MoB (Measurement of Biodiversity): a method to separate the scale-dependent effects of species abundance distribution, density, and aggregation on diversity change

Daniel McGlinn[1*†], Xiao Xiao[2*], Felix May[3], Nicholas J. Gotelli[4], Shane A. Blowes[3], Tiffany Knight[3,5,6], Oliver Purschke[3], Jonathan Chase[3,7+], Brian McGill[2+]

† corresponding author

* joint first authors

+ joint last authors

Author affiliations

1. Biology Department, College of Charleston, Charleston, SC, mcglinndj@cofc.edu

2. School of Biology and Ecology, and Senator George J. Mitchell Center of Sustainability Solutions, University of Maine, Orono, ME

3. German Centre for Integrative Biodiversity Research (iDiv), Halle-Jena-Leipzig, Deutscher Platz 5e, 04103 Leipzig, Germany

4. Department of Biology, University of Vermont, Burlington VT 05405 USA

5. Institute of Biology, Martin Luther University Halle-Wittenberg, Am Kirchtor 1, 06108, Halle (Saale), Germany

6. Dept. Community Ecology, Helmholtz Centre for Environmental Research – UFZ, Theodor-Lieser-Straße 4, 06120 Halle (Saale), Germany

7. Department of Computer Science, Martin Luther University, Halle-Wittenberg

Authors' contributions

DM, XX, JC, and BM conceived the study and the overall approach, and all authors participated in multiple working group meetings to develop and refine the approach; JC and TK collected the data for the empirical example that led to Figures 4-6; XX, DM, and FM, wrote the R package,

NG, JC, and BM provided guidance on method development, DM carried out the analysis of the empirical example, and XX carried out the sensitivity analysis; DM and XX wrote first draft of the manuscript, and all authors contributed substantially to revisions.

Abstract

1. Little consensus has emerged regarding how proximate and ultimate drivers such as abundance, productivity, disturbance, and temperature may affect species richness and other aspects of biodiversity. Part of the confusion is that most studies examine species richness at a single spatial scale and ignore how the underlying components of species richness can vary with spatial scale.
2. We provide an approach for the measurement of biodiversity (MoB) that decomposes scale-specific changes in richness into proximate components attributed to: 1) the species abundance distribution, 2) density of individuals, and 3) the spatial arrangement of individuals. We decompose species richness using a nested comparison of individual- and plot-based species rarefaction and accumulation curves.
3. Each curve provides some unique scale-specific information on the underlying components of species richness. We tested the validity of our method on simulated data, and we demonstrate it on empirical data on plant species richness in invaded and uninvaded woodlands. We integrated these methods into a new R package (`mobr`).
4. The metrics that `mobr` provides will allow ecologists to move beyond comparisons of species richness at a single spatial scale towards a more mechanistic understanding of the drivers of community organization that incorporates information on the scale dependence of the proximate components of species richness.

Key words

accumulation curve, community structure, extent, grain, rarefaction curve, species-area curve, species richness, spatial scale

Introduction

Species richness – the number of species co-occurring in a specified area – is one of the most widely-used biodiversity metrics. However, ecologists often struggle to understand the mechanistic drivers of richness, in part because multiple ecological processes can yield qualitatively similar effects on species richness (Chase and Leibold 2002, Leibold and Chase 2017). For example, high species richness in a local community can be maintained either by species partitioning niche space to reduce interspecific competition (Tilman 1994), or by a balance between immigration and stochastic local extinction (Hubbell 2001). Similarly, high species richness in the tropics has been attributed to numerous mechanisms such as higher productivity supporting more individuals, higher speciation rates, and longer evolutionary time since disturbance (Rosenzweig 1995).

Although species richness is a single metric that can be measured at a particular grain size or spatial scale, it is a response variable that summarizes the underlying biodiversity information that is contained in the individual organisms, which each are assigned to a particular species, Operational Taxonomic Unit, or other taxonomic grouping. Variation in species richness can be decomposed into three components (He and Legendre 2002, McGill 2010): 1) the number and relative proportion of species in the regional source pool (i.e., the species abundance distribution, SAD), 2) the number of individuals per plot (i.e., density), and 3) the spatial distribution of individuals that belong to the same species (i.e., spatial aggregation). Changes in species richness may reflect one or a combination of all three components changing

simultaneously. In some cases, the density and spatial arrangement of individuals simply reflect sampling intensity and detection errors. But in other cases, density and spatial arrangement of individuals may reflect responses to experimental treatments that ultimately drive the patterns of observed species richness. Thus, it is critical that we look beyond richness as a single metric, and develop methods to disentangle its underlying components that have more mechanistic links to processes (e.g., Vellend 2016). Although this is not the only mathematically valid decomposition of species richness, these three components are well-studied properties of ecological systems, and provide insights into mechanisms behind changes in richness and community structure (Harte et al. 2008, Supp et al. 2012, McGlinn et al. 2015).

The shape of the regional SAD influences local richness. The shape of the SAD is influenced by the degree to which common species dominate the individuals observed in a region, and on the total number of highly rare species. Local communities that are part of a more even regional SAD (i.e., most species having similar abundances) will have high values of local richness because it is more likely that the individuals sampled will represent different species. Local communities that are part of regions with a more uneven SAD (e.g., most individuals are a single species) will have low values of local richness because it is more likely that the individuals sampled will be the same, highly common species (He and Legendre 2002, McGlinn and Palmer 2009). The richness of the regional species pool, which is influenced by the total number of rare species, has a similar effect on local richness. As regional species richness increases, local richness will also increase if the local community is even a partly random subsample of the species in the regional pool. Because the regional species pool is never fully observed, the two sub-components –the shape of the SAD and the size of the regional species

pool – cannot be completed disentangled. Thus, we group them together, as the SAD effect on local richness.

The number of individuals in the local community directly affects richness due to the sampling effect (the More Individuals Hypothesis of richness; Hurlbert 2004). As more individuals are randomly sampled from the regional pool, species richness is bound to increase. This effect has been hypothesized to be strongest at fine spatial scales; however, even at larger spatial scales, it never truly goes to zero (Palmer and van der Maarel 1995, Palmer et al. 2008).

The spatial arrangement of individuals within a plot or across plots is rarely random. Instead most individuals are spatially clustered or aggregated in some way, with neighboring individuals more likely belonging to the same species. As individuals within species become more spatially clustered, local diversity will decrease because the local community or sample is likely to consist of clusters of only a few species (Karlson et al. 2007, Chiarucci et al. 2009, Collins and Simberloff 2009).

Traditionally, individual-based rarefaction has been used to control for the effect of numbers of individuals on richness comparisons (Hurlbert 1971, Simberloff 1972, Gotelli and Colwell 2001), but few methods exist (e.g., Cayuela et al. 2015) for decomposing the effects of SADs and spatial aggregation on species richness. Because species richness depends intimately on the spatial and temporal scale of sampling, the relative contributions of the three components are also likely to change with scale. Spatial scale can be represented both by number of individuals, which scales linearly with area when density is relatively constant, and by the number of samples (plots). We will demonstrate that this generalized view of spatial scale allows us to distinguish three different types of sampling curves: (1) (spatially constrained) plot-based accumulation; (2) non-spatial plot-based rarefaction; and (3) (non spatial) individual-based

rarefaction. Constructing these different curves allows us to parse the relative contributions of the three proximate drivers of richness and how those contributions potentially change with spatial scale. Specifically, we develop a framework that provides a series of sequential analyses for estimating and testing the effects of the SAD, individual density, and spatial aggregation on changes in species richness across scales. We have implemented these methods in a freely available R package `mobr` (https://github.com/MoBiodiv/mobr)

Materials and Methods

Method Overview

Our method targets data collected in standardized sampling units such as quadrats, plots, transects, net sweeps, or pit falls of constant area or sampling effort (we refer to these as "plots") that are assigned to treatments. We use the term treatment here generically to refer to manipulative treatments or to groups within an observational study (e.g., invaded vs uninvaded plots). The designation of plots within treatments implicitly defines the α *scale* – a single plot – and the γ *scale* – all plots within a treatment. If the sampling design is relatively balanced among treatments, the total sample area and the spatial extent (the minimum polygon encompassing all the plots in the treatment) are similar for each treatment. In an experimental study, each plot is assigned to a treatment. In an observational study, each plot is assigned to a categorical grouping variable(s). For this typical experimental/sampling design, our method provides two key outputs: 1) the relative contribution of the different components affecting richness (SAD, density, and spatial aggregation) to the observed change in richness between treatments and 2) quantifying how species richness and its decomposition change with spatial scale. We propose two complementary ways to view scale-dependent shifts in species richness and its components: a simple-to-interpret two-scale analysis and a more informative continuous scale analysis.

The two-scale analysis provides a big-picture view of the changes between the treatments by focusing exclusively on the α (plot-level) and γ (across all plots) spatial scales. It provides diagnostics for whether species richness and its components differ between treatments at the two scales. The continuous scale analysis expands the two-scale analysis by taking advantage of three distinct species richness curves computed across a range of scales: 1) plot-based accumulation curve (Gotelli and Colwell 2001, Chiarucci et al. 2009), where the order in which plots are sampled depends on their spatial proximity; 2) the non-spatial, plot-based rarefaction, where individuals are randomly shuffled across plots within a treatment while maintaining average plot density; and 3) the individual-based rarefaction curve where again individuals are randomly shuffled across plots within a treatment but in this case average plot density is not maintained. The differences between these curves are used to isolate the effects of the SAD, density of individuals, and spatial aggregation on richness and document how these effects change as a function of scale.

Detailed Data Requirements

Table 1. Mathematical nomenclature used in the study.

Treatment (or group label)	Plot	Coordinates		Species 1	…	Species S	Total abundance	Richness
1	1	$x_{1,1}$	$y_{1,1}$	$n_{1,1,1}$	…	$n_{1,1,S}$	$N_{1,1} = \sum_s n_{1,1,s}$	$S_{1,1}$
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
1	K	$x_{1,K}$	$y_{1,K}$	$n_{1,K,1}$	⋮	$n_{1,K,S}$	$N_{1,K} = \sum_s n_{1,K,s}$	$S_{1,K}$
2	1	$x_{2,1}$	$y_{2,1}$	$n_{2,1,1}$	⋮	$n_{2,1,S}$	$N_{2,1} = \sum_s n_{2,1,s}$	$S_{2,1}$
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
2	K	$x_{2,K}$	$y_{2,K}$	$n_{2,K,1}$	⋯	$n_{2,K,S}$	$N_{2,K} = \sum_s n_{2,K,s}$	$S_{2,K}$

153 Consider $T = 2$ treatments, with K replicated plots per treatment (Table 1). Within each
154 plot, we have measured the abundance of each species, which can be denoted by $n_{t,k,s}$, where $t =$
155 1, 2 for treatment, $k = 1, 2, \ldots K$ for plot number within the treatment, and $s = 1, 2, \ldots S$ for
156 species identity, with a total of S species recorded among all plots and treatments. The
157 experimental design does not necessarily have to be balanced (i.e., K can differ between
158 treatments) if the spatial extent is still similar between the treatments. For simplicity of notation
159 we describe the case of a balanced design here. $S_{t,k}$ is the number of species observed in plot k in
160 treatment t (i.e., number of species with $n_{t,k,s} > 0$), and $N_{t,k}$ is the number of individuals observed
161 in plot k in treatment t (i.e., $N_{t,k} = \sum_s n_{t,k,s}$). The spatial coordinates of each plot k in treatment t
162 are $x_{t,k}$ and $y_{t,k}$. We focus on spatial patterns but our framework also applies analogously to
163 samples distributed through time.

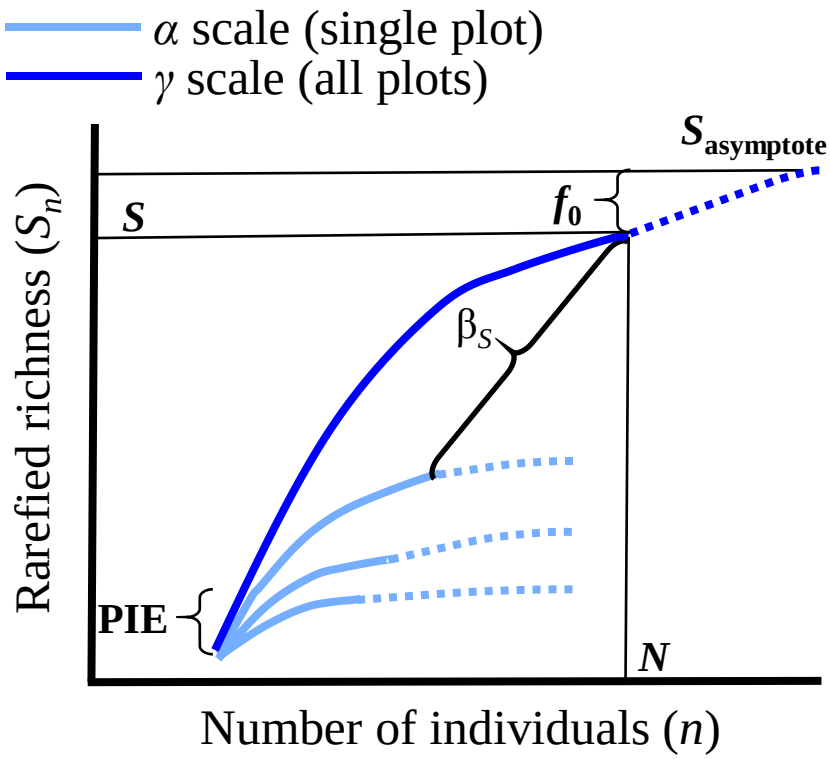
164 For clarity of explanation we focus here on a single-factor design with two (or more)
165 categorical treatment levels. The method can be extended to accommodate crossed designs and
166 regression-style continuous treatments which we describe in the Discussion and Supplement S5.

167 *Two-scale analysis*

168 The two-scale analysis is intended to provide a simple decomposition of species richness
169 while still emphasizing the three components and change with spatial scale. In the two-scale
170 analysis, we compare observed species richness in each treatment and several other summary
171 statistics at the α and γ scales (Table 2). The summary statistics were chosen to represent the
172 most informative aspects of individual-based rarefaction curves (Fig. 1). These rarefaction
173 curves plot the expected species richness S_n against the number of individuals when individuals
174 are randomly drawn from the sample at the α or γ scales. The curve can be calculated precisely

using the hypergeometric sampling formula given the SAD ($n_{t,k,s}$ at the plot level, $n_{t,+,s}$ at the treatment level) (Hurlbert 1971).

We show how several widely-used diversity metrics are represented along the individual rarefaction curve, corresponding to α and γ scales (Fig. 1, Table 2, see Supplement S1 for detailed metric description). The total number of individuals within a plot ($N_{t,k}$) or within a treatment ($N_{t,+}$) determines the endpoint of the rarefaction curves. Rarefied richness (S_n) controls richness comparisons for differences in individual density between treatments because it is the expected number of species for a random draw of n individuals ranging from 1 to N. To compute S_n at the α scale we set n to the minimum number of individuals across all samples in both treatments with a hard minimum of 5, and at the γ scale we multiplied this n value by the number of samples within a treatment (i.e., K). The probability of intraspecific encounter (PIE), $S_{\text{asymptote}}$ (via Chao1 estimator) and the number of undiscovered species (f_0) reflect the SAD component. We follow Jost (2007) and convert PIE into effective numbers of species (S_{PIE}) so that it can be more easily interpreted as a metric of diversity (See Supplement S1 for more description and justification of PIE, f_0, and associated β metrics). Whittaker's multiplicative beta diversity metrics for S, S_{PIE}, and f_0 reflect the degree of turnover between the α and γ scales. In Fig. 1, species are spatially aggregated across plots, and β_S is large.



193

194 Figure 1. Illustration of how the key biodiversity metrics are derived from the individual-based
195 rarefaction curves constructed at the α (i.e., single plot) and γ (i.e., all plots) scales. The solid
196 lines are rarefied richness derived from the randomly sampling individuals from each plot's SAD
197 and the dotted lines reflect the extrapolated richness via Chao1 estimator. The light blue curves
198 show individual rarefaction curves for each plot. The labeled metrics can also be calculated for
199 each α-scale curve (not shown). The dark blue curve reflects the individual rarefaction curve at
200 the γ-scale, with all individuals from all plots combined. *S* and *N* correspond to the ending points
201 of the rarefaction curve on the richness and individual axes, respectively. $S_{asymptote}$ is the
202 extrapolated asymptote. See Table 2 for definitions of metrics including ones not illustrated.

203 Comparison of these summary statistics between treatments identifies whether the
204 treatments have a significant effect on richness at these two scales, and if they do, the potential
205 proximate driver(s) of the change. A difference in *N* between treatments implies that differences

206 in richness between treatments may be a result of treatments changing the density of individuals.
207 Differences in S_{PIE} and/or f_0 imply that change in the shape of the SAD may contribute to the
208 change in richness, with S_{PIE} being most sensitive to changes in abundant species and f_0 being
209 most sensitive to changes in number of rare species. Differences in β-diversity metrics may be
210 due to differences in any of the three components: SAD, N, or aggregation, and each β metric
211 (Table 2) provides a different weighting on common vs rare species.
212

213 Table 2. Definitions and interpretations of the summary statistics for simplified two-scale

214 analysis

Metric	Definition	Interpretation
S	Observed richness, effective number of species of order 0 (Jost, 2007)	Number of species
N	Total abundance across all species	Measure of density of individuals
S_n	The expected richness for n randomly sampled individuals (Hurlbert 1971).	Estimate of richness after controlling for differences due to aggregation or number of individuals (i.e., only reflects SAD)
PIE	Probability of intraspecific encounter ($S_{n=2} - S_{n=1}$, Hurlbert 1971, Olszweski 2004),	Measure of evenness, slope at base of the rarefaction curve, and sensitive to common species
S_{PIE}	Number of equally abundant species needed to yield PIE (i.e., effective number of species of order 2, Jost 2007)	Effective number of species of PIE that is easier to compare with S (= 1 / (1 – PIE))
$S_{\text{asymptote}}$	Extrapolated asymptotic richness via Chao1 estimator (Chao 1984).	Richness that includes unknown species but is highly correlated with S (McGill 2011)
f_0	Richness of undetected species ($S_{\text{asymptote}} - S$, Chao et al. 2009).	Measure of rarity at top of rarefaction curve, more sensitive to rare species than S
β_S	Ratio of total treatment S and average plot S (Whittaker 1960)	More species turnover results in larger β_S which may be due to increases in spatial aggregation, N, and/or unevenness of the SAD.
β_{f_0}	Ratio of total treatment f_0 and average plot f_0	Like β_S but emphasizes rare species
$\beta_{S_{\text{PIE}}}$	Ratio of total treatment and average plot S_{PIE} (Olszewski 2004)	Like β_S but emphasizes common species

215

216 The treatment effect on these metrics can be visually examined with boxplots (see

217 Empirical example section) at the α scale and with single points at the pooled γ-scale (unless

218 there is replication at the γ scale as well). Quantitative comparison of the metrics can be made

with t-tests (ANOVAs for more than two treatments) or, for highly skewed data, nonparametric tests such as Mann-Whitney U test (Kruskal-Wallis for more than two treatments).

We provide a non-parametric, randomization test where the null expectation of each metric is established by randomly shuffling the plots between the treatments, and recalculating the metrics for each reshuffle. The significance of the differences between treatments can then be evaluated by comparing the observed test statistic to the null expectation when the treatment IDs are randomly shuffled across the plots (Legendre and Legendre 1998). When more than two groups are compared the test examines the overall group effect rather than specific group differences. At the α scale where there are replicate plots to summarize over, we use the ANOVA F-statistic as our test statistic (Legendre and Legendre 1998), and at the γ scale in which we only have a single value for each treatment (and therefore cannot use the F-statistic) the test statistic is the absolute difference between the treatments (if more than two treatments are considered then it is the average of the absolute differences, $\overline{D}$). At both scales we use $\overline{D}$ as a measure of effect size.

Note that $N_{t,k}$ and $N_{t,+}$ give the same information, because one scales linearly with the other by a constant (i.e., $N_{t,+}$ is equal to $N_{t,k}$ multiplied by the number of plots K within treatment). However, the other metrics (S, f_0 and S_{PIE}) are not directly additive across scales. Evaluation of these metrics at different scales may yield different insights for the treatments, sometimes even in opposite directions (Chase et al. *submitted*). However, complex scale-dependence may require comparison of entire sampling curves (rather than their two-scale summary statistics) to understand how differences in community structure change continuously across a range of spatial scales.

Continuous scale analysis

While the two-scale analysis provides a useful tool with familiar methods, it ignores the role of scale as a continuum. Such a discrete scale perspective can only provide a limited view of treatment differences at different scales. We develop in this section a method to examine the components of change across a continuum of spatial scale. We define spatial scale by the amount of sampling effort, which we define as the number of individuals or the number of plots sampled. Assuming that the density of individuals is constant across plots, these measures should be proportional to each other.

The three curves

The key innovation is to use three distinct types of species accumulation and rarefaction curves that capture different components of community structure. By a carefully sequenced analysis, it is possible to tease apart the effects of SAD shape, of changes in density of individuals (*N*), and of spatial aggregation across a continuum of spatial scale. The three types of curves are summarized in Table 3. Fig. 2 shows graphically how they are constructed.

The first curve, is the spatial plot-based or sample-based accumulation curve (Gotelli and Colwell 2001 or spatially-constrained rarefaction Chiarucci *et al*. 2009). It is constructed by accumulating plots within a treatment based on their spatial position such that the most proximate plots are collected first. One can think of this as starting with a target plot and then expanding a circle centered on the target plot until one additional plot is added, then expanding the circle until another plot is added, etc. In practice, every plot is used as the starting target plot and the resulting curves are averaged to give a smoother curve. If two or more plots are of equal distance to the target plot, they are accumulated in random order.

The second curve is the non-spatial, plot-based rarefaction curve (Supplement S2). It is

264 constructed by randomly sampling plots within a treatment in which the individuals in the plots
265 have first been randomly shuffled among the plots within a treatment, while maintaining the
266 plot-level average abundance ($\overline{N_{t,k}}$) and the treatment-level SAD ($\vec{n}_{t,+} = \sum_k \vec{n}_{t,k}$). Note that this
267 rarefaction curve is very different from the traditional "sample-based rarefaction curve" (Gotelli
268 and Colwell 2001), in which plots are randomly shuffled to build the curve but individuals within
269 a plot are preserved (and consequently any within-plot spatial aggregation is retained). Our non-
270 spatial, plot-based rarefaction curve contains the same information (plot density and SAD) as the
271 spatial accumulation curve, but it has nullified any signal due to species spatial aggregation both
272 within and between plots.

273 The third curve is the familiar individual-based species rarefaction curve. It is constructed
274 by first pooling individuals across all plots within a treatment, and then randomly sampling
275 individuals without replacement. This individual-based rarefaction curve reflects only the shape
276 of the underlying SAD ($\vec{n}_{t,+}$).

277 In can be computationally intensive to compute rarefaction curves, and therefore
278 analytical formulations of these curves are desirable to speed up software. It is unlikely an
279 analytical formulation of the plot-based accumulation curve exists because it requires averaging
280 over each possible ordering of nearest sites; however, analytical expectations are available for
281 the sample- and individual-based rarefaction curves. Specifically, we used the hypergeometic
282 formulation provided by Hurlbert (1971) to estimate expected richness of the individual-based
283 rarefaction curve. To estimate the plot-based rarefaction curve we extended Hurlbert's (1971)
284 formulation (see Supplement S2). Our derivation demonstrates that the non-spatial curve is a
285 rescaling of the individual-based rarefaction curve based upon the degree of difference in density
286 between the two treatments under consideration. Specifically, we use the ratio of average

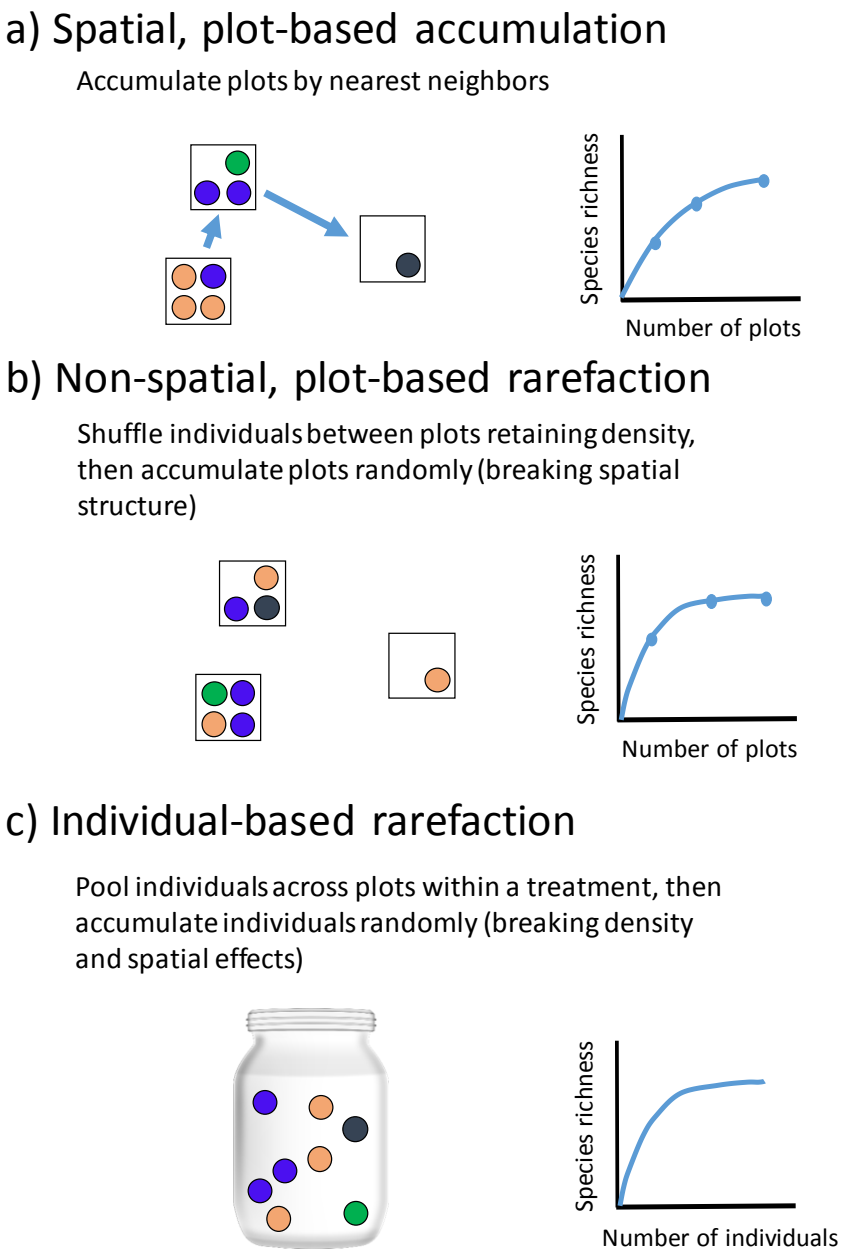
community density to the density in the treatment of interest to rescale sampling effort in the individual based rarefaction curve. For a balanced design, the individual rarefaction curve of Treatment 1 can be adjusted for density effects by multiplying the sampling effort of interest by: $\left(\sum_t \sum_k N_{t,k}\right) / \left(2 \cdot \sum_k N_{1,k}\right)$. Similarly, the Treatment 2 curve would be rescaled by $\left(\sum_t \sum_k N_{t,k}\right) / \left(2 \cdot \sum_k N_{2,k}\right)$. If the treatment of interest has the same density as the average community density then there is no density effect, and the plot-based curve is equivalent to the individual-based rarefaction curve. Here we have based the density rescaling on average number of individuals, but alternatives exist such as using maximum or minimum treatment density. Note that the plot-based curve is only relevant in a treatment comparison, which contrasts with the other two rarefaction curves that can be constructed independently of any consideration of treatment effects.

299 Table 3. Summary of three types of species sampling curves. For treatment *t*, $\vec{n}_{t,+}$ is the vector

300 of species abundances, $\vec{n}_t$ is the vector of plot abundances, and $\vec{d}_t$ is the vector of distances

301 between plots.

Curve Name	**Notation**	**Method for accumulation**	**Interpretation**
Spatial plot-based accumulation curve	$\mathrm{E}\left[S_t \middle\vert k, \vec{n}_{t,+}, \vec{N}_t, \vec{d}_t\right]$	Spatially explicit sampling in which the most proximate plots to a focal plot are accumulated first. All possible focal plots are considered and the resulting curves are averaged over.	This curve includes all information in the data including effect of SAD, effect of density of individuals, and effect of spatial aggregation.
Nonspatial, plot-based rarefaction curve	$\mathrm{E}\left[S_t \middle\vert k, \vec{n}_{t,+}, \vec{N}_t\right]$	Random sampling of *k* plots after removing intraspecific spatial aggregation by randomly shuffling individuals across plots while maintaining average plot-level abundance ($\overline{N_{t,k}}$) and the treatment-level SAD ($n_{t,+,s} = \sum_k n_{t,k,s}$). In practice, we use an analytical extension of the hypergeometric distribution that demonstrates this curve is a rescaling of the individual-base rarefaction curve based on the ratio: (average density across treatments) / (average density of treatment of interest)	This curve reflects both the shape of the SAD and the difference in density between the treatments. If density between the two treatments is identical then this curve converges on the individual-based rarefaction curve.
Individual-based rarefaction curve	$\mathrm{E}\left[S_t \middle\vert N, \vec{n}_{t,+}\right]$	Random sampling of *N* individuals from the observed SAD ($\vec{n}_{t,+}$) without replacement.	By randomly shuffling individuals with no reference to plot density, all spatial and density effects are removed. Only the effect of the SAD remains.



302

303 Figure 2. Illustration of how the three sampling curves are constructed. Circles of different colors

304 represent individuals of different species. See Table 3 for detailed description of each sampling

305 curve.

306 The mechanics of isolating the distinct effects of spatial aggregation, density, and SAD

307 The three curves capture different components of community structure that influence

308 richness changes across scales (measured in number of samples or number of individuals, both of

309 which can be easily converted to area, Table 3). Therefore, if we assume the components
310 contribute additively to richness, then the effect of a treatment on richness propagated through a
311 single component at any scale can be obtained by subtracting the rarefaction curves from each
312 other. For simplicity and tractability, we assume additivity to capture first-order effects. This
313 assumption is supported by Tjørve *et al.*'s (2008) demonstration that an additive partitioning of
314 richness using rarefaction curves reveals random sampling and aggregation effects when using
315 presence-absence data. We further validated this assumption using sensitivity analysis (see
316 "Sensitivity analysis of the method" and Table 5). Below we describe the algorithm to obtain the
317 distinct effect of each component. Figure 3 provides a graphic illustration.

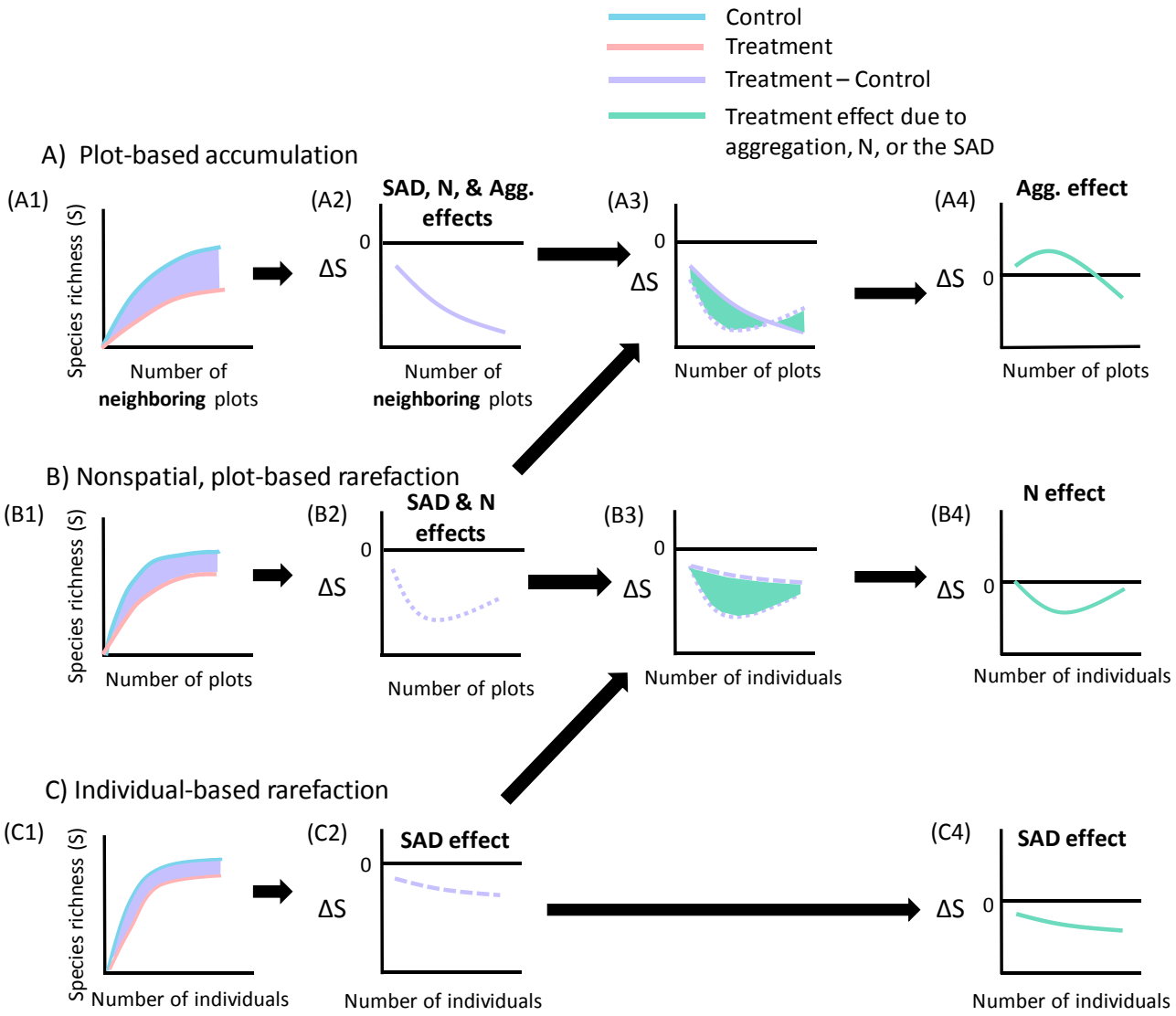


Figure 3. Steps separating the distinct effect of the three factors on richness. The experimental design has two treatments (blue and orange curves). The purple shaded area on the left and the equivalent purple curve in each plot to the right represent the difference in richness (i.e., treatment effect) for each set of curves. By taking the difference again (green shaded area and curves) we can obtain the treatment effect on richness through a single component. See text for details (Eqn 1.). The three types of curves are defined in Fig. 2 and Table 3.

i) Effect of aggregation

The difference between the plot-based accumulation curves of two treatments, $\Delta\left(S_{21}\middle|k,\vec{n}_{t,+},\vec{N}_t,\vec{d}_t\right) = \mathrm{E}\left[S_2\middle|k,\vec{n}_{2,+},\vec{N}_2,\vec{d}_2\right] - \mathrm{E}\left[S_1\middle|k,\vec{n}_{1,+},\vec{N}_1,\vec{d}_1\right]$, gives the observed difference in richness between treatments across scales (Fig. 3A2, solid purple curve). It encapsulates the treatment effect propagated through all three components: shape of the SAD, density of individuals, and spatial aggregation. Differences between treatments in any of these factors could potentially translate into observed difference in species richness.

Similarly, the difference between the non-spatial, plot-based rarefaction curves, $\Delta\left(S_{21}\middle|k,\vec{n}_{t,+},\vec{N}_t\right) = \mathrm{E}\left[S_2\middle|k,\vec{n}_{2,+},\vec{N}_2\right] - \mathrm{E}\left[S_1\middle|k,\vec{n}_{1,+},\vec{N}_1\right]$, gives the expected difference in richness across treatments when spatial aggregation is removed (Fig. 3B2, purple dotted curve). The distinct effect of aggregation across treatments from one plot to *k* plots can thus be obtained by taking the difference between the two Δ*S* values (Fig. 3A3, green shaded area), i.e.,

$$\Delta(S_{21}|\text{aggregation}) = \Delta\left(S_{21}\middle|k,\vec{n}_{t,+},\vec{N}_t,\vec{d}_t\right) - \Delta\left(S_{21}\middle|k,\vec{n}_{t,+},\vec{N}_t\right)$$

$$= \underbrace{\left(\mathrm{E}\left[S_2\middle|k,\vec{n}_{2,+},\vec{N}_2,\vec{d}_2\right] - \mathrm{E}\left[S_1\middle|k,\vec{n}_{1,+},\vec{N}_1,\vec{d}_1\right]\right)}_{\text{effect of aggregation, density, and SAD}} - \underbrace{\left(\mathrm{E}\left[S_2\middle|k,\vec{n}_{2,+},\vec{N}_2\right] - \mathrm{E}\left[S_1\middle|k,\vec{n}_{1,+},\vec{N}_1\right]\right)}_{\text{effect of density and SAD}} \quad \text{(Eqn 1)}$$

Equation 1 demonstrates that the effect of aggregation can be thought of as the difference between treatment effects quantified by the plot-based accumulation and plot-based rarefaction curves. An algebraic rearrangement of Eqn 1 demonstrates that $\Delta(S_{21}|\text{aggregation})$ can also be thought of as the difference between the treatments of the same type of rarefaction curve:

$$= \underbrace{\left(\mathrm{E}\left[S_2\middle|k,\vec{n}_{2,+},\vec{N}_2,\vec{d}_2\right] - \mathrm{E}\left[S_2\middle|k,\vec{n}_{2,+},\vec{N}_2\right]\right)}_{\text{effect of aggregation in Treatment 2}} - \underbrace{\left(\mathrm{E}\left[S_1\middle|k,\vec{n}_{1,+},\vec{N}_1,\vec{d}_1\right] - \mathrm{E}\left[S_1\middle|k,\vec{n}_{1,+},\vec{N}_1\right]\right)}_{\text{effect of aggregation in Treatment 1}} \quad \text{(Eqn 2)}$$

349 This simple duality can be extended to the estimation of the density and SAD effects, but we will
350 only consider the approach laid out in Eqn 1 below. In Fig. 3, we separate each individual effect
351 using the approach of Eqn 1 while the code in the `mobr` package uses the approach of Eqn 2.
352 One thing to note is that the effect of aggregation always converges to zero at the
353 maximal spatial scale ($k = K$ plots) for a balanced design. This is because, when all plots have
354 been accumulated, $\Delta\left(S_{21}\middle|k, \vec{n}_{t,+}, \vec{N}_t, \vec{d}_t\right)$ and $\Delta\left(S_{21}\middle|k, \vec{n}_{t,+}, \vec{N}_t\right)$ will both converge on the
355 difference in total richness between the treatments. However, for an unbalanced design in which
356 one treatment has more plots than the other, $\Delta(S_{21}|\text{aggregation})$ would converge to a nonzero
357 constant because $\mathrm{E}\left[S_t\middle|k, \vec{n}_{t,+}, \vec{N}_t, \vec{d}_t\right] - \mathrm{E}\left[S_t\middle|k, \vec{n}_{t,+}, \vec{N}_t\right]$ would be zero for one treatment but not
358 the other at the maximal spatial scale (i.e., min(K_1, K_2) plots). This artefact is inevitable and
359 should not be interpreted as a real decline in the relative importance of aggregation on richness,
360 but as our diminishing ability to detect such effect without sampling a larger region.
361 *ii) Effect of density:*
362 In the same vein, the difference between the individual-based rarefaction curves of the two
363 treatments, $\Delta\left(S_{21}\middle|N, \vec{n}_{t,+}\right) = \mathrm{E}\left[S_2\middle|N, \vec{n}_{2,+}\right] - \mathrm{E}\left[S_1\middle|N, \vec{n}_{1,+}\right]$, yields the treatment effect on
364 richness propagated through the shape of the SAD alone, with the other two components
365 removed (Fig. 3C2, purple dashed curve). The distinct effect of density across treatments from
366 one individual to N individuals can thus be obtained by subtracting the ΔS value propagated
367 through the shape of the SAD alone from the ΔS value propagated through the compound effect
368 of the SAD and density (Fig. 3B3, green shaded area), i.e.,
369

$$\Delta(S_{21}|\text{density}) = \Delta\left(S_{21}\middle|N, \vec{n}_{t,+}, \vec{N}_t\right) - \Delta\left(S_{21}\middle|N, \vec{n}_{t,+}\right)$$

$$= \left(\mathrm{E}\left[S_2\middle|N, \vec{n}_{2,+}, \vec{N}_2\right] - \mathrm{E}\left[S_1\middle|N, \vec{n}_{1,+}, \vec{N}_1\right]\right) - \left(\mathrm{E}\left[S_2\middle|N, \vec{n}_{2,+}\right] - \mathrm{E}\left[S_1\middle|N, \vec{n}_{1,+}\right]\right) \quad \text{(Eqn 3)}$$

effect of density and SAD — effect of SAD

Note that in Eqn 3, spatial scale is defined with respect to numbers of individuals sampled (*N*) (and thus the grain size that would be needed to achieve this) rather than the number of samples (*k*).

iii) Effect of SAD:

The distinct effect of the shape of the SAD on richness between the two treatments is simply the difference between the two individual-based rarefaction curves (Fig. 3B, purple dashed curve), i.e.,

$$\Delta(S_{21}|\text{SAD}) = \Delta\left(S_2\middle|p, \vec{n}_{2,+}\right) - \Delta\left(S_1\middle|p, \vec{n}_{1,+}\right) \quad \text{(Eqn 4)}$$

The scale of $\Delta(S_{21}|\text{SAD})$ ranges from one individual, where both individual rarefaction curves have one species and thus $\Delta(S_{21}|\text{SAD}) = 0$, to $N_{\min} = \min(N_{1,+}, N_{2,+})$, which is the lower total abundance between the treatments.

The formulae used to identify the distinct effect of the three factors are summarized in Table 4.

389 Table 4. Calculation of effect size curves.

Factor	Formula	Note
Aggregation	$\Delta(S_{21}\|\text{aggregation}) = \Delta\left(S_{21}\middle\|k,\vec{n}_{t,+},\vec{N}_t,\vec{d}_t\right) - \Delta\left(S_{21}\middle\|k,\vec{n}_{t,+},\vec{N}_t\right)$	Artificially, this effect always converges to zero at the maximal spatial scale (*K* plots) for a balanced design, or a non-zero constant for an unbalanced design.
Density	$\Delta(S_{21}\|\text{density}) = \Delta\left(S_{21}\middle\|N,\vec{n}_{t,+},\vec{N}_t\right) - \Delta\left(S_{21}\middle\|N,\vec{n}_{t,+}\right)$	To compute this quantity, the x-axes of the plot-based rarefaction curves are converted from plots to individuals using average individual density
SAD	$\Delta(S_{21}\|\text{SAD}) = \Delta\left(S_2\middle\|N,\vec{n}_{2,+}\right) - \Delta\left(S_1\middle\|N,\vec{n}_{1,+}\right)$	This is estimated directly by comparing the individual rarefaction curves between two treatments.

390

391 Significance tests and acceptance intervals

392 In the continuous-scale analysis, we also applied Monte Carlo permutation procedures to 1)
393 construct acceptance intervals (or non-rejection intervals) across scales on simulated null
394 changes in richness, and 2) carry out goodness of fit tests on each component (Loosmore and
395 Ford 2006, Diggle-Cressie-Loosmore-Ford test [DCLF]; Baddeley et al. 2014). See Supplement
396 S3 for descriptions of how each set of randomizations was developed to generate 95%
397 acceptance intervals (ΔS_{null}) which can be compared to the observed changes (ΔS_{obs}). Strict
398 interpretations of significance in relation to the acceptance intervals is not warranted because
399 each point along the spatial scale (x-axis) is effectively a separate comparison. Consequently, a
400 problem arises with multiple non-independent tests and the 95% bands cannot be used for formal
401 significance testing due to Type I errors. The DCLF test (see Supplement S3) provides an overall
402 significance test with a proper Type I error rate (Loosmore and Ford, 2006) but this test in turn

suffers from Type II error (Baddeley et al. 2014). There is no mathematical resolution to this and user judgement should be emphasized if formal *p*-tests are needed.

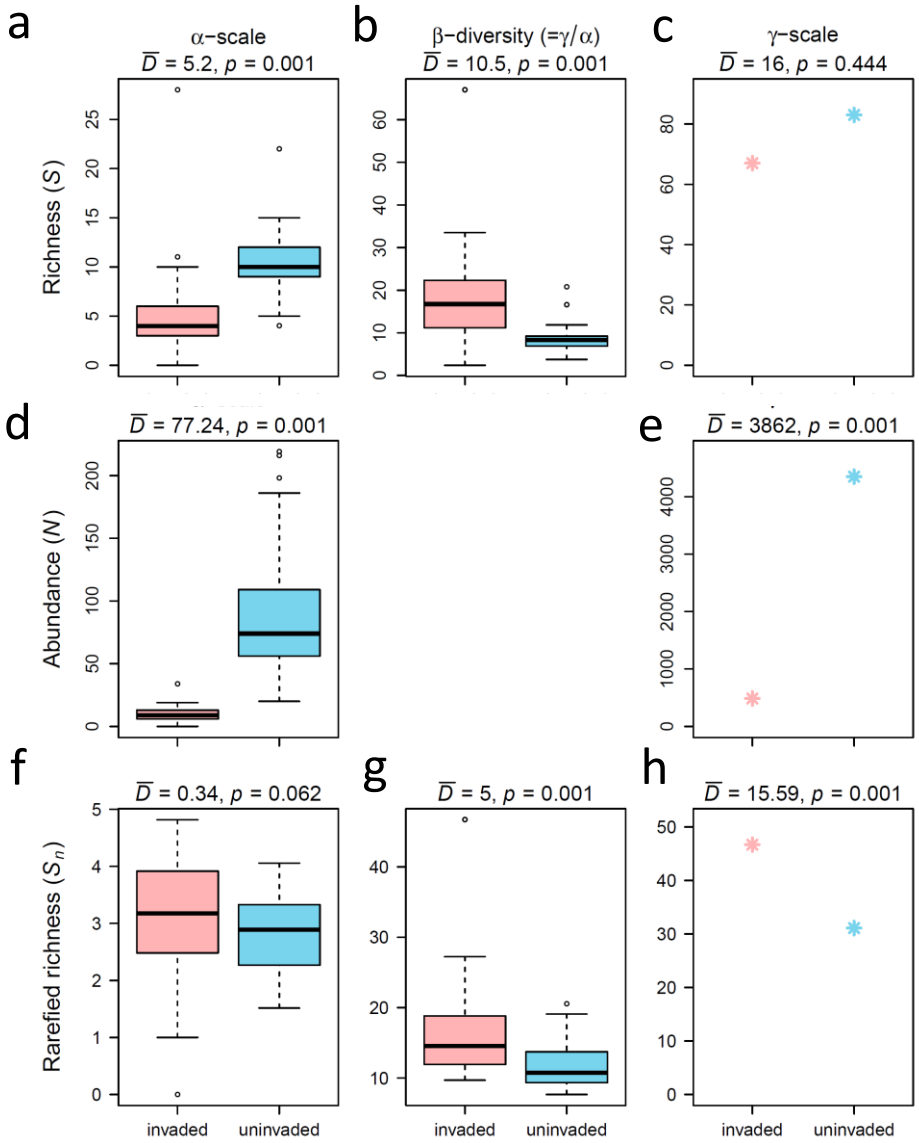
Sensitivity Analysis

Although the logic justifying the examination separating the effect of the three components is rigorous, we tested the validity of our approach (and the significance tests) by simulations using the R package `mobsim` (May et al. preprint, May 2017). The goal is to establish the rate of type I error (i.e., detecting significant treatment effect through a component when it does not differ between treatments) and type II error (i.e., nonsignificant treatment effect through a component when it does differ). This was achieved by systematically comparing simulated communities in which we altered one or more components while keeping the others unchanged (see Supplement S4). Overall, the benchmark performance of our method was good. When a factor did not differ between treatments, the detection of significant difference was low (Supplemental Table S4.1). Conversely, when a factor did differ, the detection of significant difference was high, but decreased at smaller effect sizes. Thus, we were able to control both Type I and Type II errors at reasonable levels. In addition, there did not seem to be strong interactions among the components – the error rates remained consistently low even when two or three components were changed simultaneously.

An empirical example

In this section, we illustrate the potential of our method with an empirical example, previously analyzed by Powell et al. (2013). Invasion of an exotic shrub, *Lonicera maackii*, has caused significant, but strongly scale-dependent, decline in the diversity of understory plants in eastern Missouri (Powell et al. 2013). Specifically, Powell et al. (2013) showed that the effect size of the invasive plant on herbaceous plant species richness was large at relatively plot-level

426 spatial scales (1 m^2), but the proportional effect declines with increasing windows of
427 observations, with the effect becoming negligible at the largest spatial scale (500 m^2). Using a
428 null model approach, the authors further identified that the negative effect of invasion was
429 mainly due to the decline in plant density observed in invaded plots. To recreate these analyses
430 run the R code achieved here:
431 https://github.com/MoBiodiv/mobr/blob/master/scripts/methods_ms_figures.R.
432 The original study examined the effect of invasion at multiple scales using the slope and
433 intercept of the species-area relationship. We now apply our MoB approach to data from one of
434 their sites from Missouri, where the numbers of individuals of each species were recorded from
435 50 1-m^2 plots sampled from within a 500-m^2 region in the invaded part of the forest, and another
436 50 plots from within a 500-m^2 region in the uninvaded part of the forest. Our method leads to
437 conclusions that are qualitatively similar to the original study, but with a richer analysis of the
438 scale dependence. Moreover, our new methods show that invasion influenced both the SAD and
439 spatial aggregation, in addition to density, and that these effects went in different directions and
440 depended on spatial scale.



441

442 Figure 4. Simple two scale analysis output for case study. Biodiversity statistics for the invaded
443 (red boxplots and points) and uninvaded (blue boxplots and points) for vascular plant species
444 richness at the α (i.e., single plot), beta (i.e., between plots), and γ (i.e., all plots) scales. The p-
445 values are based on 999 permutations of the treatment labels. Rarefied richness (S_n, panels f-h)
446 was computed for 5 and 250 individuals for the α (f) and γ (h) scales respectively.

447 The two-scale analysis suggests that invasion decreases average richness (S) at the α (Fig.
448 4a, $\overline{D}$ = 5.2, p = 0.001) but not γ scale (Fig. 4c, $\overline{D}$ = 16, p = 0.438). Invasion also decreased total

449 abundance (N, Fig. 4d,e, $p = 0.001$) which suggests that the decrease in S at the α scale may be
450 due to a decrease in individual abundance. Rarefied richness (S_n, Fig. 4f,h) allows us to test this
451 hypothesis directly. Specifically, we found S_n was higher in the invaded areas (significantly so $\overline{D}$
452 = 15.59, $p = 0.001$ at the γ scale defined here as $n = 250$; Fig. 4h) which indicates that once the
453 negative abundance effect was controlled for, invasion actually increased diversity through an
454 increase in species evenness.
455 To identify whether the increase in evenness due to invasion was primarily because of
456 shifts in common or rare species, we examined ENS of PIE (S_{PIE}) and the undetected species
457 richness (f_0) (see Fig.5). At the α scale, invasion did not strongly influence the SAD (Fig. 5a,d),
458 but at the γ scale, there was evidence that invaded sites had greater evenness in the common
459 species (Fig. 5f, $\overline{D} = 5.78$, $p = 0.001$). In other words, the degree of dominance by any one
460 species was reduced.
461

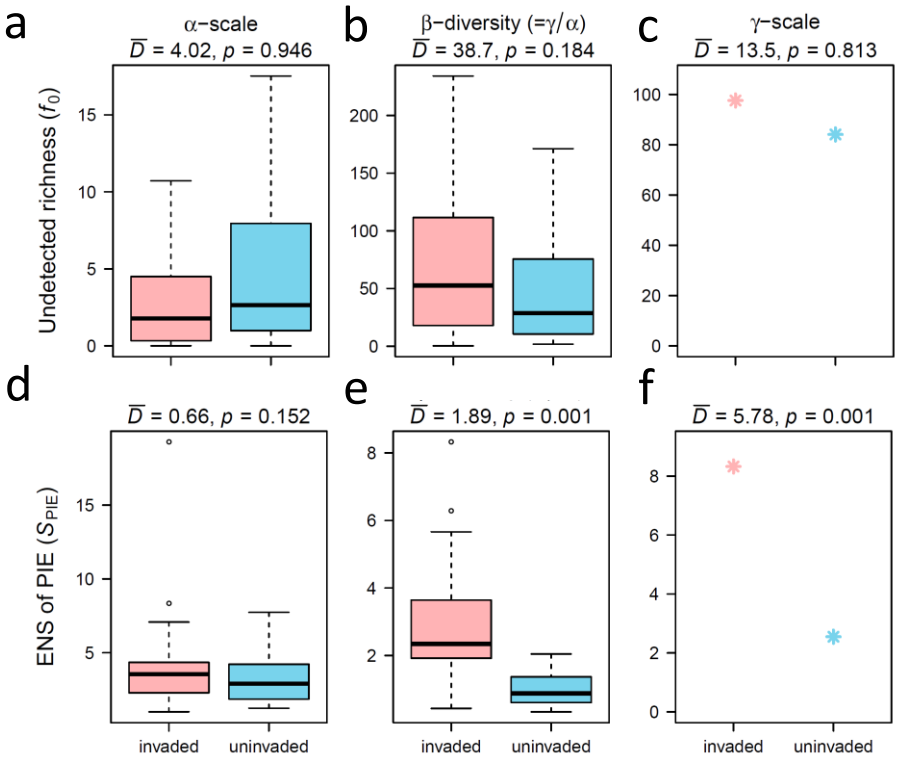


Figure 5. The two-scale analysis applied to metrics of biodiversity that emphasize changes in the SAD. Colors as described in Fig. 4. f_0 (a-c) is more sensitive to rare species, and S_{PIE} (d-f) is more sensitive to common species. The *p*-values are based on 999 permutations of the treatment labels, and outliers were removed from the f_0 plot.

The β diversity metrics were significantly higher (Fig. 4b,g, Fig. 5e, $p = 0.001$) in the invaded sites (with the exception of β_{f0}, Fig. 5b), suggesting that uninvaded sites had lower spatial species turnover and thus were more homogenous. It did not appear that changes in N were solely responsible for the changes in beta-diversity because β_{Sn} displayed a very similar, but slightly weaker, pattern as raw β_S (Fig. 4b,g).

Overall the two-scale analysis indicates: 1) that there are scale-dependent shifts in richness, 2) that these are caused by invasion decreasing N, and increasing evenness in common species, and increasing species patchiness.

475 Applying the continuous scale analysis, we further disentangled the effect of invasion on
476 diversity through the three components (SAD, density, and aggregation) across all scales of
477 interest. The results are shown in Fig. 6 which parallels the panels of the conceptual Fig. 3. Fig.
478 6a-c present the three sets of curves for the two treatments: the plot-based accumulation curve, in
479 which plots accumulate by their spatial proximity (Fig. 6a); the (non-spatial) plot-based
480 rarefaction curve, in which individuals are randomized across plots within a treatment (Fig. 6b);
481 and the individual-based rarefaction curve, in which species richness is plotted against number of
482 individuals (Fig. 6c). Fig. 6d-f show the effect of invasion on richness, obtained by subtracting
483 the red curve from the blue curve for each pair of curves (which correspond to the curves of the
484 same color in Fig. 3). The bottom panel, which shows the effect of invasion on richness through
485 each of the three factors, is obtained by subtracting the curves in the middle panel from each
486 other. The contribution of each component to difference in richness between the invaded and
487 uninvaded sites is further illustrated in Fig. 6.
488

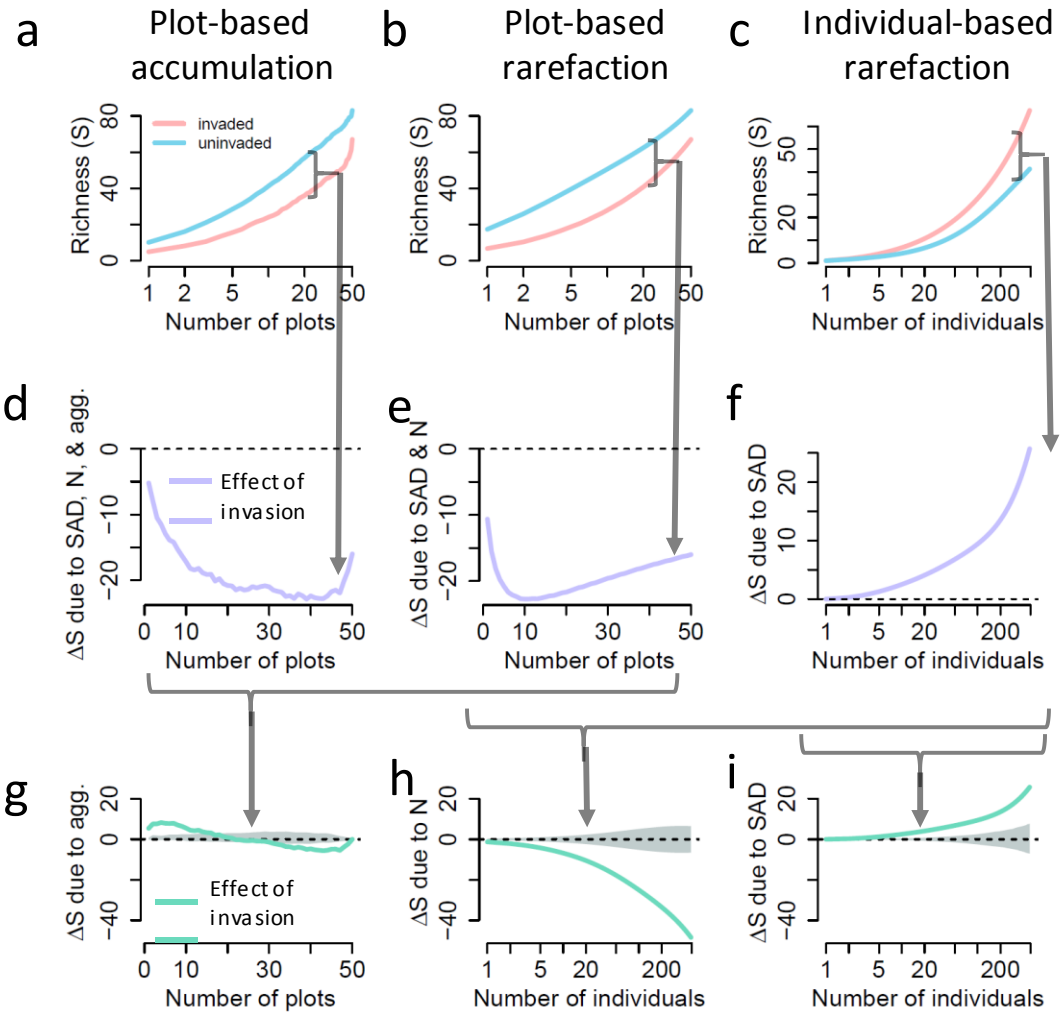


Figure 6. Applying the MoB continuous scale analysis on the invasion data set. The colors are as in Fig. 3. Panel a, shows the invaded (red) and uninvaded (blue) accumulation and rarefaction. In panel b, the purple curves show the difference in richness (uninvaded – invaded) for each set of curves. In panel c, the green curves show the treatment effect on richness through each of the three components, while the grey shaded area shows the 95% acceptance interval for the null model, the cross scale DCLF test for each factor was significant (p = 0.001). The dashed line shows the point of no-change in richness between the treatments.

Consistent with the original study, our approach shows that the invaded site had lower richness than the uninvaded site at all scales (Fig. 6a). Separating the effect of invasion into the

three components, we find that invasion actually had a positive effect on species richness through its impact on the shape of the SAD (Fig. 6i, Fig. 7a), which contributed to approximately 20% of the observed change in richness (Fig. 7b). This counterintuitive result suggests that invasion has made the local community more even, meaning that the dominant species were most significantly influenced by the invader. However, this positive effect was completely overshadowed by the negative effect on species richness through reductions in the density of individuals (Fig. 6h, Fig. 7a), which makes a much larger contribution to the effect of invasion on richness (as large as 80%, Fig. 7b). Thus, the most detrimental effect of invasion was the sharp decline in the number of individuals. The effect of aggregation (Fig. 6g), is much smaller compared with the other two components and was most important at small spatial scales. Our approach thus validates the findings in the original study, but provides a more comprehensive way to quantify the contribution to richness decline caused by invasion by each of the three components, at every spatial scale.

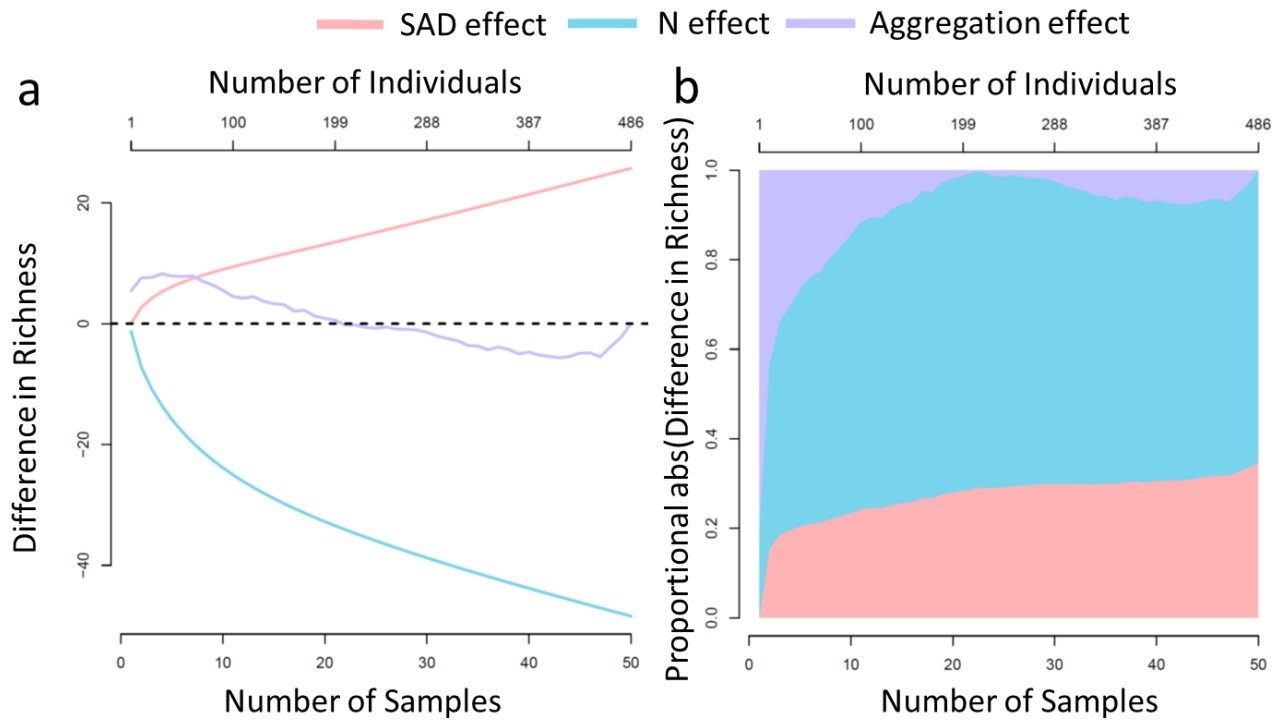


Figure 7. The effect of invasion on richness via individual effects on three components of community structure: SAD in red, density in blue, aggregation in purple across scales. The raw differences (a) and proportional stacked absolute values (b). The x-axis represents sampling effort in both numbers of samples (i.e., plots) and individuals (see top axis). The rescaling between numbers of individuals and plots we carried out by defining the maximum number of individuals rarefied to (486 individuals) as equivalent to the maximum number of plots rarefied to (50 plots), other methods of rescaling are possible. In panel (a) the dashed black line indicates no change in richness.

Discussion

How does species richness differ between experimental conditions or among sites that differ in key parameters in an observational study? This fundamental question in ecology often lacks a simple answer, because the magnitude (and sometimes even the direction) of change in

527 richness may vary with spatial scale (Chase et al. submitted, Chalcraft et al. 2004, Fridley et al.
528 2004, Knight and Reich 2005, Palmer et al. 2008, Chase and Knight 2013, Powell et al. 2013,
529 Blowes et al. 2017). Species richness is proximally determined by three underlying
530 components—N, SAD and aggregation—which are also scale-dependent (Powell, Chase &
531 Knight 2013, McGill 2011); this obscures the interpretation of the link between change in
532 condition and change in species richness.

533 The MoB framework provides a comprehensive answer to this question by taking a
534 spatially explicit approach and decomposing the effect of the condition (treatment) on richness
535 into its individual components. The two-scale analysis provides a big-picture understanding of
536 the differences and proximate drivers of richness by only examining the single plot (α) and all
537 plots combined (γ) scales. The continuous scale analysis expands the endeavor to cover a
538 continuum of scales, and quantitatively decomposes change in richness into three components:
539 change in the shape of the SAD, change in individual density, and change in spatial aggregation.
540 As such, we can not only quantify how richness changes at any scale of interest, but also identify
541 how the change occurs and consequently push the ecological question to a more mechanistic
542 level. For example, we can ask to what extent the effects on species richness are driven by
543 numbers of individuals. Or instead, whether common and rare species, or their spatial
544 distributions, are more strongly influenced by the treatments.

545 Here we considered the scenario of comparing a discrete treatment effect on species
546 richness, but clearly the MoB framework will need to be extended to other kinds of experimental
547 designs and questions (fully described in Supplement S5). The highest priority extension of the
548 framework is to generalize it from a comparison of discrete treatment variables to continuous
549 drivers such as temperature and productivity. Additionally, we recognize that abundance is

550 difficult to collect for many organisms and that there is a need to understand if alternative
551 measures of commonness (e.g., visual cover, biomass) can also be used to gain similar insights.
552 Finally, we have only focused on taxonomic diversity here, whereas other types of
553 biodiversity—most notably functional and phylogenetic diversity—are often of great interest,
554 and comparisons such as those we have overviewed here would also be of great importance for
555 these other biodiversity measures. Importantly, phylogenetic and functional diversity measures
556 share many properties of taxonomic diversity that we have overviewed here (e.g., scale-
557 dependence, non-linear accumulations, rarefactions, etc) (e.g., Chao et al. 2014), and it would
558 seem quite useful to extend our framework to these sorts of diversities. We look forward to
559 working with the community to develop extensions of the MoB framework that are most needed
560 for understanding scale dependence in diversity change.

561 MoB is a novel and robust approach that explicitly addresses the issue of scale-
562 dependence in studies of diversity, and quantitatively disentangles diversity change into its three
563 components. Our method demonstrates how spatially explicit community data and carefully
564 framed comparisons can be combined to yield new insight into the underlying components of
565 biodiversity. We hope the MoB framework will help ecologists move beyond single-scale
566 analyses of simple and relatively uninformative metrics such as species richness alone. We view
567 this as a critical step in reconciling much confusion and debate over the direction and magnitude
568 of diversity responses to natural and anthropogenic drivers. Ultimately accurate predictions of
569 biodiversity change will require knowledge of the relevant drivers and the spatial scales over
570 which they are most relevant, which MoB (and its future extensions), helps to uncover.
571

Acknowledgments

This paper emerged from several workshops funded with the support (to JMC) from the German Centre for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig funded by the German Research Foundation (FZT 118) and by the Alexander von Humboldt Foundation as part of the Alexander von Humboldt Professorship of TMK. DJM was also supported by College of Charleston startup funding. We further thank N. Sanders, J. Belmaker, and D. Storch for discussions and comments on our approach.

Data Accessibility

The data is archived with the R package on GitHub: https://github.com/mobiodiv/mobr

Literature Cited

Baddeley, A., P. J. Diggle, A. Hardegen, T. Lawrence, R. K. Milne, and G. Nair. 2014. On tests of spatial pattern based on simulation envelopes. Ecological Monographs 84:477–489.

Blowes, S. A., J. Belmaker, and J. M. Chase. 2017. Global reef fish richness gradients emerge from divergent and scale-dependent component changes. Proc. R. Soc. B 284:20170947.

Cayuela, L., N. J. Gotelli, and R. K. Colwell. 2015. Ecological and biogeographic null hypotheses for comparing rarefaction curves. Ecological Monographs 85:437–455.

Chalcraft, D. R., J. W. Williams, M. D. Smith, and M. R. Willig. 2004. Scale dependence in the species-richness-productivity relationship: The role of species turnover. Ecology 85:2701–2708.

Chao, A. 1984. Nonparametric Estimation of the Number of Classes in a Population. Scandinavian Journal of Statistics 11:265–270.

593 Chao, A., C. H. Chiu, and L. Jost. 2014. Unifying species diversity, phylogenetic diversity,
594 functional diversity, and related similarity and differentiation measures through Hill
595 numbers. Annual Review of Ecology, Evolution, and Systematics 45:297–324.
596 Chase, J. M., and T. M. Knight. 2013. Scale-dependent effect sizes of ecological drivers on
597 biodiversity: why standardised sampling is not enough. Ecology Letters 16:17–26.
598 Chase, J. M., and M. A. Leibold. 2002. Spatial scale dictates the productivity-biodiversity
599 relationship. Nature 416:427–430.
600 Chase, J. M., B. . J. McGill, D. J. McGlinn, F. May, S. A. Blowes, X. Xiao, T. M. Knight, O.
601 Purschke, and N. J. Gotelli. submitted. A scale-explicit guide for comparing biodiversity
602 across communities.
603 Chiarucci, A., G. Bacaro, D. Rocchini, C. Ricotta, M. W. Palmer, and S. M. Scheiner. 2009.
604 Spatially constrained rarefaction: incorporating the autocorrelated structure of biological
605 communities into sample-based rarefaction. Community Ecology 10:209–214.
606 Collins, M. D., and D. Simberloff. 2009. Rarefaction and nonrandom spatial dispersion patterns.
607 Environmental and Ecological Statistics 16:89–103.
608 Fridley, J. D., R. L. Brown, and J. E. Bruno. 2004. Null models of exotic invasion and scale-
609 dependent patterns of native and exotic species richness. Ecology 85:3215–3222.
610 Gotelli, N. J., and R. K. Colwell. 2001. Quantifying biodiversity: procedures and pitfalls in the
611 measurement and comparison of species richness. Ecology Letters 4:379–391.
612 Harte, J., T. Zillio, E. Conlisk, and A. B. Smith. 2008. Maximum entropy and the state-variable
613 approach to macroecology. Ecology 89:2700–2711.
614 He, F. L., and P. Legendre. 2002. Species diversity patterns derived from species-area models.
615 Ecology 83:1185–1198.

616 Hubbell, S. P. 2001. The Unified Neutral Theory of Biodiversity and Biogeography. Princeton
617 University Press, Princeton, NJ, USA.
618 Hurlbert, A. H. 2004. Species-energy relationships and habitat complexity in bird communities.
619 Ecology Letters 7:714–720.
620 Hurlbert, S. H. 1971. The nonconcept of species diversity: a critique and alternative parameters.
621 Ecology 52:577–586.
622 Karlson, R. H., H. V. Cornell, and T. P. Hughes. 2007. Aggregation influences coral species
623 richness at multiple spatial scales. Ecology 88:170–177.
624 Knight, K. S., and P. B. Reich. 2005. Opposite relationships between invasibility and native
625 species richness at patch versus landscape scales. Oikos 109:81–88.
626 Legendre, P., and L. Legendre. 1998. Numerical ecology. 2nd English Edition. Elsevier, Boston,
627 Mass., USA.
628 Leibold, M. A., and J. M. Chase. 2017. Metacommunity ecology. Princeton University Press,
629 Princeton, NJ.
630 Loosmore, N. B., and E. D. Ford. 2006. Statistical Inference Using the G or K Point Pattern
631 Spatial Statistics. Ecology 87:1925–1931.
632 May, F. 2017. mobsim: Spatial Simulation and Scale-Dependent Analysis of Biodiversity
633 Changes.
634 May, F., K. Gerstner, D. J. McGlinn, X. Xiao, and J. M. Chase. preprint. mobsim: An R package
635 for the simulation and measurement of biodiversity across spatial scales | bioRxiv.
636 bioRxiv 209502.
637 McGill, B. J. 2010. Towards a unification of unified theories of biodiversity. Ecology Letters
638 13:627–642.

639 McGill, B. J. 2011. Species abundance distributions. Pages 105–122 Biological Diversity:
640 Frontiers in Measurement and Assessment, eds. A.E. Magurran and B.J. McGill.
641 McGlinn, D. J., and M. W. Palmer. 2009. Modeling the sampling effect in the species-time-area
642 relationship. Ecology 90:836–846.
643 McGlinn, D. J., X. Xiao, J. Kitzes, and E. P. White. 2015. Exploring the spatially explicit
644 predictions of the Maximum Entropy Theory of Ecology. Global Ecology and
645 Biogeography 24:675–684.
646 Palmer, M. W., and E. van der Maarel. 1995. Variance in species richness, species association,
647 and niche limitation. Oikos 73:203–213.
648 Palmer, M. W., D. J. McGlinn, and J. F. Fridley. 2008. Artifacts and artifictions in biodiversity
649 research. Folia Geobotanica 43:245–257.
650 Powell, K. I., J. M. Chase, and T. M. Knight. 2013. Invasive Plants Have Scale-Dependent
651 Effects on Diversity by Altering Species-Area Relationships. Science 339:316–318.
652 Rosenzweig, M. L. 1995. Species Diversity in Space and Time. Cambridge University Press,
653 Cambridge, UK.
654 Simberloff, D. 1972. Properties of the Rarefaction Diversity Measurement. The American
655 Naturalist 106:414–418.
656 Supp, S. R., X. Xiao, S. K. M. Ernest, and E. P. White. 2012. An experimental test of the
657 response of macroecological patterns to altered species interactions. Ecology 93:2505–
658 2511.
659 Tilman, D. 1994. Competition and biodiversity in spatially structured habitats. Ecology 75:2–16.

660 Tjørve, E., W. E. Kunin, C. Polce, and K. M. C. Tjørve. 2008. Species-area relationship:
661 separating the effects of species abundance and spatial distribution. Journal of Ecology
662 96:1141–1151.
663 Vellend, M. 2016. The Theory of Ecological Communities. Princeton University Press,
664 Princeton, NJ, USA.
665 Whittaker, R. H. 1960. Vegetation of the Siskiyou Mountains, Oregon and California. Ecological
666 Monographs 30:279–338.
667