below). Interest in Hutchinson's model is resurging recently (Barros et al., 2016; Díaz et al., 2016; Blonder, 2018). This growing popularity links to research contexts such as niche envelope models (Thomas et al., 2004; Soberón, 2010; Barros et al., 2016) and functional ecology (Lamanna et al., 2014; Díaz et al., 2016; Pigot et al., 2016). It drives considerable methodological research, mainly in the description and parameterisation of the hypervolumes (Blonder et al., 2018; Carvalho and Cardoso, 2021). There has also been progress in relating the niche to ideas of coexistence. Concepts of stable coexistence have evolved from strict associations with stable equilibria in the Lotka–Volterra system to concepts of invasibility (Chesson and Warner, 1981; Chesson, 2000). Theoretical developments have refined the relation between limiting similarity and competitive exclusion (Leibold, 1995; Chase and Leibold, 2003; Meszéna et al., 2006; Barabás et al., 2018).

In this paper, we develop a dynamic model, using Hutchinson's hypervolume concept in a hypercube configuration. Much of the research on niche-related issues, as in community ecology generally, focuses on non-dynamic issues, seeking to explain patterns in space or in community organization. For example, in McGill's (2010) survey of unified biodiversity models (McGill, 2010), only two of the six models feature dynamic descriptions of biodiversity. In seeking a model that is dynamic, we mean a model that can predict the time evolution of communities explicitly, something that has also been noted by others generally (Hastings, 2004; Magurran, 2007; Engen et al., 2017) and in hypervolume theory in particular (Holt, 2009). Many of the emerging problems in community ecology, such as extinction debt and colonization credit or the response of communities to climate change are dynamic in nature. They require the model to predict the state of the community at some specified time into the future. In addition, we sought a model that would generate several desirable features.

The Lognormal species-abundance distribution. Within a community, there are many rare species and a few common ones. The ensemble species abundance distribution at any time is often observed to be lognormal. Preston (1962) argued that a specific lognormal predominates — the "canonical" lognormal. It means, broadly, that most individuals in a community belong to the most abundant species, rather than, for example, to a large set of less abundant ones. Preston showed that this was consistent with the Arrhenius species-area curve having an exponent of 1/4, — a pattern frequently observed (Sugihara, 1980). McGill (2003) and others have shown that many theories generate such distributions, such as the zero-sum multinomial for the neutral model.

Lognormal distributions in time. Species abundance varies through time because of interactions with other species, demographic stochasticity, and in response to environmental stochastic factors (Halley and Iwasa, 1998; Engen et al., 2005). This variation can be considerable. Not surprisingly, species that vary most are at greater risk of extinction than those that are most constant. This is unless the variability is non-stationary (Halley and Kunin, 1999). Species abundances over time are typically distributed lognormally (Halley and Inchausti, 2002). Thus, as a species traces its trajectory of abundance $X(t)$, its histogram tends to be lognormal.

More time, more variance. In this model, each species' abundance responds entirely to the environmental pressures on that niche. According to Pimm (1991), the ecological community is fundamentally non-stationary. This is equivalent to saying that

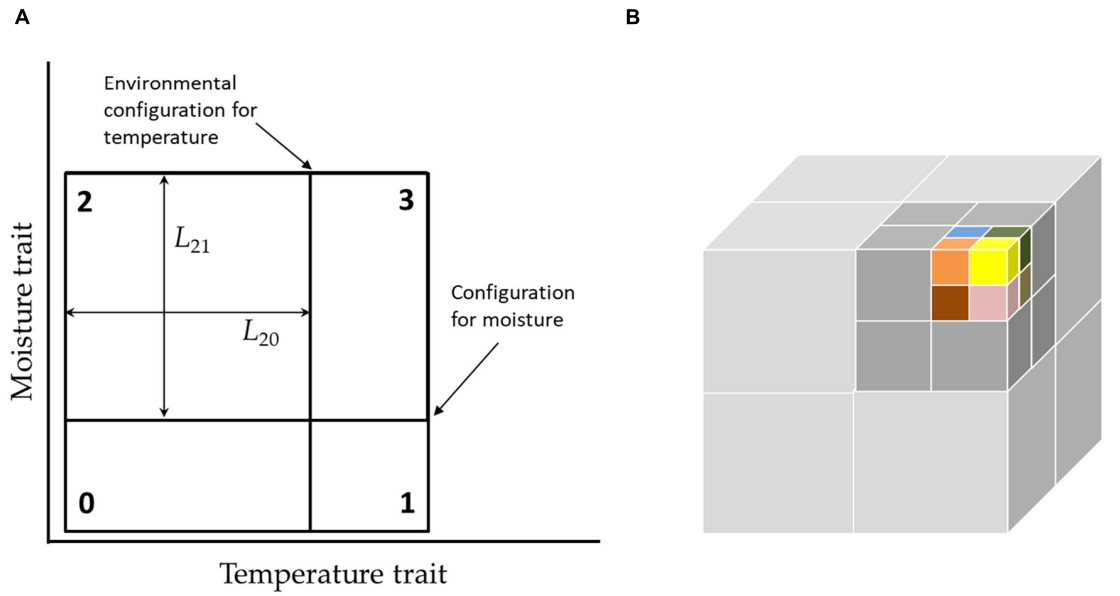


FIGURE 1 | (A) Environmental configurations and niche widths in a two-axis hypercube. In this picture, the current configuration (position of partitions) favors species-2, which likes a cool but moist environment, since this has the largest area. In our graphic notation, note that the position of the environmental configuration, does not mean exactly temperature relative to the axis, but means that degree to which the environment favors hot or cold adapted species. **(B)** While the full niche hypercube cannot be depicted for more than 3 dimensions, the structure of the community can be depicted by "unfolding" the higher-dimensional hypercube into lower dimensions. Here, a six-dimensional community with 64 species is represented by 8 by $2 \times 2 \times 2$ sub-cubes stacked to form a single "super-cube" of size $4 \times 4 \times 4$. In turn this may be nested within another cube of size $8 \times 8 \times 8$, and so on. This requires, that changes in the faster processes are mirrored in all super-cubes.

variability increases with time — the "more time means more variation" effect (Pimm and Redfearn, 1988; Inchausti and Halley, 2002). A series of studies using the Global Population Dynamics Database (GPDD) that examined the relation of variance to time systematically found that this increase of variance was universal (Inchausti and Halley, 2001, 2002). The spectral analysis showed that the model most consistent with this was a $1/f$-noise model for which variance increases linearly with the logarithm of time (Halley, 1996; Halley and Inchausti, 2004).

Energy conservation: zero-sum properties. Since only a fixed amount of energy arrives in an ecosystem of fixed area, there is a limit on the number of organisms of a given size that can live there. This is an implied property of many models through the assumption of a constant number of individuals. This leads to a corresponding zero-sum principle, that populations add up to a constant community size in the absence of extinction-speciation and that all changes in populations sum to zero.

Below, we describe its outline and some properties. We then apply it to four different situations, showing how it can be useful in ecological research and applications.

MATHEMATICAL DESCRIPTION AND PROPERTIES

We assume the simplified geometry of Hutchinson's hypervolume to have the geometry of a hypercube, where there are many axes, each of which describes an environmental resource or limitation (We call these *niche factors*). The set of preferences for the various factors defines each niche. Only one species can occupy it, following Gause's principle of competitive exclusion. After Hutchinson's conception (Hutchinson, 1978), the niche is usually defined as the environmental conditions under which a given population has a positive growth rate. The volume of that niche is the probability that an individual can reproduce, if the state of the environment is picked randomly. However, if the area is large and it contains a large community, then we may interpret hypervolume as a number of individuals that can reproduce, and hence as a carrying capacity or ceiling for population. The intended target of our model is the community in a relatively large area. So, for both the niche and community hypervolumes (referred to J) the units of hypervolume (usually shortened hereafter to *volume*) are individual organisms. We interpret it as a ceiling on population for a species or for the community overall. This may be an issue when we have species of different sizes in the community (and hence different energy requirements) but refinements of this type are beyond the scope of this paper.

To simplify this model, our model has additional assumptions. (a) The world is binary. This includes binary niche factors: each factor can only be either "high" or "low." Species responses are also binary: each species occupies only the position on the axis associated with its preferred level. Recognizing the existence of more developed concepts of limiting similarity (Leibold, 1995; Meszéna et al., 2006), we assume limited niche overlap is necessary and sufficient to avoid competitive exclusion. Thus,

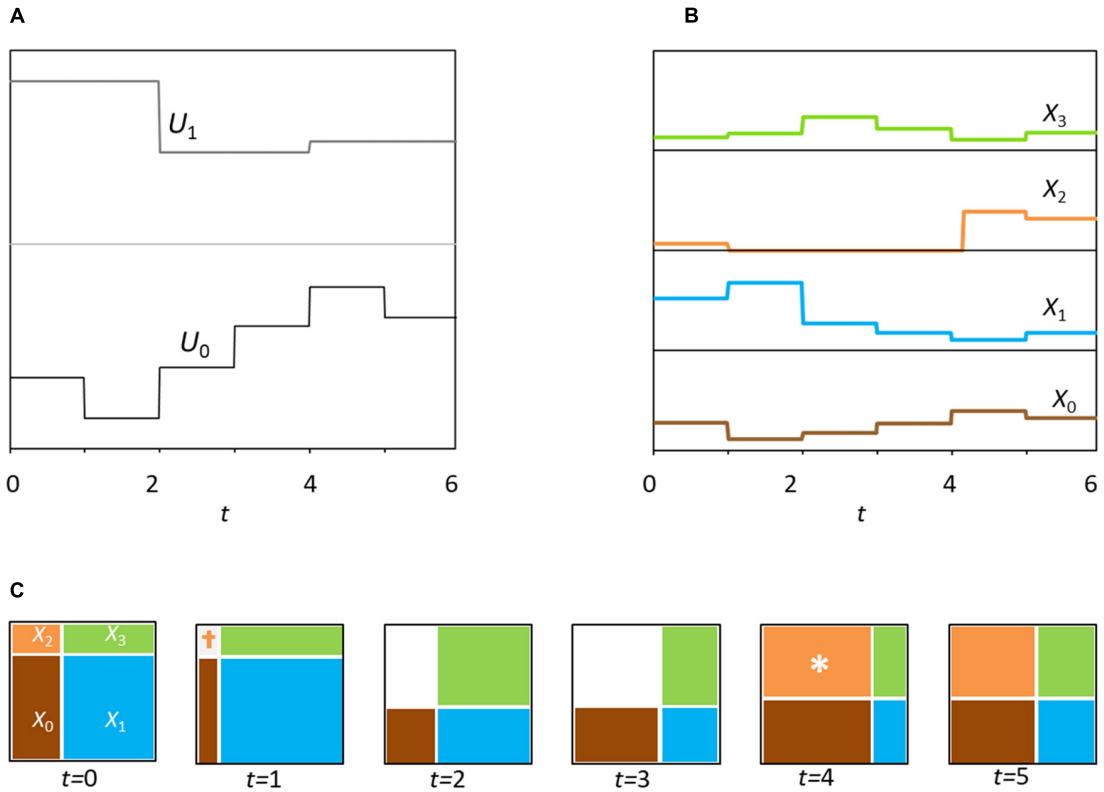


FIGURE 2 | (A) Dynamics of the environmental variation in time for six steps (position of partitions, as in **Figure 1A**). Note that the partition for axis k assumes a new state at random every 2^k units of time, in this case one and two time steps, respectively. **(B)** Six time-steps of model dynamics for the four species. **(C)** Six panels showing the changes in niche-size (V_S) for consecutive time steps. Colors denote occupation by the corresponding population (X_S). Note the extinction of X_2 (species 2) at time $t = 1$ due to population falling below viability ('†'), followed by two steps at zero and then at time $t = 4$, recolonization ('*').

we do not consider hypervolume overlaps nor the possibility of morphological displacements as, for example, in Carvalho and Cardoso (2021). This also means that species do not evolve within our community, a statement of strong *niche conservatism* (Holt, 2009). (b) In addition, we assume that each factor fluctuates independently and on a different timescale. (c) The different timescales are spaced equally in log-time. (d) Population responds proportionally to changes in niche volume. These are the only drivers of population. Thus, we do not explicitly include population growth, species interactions or density dependence. The exception is that if the population goes extinct it remains extinct until recolonisation. (e) Species do not influence each other nor the niche structure.

The Niche as a Product of Traits

The axes of the niche hypercube are associated with niche factors. There are only two available positions on any axis, and these reflect corresponding traits to respond to this factor. The axis may be a niche factor such as moisture, in which case the organism is either moisture-loving or dry-loving, but it cannot be both. In Hutchinson's perspective, each organism "chooses" an approach for each axis. We assume that each combination of all traits together generates one species associated with a single niche. Every such niche can be modeled by a corner on the hypercube. By Gause's principle of competitive exclusion, two different species with the same traits cannot occupy the same niche. In our model, two species with the same traits are considered as the same species having the same traits.

Every niche, and so every species, is defined by K separate traits, one for each of the K dimensions. Thus, the species identity $\mathbf{s}$ can be associated with a binary number:

$$\mathbf{s} = (o_0, o_1, o_2, \ldots o_{K-1})$$

The index o_k is 0 or 1 depending on the trait, the position of the species relative to axis k, which can also be associated with a unique decimal number s, where $0 \leq s \leq 2^K - 1$.

Thus, for any species, s, its functional relationships can be obtained by converting to binary. If the binary kth digit is '0' then species-s likes a low value of niche factor k and if it is a '1' then species-s likes a high value of niche factor k. When drawing diagrams, we follow matrix convention where the first index refers to the row number and the second to the column number. Thus, for $K = 2$, 01 is found on the lower row, rightmost column.

While the corners of the hypercube define the species traits, the dividing lines define the volumes of niches and reflect the environment (**Figure 1A**), which changes with time (see next section). The most obvious "visual" representation of the

TABLE 1 | Niche representations for two niche factor axes (K = 2) in the example of **Figure 1** and the corresponding niche sizes when the environment is warm and moist.

Species number, s	**Species name (Binary, s)**	**Description of traits**	**Niche sizes under given environmental conditions**
0	00	Prefers cold, prefers dry	(0.9) × (0.8) = 0.72
1	01	Prefers cold, prefers wet	(0.9) × (0.2) = 0.18
2	10	Prefers warm, prefers dry	(0.1) × (0.8) = 0.08
3	11	Prefers warm, prefers wet	(0.1) × (0.2) = 0.02

Suppose we specify B_0 = 3 and B_1 = 2. From this it follows that Γ_0 = 0.3 and Γ_1 = 0.4 that U_0 = 0.9 and U_1 = 0.8, respectively. This clearly defines one large and one small niche width on each axis, one "friendly" the other "unfriendly" to a given species.

niche is as a hypercube in K-dimensional space. However, this excludes actual visual representation for $K > 3$. An alternative representation is through a series of nested cubes (**Figure 1B**).

The *community hypercube* is assumed to have a hypervolume of unity that is the sum of all the different *niche hypercubes*. For species-s the width on axis k is L_{sk} (see **Appendix A** for a list of symbols). The niche hypervolume of a species, is the product of its widths on all axes:

$$V_s = L_{s0} \cdot L_{s1} \cdot L_{s2} \cdot \ldots \cdot L_{s,K-1} = \prod_{k=0}^{K-1} L_{sk} \tag{1}$$

Given this potential community of 2^K species, the environment will select the species that flourish. At any time, the environment has a *configuration* for each niche factor k. Axes may also represent stresses, such as radiation, or chemical pollution. A positive trait for a stress means the species is tolerant and has a potential advantage over other species. An example with two factors is given in **Table 1**.

The niche width, L_{sk}, is the outcome of the interaction of species orientation with the environmental configuration. We assume each niche's width on axis k has two possibilities:

$$L_{sk} = \begin{cases} U_k \\ 1 - U_k \end{cases} \tag{2}$$

Since we are working on the *unit hypercube* the total width of the two traits on any axis must add to unity. In general, U_k is different for each axis.

As an alternative way to quantify the two states on an axis, we introduce the constant B, such that $U_k \propto B$ and $1-U_k \propto 1/B$. This, together with (2), and introducing a normalizing constant Γ_k, leads to the values for traits on each axis k:

$$L_{sk} = \begin{cases} \Gamma_k \cdot B_k \\ \Gamma_k / B_k \end{cases} \quad , \quad \Gamma_k = 1/(B_k + 1/B_k) \tag{3}$$

The reason for this alternative representation is that when dealing with populations it is often preferable to work with logarithms. When we take logarithm of Equation (3), it assumes a simple symmetric form:

$$\lambda_{sk} = \ln L_{sk} = \begin{cases} \gamma_k + a_k \\ \gamma_k - a_k \end{cases} \tag{4}$$

We calculate the factors $a_k = \ln B_k$ and $\gamma_k = \ln \Gamma_k$ through the formulas

$$\gamma_k = \ln \sqrt{U_k(1-U_k)} \ , \ \ a_k = \ln \sqrt{\frac{U_k}{1-U_k}} \tag{5}$$

Niche Factors and the Environmental Configuration

Clearly the configuration of the niche factors $U_0, U_1, U_2, \ldots, U_{K-1}$ depends on the environment. In this paper, we will assume that U_k is an instance of a random variable taken from a distribution $f(u)$. All the U_k's are assumed independent and identically distributed. Note that even if the distribution is the same for all axes, any realization of the hypercube will in general have different configurations on each axis.

Special Case A: Broken stick model. If U is a realization of a uniform random variable:

$$\begin{aligned} f(u) &= 1, \ \forall 0 \le u \le 1 \\ &= 0, \ \text{elsewhere} \end{aligned}$$

This has important similarities to the "broken stick" model used by MacArthur (1957) but differs in that we break each axis only once. Sugihara introduced a "sequential broken stick" model to model the community that would have a lognormal species abundance distribution (Sugihara, 1980). Our model also has important similarities to Sugihara's model. The corresponding distribution for the niche factor is:

$$f_\lambda(x) = e^x \quad x \le 0$$

Note that here λ is the niche width in log space, so that this distribution comes simply from a change of variable.

Special Case B: Symmetric alternating states for all axes. An important special case, that will be explored in this paper, is when all traits have two possible values, so $f(u)$ is a dichotomous distribution. This may be written as follows:

$$f(u) = \frac{\delta(u-u_0)}{2} + \frac{\delta(u-(1-u_0))}{2}$$

where the function $\delta(x)$ is the Dirac delta function. In this case, u_0 is the same for all axes, which means that also B is the same for all axes, so that $u_0 = \Gamma B$ and $1–u_0 = \Gamma/B$. Also, the distribution of log of niche factor, λ_{sk}, will be similarly dichotomous:

$$f_\lambda(x) = \frac{\delta(x-\gamma_0+a_0)}{2} + \frac{\delta(u-\gamma_0-a_0)}{2}$$

Since B is the same for all axes, the unit volume of the community hypercube can be written as:

$$1 = \left[\Gamma B + \frac{\Gamma}{B}\right]^K = \Gamma^K\left[B^K + KB^{K-2} + K(K-1)B^{K-4} + ...+ C_k^K B^{K-2k} + ... + \frac{1}{B^K}\right]$$

revealing the states of the community: the volume of any niche is $\Gamma^K B^{K-2k}$ for any $k \in \{0, 1, 2, ..., K\}$, while there are C_k^K niches of that size in the community hypercube. There are thus maximum and minimum values for niche volume:

$$\left(\frac{\Gamma}{B}\right)^K \leq V_S \leq (\Gamma B)^K$$

Special Case C: Asymmetric alternating states for all axes. This is the same as Special Case B, except that the two states between which a niche factor can alternate, are not equally probable. This may be written as follows:

$$f(u) \;=\; q \cdot \delta(u - u_0) \;+\; (1-q) \cdot \delta(u - (1-u_0))$$

In this case too, B will be the same for all axes, with $u_0 = \Gamma B$ and 1-$u_0 = \Gamma/B$. For $q = 0.5$, special case C is the same as Special Case B.

In general, to relate the niche widths of Equation (4) in a more precise way to the environment, we introduce the matrix term c_{sk} and the vector element a_k:

$$\lambda_{sk} \;=\; \gamma_k \;+\; c_{sk} \cdot a_k \tag{6}$$

Here, we have defined a_k as the environmental configuration of the kth niche factor, $a_k = \ln\sqrt{U_k/(1-U_k)}$, with U_k a realization of the random variable $f(x)$. We define c_{sk} as the *orientation* of species-s to niche factor-k: if species-s responds positively to niche factor-k then $c_{sk} = +1$ and if it responds negatively to niche factor-k then $c_{sk} = -1$. Specifically, $c_{sk} = +1$ if the kth digit of s (in binary) is 1, otherwise it is -1.

In the logarithmic domain, the niche volume of Equation (1), becomes a sum and, for species-s, the overall niche volume can be written by substituting (6) as follows:

$$v_s = \ln V_s = \sum_{k=0}^{K-1} \lambda_{sk} = \sum_{k=0}^{K-1} \gamma_k + \sum_{k=0}^{K-1} c_{sk} a_k \tag{7}$$

The community niche volume can be written in matrix form:

$$\mathbf{v} = K\bar{\gamma} + \mathbf{C} \cdot \mathbf{a} \tag{8}$$

Where $\bar{\gamma} = {}^1\!/_K \sum \gamma_k$. Thus, the log-hypervolume of the niche is a constant plus the product of a matrix representing species traits with a vector representing the environmental configuration.

Population Dynamics and Turnover

We associate the niche hypervolume with population size. So, if the entire community size is J, the number of individuals in species-s is the niche hypervolume times J, so $X_s = JV_s$ and for the dynamics of the log-population vector $\mathbf{x} = \ln J + \mathbf{v}$ we have:

$$\mathbf{x}(t) \;=\; \ln J + \mathbf{v}(t) \;=\; \ln J + K\bar{\gamma}(t) \;+\; \mathbf{C} \cdot \mathbf{a}(t) \tag{9}$$

The time-dependent *environmental configuration vector*, $\mathbf{a}(t)$ of size K, expresses the changing environment, while the *orientation matrix* $\mathbf{C}$ (of size $S \times K$) describes the association between the species traits and this changing environment. This matrix is constant and is easy to construct. Its rows are simply all the numbers 0 to 2^K, written in binary form, with the ones remaining as +1 and the zeroes replaced by -1. The dimensions of this matrix are $2^K \times K$. Each species occupies one row and each column is a niche factor axis. Each column of $\mathbf{C}$ (or each row of $\mathbf{C}^T$) has a sum of zero.

An important question is how we describe environmental changes that drive $\mathbf{a}(t)$, the niche factor axes. The elements of $\mathbf{a}$ are the configurations of the niche factors of each axis. At each time-step, each component of $a_k(t)$ is based on U_k a realization of the random variable from the distribution $f_U(x)$, specifically $a_k(t) = \ln\sqrt{U_k/(1-U_k)}$. We assume that axes have separate timescales, namely each axis has different timescale of change, τ_k, see **Figure 2A**. Moreover, we will assume that timescales are equally spaced logarithmically, such as $\tau_k \in \{1, 2, 4, 8, \ldots, 2^{K-1}\}$ time units. We label the axes with indices 0,1,2,3,. . .,$K-1$ associated with the respective timescales 1, 2, 4, 8, . . ., 2^{K-1} units. We will write the vector as follows:

$$\mathbf{a}(t) = [a_0(t), a_1(t), ..., a_{K-2}(t), a_{K-1}(t)]$$

Thus, the element $a_k(t)$ of $\mathbf{a}(t)$ assumes a new state every 2^k units of time.

In this theory, we suppose that species' own population dynamics are not important — they respond to the changes in the environment and reach equilibrium rapidly. Certainly, within this framework, starting from Equation (8), more detailed population dynamics could be constructed at the price of greater complexity.

Figure 3 shows the typical pattern of fluctuations of the model. This is significantly different from the pattern where populations fluctuate within a basin of attraction around a fixed equilibrium. Here, the population remains within a "basin" only within a timescale. Long-term it departs from any basin of attraction. It is also different from populations in the neutral theory, where population trajectories that are essentially random walks (Brownian motion) that can drift anywhere within the bounds [0,J]. In the $1/f$ model (Halley and Inchausti, 2004), the populations can also drift but require a much longer time to depart far from the initial value.

The abundance for species-s is given by (8) as:

$$x_s(t) = K\bar{\gamma}(t) + \ln J + \sum_{k=0}^{K-1} c_{sk} a_k(t) \tag{10}$$

Equations (8) and (10) describes the dynamics of all the populations in the community in the absence of extinction and colonization or speciation. However, we can include local extinction or extirpation, which plays a major role in the ecological community, as a form of ecological turnover, offset by colonization. The population is extirpated if it falls below a minimum viable population, if $x_s(t) < m$, creating an empty niche. Recolonisation of an empty niche can be modeled by a

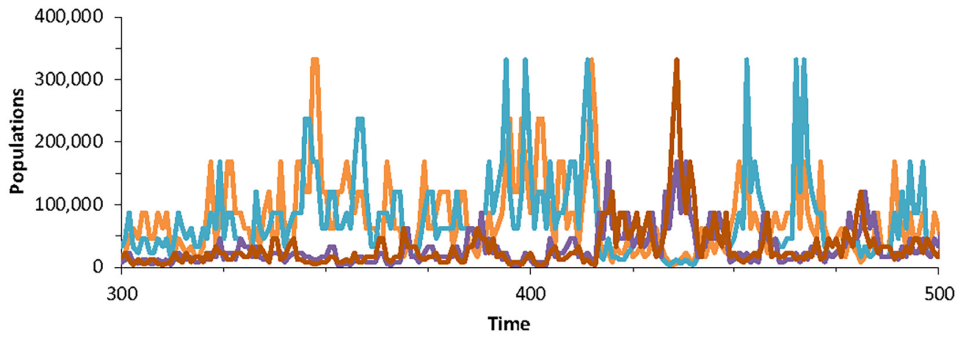


FIGURE 3 | Population dynamics of four randomly chosen species in the community model generated using Special case B, with parameters $K = 7$, with $J = 5 \times 10^6$ for 200 time-steps. The dynamics are according to Equation (10) and thus do not include extinction-colonization.

constant probability. Thus, the overall population dynamics can be expressed in island-biogeographical terms. For any species s, the probability that the species is present at any time t changes as follows:

$$p(t+1) = c + [1 - c - e_s(t)]\, p(t) \tag{11}$$

Here c is the probability of colonization per unit time and p is the probability that the species is not extinct at time t. The probability of extirpation at time t is $e_s(t)$ is the probability of the log-population being less than m, namely:

$$e_s(t) = \int_{-\infty}^{m} f_{xs}(z)dz \tag{12}$$

This equation differs from the usual equation of island-biogeography because e_s is time dependent. If extinction occurs, then the niche becomes empty. Thus, in the presence of extinction, we rewrite Equation (10) as:

$$x_s(t) = \begin{cases} K\bar{\gamma}(t) + \ln J + \sum_{k=0}^{K-1} c_{sk} a_k(t)\ , & \forall x_s(t-1) \geq m \\ < m\ , & \forall x_s(t-1) < m \end{cases} \tag{13}$$

Example. Suppose $K = 3$, $J = 1000$, $B_0 = B_1 = B_2 = 3$ (*Special Case B*). Then there are 8 species, $\Gamma = B/(1+B^2) = 0.3$, and the two values of niche width are $L_s(t_0) = 0.9$ or 0.1, from Equation (2A). The 8 niches have 4 possible sizes: 729, 81, 9 and 1 with one, three, three and one species, respectively. Thus, at any time there will be one niche with 729 individuals, three with 81, three with 9 and a single niche with one individual. The identities of the niches will move around according to the environmental signal, which we assume to be at $t = t_0$ and $\boldsymbol{a}(t_0) = \ln B \times (-1, -1, -1)^T$. In the log domain, we have $\ln B = 1.10$ and $\gamma \approx -1.20$, while the orientation matrix $\mathbf{C}$ is:

$$\mathbf{C} = \begin{bmatrix} -1 & +1 & -1 & +1 & -1 & +1 & -1 & +1 \\ -1 & -1 & +1 & +1 & -1 & -1 & +1 & +1 \\ -1 & -1 & -1 & -1 & +1 & +1 & +1 & +1 \end{bmatrix}^T$$

Thus, using Equation (8), the population vector $\mathbf{x}$ at time t, is given by:

$$\mathbf{x}(t) \ \cong\ 3.30 + (1.10)\,[-3, -1, -1, +1, -1, +1, +1, +3]^{\mathrm{T}}$$

Note, that no reference is required to the previous states of the population here. Thus, in the absence of extinction, the states of the population are given entirely by the environmental process vector $\mathbf{a}(t)$.

The Zero-Sum Property and Linearity

One of the desirable properties for a model of the ecological community is the zero-sum property. This means that changes in populations add up to a constant community size. For example, this is also an implied property of neutral model that also has the attribute of a constant number of individuals. Here, as all niches are a partition of the unit hypercube then, by definition, the total hypervolume of all niches is unity. Thus, in Equation (8) even though the relative sizes of niches change, the overall community size always adds up to J individuals.

Surprisingly, the log-populations in this niche-space may also have the zero-sum property. Even if the abundances themselves are zero-sum, there is usually no reason for the log-populations to follow suit. To see how this second zero-sum property may emerge, we find the total community size by summing Equation (10) over all species:

$$\begin{aligned} \sum_{s=0}^{S-1} x_s(t) \ &= \ \sum_{s=0}^{S-1} \left(K\bar{\gamma}(t) + \ln J\right) \ + \ \sum_{s=0}^{S-1} \sum_{k=0}^{K-1} c_{sk} a_k(t) \\ &= \left(K\bar{\gamma}(t) + \ln J\right) S \ + \ \sum_{s=0}^{K-1} a_k(t) \sum_{k=0}^{s-1} c_{sk} \\ &= K \cdot S \cdot \bar{\gamma}(t) \ + \ S \cdot In\, J \end{aligned} \tag{14}$$

The final term, in the second line of (14), disappears because we have a summation over a column of the orientation matrix, which sums to zero (see section "Population Dynamics and Turnover"), something that pertains irrespective of which distribution $f(x)$ holds for niche factors. The second term on the last line of (14) is constant in time, but not the first term. This means that, in general, the total community size changes with time and therefore there is no zero-sum property. For Equation (14) not to change, γ_k should remain constant. The form of γ_k, given by Equation (5), shows that γ_k will remains constant in time provided $U(t)[1-U(t)]$ does not change. This happens

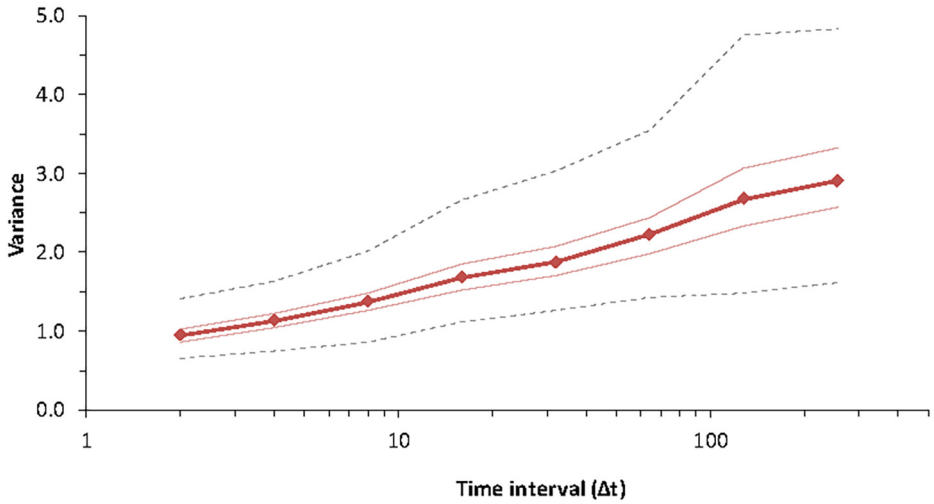


FIGURE 4 | Pattern of variance growth in the model with observation time for Special Case B for species #3 (chosen at random) of the community, with parameters $K = 7$ and $B = 1.9$, $J = 4 \times 10^4$ for 256 time steps. The lines indicate the median, minimum, maximum and the two hinges (25th and 75th percentiles) for 1,000 replicates. The dynamics proceed as for Equation (10) and do not include extinction or colonization.

for a restricted class of distributions, including the dichotomous distributions of Special cases B and C. For such distributions, all population changes can be expressed relative to the mean log-population, $K\bar{\gamma} + \ln J$, so that Equation (10) can be written as:

$$\mathbf{z}(t) = \mathbf{C} \cdot \mathbf{a}(t), \quad \mathbf{z}(t) = \mathbf{x}(t) - \left(K\bar{\gamma} + \ln J\right) \tag{15}$$

This is important because it means that for this class of distributions of environmental change, population changes in the community can be related *linearly* to changes in the environment through Equation (15), which is a major mathematical simplification.

Variance in Time

According to Pimm (1991), the ecological community is fundamentally non-stationary. Subsequent studies of variance as a function of time in ecological populations found that this increase of variance was universal (Inchausti and Halley, 2001, 2002). The model of variability most consistent with ecological time-series was a $1/f$-noise model for which variance increases linearly with the logarithm of time (Halley, 1996; Halley and Inchausti, 2004). It is thus important to see if such a pattern can arise in a community model (**Figure 4**).

Here, the relevant question is how the population will fluctuate on short timescales ($\Delta t << 2^K$). Let us now consider the time series of length $T << 2^K$ for species-s with log-population $\{x_s(1), x_s(2),\ldots, x_s(t),\ldots,x_s(T)\}$. Let us assume that each niche-factor has the same variance $\mathrm{Var}\{U_k(t)\} = V_0$. The variance of this time series is the expected value of the variance of (10), which is:

$$\begin{aligned} Var\{x_s(t)\} &= Var\left\{\ln J + K\bar{\gamma}(t) + \sum_{k=0}^{K-1} c_{sk}a_k(t)\right\} \\ &= \sum_{k=0}^{K-1} Var\{\gamma_k(t) + c_{sk}a_k(t)\} \end{aligned} \tag{16}$$

Our problem reduces to finding the variance of one of the components $\gamma_k(t) + c_{sk}a_k(t)$. This is done rigourously for Special Case B in **Appendix B** but the following more general argument holds for all kinds of distributions.

If the cycle time is very long there is no variance because the series length is much shorter than the cycle time, that is $T << \tau_k$, so the components associated with axis-k are unchanging and so the variance contribution is zero. At the other extreme, if the cycle time is very short, shorter than the distance between successive samples, then $\tau_k << T$ and we have:

$$\begin{aligned} &Var\{\gamma_k(t) + c_{sk}a_k(t)\} \\ &= \tfrac{1}{2} Var\left\{\ln U_k + \ln(1-U_k) + c_{sk}\ln U_k - c_{sk}\ln(1-U_k)\right\} \\ &= \tfrac{1}{2} Var\left\{(1+c_{sk})\ln U_k) + (1-c_{sk})\ln(1-U_k)\right\} \\ &= Var\left\{\ln U_k\right\} = V_0 \end{aligned}$$

The last line results because, since all the c's are ± 1, then either we have $\mathrm{Var}\{\ln U_k\}$ or $\mathrm{Var}\{\ln(1-U_k)\}$, so provided the distribution of $\mathrm{Var}\{\ln U_k\}$ and $\mathrm{Var}\{\ln(1-U_k)\}$ are the same, we have the simple result for $\tau_k << \mathrm{T}$. Thus,

$$\begin{aligned} Var\{\gamma_k(t) + c_{sk}a_k(t)\} &= 0, \quad \tau_k >> T \\ &= V_0, \ \tau_k << T \end{aligned}$$

Thus, as a series gets longer, as it reaches each time-constant $\tau_k = 2^k$, its variance increases by V_0. In the case where the time-constants are arranged uniformly on the log-axis, then the variance increases linearly in logarithmic time, as is the case for $1/f$ noise (Inchausti and Halley, 2001):

$$W_s \approx \sum_{2^k < T} V_0 \approx V_0 \log_2 T \tag{17}$$

This is equivalent to saying that variability increases with time — the "more time means more variation" effect

(Pimm and Redfearn, 1988; Inchausti and Halley, 2002). This log-pattern is characteristic of $1/f$-noise, which was the model most consistent with ecological time series, for which variance increases linearly with the logarithm of time (Halley, 1996; Halley and Inchausti, 2004). This result can be used, if we have an empirical time-series, to estimate V_0 from the rate of increase of variance. However, if the time constants are arranged otherwise, we could observe other patterns of variance increase (Halley and Inchausti, 2004).

The Species Abundance Distribution

Within real communities, there are many rare species and a few common ones. The ensemble species abundance distribution at any fixed time is often observed to be lognormal. Some have questioned the suitability of the lognormal (Williamson and Gaston, 2005). McGill (2003) argued that many theories generate such bell-shaped distributions and having a lognormal species abundance distribution is a weak test for a model. Nevertheless, the lognormal pattern is so frequently encountered in field data (Ulrich et al., 2010) that it constitutes an important model property. The niche-hypercube model leads naturally to a lognormal species abundance distribution.

In our model, the cube is divided into two pieces at a random point, $0 < U_1 < 1$. Then the process is repeated independently along a second axis to get $0 < U_2 < 1$ (**Figure 1**). At this point, the cube contains four sub-cubes. Repeating the process in higher dimensions creates new sub-cubes, 8, 16 and so on. For a process of k repetitions, we have 2^k segments in total. Without loss of generality, a sub-cube has a volume given by Equation (1). Since the U_k's are random variables, it follows from Equation (2) that each L_{sk} is a random variable on [0,1]. If we take logarithms, we get Equation (7), where each $\lambda_{sk} = \ln(L_{sk})$ is a random variable with a distribution $g(x)$ on $(-\infty,0)$. Suppose this distribution has an expected mean of μ and an expected variance of σ^2. Then we can use the central limit theorem, provided all the moments of the distribution $g(x)$ are finite, to show that the sum of k of these (for large k) will be approximately normal, with mean $k\mu$ and variance $k\sigma^2$, so that:

$$v_s \sim N(k\mu, k\sigma^2) \tag{18}$$

The fragments may be considered as species and their size the number of individuals of that species. If there are J individuals in the community any species has a population Je^z. This is a lognormal random variable.

In *Special Case A*, the distribution $f(x)$ is uniform on [0,1], while the distribution of the logarithm of this is exponential with a mean of -1 and a variance of $+1$. Thus, the distribution of the niche hypervolume v_s is expected to be normal:

$$f_{vs}(x) = \sqrt{\frac{2}{\pi K}} \exp\left[-\frac{(x-K)^2}{2K}\right] \tag{19}$$

This means, from Equation (9), that x_s is also distributed normally and hence that N_s has a lognormal distribution.

In *Special Case B*, the distribution $f(x)$ has an expected mean of γ and an expected variance β^2. Thus, the distribution of the niche hypervolume is expected to be:

$$f_{vs}(x) = \sqrt{\frac{2}{\pi\beta^2 K}} \exp\left[-\frac{(x-K\gamma)^2}{2K\beta^2}\right] \tag{20}$$

Again, x_s will be distributed normally, so N_s has a lognormal distribution.

Both the species abundance distribution and the distribution of abundance in time are shown in **Figure 5**. The time series for a fixed species s at different times, namely of the set $\{x_s(1), x_s(2), x_s(3), \ldots, x_s(T)\}$ is found to be lognormal. This might be expected because each species in time traces out a sort of random course within the niche space defined by the species abundance distribution. However, in this case, the variables are not independent between different time-steps. Clearly, if T is very large, then the time-series will yield a distribution following (23), so we should expect this limit of the distribution in time to follow a lognormal.

APPLICATIONS USING SIMULATION

The examples in this section all use Special Case B. The biodiversity of the community typically determines the number of dimensions K that must be used, since $S = 2^K$. The central parameter B is the measure of the fundamental variability of each niche factor and of the community. We can estimate this from the scaling behavior of population fluctuations. Together, these parameters describe the community and how it interacts with the environment.

In problems involving population dynamics, we can define the number of individuals in the community, J. We can calculate this for various types of organisms as a density of individuals per unit area. Thus $J_0 = \rho A_0$. However, the magnitude of J becomes important mainly when relating problems such as extinction, for which we must also choose a value for the minimum viable population, m. Given the choices of the parameters above, this fixes the number of extinctions expected in each time-step of the model. It is also possible to choose the colonization probability.

Once the foregoing parameters have been chosen, the others follow and further changes in the model relate to the specific questions being asked. In problems involving community dynamics, one of the most difficult tasks is fitting the community observed to the niche hypercube. Although the explicit correspondence of the axes to specific species traits is not needed, it is seldom possible to find an integer K corresponding to the known species in the community. Provided that $2^K \geq S_0$, there are more niches in the hypercube than species, so some cells have zero population (empty niches).

For the applications which follow, we parameterised simulation models in C++ and an abridged version on an Excel spreadsheet.

Response to Habitat Reduction and Extinction Debt

When human actions reduce the size of a habitat, there follows a decrease in the overall population in the community. This

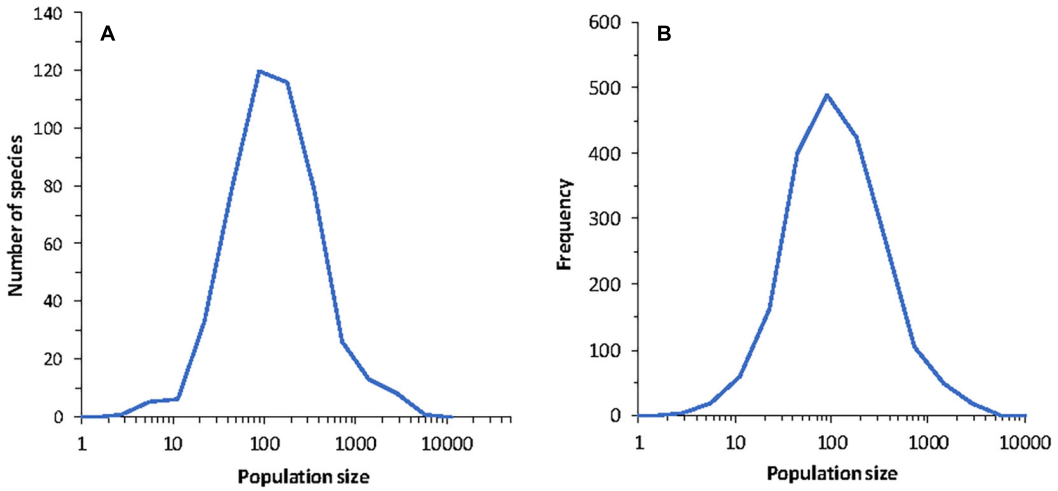


FIGURE 5 | (A) Species abundance distribution for a single community in one run of the model. **(B)** Histogram of population size in time for a run covering 2,000 years for a randomly chosen species from the community. The parameters of the model were $J = 10^5$, $K = 9$, $B = 1.5$ and MVP = 2. The dynamics are according to Equation (10) and do not include extinction or colonization.

leads to a decrease in the number of species, as reflected in the species-area relationship. This loss of species does not follow immediately. The community may appear to get along for a while without the loss of any species. Later, however, these extinctions will happen – this is the phenomenon of extinction debt. Extinction debts have been observed or inferred over a range of scales from months to thousands of years and depend on various covariates of which the strongest is the area of the habitat (Halley et al., 2016). One of the challenges of modeling extinction debt is this range of timescales, from a few weeks to thousands of years, virtually into evolutionary time.

From a theoretical perspective, two types of community mechanism explain extinction debt. One of these features in the original extinction debt study by Tilman et al. (1994) based on a Tilman's spatial competition model (Tilman, 1994). In this model, the delay arises from the combined decline of various species found with metapopulations below the threshold when habitat is lost. An alternative approach was based on neutral drift (Halley and Iwasa, 2011). In a neutral model, all species perform random walks due to demographic stochasticity. When habitat is lost and the community shrinks, the locus of each species' walk is forced closer to zero. As a result, species all have a higher probability of extinction, not matched by colonization. While this forces the community to lose species, it takes time. The dynamic hypercube model in this paper provides a third theoretical mechanism that leads to delayed extinctions. We can view the loss of area of a community as a general loss of hypercube volume. This reduces the populations of all species proportionally, including a loss of the lower limit. In our model, at any time, there is a species at the minimum population size, where it is vulnerable to extinction. As in the neutral-drift mechanism, this is a stochastic mechanism.

Delayed extinction has major implications for conservation. For example, how much time will pass before the extinctions start to happen? How long before the extinction debt is "paid"? Typically, we try to predict the rate of loss of species after a sudden loss of area, assuming we know the initial number of species. This problem has been addressed for the neutral model several times (Gilbert et al., 2006; Halley and Iwasa, 2011; Hugueny, 2017). The advantage of the model presented here for this kind of problem is that it is suitable for looking at a wide range of timescales.

We can model extinction debt by simulation. For this, the key question is the relationship between minimum viable population, m, and the minimum niche size, x_{min}. If $m > x_{min}$, then the characteristic timescale is a single time-step since every time-step a different species is potentially subject to the minimum niche size. On the other hand, if $m < x_{min}$ then extinction never happens (without an extra noise term). It thus makes sense that m is related to x_{min}. We will assume that if $m = x_{min}$ then there is a probability of extinction ε, the probability of extinction at the minimum viable population.

Figure 6 shows the results for a specific set of parameters, depicting the decay in species number over time. Compared to the simplest expectation — exponential decrease in the number of species — the trajectories we predict are substantially different. An exponential curve that had the same expected number of species losses at the mid-point of the time considered (i.e., 50 years) suggests more species would survive in the short-term, but few in the long-term. Put another way, our model suggests a transient rapid loss of species, followed by a deceleration with eventually few losses. The trajectory is also different from the neutral model because the dynamic hypercube model yields losses on a relatively short timescale. It is notable that, overall, the pattern of biodiversity decrease follows the *logarithm of time*.

It is worth noting that our model is like that Hubbell's neutral theory, as model of *unstable* coexistence, where species eventually drift to zero unless there is a process of speciation or colonization (Huston, 1979; Chesson, 2000).

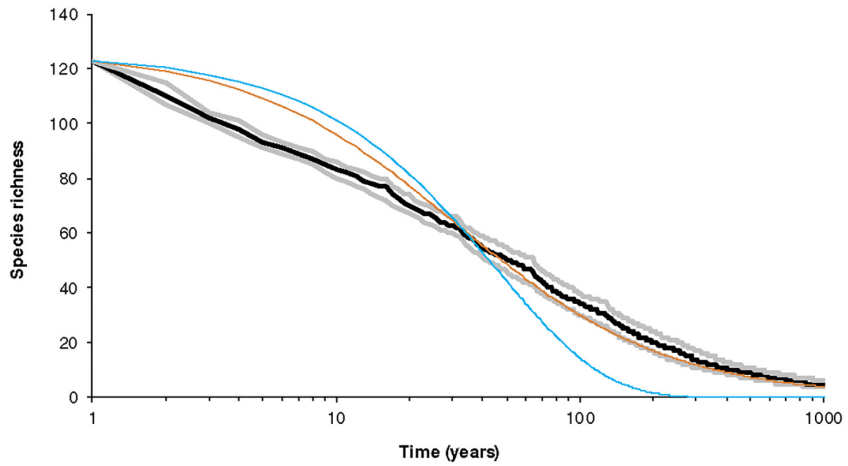


FIGURE 6 | Pattern of survival as a function of time after habitat is reduced suddenly (at $t = 0$). The area initially has 123 species in broken-stick configuration with each species having a large population. The reduction in area is such that the reduced total number of individuals is 13,000 so the average population falls to just above 100. The model has $K = 8$ with $B = 1.5$ and the minimum viable population is set at $M = 2$. This figure summarizes 1,000 simulations. The curves correspond to the five quartiles of species richness at each year. Notice that decay occurs approximately linearly on a logarithmic scale. The black line is the median and the gray lines the hinges (25th and 75th percentiles). The other curves are the equivalent exponential decay (blue) and hyperbolic decay (brown, for neutral theory) with the same S_0 and the same t_{50}. The dynamics are according to Equations (11)–(13) and thus includes extinction. However, colonization is assumed to be zero.

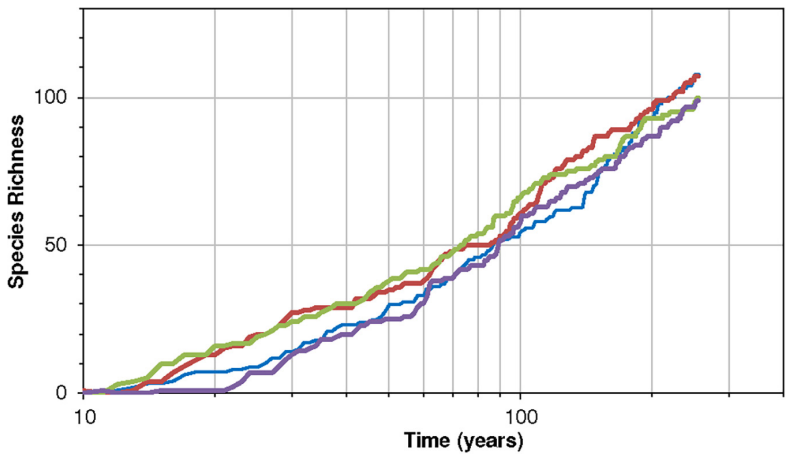


FIGURE 7 | Four realizations of the species-time relationship for the community, with parameters $K = 7$ and $B = 1.7$, $J = 4 \times 10^4$ for 256 time-steps and the minimum viable population is $M = 11$. Notice the near linearity of the rise on a logarithmic scale of time (years). The probability of colonization is constant at 0.007 per species per year. The dynamics are according to Equations (11)–(13) and thus include extinction-colonization.

A Species-Time Relationship

The species-time relationship describes the increase in the number of species recorded at a site as that site is observed for increasingly long periods of time (Rosenzweig, 1995). In some ways the species-time relationship is the reverse of the species relaxation curve discussed above, though typically the emphasis is on the sampling aspects of the problem rather than recolonisation issues. This is because a process of Poisson sampling a fixed community at regular intervals will yield a steadily increasing cumulative number of observed species. For example, the corresponding author's interest in Greek endemic orchids resulted in a collection of species on the University of Ioannina campus beginning with 5 in 2008. Thereafter sequence of annual new species was {0, 0, 4, 0, 1, 0, 4, 1, 1, 1} reaching a cumulative total of 17 in 2018. The randomness of a Poisson process does not convincingly explain this increase. The increase in species richness is clearly being driven by a combination of sampling processes and ecological processes on various scales (Preston, 1960; White et al., 2005). In contrast to the well-known species-area relationship, the species-time relationship has received relatively little attention. The evidence from wider studies of ecological communities suggests that species-time relationships are well fit either by power functions or by logarithmic functions. Large-scale analyses of species-time relationships have revealed both of these (White et al., 2005).

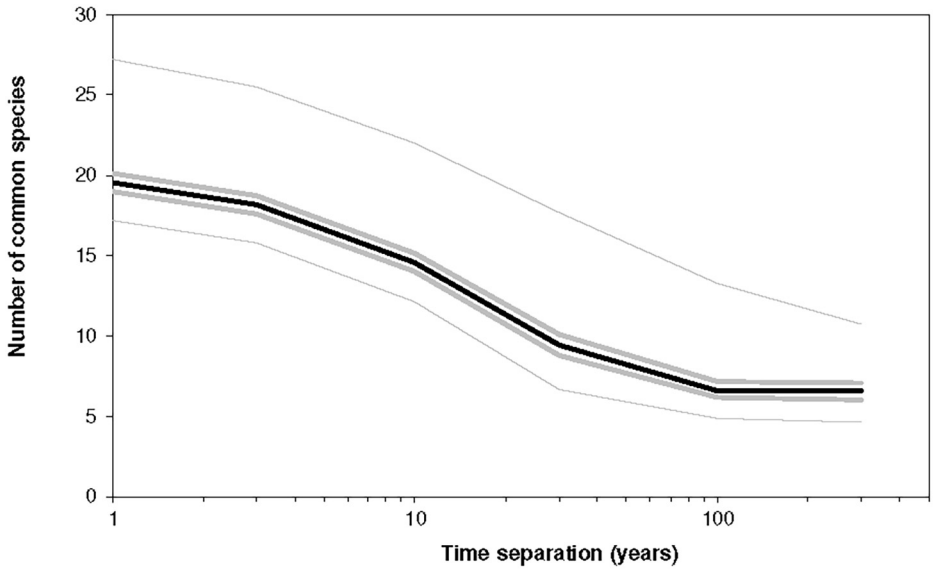


FIGURE 8 | The number of common species between two instances of a community separated by Δt years for different values of Δt. The parameters were $K = 6$ and $B = 1.5$, $J = 10^4$, and the simulation covered 750 time-steps. Initially, the community was unoccupied, but each niche was subject to a colonization probability of 0.02/year with an MVP of 22. The average diversity in any year (for the interval $t = 50$ to $t = 750$) was 21.3 ± 4.4 species, and the population was on average 4,000. The lines indicate the median, minimum, maximum and the two hinges (25th and 75th percentiles) for 1,000 replicates. The dynamics are according to Equations (11)–(13) and thus include extinction-colonization.

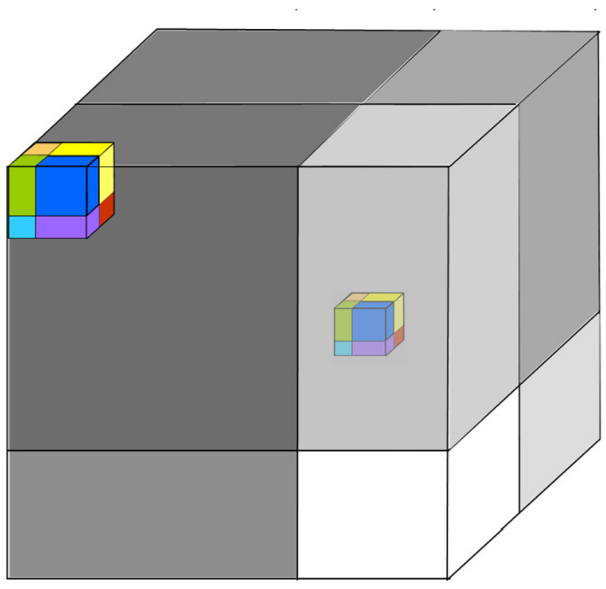


FIGURE 9 | On a short timescale, the initial three axes are represented by a set of species in a compact niche, this changes and the cube moves around the hypercube on longer timescales.

The species-time relationship is also very important in cases of real ecosystem restoration. For example, taking the case of the restoration of European forests after the ice age, Clark and Mclachlan (2003) showed that variance associated with population and species drift slows down or appears to reach an asymptote. In contrast the neutral model exhibits variance increasing almost linearly with time. A more niche-based model is thus needed to investigate problems like this. Clark and Mclachlan (2003) attributed the diminishing growth to a mechanism of density dependence and stabilization, other mechanisms may also contribute to this effect. Our model can explore this — it provides a powerful but simple niche-based dynamic framework to explore problems of this type. **Figure 7** here shows a species-time curve for an environment that begins with no species but allows colonization to happen.

The above pattern exhibits a logarithmic type of increase in species richness with time. It demonstrates that a decelerating incidence of new species with time need not only imply a stabilization toward an asymptote or an equilibrium of the logistic variety. A slowing increase in species numbers also has a non-stationary interpretation because the $1/f$ community embodied in this model is non-stationary.

Community Turnover, Speciation and the Drift Into Evolutionary Time

One of the most interesting predictions of MacArthur and Wilson's theory of island biogeography is that island communities, due to their rapid rates of extinction and recolonisation, are in a permanent state of flux regarding community composition. An observer returning to a small island, after a long absence, should find the new community virtually unrelated to the previous one. Similar community drift is also predicted in the neutral model because the species are constantly drifting in abundance relative to one another. Thus, in neutral theory, due to internal drift and lack of niche definition,

any species can come to dominate the community. In fact, the drift associated with neutral dynamics has been used as a null model to calibrate the significance of changes attributed to global warming (Sgardeli et al., 2016). Both models feature community drift by virtue of their neutrality by virtue of not having niches. The absence of niches means that the community is not fixed and may drift between different structures.

The current model also generates community drift, but the drift means something different in this niche-based case. The model is neutral in a sense that no trait is associated with dominance, and so any of its species can dominate. It depends on which niche is pushed to the largest size by the environment. The drift of the community between different species compositions is also here. **Figure 8** shows the average overlap as a function of time-separation.

In this figure $K = 6$, so In this case the trait vector has the form $(c_1,c_2,c_3,\ldots,c_6)$ and the community has a potential size of 64 species and 10,000 individuals. Due to a high value of B, which leads to a high variability, coupled with a high value of minimum viable population ($M = 22$), the extinction rate was high. As a result, there were only ∼20 species present at any time, and so the average population was only ∼4,000, well below the carrying capacity of 10,000.

As the environment changes, different groups of species are favored. We can associate the twenty species in the community with a cluster of niches with larger volume. Each year the environment changes and causing some niches to shrink and others to grow. As this happens, those that shrink become more likely to lose their population while those that grow can be recolonised. The slowest of the processes changes on a scale of 64 years. Thus, on scales greater than 64 years, changes in the environment just retrace at random the previous configurations; there is no real evolution in the environment. So, for time separations above 64 the difference between the two communities does not decline further.

However, we can extend the current model to include variability on longer timescales. We can associate the still-longer timescales of environmental forcing in the evolution of other traits, ones that appear fixed on shorter timescales (Jackson and Overpeck, 2000). Thus, a system with a 25 traits and a vector $(c_1,c_2,c_3,\ldots,c_{25})$ and a maximum timescale of 3.35 M years, reaches deep into evolutionary time. This includes a space of many millions of possible niches, which may be much larger than the community that interests us. The active community may be large enough to support only a thousand species. How can we visualize this? The initial community is inhabiting a hypercube in 10 dimensions, in which the traits are all changing, resulting in changing dominance hierarchies. However, on timescales above a thousand years, other traits will start to change too. What will happen, is that the initial community is no longer so competitive relative to some other species whose niche lies outside the initial hypercube. These species were not initially present in the community, but now there is an incentive for them to colonize or be created. Thus, the community will tend to drift according to the changing environment. Thus, if we fast-forward a few million years, the initial community moves within the larger hypercube to a different center (**Figure 9**). We still have a community whose short-term variability involves traits $c1$–$c10$ but other traits $c10$–$c25$ will also have changed so that very few species remain the same as for $t = 0$. This reflects the adaptive landscape of Wright as noted by Holt (2009) and Blonder (2018), but this is a dynamic adaptive landscape.

Thus, we can see the ecological community slowly drifts away from its initial configuration. This shows that the current model can look at the ecological community in evolutionary time also. Hutchinson's concept unifies ecological and evolutionary time. So, the community $1/f$ model described in this paper also predicts community drift on a longer evolutionary timescale. In contrast to the McArthur–Wilson and neutral model models, it is possible to identify traits. Part of this model is that there are processes in the environment that impact organisms that are stationary on a short timescale. They only change on longer timescales. But as they do, their influence is to change the balance toward organisms that were previously marginal. In contrast, organisms that were previously well-suited to the environment now find themselves "on the margins."

GENERALITY OF RESULTS

An important question that should be asked is how robust are the results we have obtained in the face of structural changes? The model we have developed rests on some fundamental assumptions that cannot be changed and other assumptions that are less critical which affect in different ways the results obtained for lognormality, variance growth and linearity.

The model is based on four fundamental structural assumptions. Foremost is the conception of Hutchinson's niche itself, namely that the space in which a species operates can be defined by an area of hyperspace. A discussion by Leibold covers many of the difficulties this implies (Leibold, 1995). Most important among these is: since so many aspects of any species requirements are tied to interactions with other organisms, how can it be tied instead to a relatively fixed structure like a hypervolume? The main justification has always been that the simplification enabled by Hutchinson's conception brings worthwhile results. Secondly, we have assumed a cuboidal geometry for niches. This contrasts with much of the work being carried out by researchers who are trying to build up a picture of the niche in more applied contexts such as discussed in the Section "Introduction." Thirdly, we have assumed that organisms respond to each niche factor in a binary way: either a positive niche factor helps or hinders each species. This is obviously a simplification for the nuanced way that organisms respond to changing environment – real responses are rarely binary in this way. Finally, we assume that the states of the axes of the hypercube change independently. This too might seem restrictive for things like temperature and moisture that clearly influence one another. We argue that all these fundamental assumptions enable major simplifications in tractability, in parameterization and also enable intuitive understanding and visualization.

The most important additional assumption of the model is the choice of the distribution of niche factors, U_k on

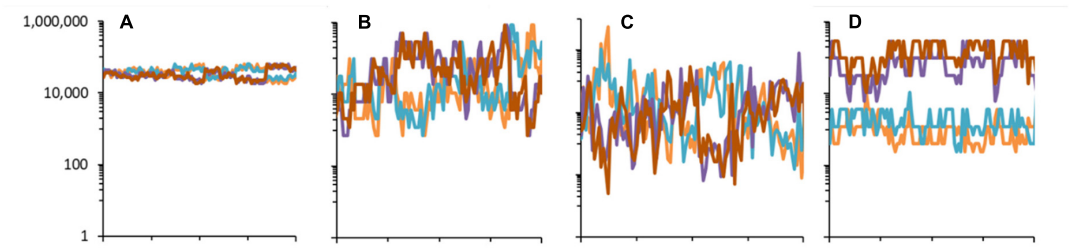


FIGURE 10 | Population dynamics of four randomly chosen species in the community model generated using different distributions with parameters $K = 7$, with $J = 5 \times 10^6$ for 512 time-steps. **(A)** Special case B with $B = 1.15$. **(B)** Special case B with $B = 1.73$. **(C)** Special case A (uniform distribution for U). **(D)** Special case C with $B = 1.73$, with asymmetry $q = 0.2$. The dynamics are according to Equation (10) without extinction-colonization.

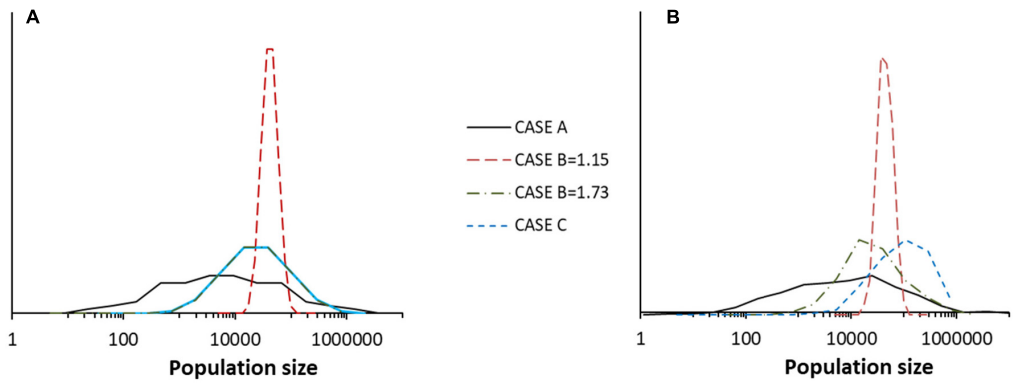


FIGURE 11 | (A) Species abundance distribution for a single community in one run of four different models taken at time $t = 511$ in each case. The models were: Special case A, Special Case B with $B = 1.15$, Special Case B with $B = 1.73$ and Special Case C with $q = 0.2$ and $B = 1.73$. **(B)** For the same four models, the histogram of population size over time for a run covering 512 years for species number 126 in each case. The other parameters of the models were $J = 10^5$, $K = 7$, and MVP = 2. The dynamics are according to Equation (10) without extinction-colonization.

TABLE 2 | The main additional assumptions required to produce the results in this paper.

Result	Required assumption				
	Identical *f(u)*	**Dichotomous *f(u)***	**Timescale separation**	**Timescale geometric spacing**	**Extinction-colonization**
Lognormal SAD (Equation 18)	X				
Lognormal in time (**Figure 5B**)	X				
Zero sum for log-populations. Linearity (Equation 15)	X	X			
1/*f* noise type variance increase			X	X	
Extinction debt (see section "Response to Habitat Reduction and Extinction Debt")					X
Log species time increase (see section "A Species-Time Relationship")			X	X	X
Community turnover (see section "Community Turnover, Speciation and the Drift Into Evolutionary Time")					X

An 'X' in any box implies that the associated assumption is required to produce the result.

[0,1], for the different axes. The distributions have been assumed to be the same for all axes, though this could be relaxed at the expense of losing some results (e.g., for lognormality). The choice of distribution affects the dynamics considerably.

From **Figure 10**, It is notable that in the case of the models of Special Case B, the parameter B is an asymmetry parameter, increasing B increases the difference between species population sizes and the variability in time. In special cases A and B, there is no asymmetry between states of the niche factors. This means

that every species, long-term, has the same average population; there are no privileged species. Introducing asymmetry between states ($q \neq 0.5$) means that the environment spends longer in some states than in others. Ecologically, this is clearly possible, as with moisture in an extreme desert, for example. Asymmetry, in this model, causes some states to have a greater average population size than others.

When deriving Equations (1–12), we made no special assumptions about the distribution of U_k, so these will hold in general. This is also true for Equation (18), the lognormality for species abundance distributions, which emerges from the central limit theorem. This only requires that there be a large number of axes and that the moments of $\ln(U_k)$ are finite. Thus, lognormality is a very general property of this model. **Figure 11** shows this through the results of running different models.

In **Figure 11**, it is worth noting that the species abundance distribution (distribution of species population sizes in the community) conforms to a lognormal as shown by Equation (18). In simulations, this was also found for the distribution of species population sizes in time. It is also worth noting that in simulations, these two types of distribution were indistinguishable except in Special Case C, for which the distribution in time has a higher average value. This can be explained by the fact that in this asymmetric model, species have different distributions; some species have higher average populations than the overall community average.

The zero-sum property of log populations and the linear dependence of population on the environment, Equation (15), represent an important analytical simplification. This depends on the term $\bar{\gamma}(t)$ being constant or not changing much. This only holds for a restricted set of models. For example, it does not hold for Special Case A but it does for Special cases B and C. In general, to secure constancy in time for $\bar{\gamma}(t)$, we need $U(t)[1 - U(t)]$ to remain constant in time, which is true for models where $f(\mathrm{u})$ has a dichotomous distribution. From a biological perspective there is usually no reason to expect environmental variables to have a dichotomous distribution, although we often experience things this way (temperature is "cold" or "hot", weather is "wet" or "dry"). The dichotomous distribution is thus a gross simplification for any environmental variable, but because of the Central Limit Theorem this may not matter too much. Even with such unrealistic distributions, model trajectories for populations are not so different from one using more biologically reasonable distributions.

The logarithmic increase of variance with time, as explained in Section "Variance in Time," does not arise from intrinsic structure of the hypercube but is related to the arrangement of time-constants on the different axes, with each axis having a separate and unique timescale of switching and with timescales spaced geometrically. This assumption reflects that fact that often environmental changes themselves occur evenly across octaves. Other arrangements, with the time-constants spaced differently would produce patterns of variance increase more like fractional Brownian motion or fractional Gaussian noise (Halley and Inchausti, 2004). This is also the case for the manner of increase of the species-time curve.

A summary of the results in this paper, regarding the model, and the assumptions required to produce them, is given in **Table 2** below.

While the simple switching between two states, producing "square" pulses (**Figure 2B**), may seem highly artificial, the resulting population changes, once all the niche factors are acting together, the artificial geometry of square pulses is no longer in evidence (**Figure 3**).

Equation (10) describes the population constraint caused by the environment but not the population response. The only part of this we have included is the extinction-colonization process modeled through Equations (11)–(13). Apart from these processes, this model does not include population dynamics, either interspecific or intraspecific. An easily envisaged extension of the model would be to include a demographic model in each niche, such as a simple discrete growth process with a carrying capacity, where the capacity value is equal to the niche hypervolume. The population dynamics within the niche would respond to changes in carrying capacity but there would be a delay due to the population dynamics. Interspecific interactions would be more challenging to incorporate but could be incorporated by including interactions between the species in different niches.

DISCUSSION AND CONCLUSION

We extend the niche-hypercube model of Hutchinson to form a niche-based dynamic hypercube (DH) model. This model is well-suited to exploring the dynamics of ecological communities from a niche-based perspective. The model has several features that make it desirable as a model of the community.

By dividing an ecosystem into a hypercube of finite size with a fixed number of organisms, the model reflects the fundamental limitation of energy input. Since only a fixed amount of energy arrives in an ecosystem, only a limited number of organisms of a given size can live there. Thus, when all niches are occupied, the populations add up to a constant community size. For a restricted class of niche fluctuations, this also holds for the logarithm of population; the fluctuations of the logarithm of population can then be expressed as a linear combination of the environmental fluctuations: $\mathbf{z}(t) = \mathbf{C}.\mathbf{a}(t)$, where $\mathbf{z}$ is a vector of all species' populations and $\mathbf{a}$ represents environmental fluctuations and $\mathbf{C}$ is a constant orientation matrix.

The zero-sum principle is shared with the Hubbell's neutral model among others. However, this model is not neutral in the same way as Hubbell's, since each species occupies a fixed place in the community "hypercube." Our model is still neutral with respect to biotic interactions. That is, competition is completely absent from this model and species do not interact in any other way. Their fitness is simply a by-product of their fundamental niche area, and there are no direct interactions between species.

For a very wide range of environmental fluctuations, this dynamic hypercube model generates a lognormal species-abundance distribution. Many theories generate similar distributions, bell-shaped on an octave scale, such as the zero-sum multinomial of the neutral model. Importantly, the dynamic

hypercube reproduces the widespread observation that most individuals in the community belong to a small number of abundant species. This model also produces time-series for populations that are lognormally distributed in time, something observed as the most common distribution for a large number of ecological communities. Thus, as a species traces out its trajectory of abundance, the resulting histogram will tend to be lognormal.

As presented here, one of the most important features of the dynamic hypercube model is that it need not be stationary in time, depending on the environmental fluctuations on the niche factors. Studies of ecological time-series have shown that the "more time means more variation" effect is almost universal and is most consistent with a $1/f$-noise model, where variance increases linearly with the logarithm of time. This type of environmental variability was used in the examples of this paper. The dynamic hypercube model provides a means to describe this variance growth in a community context. It shows that there is no balance of nature in one sense — as there is no permanent community. Nevertheless, neither is balance "discredited," because communities remain similar for a long time and because certain quantities such as species richness and energy flow are conserved long-term.

This paper outlines three different applications in ecology, where the dynamic hypercube model provides results and new insights. One of the obvious applications of such as model, with conservation implications, is studying the dynamics of extinction. In recent years, because of the biodiversity crisis, there has been great interest in extinction risk. For example, species distribution models may be used to estimate extinction risk without a population-level model of the actual process of extinction. This has meant that things like extinction debt get overlooked. Certainly, the neutral model has been used to address the problem. Nevertheless, arguably the neutral model and its modifications cannot address issues on all timescales because they do not take environmental variability into account. The model described here spans both ecological and evolutionary time by taking into account the drift of the environment itself. Theoretical studies such as ours still have "barely scratched the surface of the universe of possible genetic architectures, landscape geometries, and demographic scenarios at play in niche evolution" (Holt, 2009). Nevertheless, by characterizing the drift of selective forces by an environment fluctuating on many scales as (Jackson and Overpeck, 2000) we are opening a bridge.

It is also worth noting that the dynamic hypercube model may be used to throw light on functional diversity. Since each axis of the model can be associated with a trait, it fits in with this perspective by dividing the ecosystem according to functional axes. The usefulness of a trait to any species changes in response to environmental changes. In this picture, the importance of different ecological traits rises and falls on different timescales. Some do so very rapidly but others change so slowly as the be imperceptible in most ecological studies. They correspond to the invisible long-term axes of our model, which only reveal themselves on longer scales, reaching into evolutionary time.

In constructing our model, we made several assumptions. The binary aspect, that each niche factor changes in a strictly binary fashion, is an obvious limitation relative to the real world. Another important limitation is the requirement that that each niche factor axis changes according to a specific and limited timescale. As we know that all niche factors tend to fluctuate over a range of timescales, a more realistic assumption would be for this to describe all axes, but that would severely limit the mathematical tractability of the system. Also, there are challenges with the interpretation of the size of the system we are describing. It is not obvious how the hypercube should interact with population size. Suppose we change the ecosystem's size. Should this change the number of niches or just the size of number of individual organisms in the community? A change in the size of an ecosystem should immediately cause a loss of population, but should it lead to a contraction of the number of axes? Leaving aside the feedback element, we could associate a reduction of the axes with some physical reduction of the size of an ecosystem (e.g., a sea-level rise on an island), whereas the loss of the community within the hypercube reflects the killing of organisms without loss of habitat (e.g., a forest fire).

This model is a natural extension of the classical theory of the niche and of functional traits and brings many of the advantages of other dynamical theories such as the neutral model. It can be applied to numerous problems of community ecology to improve both explanatory and predictive power. It reveals how a community can have 1/f noise fluctuations and still be zero-sum. It is perhaps the first dynamic theory that links ecological time with evolutionary deep time.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author/s.

AUTHOR CONTRIBUTIONS

JH derived theoretical results and prepared figures. JH and SP wrote the manuscript. Both authors contributed to the article and approved the submitted version.

ACKNOWLEDGMENTS

The authors would like to thank reviewers DM and TK, also editor GB for helpful comments and insights.

REFERENCES

Adler, P. B., HilleRislambers, J., and Levine, J. M. (2007). 'A niche for neutrality'. *Ecol. Lett.* 10, 95–104. doi: 10.1111/j.1461-0248.2006.00996.x

Barabás, G., D'Andrea, R., and Stump, S. M. (2018). 'Chesson's coexistence theory'. *Ecol. Monogr.* 88, 277–303. doi: 10.1002/ecm.1302

Barros, C., Thuiller, W., Georges, D., Boulangeat, I., and Münkemüller, T. (2016). 'N-dimensional hypervolumes to study stability of complex ecosystems'. *Ecol. Lett.* 19, 729–742. doi: 10.1111/ele.12617

Bertuzzo, E., Suweis, S., Mari, L., Maritan, A., Rodríguez-Iturbe, I., and Rinaldo, A. (2011). 'Spatial effects on species persistence and implications for biodiversity'. *Proc. Natl. Acad. Sci. U.S.A.* 108, 4346–4351. doi: 10.1073/pnas.1017274108

Blonder, B. (2018). 'Hypervolume concepts in niche- and trait-based ecology'. *Ecography* 41, 1441–1455. doi: 10.1111/ecog.03187

Blonder, B., Morrow, C. B., Maitner, B., Harris, D., Violle, C., Enquist, B. J., et al. (2018). 'New approaches for delineating n-dimensional hypervolumes'. *Methods Ecol. Evol.* 9, 305–319. doi: 10.1111/2041-210X.12865

Carvalho, J. C., and Cardoso, P. (2021). Decomposing the causes for niche differentiation between species using hypervolumes. *Front. Ecol. Environ.* doi: 10.3389/fevo.2020.00243

Chase, J. M., and Leibold, M. A. (2003). *Ecological Niches: Linking Classical and Contemporary Approaches*. Chicago, IL: University of Chicago Press. doi: 10.7208/chicago/9780226101811.001.0001

Chave, J. (2004). 'Neutral theory and community ecology'. *Ecol. Lett.* 7, 241–253. doi: 10.1111/j.1461-0248.2003.00566.x

Chave, J., and Leigh, E. G. (2002). 'A spatially explicit neutral model of beta-diversity in tropical forests'. *Theor. Popul. Biol.* 62, 153–168. doi: 10.1006/tpbi.2002.1597

Chesson, P. (2000). 'Mechanisms of maintenance of species diversity'. *Annu. Rev. Ecol. Syst.* 31, 343–366. doi: 10.1146/annurev.ecolsys.31.1.343

Chesson, P. L., and Warner, R. R. (1981). Environmental variability promotes coexistence in lottery competitive systems. *Am. Nat.* 117, 923–943. doi: 10.1086/283778

Chisholm, R. A., and Pacala, S. W. (2010). 'Niche and neutral models predict asymptotically equivalent species abundance distributions in high-diversity ecological communities.'. *Proc. Natl. Acad. Sci. U.S.A.* 107, 15821–15825. doi: 10.1073/pnas.1009387107

Clark, J. S., and Mclachlan, J. S. (2003). 'Stability of forest biodiversity'. *Nature* 423, 635–638. doi: 10.1038/nature01632

Cuddington, K. (2001). 'The "balance of nature" metaphor and equilibrium in population ecology'. *Biol. Philos.* 16, 463–479. doi: 10.1023/A:1011910014900

Díaz, S., Kattge, J., Cornelissen, J. H., Wright, I. J., Lavorel, S., Dray, S., et al. (2016). 'The global spectrum of plant form and function'. *Nature* 529, 167–171. doi: 10.1038/nature16489

Dornelas, M., Gotelli, N. J., McGill, B., Shimadzu, H., Moyes, F., Sievers, C., et al. (2014). 'Assemblage time series reveal biodiversity change but not systematic loss'. *Science* 344, 296–299. doi: 10.1126/science.1248484

Engen, S., Lande, R., Aether, B. E., and Weimerskirch, H. (2005). Extinction in relation to demographic and environmental stochasticity in age-structured models. *Math. Biosci.* 195, 210–227. doi: 10.1016/j.mbs.2005.02.003

Engen, S., Solbu, E. B., and Sæther, B. E. (2017). 'Neutral or non-neutral communities: temporal dynamics provide the answer'. *Oikos* 126, 318–331. doi: 10.1111/oik.03707

Ergazaki, M., and Ampatzidis, G. (2012). 'Students' reasoning about the future of disturbed or protected ecosystems & the idea of the "balance of nature"'. *Res. Sci. Educ.* 42, 511–530. doi: 10.1007/s11165-011-9208-7

Fisher, C. K., and Mehta, P. (2014). 'The transition between the niche and neutral regimes in ecology'. *Proc. Natl. Acad. Sci. U.S.A.* 111, 13111–13116. doi: 10.1073/pnas.1405637111

Forte, B., and Vrscay, E. R. (1996). "No Title," in *Theory of Generalized Fractal Transforms*, ed. Y. Fisher (New York, NY: Springer-Verlag), 36.

Franzke, C., Barbosa, S., Blender, R., Fredriksen, H., Laepple, T., Lambert, F., et al. (2020). The structure of climate variability across scales. *Rev. Geophys.* 58:e2019RG000657. doi: 10.1029/2019RG000657

Gilbert, B., Laurance, W. F., Leigh, E. G. Jr., and Nascimento, H. E. (2006). 'Can neutral theory predict the responses of amazonian tree communities to forest fragmentation?'. *Am. Nat.* 168, 304–317. doi: 10.1086/506969

Halley, J., and Inchausti, P. (2002). Lognormality in ecological time series. *Oikos* 99, 518–530.

Halley, J. M. (1996). 'Ecology, evolution and 1/f-noise'. *Trends Ecol. Evol.* 11, 33–37. doi: 10.1016/0169-5347(96)81067-6

Halley, J. M., and Inchausti, P. (2004). 'The increasing importance of 1/f-noises as models of ecological variability'. *Fluctuation Noise Lett.* 4, R1–R26. doi: 10.1142/S0219477504001884

Halley, J. M., and Iwasa, Y. (1998). 'Extinction rate of a population under both demographic and environmental stochasticity'. *Theor. Popul. Biol.* 53, 1–15. doi: 10.1006/tpbi.1997.1336

Halley, J. M., and Iwasa, Y. (2011). 'Neutral theory as a predictor of avifaunal extinctions after habitat loss'. *Proc. Natl. Acad. Sci. U.S.A.* 108, 2316–2321. doi: 10.1073/pnas.1011217108

Halley, J. M., and Kunin, W. E. (1999). Extinction risk and the 1/f family of noise models. *Theor. Popul. Biol.* 56, 215–230. doi: 10.1006/tpbi.1999.1424

Halley, J. M., Monokrousos, N., Mazaris, A. D., Newmark, W. D., and Vokou, D. (2016). 'Dynamics of extinction debt across five taxonomic groups'. *Nat. Commun.* 7:12283. doi: 10.1038/ncomms12283

Halley, J. M., Sgardeli, V., and Triantis, K. A. (2014). 'Extinction debt and the species-area relationship: a neutral perspective'. *Glob. Ecol. Biogeogr.* 23, 113–123. doi: 10.1111/geb.12098

Hastings, A. (2004). 'Transients: the key to long-term ecological understanding?'. *Trends Ecol. Evol.* 19, 39–45. doi: 10.1016/j.tree.2003.09.007

Holt, R. D. (2009). 'Bringing the Hutchinsonian niche into the 21st century: ecological and evolutionary perspectives'. *Proc. Natl. Acad. Sci. U.S.A.* 106(Suppl. 2), 19659–19665. doi: 10.1073/pnas.0905137106

Hubbell, S. P (ed.) (2001). *The Unified Neutral Theory Of Biodiversity And Biogeography*. Princeton, NJ: Princeton University Press.

Hugueny, B. (2017). 'Age-area scaling of extinction debt within isolated terrestrial vertebrate assemblages'. *Ecol. Lett.* 20, 591–598. doi: 10.1111/ele.12756

Huston, M. (1979). 'A general hypothesis of species diversity'. *Am. Nat.* 113, 81–101. doi: 10.1086/283366

Hutchinson, G. E. (1957). 'Cold spring harbor symposium on quantitative biology'. *Concluding Remarks* 22, 415–427. doi: 10.1101/SQB.1957.022.01.039

Hutchinson, G. E. (1978). *An Introduction to Population Ecology*. Yale University Press.

Inchausti, P., and Halley, J. (2001). 'Investigating long-term ecological variability using the global population dynamics database'. *Science* 293, 655–657. doi: 10.1126/science.293.5530.655

Inchausti, P., and Halley, J. (2002). 'The long-term temporal variability and spectral colour of animal populations'. *Evol. Ecol. Res.* 4, 1033–1048.

Jackson, S. T., and Overpeck, J. T. (2000). 'Responses of plant populations and communities to environmental changes of the late Quaternary'. *Paleobiology* 26(4 Suppl.), 194–220. doi: 10.1666/0094-8373(2000)26[194:ROPPAC]2.0.CO;2

Kalyuzhny, M., Kadmon, R., and Shnerb, N. M. (2015). 'A neutral theory with environmental stochasticity explains static and dynamic properties of ecological communities'. *Ecol. Lett.* 18, 572–580. doi: 10.1111/ele.12439

Kessler, D. A., and Shnerb, N. M. (2015). 'Generalized model of island biodiversity'. *Phys. Rev. E Stat. Nonlinear Soft Matter. Phys.* 91, 1–11. doi: 10.1103/PhysRevE.91.042705

Kimura, M. (1955). 'Random genetic drift in a multi-allelic locus'. *Evolution* 9, 419–435. doi: 10.1111/j.1558-5646.1955.tb01551.x

Lamanna, C., Blonder, B., Violle, C., Kraft, N. J., Sandel, B., Šímová, I., et al. (2014). 'Functional trait space and the latitudinal diversity gradient'. *Proc. Natl. Acad. Sci. U.S.A.* 111, 13745–13750. doi: 10.1073/pnas.1317722111

Leibold, M. A. (1995). 'The niche concept revisited: mechanistic models and community context'. *Ecology* 76, 1371–1382. doi: 10.2307/1938141

Leigh, E. G. (2007). 'Neutral theory: a historical perspective'. *J. Evol. Biol.* 20, 2075–2091. doi: 10.1111/j.1420-9101.2007.01410.x

MacArthur, R. H. (1957). 'On the relative abundance of bird species'. *Proc. Natl. Acad. Sci. U.S.A.* 43, 293–295. doi: 10.1073/pnas.43.3.293

MacArthur, R. H., and Wilson, E. O. (1967). *The Theory of Island Biogeography*. Princeton, NJ: Princeton University Press.

Magurran, A. E. (2007). 'Species abundance distributions over time'. *Ecol. Lett.* 10, 347–354. doi: 10.1111/j.1461-0248.2007.01024.x

Matter, S. F., Hanski, I., and Gyllenberg, M. (2002). A test of the metapopulation model of the species-area relationship. *J. Biogeogr.* 29, 977–983. doi: 10.1046/j.1365-2699.2002.00748.x

May, R. M. (1973). *Stability And Complexity In Model Ecosystems*. Princeton, NJ: Princeton University Press.

McGill, B. J. (2003). 'A test of the unified neutral theory of biodiversity'. *Nature* 422, 881–885. doi: 10.1038/nature01583

McGill, B. J. (2010). 'Towards a unification of unified theories of biodiversity'. *Ecol. Lett.* 13, 627–642. doi: 10.1111/j.1461-0248.2010.01449.x

Meszéna, G., Gyllenberg, M., Pásztor, L., and Metz, J. A. (2006). 'Competitive exclusion and limiting similarity: a unified theory'. *Theor. Popul. Biol.* 69, 68–87. doi: 10.1016/j.tpb.2005.07.001

Pigot, A. L., Trisos, C. H., and Tobias, J. A. (2016). 'Functional traits reveal the expansion and packing of ecological niche space underlying an elevational diversity gradient in passerine birds'. *Proc. R. Soc. B Biol. Sci.* 283, 1–9. doi: 10.1098/rspb.2015.2013

Pimm, S. L. (1991). *The Balance of Nature?: Ecological Issues in the Conservation of Species and Communities*. Chicago, IL: University of Chicago Press.

Pimm, S. L., and Redfearn, A. (1988). 'The variability of population densities'. *Nature* 334, 613–614. doi: 10.1038/334613a0

Preston, F. W. (1960). 'Time and space and the variation of species'. *Ecology* 41, 611–627. doi: 10.2307/1931793

Preston, F. W. (1962). 'The canonical distribution of commonness and rarity: part 1'. *Ecology* 43, 185–215. doi: 10.2307/1931976

Ricklefs, R. E. (2006). 'The unified neutral theory of biodiversity: do the numbers add up?'. *Ecology* 87, 1424–1431. doi: 10.1890/0012-9658(2006)87[1424:TUNTOB]2.0.CO;2

Ripa, J., and Lundberg, P. (2000). 'The route to extinction in variable environments'. *Oikos* 90, 89–96. doi: 10.1034/j.1600-0706.2000.900109.x

Rosenzweig, M. L. (1995). *Species Diversity In Space And Time*. Cambridge: Cambridge University Press. doi: 10.1017/CBO9780511623387

Schoener, T. W. (1973). 'Effects of density- restricted food encounter on some single- level competition models'. *Theor. Popul. Biol.* 13, 365–381. doi: 10.1016/0040-5809(78)90052-7

Sgardeli, V., Zografou, K., and Halley, J. M. (2016). 'Climate change versus ecological drift: assessing 13 years of turnover in a butterfly community'. *Basic Appl. Ecol.* 17, 283–290. doi: 10.1016/j.baae.2015.12.008

Simberloff, D. (2014). 'The "balance of nature"—evolution of a panchreston'. *PLoS Biol.* 12:e1001963. doi: 10.1371/journal.pbio.1001963

Soberón, J. M. (2010). 'Niche and area of distribution modeling: a population ecology perspective'. *Ecography* 33, 159–167. doi: 10.1111/j.1600-0587.2009.06074.x

Sugihara, G. (1980). 'Minimal community structure: an expanation of species abundance patterns'. *Am. Nat.* 116, 770–787. doi: 10.1086/283669

Thomas, C. D., Cameron, A., Green, R. E., Bakkenes, M., Beaumont, L. J., Collingham, Y. C., et al. (2004). Extinction risk from climate change. *Nature* 427, 145–148. doi: 10.1038/nature02121

Tilman, D. (1994). 'Competition and biodiversity in spatially structured habitats'. *Ecology* 75, 2–16. doi: 10.2307/1939377

Tilman, D., May, R. M., Lehman, C. L., and Nowak, M. A. (1994). 'Habitat destruction and the extinction debt'. *Nature* 371, 65–66. doi: 10.1038/371065a0

Ulrich, W., Ollik, M., and Ugland, K. I. (2010). 'A meta-analysis of species-abundance distributions'. *Oikos* 119, 1149–1155. doi: 10.1111/j.1600-0706.2009.18236.x

White, E. P., Adler, P. B., Lauenroth, W. K., Gill, R. A., Greenberg, D., Kaufman, D. M., et al. (2005). 'A comparison of the species-time relationship across ecosystmes and taxonomic groups'. *Oikos* 112, 185–195. doi: 10.1111/j.0030-1299.2006.14223.x

Williamson, M., and Gaston, K. J. (2005). 'The lognormal distribution is not an appropriate null hypothesis for the species-abundance distribution'. *J. Anim. Ecol.* 74, 409–422. doi: 10.1111/j.1365-2656.2005.00936.x

Wunsch, C. (2003). 'The spectral description of climate change including the 100 ky energy'. *Clim. Dyn.* 20, 353–363. doi: 10.1007/s00382-002-0279-z

Zillio, T., and Condit, R. (2007). 'The impact of neutrality, niche differentiation and species input on diversity and abundance distributions'. *Oikos* 116, 931–940. doi: 10.1111/j.2007.0030-1299.15662.x

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

APPENDIX A. LIST OF SYMBOLS

a, Vector of environmental configurations of niche factors in logarithmic space.
a_k, Environmental configuration of the k^{th} niche factor in log-space ($k \in 0,1,2,\ldots, K-1$).
B_k, A constant related to U_k, such that $U_k = \Gamma_k B_k$ and 1-$U_k = \Gamma_k/B_k$ and with $\Gamma_k = 1/(B_k + 1/B_k) = \sqrt{U_k(1-U_k)}$.
C, The orientation matrix ($S \times K$).
c_{sk}, Element of the orientation matrix **C** for species (row) s and niche factor (column) k. All $c_{sk} = \pm 1$.
$e_s(t)$, Probability of extirpation of the species s at time t.
$f(u)$, Probability distribution function of the random variable U_k (assumed independent and identically-distributed on all axes).
$f_{xs}(x)$, Probability distribution function of the log-population of species s.
$f_{vs}(x)$, Probability distribution function of the log-hypervolume of species s.
$g(x)$, Probability distribution function of the niche width $\lambda_{sk} = \ln(L_{sk})$.
J, The size of the community (sum of all individuals from all species).
K, Number of niche factors (dimension of niche hypercube, number of axes)
k, Niche-factor index, $k \in 0,1,2,\ldots, K-1$.
L_{sk}, The niche width of species s for factor k.
M, Minimum viable population (assumed same for all species in the community).
m, Natural logarithm of M.
o_k, The k^{th} digit of the binary representation of species.
p, Probability that a given species is present in the community.
q, Probability that U_k takes value u_0 for the asymmetric dichotomous distribution of Special Case C, while $1-q$ is the probability that U_k takes the value $1-u_0$.
S, The total number of possible species in the community ($S = 2^K$).
s, Species index (in decimal), $s \in 0,1,2,\ldots, S-1$.
s, Species index (in binary).
T, Length of time series $T \in 1,2,3\ldots$
t, Time: $t \in 0,1,2,3\ldots$
U_k, The random variable representing the division of axis for factor k.
u_0, The value defining the dichotomous distribution on Special Case B. In this case the random variable. U_k, can take only the values u_0 and $1-u_0$.
V_0, Variance of random variable U_k.
V_s, Niche hypervolume of species s.
v, Vector of niche hypervolumes for all species in the community.
v_s, Natural logarithm of niche hypervolume of species s.
W_s, Total variance for time series of species s.
X_s, Population of species s.
x, Vector of log-populations of species in the community.
x_s, Logarithm of the population of species s.
z, Vector of deviations from the mean log-population, $K\bar{\gamma} + \ln J$, for the case where the mean is constant in time.
β^2, Expected variance of $f(x)$ for Special Case B.
Γ_k, A constant related to U_k and B_k. It is the geometric mean of the two partition sizes (U_k and $1-U_k$) on axis-k:. i.e., $\Gamma_k = \sqrt{U_k(1-U_k)}$.
γ_k, A constant $\gamma_k = \ln(\Gamma_k)$. It is mean of the logarithms of the two partition sizes (U_k and $1-U_k$) of the axis k.
$\bar{\gamma}$, The average value of γ_k over all niche factors, namely $\bar{\gamma} = {}^1/_K \sum \gamma_k$.
$\lambda_{sk} = \ln L_{sk}$.
μ, Expected mean of $g(x)$.
ρ, Number of individuals per unit area.
τ_k, Timescale of changes happening for niche factor k.
σ^2, Expected variance of $g(x)$.

APPENDIX B. MATHEMATICAL PROOFS

Relation Between the Spectrum of Population and Environmental Forcing

We are interested in the relationship between the spectra of the environmental vector $a_k(t)$ and that of the population response $x_s(t)$. In the case of distributions for which $K\bar{\gamma}(t)$ is constant (e.g., Special Cases B and C) we may express the dynamics in Equation (15) as:

$$z_s(t) = \sum_k c_{sk} a_k(t), \quad z_k(t) = x_k(t) - \left(K\bar{\gamma} + \ln J\right) \tag{A1}$$

If we take the Fourier transform of both sides, we get the Fourier coefficient of z in the domain of frequency, ω, namely:

$$\tilde{z}_s(\omega) = \sum_k c_{sk} \tilde{a}_k(\omega) \tag{A2}$$

Since both a and z are random variables, it makes sense to find the power spectrum z. The power spectrum of the Fourier coefficient of $z_s(t)$ is found by finding the norm of it. Thus:

$$S_{z_s}(\omega) = \tilde{z}_s^*(\omega) \cdot \tilde{z}_s(\omega) = ||\tilde{z}_s(\omega)||^2 \tag{A3}$$

Where the z^* refers to the complex conjugate of z. If we substitute into this the Equation (A2) we get:

$$\begin{aligned} S_z(\omega) &= \sum_i \sum_k c_{si} \cdot c_{sk} \cdot \tilde{a}_i^*(\omega) \cdot \ \tilde{a}_k(\omega) \\ &= \sum_k c_{sk}^2 \cdot \left|\left|\tilde{a}_k^2(\omega)\right|\right| \end{aligned} \tag{A4}$$

In A4, the interaction terms are zero because the components are independent. Noting also that all components of the orientation matrix are ± 1, we get:

$$S_z(\omega) = \sum_k S_{a_k}(\omega) \tag{A5}$$

Thus, at least in cases where $K\bar{\gamma}(t)$ is constant (e.g., Special Cases B and C), the spectrum of the species population is simply the sum of the spectra of its niche factors.

Variance of Sequence of Square Pulses

This shows, for Special Case B, that the variance goes to zero if the length of the series T is much less than the time-constant of the pulses and that it converges to a finite value as T heads to infinity.

We note that here $\mathrm{Var}\{a_k(t)\}$ is a sample variance in time rather than an ensemble variance, so the expected value of the variance estimator for the time series in question is:

$$E\left[\frac{1}{T-1}\sum_{t=1}^{T}(a_t - \bar{a})^2\right], \ \bar{a} = \frac{1}{T}\sum_{t=1}^{T} a_t$$

Here, we use the notation $a_k(t) = a_t$ and $\tau_k = \tau$ for simplicity. Following the approach of Halley and Kunin (1999), this can be expressed as:

$$Var\{a\} = \frac{1}{T^2(T-1)}\sum_{i,j,k}^{T} E\left[\Delta a_{ji}\Delta a_{jk}\right], \ \Delta a_{ji} = a_j - a_i$$

Note that the random variable $a_j = \pm 1$ and that a_j changes state at random every τ time steps (at times $j = \tau, 2\tau, 3\tau, \ldots$). So, when *they are well separated in time the expected correlation is zero*. Conversely, if the time difference is small enough then a_j and a_i are the same then $\Delta a_{ji} = 0$ and similarly if a_j and a_k are the same. Thus:

$$Var\{a\} = \frac{1}{T^2(T-1)}\sum_{i,j,k}^{T}\left[E[a_j^2] - E[a_i a_j] - E[a_j a_k] + E[a_i a_k]\right]$$

We note that:

$$\sum_{i,j,k}^{T} E\left[\Delta a_{ji}\Delta a_{jk}\right] \ = \ \sum_{i,j,k}^{T}\left[E[a_j^2] - E[a_i a_j] - E[a_j a_k] + E[a_i a_k]\right]$$

$$= \sum_{i,j,k}^{T} E[a_j^2] - \sum_{i,j,k}^{T} E[a_i a_j] - \sum_{i,j,k}^{T} E[a_j a_k] + \sum_{i,j,k}^{T} E[a_i a_k]$$

$$= T^3 - T\sum_{i,j}^{T} E[a_i a_j] - T\sum_{i,j}^{T} E[a_j a_k] + T\sum_{i,j}^{T} E[a_i a_k]$$

Note that the last three terms on the r.h.s. are identical by symmetry among the indices. Thus, we can write:

$$\sum_{i,j,k}^{T} E\left[\Delta a_{ji}\Delta a_{jk}\right] = T^3 - T\sum_{i,j}^{T} E[a_i a_j]$$

The value of $E(a_i.a_j)$ depends on $|i\text{-}j|$ because for a square-wave of the type of a_i, the following holds for a wave with period length τ :

$$\begin{aligned} E[a_i a_j] &= 1 - \tfrac{|i-j|}{\tau} \ \forall |i-j| < \tau \\ &= 0, \text{ otherwise} \end{aligned}$$

Therefore, we arrange the summation by $|i\text{-}j|$. There are T diagonal elements so $E(a_i.a_j) = 1$ here. There are $2(T-1)$ off-diagonal by one step, for which $E(a_i.a_j) = 1-1/\tau$, then $2(T-2)$ off-diagonal by two steps, for which $E(a_i.a_j) = 1-2/\tau$ and so on.

$$\begin{aligned} \sum_{i,j,k}^{T} E\left[\Delta a_{ji}\Delta a_{jk}\right] &= T^3 - T\left[T + 2(T-1)(1-1/\tau) + 2(T-2)(1-2/\tau) + +2(T-3)(1-3/\tau) + ...+\right] \\ &= T^3 - T\left[T\tau + 2(T-1)(\tau-1) + 2(T-2)(\tau-2) + 2(T-3)(\tau-3) + ...+\right]/\tau \\ &= T^3 - 2\left(\tfrac{T}{\tau}\right)\sum_{j=0}^{\min(T,\tau)}(T-j)(\tau-j) \ - \ T^2 \\ &= T^2 - (T-1)2\left(\tfrac{T}{\tau}\right)\sum_{j=0}^{\min(T,\tau)}(T-j)(\tau-j) \end{aligned}$$

Thus,

$$Var\{a\} = \tfrac{1}{T^2(T-1)}\left[T^2(T-1) - 2\left(\tfrac{T}{\tau}\right)\sum_{j=0}^{\min(T,\tau)}(T-j)(\tau-j)\right]$$

$$= 1 - \left(\tfrac{2}{T(T-1)\tau}\right)\sum_{j=0}^{\min(T,\tau)}(T-j)(\tau-j)$$

For short series, $T << \tau$ so then:

$$Var\{a\} \approx 1 - \left(\frac{2}{T(T-1)\tau}\right)\sum_{j=0}^{T}(T-j)\tau = 1 - \left(\frac{2}{T(T-1)}\right)\sum_{j=1}^{T-1} j = 0$$

On the other hand, if $T >> \tau$ then the summation

$$Var\{a\} \approx 1 - \left(\frac{2}{T(T-1)\tau}\right)T\sum_{j=0}^{\tau}(\tau-j) = 1 - \left(\frac{2}{(T-1)\tau}\right)\frac{\tau(\tau-1)}{2} \approx 1 - \frac{\tau-1}{T}$$

When $T = \tau$ then,

$$Var\{a\} = 1 - \left(\frac{2}{T^2(T-1)}\right)\sum_{j=1}^{T-1} j^2 = 1 - \left(\frac{2}{T^2(T-1)}\right)\left(\frac{T(2T-1)(T-1)}{6}\right) = 1 - \frac{2T-1}{3T} = \frac{T+1}{3T}$$

Advantages of publishing in Frontiers

OPEN ACCESS
Articles are free to read for greatest visibility and readership

FAST PUBLICATION
Around 90 days from submission to decision

HIGH QUALITY PEER-REVIEW
Rigorous, collaborative, and constructive peer-review

TRANSPARENT PEER-REVIEW
Editors and reviewers acknowledged by name on published articles

Frontiers
Avenue du Tribunal-Fédéral 34
1005 Lausanne | Switzerland

Visit us: www.frontiersin.org
Contact us: frontiersin.org/about/contact

REPRODUCIBILITY OF RESEARCH
Support open data and methods to enhance research reproducibility

DIGITAL PUBLISHING
Articles designed for optimal readership across devices

FOLLOW US
@frontiersin

IMPACT METRICS
Advanced article metrics track visibility across digital media

EXTENSIVE PROMOTION
Marketing and promotion of impactful research

LOOP RESEARCH NETWORK
Our network increases your article's readership