

Using coverage-based rarefaction to infer non-random species distributions

THORE ENGEL ^{1,2,†} SHANE A. BLOWES ^{1,2} DANIEL J. MCGILN ³ FELIX MAY ⁴
NICHOLAS J. GOTELLI ⁵ BRIAN J. MCGILL ⁶ AND JONATHAN M. CHASE ^{1,2}

¹German Centre for Integrative Biodiversity Research (iDiv), Halle-Jena-Leipzig, 04103 Leipzig, Germany

²Institute of Computer Science, Martin Luther University Halle-Wittenberg, 06120 Halle (Saale), Germany

³Department of Biology, College of Charleston, Charleston, 29424 South Carolina, USA

⁴Institute of Biology, Freie Universität Berlin, 14195 Berlin, Germany

⁵Department of Biology, University of Vermont, Burlington, 05405 Vermont, USA

⁶School of Biology and Ecology, and Senator George J. Mitchell Center of Sustainability Solutions, University of Maine, Orono, 04469 Maine, USA

Citation: Engel, T., S. A. Blowes, D. J. McGlinn, F. May, N. J. Gotelli, B. J. McGill, and J. M. Chase. 2021. Using coverage-based rarefaction to infer non-random species distributions. *Ecosphere* 12(9):e03745. 10.1002/ecs2.3745

Abstract. Understanding how species are non-randomly distributed in space and how the resulting spatial structure responds to ecological, biogeographic, and anthropogenic drivers is a critical piece of the biodiversity puzzle. However, most metrics that quantify the spatial structure of diversity (i.e., community differentiation), such as Whittaker's β -diversity, depend on sampling effort and are influenced by species pool size, species abundance distributions, and numbers of individuals. Null models are useful for identifying the degree of differentiation among communities due to spatial structuring relative to that expected from sampling effects, but do not accommodate the influence of sample completeness (i.e., the proportion of the species pool in a given sample). Here, we develop an approach that makes use of individual- and coverage-based rarefaction and extrapolation, to derive a metric, β_C , which captures changes in intraspecific aggregation independently of changes in the species pool size. We illustrate the metric using spatially explicit simulations and two case studies: (1) a re-analysis of the "Gentry" plot data set consisting of small forest plots spanning a latitudinal gradient from North to South America and (2) comparing a large plot in high diversity tropical forests of Barro Colorado Island, Panama, with a plot in a lower diversity temperate forest in Harvard Forest, Massachusetts, USA. We find no evidence for systematic changes in spatial structure with latitude in these data sets. As it is rooted in biodiversity sampling theory and explicitly controls for sample completeness, our approach represents an important advance over existing null models for spatial aggregation. Potential applications range from better descriptors of biogeographic diversity patterns to the consolidation of local and regional diversity trends in the current biodiversity crisis.

Key words: β -deviation; β -diversity; coverage; null models; rarefaction; sample completeness; spatial aggregation; species abundance distribution; species pool; turnover.

Received 8 March 2021; revised 29 April 2021; accepted 6 May 2021. Corresponding Editor: Debra P. C. Peters.

Copyright: © 2021 The Authors. This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

† **E-mail:** thore.engel@posteo.de

INTRODUCTION

Species are non-randomly distributed across the globe, and understanding spatial patterns of

species diversity from ecological samples remains a central challenge (Gaston 2000, McGill 2011, Worm and Tittensor 2018). Spatial structure in species diversity (e.g., species turnover

from site to site) is typically quantified by one or more metrics of compositional dissimilarity or spatial β -diversity (Anderson et al. 2011). Measures of β -diversity offer a mathematical link between local (i.e., α) and regional (i.e., γ) species diversity and can, for example, shed light onto the metacommunity processes that shape biological assemblages (Chase and Myers 2011) inform biodiversity conservation (Socolar et al. 2016), and help understand the provisioning of ecosystem functions and services (Mori et al. 2018).

Although β -diversity is conceptually appealing, its quantification and interpretation are often ambiguous (Tuomisto 2010a, b, Anderson et al. 2011). Whittaker's (1960) multiplicative β -diversity (γ/α), for example, is commonly thought to represent the sort of community differentiation that arises due to non-random distributions of species (e.g., species turnover or intraspecific spatial aggregation). However, this and other measures of β -diversity are also influenced by the size and number of samples, the size of the regional species pool, the shape of the regional species abundance distribution (SAD), and the number of individuals captured by the samples (McGill 2011, Chase and Knight 2013, Chase et al. 2018). This makes it a challenge to compare and interpret patterns of β -diversity and related measures along biogeographic gradients (e.g., Kraft et al. 2011). To disentangle the effect of spatial aggregation from the non-spatial components that influence β -diversity (SAD or relative proportion of rare species, species pool size), empirical studies have frequently adopted null-modeling approaches that compare the observed patterns against a null expectation that simulates spatial randomness by shuffling individuals among sites (Chase et al. 2011, Kraft et al. 2011). However, the exact formulation of the null expectation and its deviation (β -deviation) remains debated (Kraft et al. 2012, Qian et al. 2012, Tucker et al. 2016, Mori et al. 2018, Xing and He 2021). In particular, the null model approach has been criticized because it overlooks the influence of the completeness of the samples (Ulrich et al. 2017, Sreekar et al. 2018).

Much of the ambiguity surrounding measures of β -diversity and its null expectations can be understood in terms of sampling effects and

sample completeness (i.e., the proportion of species in the species pool captured by sampling). For instance, regions with large species pools are expected to exhibit high β -diversity simply because local samples only capture a small and incomplete portion of the total diversity; this can lead to strong, but spurious differentiation among local samples (Chase and Myers 2011, Kraft et al. 2011). This is not to say that this kind of sample differentiation is not meaningful, but it reflects the species pool (or the inability of local samples to sample it) rather than non-random species distributions. Similarly, sampling effects can "inflate" metrics of β -diversity when there are many rare species in an assemblage, or when the total community density is relatively low (i.e., widespread species remain undetected in most samples) (Barwell et al. 2015). Although such sampling effects are ubiquitous in ecological studies (Colwell and Coddington 1994, Gotelli and Colwell 2001), sampling theory is not well developed with respect to β -diversity (Wolda 1981, Beck et al. 2013). For example, Chao and Chiu (2016) developed a framework to unify different approaches to community differentiation, but they state clearly that their approach ignores such sampling issues. There have also been attempts to develop asymptotic β -diversity metrics (Chao et al. 2005), but these have been found to show strong biases when tested on simulated and empirical data (Cardoso et al. 2009, Beck et al. 2013). While rarefaction-based approaches are commonly used to address sampling effects at a single scale by standardizing diversity to a common number of individuals (Gotelli and Colwell 2001) or to equal levels of sample completeness (Chao and Jost 2012), these approaches have been rarely applied to concepts related to β -diversity (but see Olszewski 2004, Dauby and Hardy 2012, Stier et al. 2016, Chase et al. 2018). This is despite the fact that the null models used in detecting deviations from random expectations in β -diversity (e.g., Kraft et al. 2011, Xing and He 2021) are based on largely similar concepts (i.e., difference between observed and expected measures of diversity).

In what follows, we consider non-random spatial distributions through the lens of the individual-based rarefaction curve and combine existing approaches (Chase et al. 2018, McGlenn et al. 2019) with coverage-based standardization.

Specifically, we compare rarefaction curves taken from subsets of samples (i.e., an α -scale curve) to those from the entire set of samples (i.e., a γ -scale rarefaction curve), using a constant γ -scale coverage (i.e., an estimate of sample completeness). From this, we obtain a metric, which we call β_C , that estimates the degree of spatial structure in the assemblage independently of the species pool size and the SAD. We emphasize that our goal here is not to develop a better measure of β -diversity per se, as it is true that Whittaker's β -diversity and relatives have many useful properties for discerning biodiversity scaling (e.g., Jost 2007, Tuomisto 2010a, b, Chao and Chiu 2016). Instead, our goal is to develop a measure that allows us to discern the magnitude of spatial structuring within a given regional assemblage. Building on rarefaction and sampling theory has the advantage that we can evaluate sample completeness and bypass the shuffling algorithms and estimates of beta-deviation inherent to previous null-modeling approaches. We test our method on simulated spatial point patterns with different degrees of spatial structure (intraspecific spatial aggregation) and varying species pool sizes and apply it to two empirical data sets to examine variation in spatial structure along a latitudinal gradient of tree diversity.

INDIVIDUAL-BASED RAREFACTION AND EXTRAPOLATION (IBRE)

Our approach is based on individual-based rarefaction and extrapolation (IBRE), which is a common method to standardize species richness estimates (Hurlbert 1971, Gotelli and Colwell 2001, Chao and Jost 2012). IBRE curves describe the nonlinear scaling relationship between the number of individuals in a sample and expected species richness (i.e., rarefied richness). The shape of the curve is determined by the size of the species pool and the relative abundances of species in that pool, which is often referred to as the species abundance distribution (SAD, McGill et al. 2007). The slope at any point along the curve is related to the estimated sample completeness for the number of individuals sampled at that point (Chao and Jost 2012). For smaller than observed sample sizes, the expected number of species can be interpolated using the following formula (Hurlbert 1971):

$$S_n = S_{\text{obs}} - \sum_{X_i \geq 1} \frac{\binom{N - X_i}{n}}{\binom{N}{n}} \quad (1)$$

where S_n is the rarefied richness, or the expected number of species for n individuals ($n < N$), S_{obs} is the observed number of species, N is the observed number of individuals in the sample, and X_i is the number of individuals of the i th species.

For larger than observed sample sizes, the expected number of species can be estimated using the following extrapolation formula (Chao and Jost 2012):

$$S_n = S_{\text{obs}} + \hat{f}_0 \left[1 - \left(1 - \frac{f_1}{N\hat{f}_0 + f_1} \right)^{n-N} \right] \quad (2)$$

where \hat{f}_0 is the estimated number of unseen species, estimated as follows:

$$\hat{f}_0 = \begin{cases} \frac{(N-1)f_1^2}{N \cdot 2f_2}, & \text{if } f_2 > 0 \\ \frac{(N-1)f_1(f_1-1)}{N \cdot 2}, & \text{if } f_2 = 0 \end{cases} \quad (3)$$

and f_1 and f_2 are the observed numbers of singletons and doubletons (i.e., species represented by one or two individuals), respectively. Extrapolation of species richness is considered unbiased, though only recommended for sample sizes up to two times the observed sample size (Chao et al. 2014).

SPATIAL STRUCTURE THROUGH THE LENS OF THE IBRE CURVE

By constructing IBRE curves from samples at two or more nested spatial scales, we can assess intraspecific spatial aggregation (Olszewski 2004, Dauby and Hardy 2012, Chase et al. 2018, McGlenn et al. 2019). Like most classical approaches to diversity partitioning, we define α -diversity as the mean number of species within a given sample or subset of localized samples, and γ -diversity as the total number of species from multiple pooled samples or local subsets of samples (Tuomisto 2010a). Accordingly, the α -scale IBRE curve is derived by calculating the

IBRE curve from each individual sample and then averaging the all samples ($\alpha\overline{S}_n$) while the γ -scale curve consists of S_n values calculated from the pooled sample ($\gamma\overline{S}_n$). The α -scale is influenced by turnover (i.e., spatial structure) among samples within the assemblage, whereas the γ -scale breaks up any spatial structure by randomly accumulating individuals from all samples. If species are distributed randomly among samples (i.e., there is no aggregation), the α - and γ -scale IBRE curves sit on top of each other (Fig. 1). Downward and upward deviations of the α -scale curve, then, would be interpreted as intraspecific aggregation and overdispersion, respectively (Chase et al. 2018, McGlinn et al. 2019). The γ -scale IBRE is conceptually very similar to abundance-based null expectations (Kraft et al. 2011), but it uses an analytical formula rather than a shuffling algorithm. Furthermore, rather

than comparing the observed β -diversity to a null distribution of β -diversity, it directly compares the observed α -scale IBRE curve to the null expectation given by the γ -scale IBRE curve.

Using IBRE curves (only interpolation shown for simplicity), Fig. 1 illustrates how β -diversity of a reference assemblage (Fig. 1A) responds to changes in the size of the species pool (Fig. 1B), the numbers of individuals (Fig. 1C), and intraspecific spatial aggregation (Fig. 1D). Whittaker's β -diversity ($\beta = \frac{\gamma}{\alpha}$) is represented as the height ratio of the two curves at the respective right-hand end of the curves (dashed horizontal lines). In each of the four examples, $\beta > 1$, but for very different underlying reasons. Only the assemblage underlying Fig. 1D exhibits spatial turnover in species composition (due to aggregation). In the other cases (A-C), differentiation only emerges due to a sampling effect (i.e., a

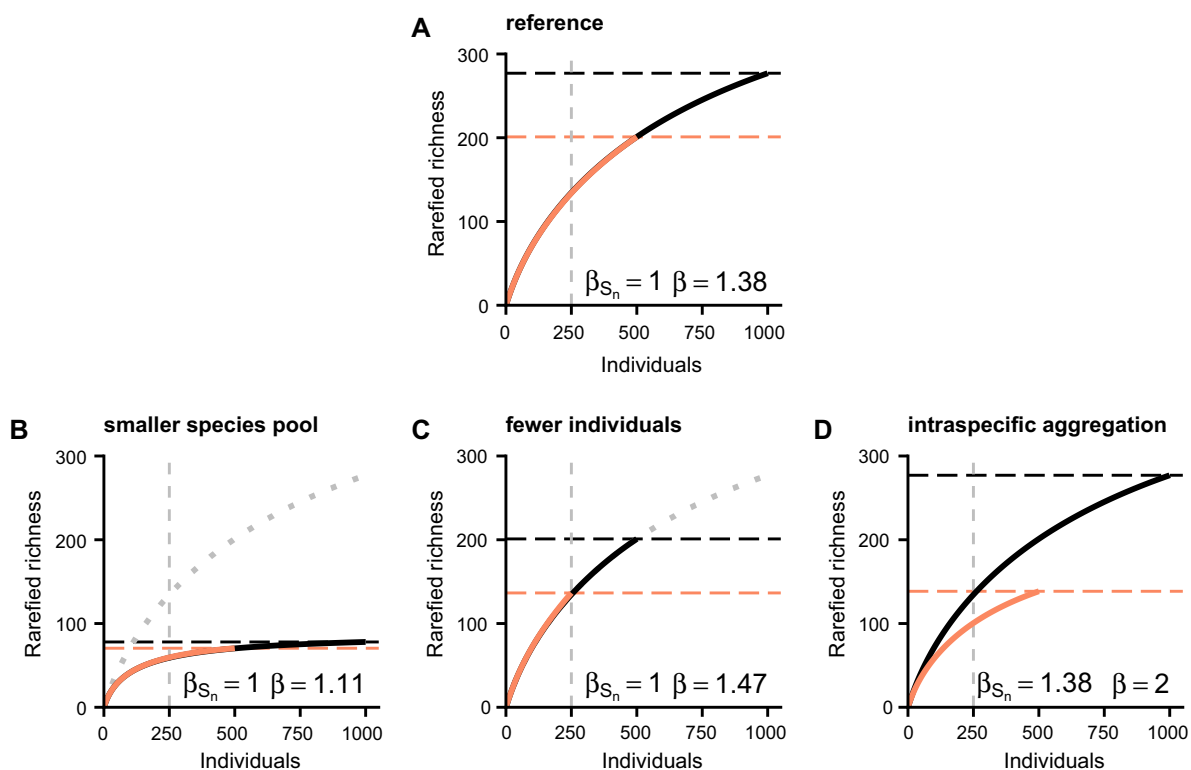


Fig. 1. Examples of two-scale individual-based rarefaction curves for (A) a hypothetical species pool size of 450, and how they respond to, (B) reduced species pool size/altered SAD (species pool size of 80), (C) reduced numbers of individuals, (D) and changes to patterns of within species aggregation. Orange curve: α -scale, black curve: γ -scale. Dashed lines represent observed species richness at α - and γ -scales, and Whittaker's β -diversity (β) can be illustrated as the height ratio of the two. β_{S_n} values are calculated for $n = 250$ individuals on all panels (dashed vertical lines). Dotted gray curve in panel C: reference curve (from A) to aid comparison.

“more-individuals” effect between the α -scale and γ -scale).

Due to the nonlinear shape of the IBRE curve, the sampling effect depends on the regional SAD (Fig. 1B) and the number of individuals sampled (Fig. 1C). Chase et al. (2018) suggested that when calculated at a common number of individuals (n), the ratio of rarefied richness (S_n) calculated between the γ - and α -scales, termed β_{S_n} , could provide an indication of the degree of intraspecific aggregation, or non-randomness in the distribution of species in the assemblage, independent of any sampling effect (see also McGlenn et al. 2019). β_{S_n} is related to metrics developed by Olszewski (2004) and Dauby and Hardy (2012) who also assess the differences between γ and α IBRE curves. When assemblages have a random spatial structure, β_{S_n} is expected to equal 1 regardless of species pool and sample size (Fig. 1A-C). Conversely, β_{S_n} values larger than 1 reflect spatial aggregation or species turnover among sites in the region (Fig. 1D).

While the deviation between α - and γ -scale IBRE curves (i.e., $\beta_{S_n} \neq 1$ is due to spatial structure, its magnitude is contingent on the value of n and the shape of the curves (i.e., the size and evenness of the species pool). Thus, as we will illustrate below, β_{S_n} is biased when comparing the degree of aggregation among regions where species pools and shapes of the γ -scale IBRE curves change (e.g., along biogeographical gradients). To visualize this problem, consider two assemblages each composed of two patches, but which differ in the size of their regional species pool (500 vs. 100 species Fig. 2A, B, respectively). Supposing that both assemblages have complete species turnover between their respective patches, Fig. 2 shows the IBRE curves that we would expect if we sampled 500 individuals from each patch in the large (Fig. 2A) and small (Fig. 2B) species pools. Note how the γ -curve from the small species pool is much closer to its asymptote than the one from the large species pool (slope of gray tangential lines). This

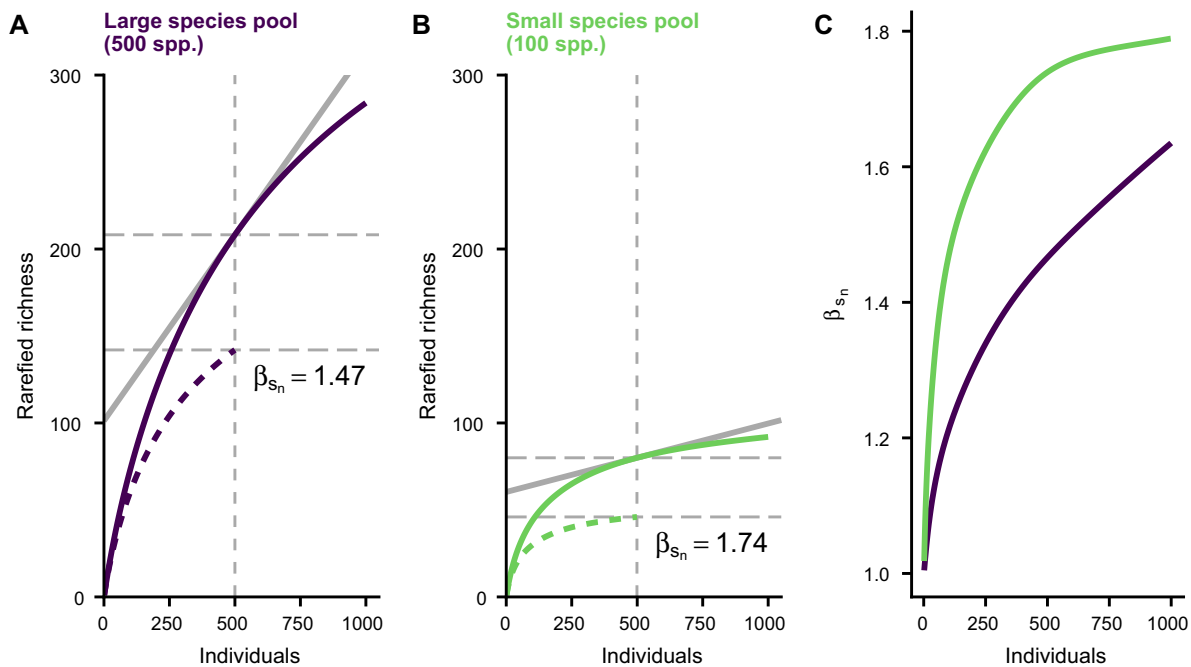


Fig. 2. β_{S_n} is affected by species pool size in aggregated communities. Two-scale IBRE curves for (A) a large species pool and (B) a small species pool. Solid curve: γ -scale; dashed curve: α -scale; dashed vertical gray line: number of individuals ($n = 500$) used for the calculation of β_{S_n} (i.e., the ratio of the horizontal dashed lines). Gray solid line shows the slope of the γ -scale IBRE that relates to sample coverage (the steeper the slope of this line, the less complete the sample). (C) β_{S_n} plotted as a function of the corresponding sample size (n).

difference in completeness has implications for the values of β_{S_n} . Although both assemblages are maximally and equally structured at the patch scale, the relative deviation between α and γ is substantially higher in the small species pool (indicated by β_{S_n}). Asymptotically, both assemblages have a theoretical β -diversity of 2 (i.e., complete turnover), but at any common sample size, β_{S_n} differs between them due to the difference in sample completeness that is associated with species pool size (Fig. 2C).

COVERAGE-BASED RAREFACTION AND EXTRAPOLATION

To account for the differences in species pool size and sample completeness, we extend β_{S_n} to include the concept of coverage-based rarefaction (Chao and Jost 2012). Sample coverage is a measure of sample completeness that ranges from 0 to 1 and refers to the “proportion of the total number of individuals in a community that belong to the species represented in the sample” (Chao and Jost 2012). Sample coverage depends on the sample size and the species abundance distribution of the underlying assemblage. It can be estimated from the number of rare species in a sample (Good 1953, Chao and Shen 2010). As it is directly related to the steepness of the IBRE curve, expected coverage can also be estimated for any sample size along the curve using the following equations (Chao and Jost 2012).

$$C_n = \begin{cases} 1 - \sum_{X_i \geq 1} \frac{X_i}{N} \binom{N-X_i}{n} \binom{N-1}{n-1}, & \text{for interpolation } (n < N) \\ 1 - \frac{f_1}{N} \left[\frac{(N-1)f_1}{(N-1)f_1 + 2f_2} \right]^{n-N+1}, & \text{for extrapolation } (n > N) \end{cases} \quad (4)$$

where C_n is the expected coverage for a subsample of sample size n . N is the total number of individuals in the sample, X_i is the number of individuals of the i th species, and f_1 and f_2 are the numbers of singletons and doubletons.

For coverage-based standardization, Eq. 4 can be solved numerically to determine how many individuals, n , are necessary to obtain a given target coverage, C_{target} . This is computed by calculating C_n for every possible n and choosing the

one that minimizes the difference between C_n and C_{target} . Subsequently, IBRE can be used to standardize the diversity estimate to a sample size of n , and thus the desired coverage level C_{target} (Hsieh et al. 2016).

INTRODUCING β_C

By calculating β_{S_n} for equal γ -scale coverage, we can resolve the species pool dependence when making comparisons across assemblages. Specifically, rather than keeping sample size (n) constant when comparing across assemblages, we instead maintain a consistent sample coverage at the γ -scale (C_{target}) and refer to β_{S_n} standardized by sample coverage as β_C . Fig. 3 illustrates this approach using the same example with large (Fig. 3A) and small (Fig. 3B) species pools. By allowing n to vary between scenarios so that we maintain a constant γ -scale coverage (indicated by the slope of the tangential lines), the resulting pair of β_C values become practically identical (compare with Fig. 2), which accurately reflects that both scenarios are equally aggregated at the patch scale. The advantage of standardizing γ -scale coverage becomes particularly clear when we consider the entire scaling relationship of β_C . If we quantify β_C for every possible value of n and plot them against γ -scale expected coverage, the values from large and small species pool fall on approximately the same line and the species pool dependence vanishes (Fig. 3C, compare with Fig. 2C).

To compare β_C across multiple assemblages (e.g., with different species pools), we suggest the following protocol that makes use of interpolation and extrapolation.

1. Determine the appropriate target coverage value C_{target} for the standardization:
 - 1.1. For each assemblage j , determine the smallest number of individuals observed at the α -scale and call it $N_{\text{min},j}$.
 - 1.2. Using Eq. 4, estimate the expected γ -scale coverage C_n corresponding to $N_{\text{min},j}$ individuals, or up to $2N_{\text{min},j}$ individuals if you wish to use extrapolation (Chao et al. 2014).
 - 1.3. Let C_{target} be the smallest of the C_n values across all assemblages.

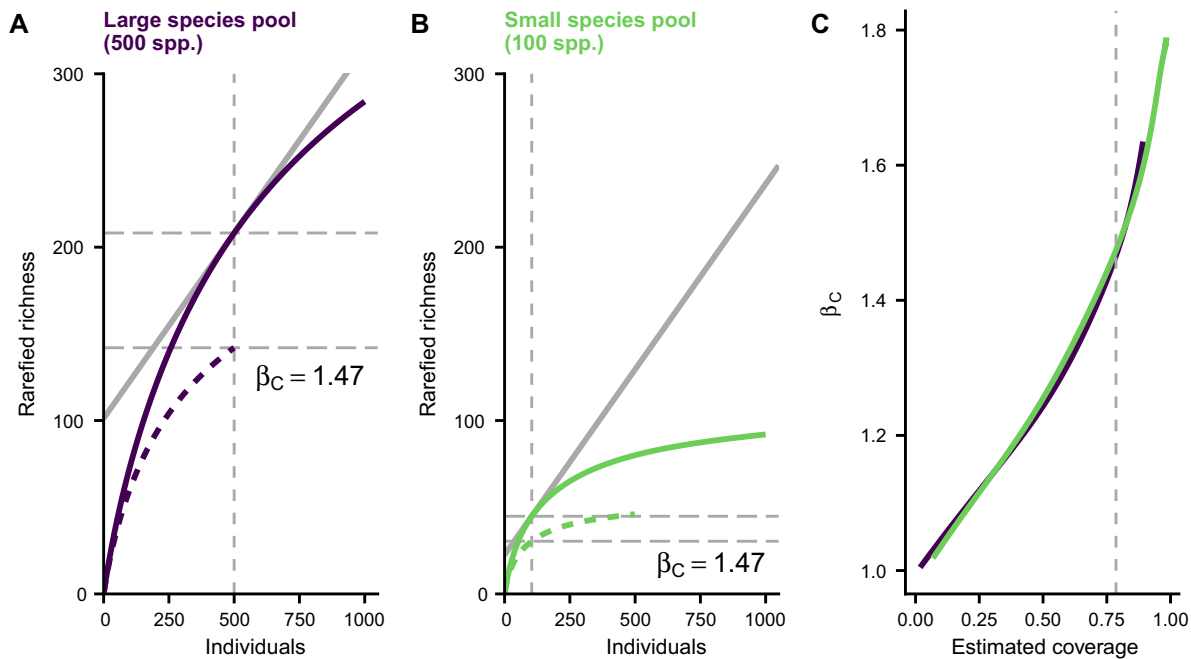


Fig. 3. β_C is unaffected by species pool size in aggregated communities (compare with Fig 2). Two-scale IBRE curves for (A) a large species pool and (B) a small species pool. Solid curve: γ -scale; dashed curve: α -scale; dashed vertical gray line: number of individuals used for the calculation of β_{S_n} (i.e., the ratio of the horizontal dashed lines). Gray solid line shows the slope of the γ -scale IBRE that relates to sample coverage ($C = 0.79$ in both panels). (C) β_C plotted as a function of expected coverage calculated at the γ -scale. Dashed line marks the coverage value of 0.79 used in other panels.

2. Calculate β_C :

- 2.1. For each assemblage, use the inverse of Eqn 4 to estimate the sample size n_j corresponding to a γ -scale coverage of C_{target} . Exclude all assemblages for which $n_j \leq 1$.
- 2.2. Standardize γ - and α -scale species richness to n_j individuals (using IBRE) to get ${}^\gamma S_{n,j}$ and ${}^\alpha S_{n,j}$.
- 2.3. Calculate β_C as follows:

$$\beta_{C_j} = \frac{{}^\gamma S_{n,j}}{{}^\alpha S_{n,j}} \quad (5)$$

We implemented this procedure in an R package available on GitHub (<https://github.com/t-engel/betaC>) and archived on Zenodo (<https://doi.org/10.5281/zenodo.4727184>). It provides the function “C_target” that can be used for steps 1.1 and 1.2, and the function “beta_C” that carries out step 2.

This approach requires spatially replicated samples with abundance data (i.e., site-by-

species matrices with abundance data), so that one can define at least two nested sampling scales (α and γ). We assume that the sampling design is standardized across all assemblages. This means there should be a consistent number of samples per assemblage and every sample should have the same effort (e.g., plot size and trap nights). Furthermore, we assume a consistent spatial extent at the γ -scale. If the number of samples, or their spatial extent changes from one assemblage to the next, users should take a spatially constrained subset of the samples to keep extents as consistent as possible.

Like most measures of community differentiation, β_C does not have an analytical variance estimator because there are no replicates at the γ -scale. Nevertheless, such variance is often desired, for example, when comparing spatial structure among different regions. To do so, we recommend calculating a distribution of β_C for repeated random subsets of the samples. For

example, one could use a Jackknife approach (i.e., systematically leaving out one sample at a time), use pairwise comparisons of samples, or comparisons among larger subsets of samples. While such variance can help to contextualize the observed values of β -diversity, caution should be taken because these calculations incorporate some degree of non-independence. Nevertheless, we provide an R function that carries out such resampling procedures for any number of samples (i.e., “betaC::beta_stand”).

PROOF OF CONCEPT USING SIMULATIONS

To test the properties of β_C , we simulated spatially explicit assemblages that varied in the size of the species pool and the degree of intraspecific aggregation using multivariate spatial point patterns. We used the R package *mobsim* (May et al. 2018) to carry out the simulations. Each simulated assemblage had 4000 individuals drawn from a lognormal SAD that was parameterized with a given species pool size, and a coefficient of variation equal to one. Then, we used the Thomas cluster process to distribute individuals in space, varying the degree of intraspecific spatial aggregation through the parameter that determines the number of conspecific clusters. The simulation was parameterized in a full-factorial design, where the species pool size encompassed every integer between 10 and 500, and the number of clusters was set to 1 (i.e., extreme intraspecific aggregation), 4, 10, and 20. To include a level that had no within species aggregation, we also implemented a random Poisson process to simulate completely random spatial distributions for species. Each combination of species pool and aggregation was replicated 3 times yielding a total of 7365 simulated communities. To sample from the regional communities, we placed 4 sample quadrats into each simulated community (Fig. 4A, B). We calculated Whittaker’s β ($=\frac{\gamma}{\alpha}$), β_{S_n} and β_C among the four samples and examined their response to changes in species pool size and aggregation. Following the protocol above, C_{target} was set to 0.55 and the sample size for β_{S_n} was 50 individuals. The R code for the simulation is available on Zenodo (<https://doi.org/10.5281/zenodo.4727184>).

All three indices (Whittaker’s β , β_{S_n} , and β_C) were influenced by intraspecific community

aggregation. Additionally, Whittaker’s β and β_{S_n} were affected by the changes in species pool size, whereas β_C was insensitive to this parameter. While both Whittaker’s β and β_{S_n} responded to the degree of spatial aggregation and the species pool, they did so in contrasting ways, which is consistent with theoretical expectations (Fig. 4C). For Whittaker’s β , the effect of the species pool decreased with increasing aggregation. This reflects that for strongly aggregated species distributions, the samples will always show high turnover regardless of the species pool. In contrast, under random species distributions, “spurious” (i.e., SAD related) sample differentiation is more likely to occur when there are many rare species that only occur in some of the samples (i.e., in large species pools). For β_{S_n} , the effect of the species pool increased with increasing aggregation; as long as species are randomly distributed, β_{S_n} is always one because the α IBRE curve falls onto the γ -scale. However, when there is a deviation between the curves (as a result of aggregation), its magnitude for a given number of individuals (n) depends on the shape of the IBRE curves, which in turn depend on the species pool (Fig. 3). Only β_C captures the spatial structure of the simulated communities independently of the species pool size because, by incorporating sample coverage, it adjust for the species pool dependence.

We examined the robustness of β_C using an alternative SAD (log-series) and by simulating spatial aggregation using the mean displacement length of the Thomas process (Appendix S1). The results were qualitatively similar: β_C but not Whittaker’s β or β_{S_n} , responded to the changes of aggregation independently to changes in the SAD parameter (i.e., Fisher’s α). Additionally, we applied the null model by Kraft et al. (2011) to the simulated data and found that the measure of β -deviation, like β_C , responded to the aggregation, but not to the species pool. Spearman’s rank correlation between β_C and β -deviation was 97.7% which suggests that both approaches are measuring the same effect (Appendix S2).

EMPIRICAL CASE STUDIES

Next, we applied our approach to two forest data sets with varying species pool sizes. First, we reanalyzed the Gentry Forest plot data set (Gentry 1988, Phillips and Miller 2002). This data

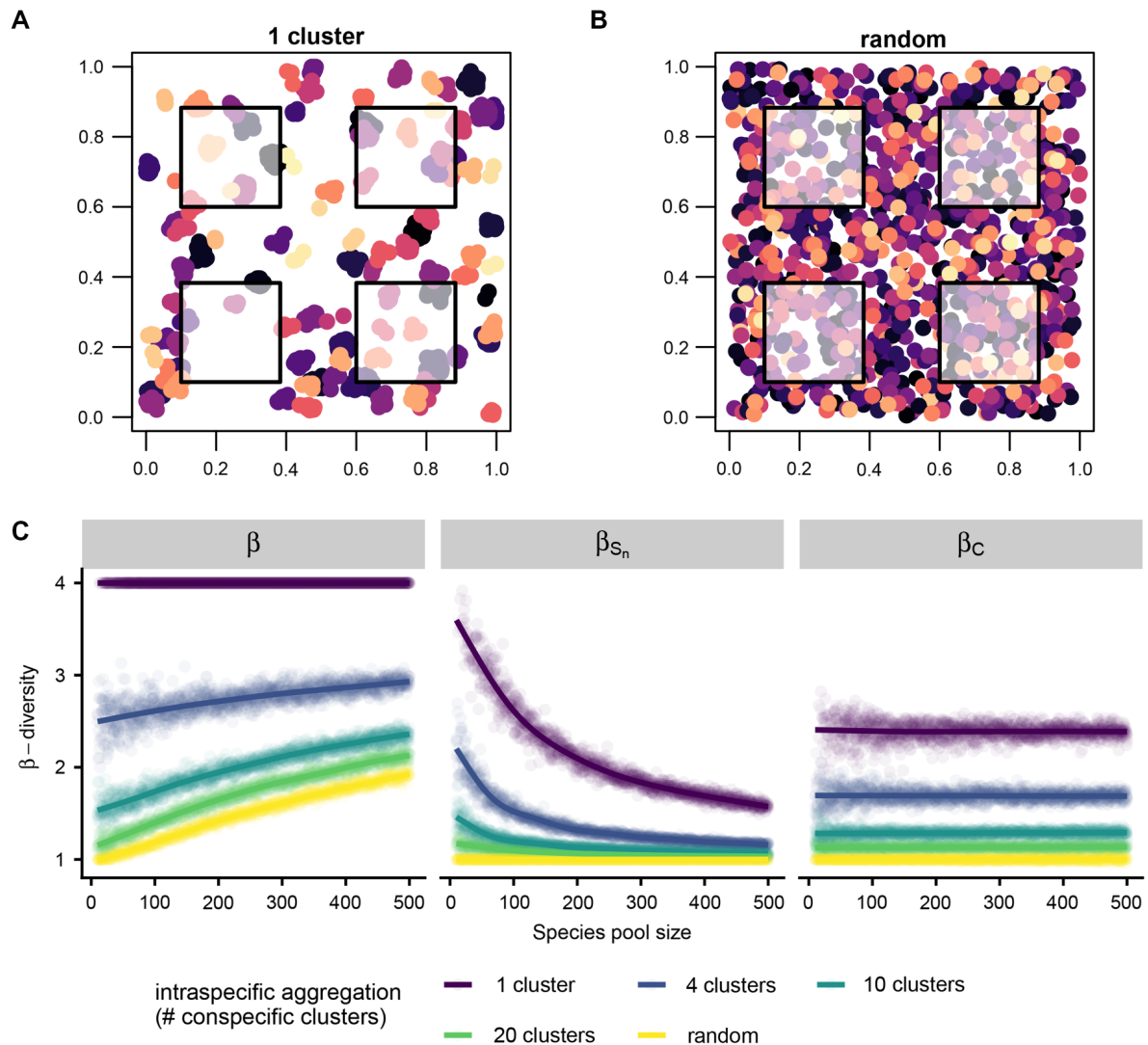


Fig. 4. Simulated assemblages and the response of Whittaker's β , β_{S_n} , and β_C to changes in aggregation and species pool size. (A) Assemblage with extreme intraspecific aggregation (1 cluster per species) and (B) assemblage where species have random spatial distribution. Species pool, SAD, and numbers of individuals are constant between (A) and (B). Squares represent sample quadrats. (C) β -diversity metrics and their response to species pool size and intraspecific aggregation. Dots show the data. Lines show GAM fit to each metric with species pool size as the predictor; the GAM estimated separate smoothers for each level of intraspecific aggregation.

set has frequently been used in the debate on how to formulate appropriate null models for β -diversity and how spatial aggregation varies with latitude (Kraft et al. 2011, Qian and Song 2013, Qian et al. 2013, Xu et al. 2015, Xing and He 2021). We computed Whittaker's β and β_C among the subplots of the sites located in the Americas. As expected from the difference in

species pool size along this gradient and shown by previous studies (e.g., Kraft et al. 2011), Whittaker's β declined with latitude (Fig. 5A). In contrast, β_C , which controls for species pool-related sampling effects, showed no significant change along the latitudinal gradient (Fig. 5B). Given this, we conclude that there is no evidence for a change in spatial aggregation along this gradient

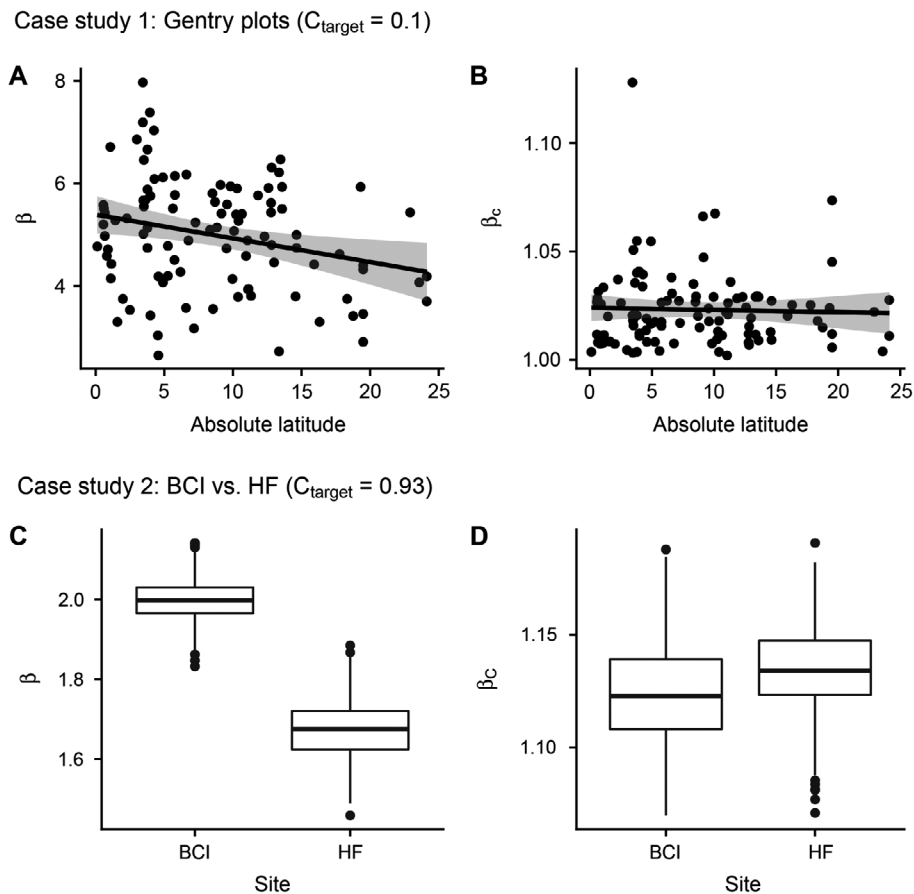


Fig. 5. Case studies exploring β -diversity along (A, B) a latitudinal gradient of Gentry forest plots and (C, D) comparing Barro Colorado Island (BCI) and Harvard Forest (HF). Both examples show significant changes in Whittaker's β (A, C) while β_c (i.e., β standardized for sample coverage) showed no significant change. Solid lines show simple linear regressions. Boxplots show distribution of values from 1000 redraws of 10 samples, respectively. Whiskers: non-outlier range; box: interquartile range; bar: median.

in these samples. Declines in Whittaker's β with increasing latitude appear to be mostly driven by changes in the size of the regional species pool rather than changes in the spatial structuring of individuals. Importantly, although we come to qualitatively similar results as the abundance-based null models of Kraft et al. (2011) and Xu et al. (2015), our method has the advantage of explicitly incorporating an estimate of sample completeness. Rather than simply shuffling site-by-species matrices, users are confronted with the completeness of their samples as part of the analytical workflow, and prior to interpreting any results. In this case, our analysis provides quantitative evidence for the argument that these

types of forest plots may be too small to robustly compare spatial patterns of diversity and any associated differences in community assembly, especially in the tropics (Tuomisto and Ruokolainen 2012). C_{target} is set by the site with the lowest coverage (C_n), which for the Gentry Forest Plot data were $C_{\text{target}} = 0.1$. This means that inferences are being drawn from a sample of only approximately 10 percent of the individuals in the assemblage.

For a second case study, we explored a similar question along the latitudinal gradient, but with larger plots, so that a more substantial fraction of the species pool would be sampled. To do so, we compare the spatial structuring of trees within

the temperate forest plot at Harvard Forest (HF), in the northeastern United States (Orwig et al. 2015), with the well-studied tropical rainforest plot at Barro Colorado Island (BCI), Panama (Condit et al. 2019). Because the data from HF were from only 35 hectares, we took a 35 ha subsection of the 50 ha plot at BCI and found a considerable difference in γ -diversity; 38 species were present at HF and 217 species were in the analyzed section of BCI. For both locations, we used our resampling approach to calculate a distribution of β -diversity at the scale of ten 1 ha subplots (using 1000 random draws with replacement). Again, we found that Whittaker's β was consistently higher in the more diverse site (BCI). While HF had an expected α -scale coverage of 99%, it was 90% at BCI. By interpolating HF and extrapolating BCI (following the protocol above), we standardized both sites to a target coverage of 0.93 and found no meaningful difference in the corresponding β_C (Fig. 5 D). In short, as with our analyses of the Gentry data above, but with samples that more adequately characterize the assemblages, we find that the observed differences in Whittaker's β from temperate to tropical forests were largely expected given the differences in the species pool. That is, once species pool-related sampling effects were taken into account, there do not appear to be any meaningful differences in the spatial structuring of these two forests.

DISCUSSION AND CONCLUSIONS

Building on previous work using rarefaction and coverage-based approaches (Chao and Jost 2012, Chao et al. 2014, Chase et al. 2018, McGlenn et al. 2019), we developed a metric standardized by sample coverage to quantify the degree of intraspecific spatial aggregation, independent of changes in the size of the species pool and the regional SAD. Our theoretical considerations and simulations of spatially explicit assemblages show that β_C remains unaffected by changes in the species pool, which allows for comparisons of intraspecific aggregation along large biogeographic gradients. Our empirical case studies suggest that the magnitude of intraspecific aggregation does not change along a latitudinal gradient of forest plots. Importantly, our method requires analysts to determine the target

completeness using information contained in the samples from their study. In the case of the commonly used Gentry plots, this shows that the samples cover only a small fraction (10%) of the individuals in the underlying assemblages, and may therefore be of limited use for making inferences about their small-scale spatial structure (Tuomisto and Ruokolainen 2012).

Our approach represents an important advance over existing methods to measuring spatial aggregation because of its strong link to existing biodiversity sampling theory. Specifically, we use the γ -scale rarefaction curve as an analytical null model for the expected α -diversity in the absence of spatial structure. While conceptually similar to existing null models, our approach has several advantages. For example, it bypasses the computation of Whittaker's beta and subsequently β -deviation. Instead, we measure the deviation between γ - and α -scales directly from the IBRE curves. β_C can be thought of as the factor by which spatial structure has reduced α diversity compared with the random expectation. This makes it more intuitive than β -deviation (Kraft et al. 2011, Xing and He 2021) because it can be directly interpreted as an effective number of distinct communities (Jost 2007), conditional on the estimated sample coverage. Additionally, our approach explicitly incorporates an estimate of sample completeness into the analytical workflow, which means that the analyst is confronted with the limitations of the data.

Ulrich et al. (2017) have argued that null model approaches are limited in their ability to disentangle species pool and aggregation effects, unless they incorporate external data on the sizes of the relevant species pools. Here, we make use of the idea that the sample itself can also provide an estimate of its completeness (Good 1953, Chao and Jost 2012). Our method uses the shape of the IBRE curve, itself determined by the SAD, to draw inferences for an estimated constant fraction of the individuals in the underlying community (i.e., a constant sample coverage). This approach to standardization implicitly assumes that there is an asymptote in species diversity (Chao and Jost 2012). While this assumption is mathematically convenient, it cannot strictly be true; due to species aggregation at higher scales, we will always find more species with more samples, until the entire global pool is sampled

(Williamson et al. 2001). Nevertheless, β_C does not extend to the asymptote itself, but merely employs a useful approximation of sample completeness via sample coverage. As spatial structure is an inherently scale-dependent phenomenon, with this approach we can only measure it at the spatial grain prescribed by the samples (i.e., spatial structure within the γ -scale). Even when extrapolation is used, the spatial grain of the data remains unchanged. For a scale-dependent examination of spatial aggregation, we recommend comparing results obtained from different spatial grains.

While β_C isolates the degree of spatial aggregation, it does not have some of the properties that are sometimes considered essential for β -diversity metrics (Jost 2007, Tuomisto 2010a). For example, traditional β -diversities range between unity and the total number of sampling units, and they can be transformed into N-community (dis-)similarities in the range [0,1] (Jost 2007, Chao and Chiu 2016). In contrast, β_C can only reach the number of samples when the samples have a coverage of 1 (i.e., C_{target} is 100%) (see Appendix S3). In such cases, where the curves have reached an asymptote, β_C equals Whitaker's β and one can derive the corresponding (dis-)similarity. However, for incomplete samples, the N-community transformation of β -diversity is generally not recommended (Chao and Chiu 2016), and we consider it a strength of our method that it exposes such situations. Only in the rare cases where C_{target} is 100%, classical β -diversity metrics and pairwise (dis-)similarities will not be affected by the species pool.

Although our approach accounts for differences in sample completeness, it still requires standardized sampling or post hoc standardizations (i.e., rarefying samples to the same number and spatial extent) to make valid inferences. The sampling effects we treat here arise passively due to differences in species pools, and not as a result of different sampling strategies and/or effort. For example, the forest plots in our case studies were completely sampled in the sense that every tree was counted in a given area or subplot. However, with respect to the regional species pool or even just the observed γ diversity, a subplot of a given size in the tropics is likely a much less complete sample, compared to an equally sized subplot a temperate region, even

when all individuals are counted in each subplot. It is this interaction of sampling effort and the size of the species pool that leads to the null expectation of increasing sample differentiation with increasing species pool size, and for which our method adjusts.

In conclusion, our approach allows us to explicitly disentangle non-random spatial patterns of species diversity (e.g., intraspecific aggregation) amidst variation in species pool size and associated sampling effects. Together with other diversity metrics sensitive to diversity components such as the SAD and total community abundance (Chase et al. 2018, McGlenn et al. 2019), β_C allows deeper insights into how spatial structuring within communities influences patterns of biodiversity and its change. Applications could, for example, shed light onto the assembly processes that govern (meta-)communities along biogeographic gradients and contribute to a better understanding of the spatial diversity patterns that underlie the scale-dependent biodiversity trends observed during the current biodiversity crisis.

ACKNOWLEDGMENTS

TE, SAB, and JMC gratefully acknowledge the support of iDiv funded by the German Research Foundation (DFG–FZT 118, 202548816). In addition, we thank Xiao Xiao, Tiffany Knight, Leana Gooriah, Petr Keil, and Ingmar Staude for discussions and feedback on the approach.

LITERATURE CITED

- Anderson, M. J., et al. 2011. Navigating the multiple meanings of β diversity: a roadmap for the practicing ecologist. *Ecology Letters* 14:19–28.
- Barwell, L. J., N. J. B. Isaac, and W. E. Kunin. 2015. Measuring β -diversity with species abundance data. *Journal of Animal Ecology* 84:1112–1122.
- Beck, J., J. D. Holloway, and W. Schwanghart. 2013. Undersampling and the measurement of beta diversity. *Methods in Ecology and Evolution* 4:370–382.
- Cardoso, P., P. A. V. Borges, and J. A. Veech. 2009. Testing the performance of beta diversity measures based on incidence data: the robustness to under-sampling. *Diversity and Distributions* 15:1081–1090.
- Chao, A., R. L. Chazdon, R. K. Colwell, and T.-J. Shen. 2005. A new statistical approach for assessing

- similarity of species composition with incidence and abundance data. *Ecology Letters* 8:148–159.
- Chao, A., and C.-H. Chiu. 2016. Bridging the variance and diversity decomposition approaches to beta diversity via similarity and differentiation measures. *Methods in Ecology and Evolution* 7:919–928.
- Chao, A., N. J. Gotelli, T. C. Hsieh, E. L. Sander, K. H. Ma, R. K. Colwell, and A. M. Ellison. 2014. Rarefaction and extrapolation with Hill numbers: a framework for sampling and estimation in species diversity studies. *Ecological Monographs* 84:45–67.
- Chao, A., and L. Jost. 2012. Coverage-based rarefaction and extrapolation: standardizing samples by completeness rather than size. *Ecology* 93:2533–2547.
- Chao, A., and T. J. Shen. 2010. User's guide for program SPADE (Species prediction and diversity estimation). National Tsing Hua University, Hsin-Chu, Taiwan.
- Chase, J. M., and T. M. Knight. 2013. Scale-dependent effect sizes of ecological drivers on biodiversity: Why standardised sampling is not enough. *Ecology Letters* 16:17–26.
- Chase, J. M., N. J. B. Kraft, K. G. Smith, M. Vellend, and B. D. Inouye. 2011. Using null models to disentangle variation in community dissimilarity from variation in α -diversity. *Ecosphere* 2:art24.
- Chase, J. M., B. J. McGill, D. J. McGlinn, F. May, S. A. Blowes, X. Xiao, T. M. Knight, O. Purschke, and N. J. Gotelli. 2018. Embracing scale-dependence to achieve a deeper understanding of biodiversity and its change across communities. *Ecology Letters* 21:1737–1751.
- Chase, J. M., and J. A. Myers. 2011. Disentangling the importance of ecological niches from stochastic processes across scales. *Philosophical Transactions of the Royal Society B: Biological Sciences* 366:2351–2363.
- Colwell, R. K., and J. A. Coddington. 1994. Estimating terrestrial biodiversity through extrapolation. In Hawksworth, D.L.: Biodiversity. Measurement and estimation. *Philosophical Transactions of the Royal Society of London Biological Sciences* 345:101–118.
- Condit, R., R. Perez, S. Aguilar, S. Lao, R. Foster, and S. P. Hubbell. 2019. Complete data from the Barro Colorado 50-ha plot: 423617 trees, 35 years. Dryad. <https://doi.org/10.15146/5xcp-0d46>
- Dauby, G., and O. J. Hardy. 2012. Sampled-based estimation of diversity sensu stricto by transforming Hurlbert diversities into effective number of species. *Ecography* 35:661–672.
- Gaston, K. J. 2000. Global patterns in biodiversity. *Nature* 405:220–227.
- Gentry, A. H. 1988. Changes in plant community diversity and floristic composition on environmental and geographical gradients. *Annals of the Missouri Botanical Garden* 75:1–34.
- Good, I. J. 1953. The population frequencies of species and the estimation of population parameters. *Biometrika* 40:237.
- Gotelli, N. J., and R. K. Colwell. 2001. Quantifying biodiversity: procedures and pitfalls in the measurement and comparison of species richness. *Ecology Letters* 4:379–391.
- Hsieh, T. C., K. H. Ma, and A. Chao. 2016. iNEXT: an R package for rarefaction and extrapolation of species diversity (Hill numbers). *Methods in Ecology and Evolution* 7:1451–1456.
- Hurlbert, S. H. 1971. The nonconcept of species diversity: a critique and alternative parameters. *Ecology* 52:577–586.
- Jost, L. 2007. Partitioning diversity into independent alpha and beta components. *Ecology* 88:2427–2439.
- Kraft, N. J. B., et al. 2011. Disentangling the drivers of β diversity along latitudinal and elevational gradients. *Science* 333:1755–1758.
- Kraft, N. J. B., et al. 2012. Response to comments on “Disentangling the drivers of diversity along latitudinal and elevational gradients”. *Science* 335:1573.
- May, F., K. Gerstner, D. J. McGlinn, X. Xiao, and J. M. Chase. 2018. mobsim: An r package for the simulation and measurement of biodiversity across spatial scales. *Methods in Ecology and Evolution* 9:1401–1408.
- McGill, B. J., et al. 2007. Species abundance distributions: moving beyond single prediction theories to integration within an ecological framework. *Ecology Letters* 10:995–1015.
- McGill, B. J. 2011. Linking biodiversity patterns by autocorrelated random sampling. *American Journal of Botany* 98:481–502.
- McGlinn, D. J., X. Xiao, F. May, N. J. Gotelli, T. Engel, S. A. Blowes, T. M. Knight, O. Purschke, J. M. Chase, and B. J. McGill. 2019. Measurement of Biodiversity (MoB): a method to separate the scale-dependent effects of species abundance distribution, density, and aggregation on diversity change. *Methods in Ecology and Evolution* 10:258–269.
- Mori, A. S., F. Isbell, and R. Seidl. 2018. β -Diversity, community assembly, and ecosystem functioning. *Trends in Ecology and Evolution* 33:549–564.
- Olszewski, T. D. 2004. A unified mathematical framework for the measurement of richness and evenness within and among multiple communities. *Oikos* 104:377–387.
- Orwig, D., D. Foster, and A. Ellison. 2015. Harvard Forest CTFs-ForestGEO Mapped Forest Plot since 2014. Harvard Forest Data Archive: HF253.
- Phillips, O., and J. S. Miller. 2002. Global patterns of plant diversity: Alwyn H. Gentry's forest transect

- data set. Missouri Botanical Garden Press, St Louis, Missouri, USA.
- Qian, H., S. Chen, L. Mao, and Z. Ouyang. 2013. Drivers of β -diversity along latitudinal gradients revisited. *Global Ecology and Biogeography* 22:659–670.
- Qian, H., and J.-S. Song. 2013. Latitudinal gradients of associations between beta and gamma diversity of trees in forest communities in the New World. *Journal of Plant Ecology* 6:12–18.
- Qian, H., X. Wang, and Y. Zhang. 2012. Comment on “Disentangling the drivers of diversity along latitudinal and elevational gradients”. *Science* 335:1573.
- Socolar, J. B., J. J. Gilroy, W. E. Kunin, and D. P. Edwards. 2016. How should beta-diversity inform biodiversity conservation? *Trends in Ecology & Evolution* 31:67–80.
- Sreekar, R., et al. 2018. Spatial scale changes the relationship between beta diversity, species richness and latitude. *Royal Society Open Science* 5:181168.
- Stier, A. C., B. M. Bolker, and C. W. Osenberg. 2016. Using rarefaction to isolate the effects of patch size and sampling effort on beta diversity. *Ecosphere* 7:e01612.
- Tucker, C. M., L. G. Shoemaker, K. F. Davies, D. R. Nemergut, and B. A. Melbourne. 2016. Differentiating between niche and neutral assembly in meta-communities using null models of β -diversity. *Oikos* 125:778–789.
- Tuomisto, H. 2010a. A diversity of beta diversities: Straightening up a concept gone awry. Part 1. Defining beta diversity as a function of alpha and gamma diversity. *Ecography* 33:2–22.
- Tuomisto, H. 2010b. A diversity of beta diversities: Straightening up a concept gone awry. Part 2. Quantifying beta diversity and related phenomena. *Ecography* 33:23–45.
- Tuomisto, H., and K. Ruokolainen. 2012. Comment on “Disentangling the drivers of diversity along latitudinal and elevational gradients”. *Science* 335:1573.
- Ulrich, W., A. Baselga, B. Kusumoto, T. Shiono, H. Tuomisto, and Y. Kubota. 2017. The tangled link between β - and γ -diversity: a Narcissus effect weakens statistical inferences in null model analyses of diversity patterns. *Global Ecology and Biogeography* 26:1–5.
- Whittaker, R. H. 1960. Vegetation of the Siskiyou Mountains, Oregon and California. *Ecological Monographs* 30:279–338.
- Williamson, M., K. J. Gaston, and W. M. Lonsdale. 2001. The species–area relationship does not have an asymptote! *Journal of Biogeography* 28:827–830.
- Wolda, H. 1981. Similarity indices, sample size and diversity. *Oecologia* 50:296–302.
- Worm, B., and D. P. Tittensor. 2018. *A theory of global diversity (MPB-60)*. Princeton University Press, Princeton, New Jersey, USA.
- Xing, D., and F. He. 2021. Analytical models for β -diversity and the power-law scaling of β -deviation. *Methods in Ecology and Evolution* 12:405–414.
- Xu, W., G. Chen, C. Liu, and K. Ma. 2015. Latitudinal differences in species abundance distributions, rather than spatial aggregation, explain beta-diversity along latitudinal gradients. *Global Ecology and Biogeography* 24:1170–1180.

DATA AVAILABILITY

Code and data (Phillips and Miller 2002, Orwig et al. 2015, Condit et al. 2019) are available from Zenodo: <https://doi.org/10.5281/zenodo.4727184>.

SUPPORTING INFORMATION

Additional Supporting Information may be found online at: <http://onlinelibrary.wiley.com/doi/10.1002/ecs2.3745/full>