# Disentangling non-random structure from random placement when estimating β-diversity through space or time

Daniel J. McGlinn (ORCID)[1*], Shane A. Blowes (ORCID)[2,3], Maria Dornelas[4], Thore Engel (ORCID)[2,4,5], Inês S. Martins[4,7], Hideyasu Shimadzu(ORCID)[8,9], Nicholas J. Gotelli (ORCID)[10], Anne Magurran[4], Brian McGill (ORCID)[11], Jonathan M. Chase (ORCID)[2,3]

[1]Department of Biology, College of Charleston; Charleston, SC, USA.

[2]German Centre for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig; Leipzig, Germany.

[3]Department of Computer Science, Martin Luther University Halle-Wittenberg; Halle (Saale), Germany.

[4]Centre for Biological Diversity, School of Biology, University of St Andrews; St Andrews, Scotland.

[5]Institute of Biodiversity, Friedrich Schiller University Jena; Jena, Germany

[6]Helmholtz Centre for Environmental Research - UFZ, Department of Ecosystem Services' Leipzig, Germany

[7]Leverhulme Centre for Anthropocene Biodiversity, Berrick Saul Second Floor, University of York; York, UK.

[8]Department of Mathematical Sciences, Loughborough University; Leicestershire, UK.

[9]Department of Data Science, Kitasato University; Kanagawa, Japan.

[10]Department of Biology, University of Vermont; Burlington, VT, USA.

[11]School of Biology and Ecology and Mitchell Center for Sustainability Solutions, University of Maine; Orono, ME, USA.

[*]Corresponding & contact author: danmcglinn@gmail.com

**Open Research statement:** all code and data used in this manuscript are available at the following link: https://github.com/MoBiodiv/beta_concept

**Keywords:** aggregation; patchiness; rarefaction; sampling effects; scaling; species-abundance distribution; species turnover;

## Abstract:

There is considerable interest in understanding patterns of β-diversity that measure the amount of change in species composition through space or time. Most hypotheses for β-diversity evoke nonrandom processes that generate spatial and temporal within species aggregation; however, β-diversity can also be driven by random sampling processes. Here, we describe a framework based on rarefaction curves that quantifies the non-random contribution of species compositional differences across samples to β-diversity. We isolate the effect of within-species spatial or temporal aggregation on beta-diversity using a coverage standardized metric of β-diversity ($\beta_C$). We demonstrate the utility of our framework using simulations and an empirical case study examining variation in avian species composition through space and time in engineered versus natural riparian areas. The primary strengths of our approach are that it provides an intuitive visual null model for expected patterns of biodiversity under random sampling that allows integrating analyses across α-, γ-, and β-scales. Importantly, the method can accommodate comparisons between communities with different species pool sizes, and can be used to examine species turnover both within and between meta-communities.

## Introduction

Ecologists are frequently interested in how the composition of species in a community changes across space or time (Scheiner et al. 2011, Magurran et al. 2019, Daskalova et al. 2020).The degree of change in species composition in assemblages across space or time is often referred to as β-diversity: localities or time periods with fewer species in common have higher β-diversity. Most conceptual explanations of β-diversity evoke processes that generate non-random spatial or temporal patterns of species aggregation (Leibold and Chase 2018). Aggregation here refers to clustering whereby individuals occur near other individuals of the same species in time and/or space. For instance, two of the most commonly discussed mechanisms underlying patterns of β-diversity are environmental filtering and dispersal limitation (Legendre et al. 2005, Vellend 2016, Leibold and Chase 2018). Both of these mechanisms increase aggregation of species distributions via conspecific clustering in space or time increasing β-diversity.

Although most attention has been focused on the non-random mechanisms underpinning β-diversity, it can also reflect random sampling effects of individuals and species taken from multiple points in space or time. Imagine we collect a sample of 40 individuals within a region or time period that supports up to 50 different species. Even in the improbable case that the numbers of individuals of each species are exactly the same (completely even), at least 10 species will be excluded from our sample because of the limited number of individuals. If we then compare that sample to another from a different location in space or time, a different set of 10 (or more) species will be excluded simply due to random sampling effects: the species composition of the two samples will differ entirely due to incomplete sampling. This phenomenon has been variously termed a “sampling effect” (e.g., Adler et al. 2005), a “rarefaction effect” (e.g., Palmer et al. 2008), and the “random placement model” (e.g., Coleman et al. 1982). The core idea is that the number of species observed in a sample is constrained by the number of individuals in that sample. Returning to our thought experiment, if the species have a more realistic abundance distribution, with many individuals of a few common species and many species with few individuals (i.e., rare species), these sampling effects on β diversity can be strong (Kraft et al. 2011, Chase et al. 2018, McGlinn et al. 2019, Engel et al. 2021). This example emphasizes that spatial or temporal β diversity is potentially underlain by two factors: a) the non-random turnover of species, due to ecological mechanisms such as environmental filtering or dispersal limitation; and, b) the random turnover of species due to incomplete sampling, especially of rare species (i.e., sampling effects).

Most metrics of β-diversity conflate variation from both random sampling effects and spatially non-random mechanisms (Stegen et al. 2013, Chase et al. 2018, McGlinn et al. 2019, Engel et al. 2021, Chao et al. 2023). This means that the same observed change in β-diversity may be due to different underlying mechanisms, sometimes referred to as a “many-to-one problem”, which are common in ecological studies (Frank 2014, Scholes 2017). Specifically, random turnover can occur where there are changes/differences in other non-spatial components of diversity, such as the species abundance distribution and size of the regional species pool and the total number of individuals. To illustrate this, consider the three hypothetical scenarios in Figure 1 using Whittaker’s (1960) β-diversity ($\beta_S = \gamma / \bar{\alpha}$ , where $\gamma$ is the regional, and $\bar{\alpha}$ is the average of local diversity). In each scenario (Fig1. a-c), a shift in a different component of community structure results in a doubling of Whitaker’s β-diversity (from 1 to 2). In the first two cases (Fig1.a, b), β-diversity increases due simply to random placement of individuals resulting either from a shift in the species-abundance distribution (SAD), for example, a decrease in evenness (Fig. 1a) or, from a decrease in the total number of individuals (*N*) (Fig. 1b). In the third case, the same magnitude of shift in $\beta_S$ is due to an increase in conspecific aggregation (Fig. 1c). This “many-to-one" effect is particularly problematic when trying to link changes in β-diversity to hypotheses that evoke changes to conspecific clustering due to environmental filtering or dispersal limitation. To link these mechanisms to β-diversity, it would make sense to focus on patterns of β-diversity that reflect only changes in conspecific aggregation rather than changes in $N$ or the SAD (which we refer to as sampling effects). One consequence of β-diversity metrics confounding both random and non-random variation is that most β-diversity metrics can increase as aggregation decreases if $N$ is decreasing or the SAD is becoming less even for example.

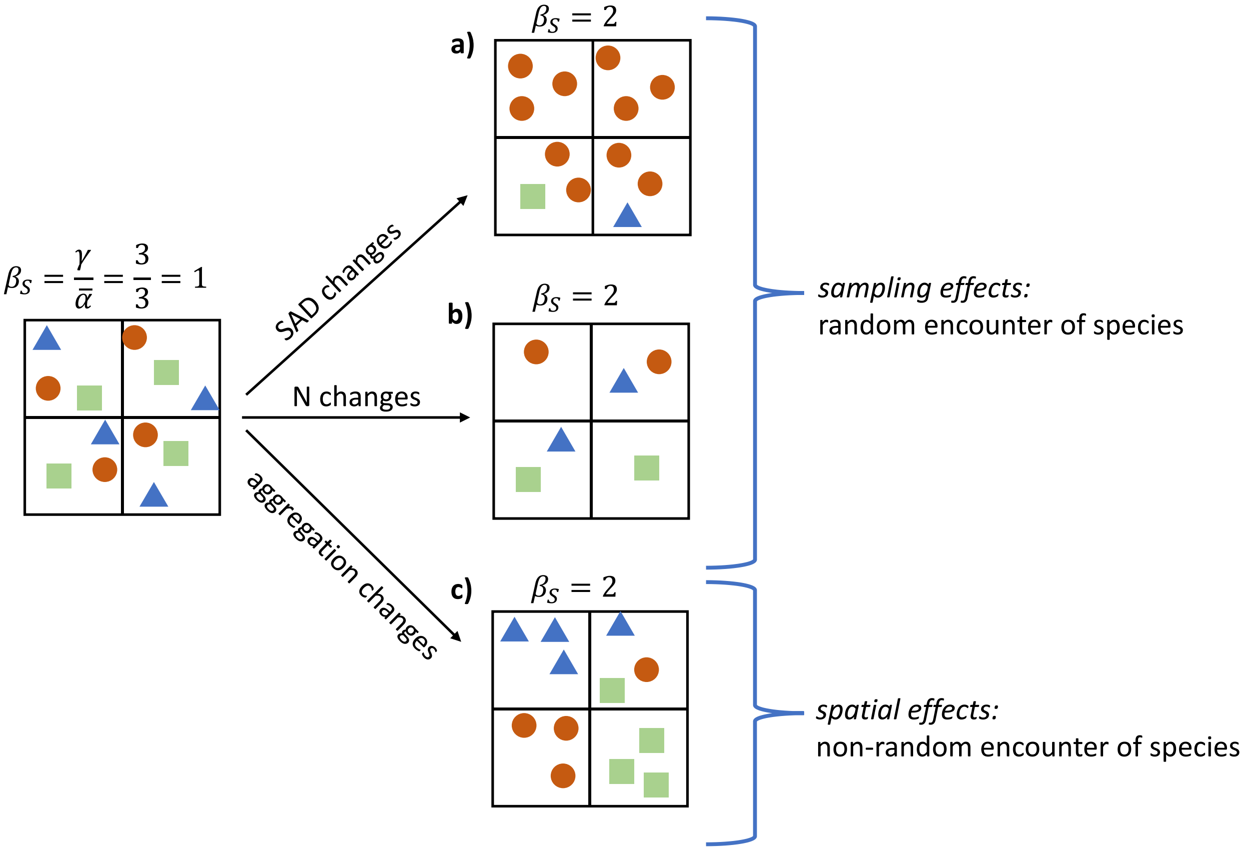

Figure 1. Cartoon communities that illustrate how random sampling effects and non-random spatial effects can result in identical values of Whittaker’s β-diversity ($\beta_S$), where $\bar{\boldsymbol{\alpha}}$ is average sample richness across plots (small boxes) and $\gamma$ is total species richness in a site (large boxes). The different symbols represent individuals of different species. Panels (a) and (b) illustrate changes in community structure that

are consistent with a random sampling model in which spatial β increases either because the regional species-abundance distribution (SAD) is less even (a) or because there are many fewer individuals (b). Panel (c), illustrates how a spatially non-random process such as environmental filtering results in conspecific aggregation producing an identical value of $\beta_S$.

It is important to emphasize that this is a problem that potentially influences all β-diversity metrics. Any metric of β-diversity that does not explicitly consider the process of sampling is sensitive to sampling effects. So regardless of whether turnover is calculated using presence-absence vs abundance data, is examined in space or time, or using pairwise vs multisite metrics, if the goal of the analysis is to link patterns of compositional change to mechanisms that generate non-random conspecific occurrence patterns, then sampling effects should be controlled for in the measurement of β-diversity. Other authors have recognized this and proposed a randomization algorithm to try to control for sampling effects on β-diversity (Kraft et al. 2011, Chase et al. 2011, Myers et al. 2013, 2015). Yet, continued debate as to exactly how to develop those randomizations, and just what the deviations mean (Kraft et al. 2012, Qian et al. 2012, 2013, Xu et al. 2015, Tucker et al. 2016) indicates that a more general solution is necessary.

In this paper, we describe a framework for quantifying the non-random contribution of species compositional differences across samples to β-diversity. This framework can be applied to any question related to measuring compositional variation (i.e., β-diversity) across samples, whether it be within a given (relatively homogeneous) metacommunity, across an environmental gradient, or through time. The approach allows us to differentiate the contribution of non-random species compositional shifts from the effects of sampling properties due to random placement to changes in β-diversity. As a result, we can quantify and compare compositional shifts among samples through space or time, and potentially relate these to other features of the system (e.g., changing spatial or environmental conditions). Here, our primary purpose is not to review and/or unify all metrics and measures of β-diversity, nor to advocate for a single superior metric, both of which have been attempted (Tuomisto 2010, Chao et al. 2012, 2023). Rather, we promote a framework for measuring the relative influence of sampling and non-random associations that underlie β-diversity among samples, regardless of whether it is measured within or across landscapes, through time, or any combination thereof. Furthermore, rather than using different concepts and tools, we show how a single conceptual framework can identify the key components underlying variation in species composition.

First, we describe a simple framework that uses rarefaction curves to decompose β-diversity into components due to sampling effects, and those that are due to non-random aggregations of species. Second, we show the framework can be applied to multiple, related questions about how species composition varies across samples.

## A unified framework for dissecting the non-random contribution of species compositional variation to β-diversity in space and time

The components of our framework are not new. The framework is based on a long history of rarefaction and accumulation curves that depict how species numbers increase with increasing sampling effort (Preston 1960, Sanders 1968). For example, Kobayashi (1982, 1983) showed how spatial aggregation could be quantified from rarefaction curves by comparing subsets of spatially explicit samples to the entire range of spatially randomized samples. Likewise, Gotelli and Colwell (2001) showed how comparing accumulation or ‘collectors’ curves that retain spatial information about the distributions of individuals to individual-based rarefaction curves could provide an indicator of the degree to which aggregation influenced spatial patterns of species accumulations (see also Crist and Veech 2006, Chiarucci et al. 2009, Cayuela et al. 2015, Chase et al. 2018, McGlinn et al. 2019). Finally, Olszewski (2004) explicitly discussed how the comparisons between spatially explicit and randomized rarefaction curves could be used as an index of β-diversity (see also Crist and Veech 2006, Dauby and Hardy 2012). These perspectives have been more recently formalized using individual-based rarefaction curves (and related diversity curves) to disentangle non-random structure from random placement underlying β-diversity within a given set of environmental conditions (i.e., a metacommunity) (Chase et al. 2018, McGlinn et al. 2019, 2021, Engel et al. 2021). Here, we generalize this approach and apply it to questions examining β-diversity among different kinds of samples, such as sites across a strong environmental gradient, or when quantifying temporal β-diversity.

Our framework is designed for one of the most common data types available to community ecologists - a sample-by-species matrix. Each sample contains a vector of abundances of all species sampled from a given assemblage and comes from a given local site. Samples can be collected across multiple sites (a site-by-species matrix) or across multiple time periods (a time-by-species matrix), or a combination of the two. For simplicity, we illustrate the different steps of the approach with samples taken from two spatial locations or time points in Figure 2, but it can be generalized to any number of samples. We assume that the communities being compared are sampled in such a way that they have the same sample effort, i.e., the spatial and temporal grain, extent, and sample arrangement are equal across communities (or can be standardized to such). Here, we define a single sample as the α-scale, and the sum of samples as the γ-scale; however, other accumulation schemes are also possible, so long as the α-scale is a subset of the γ-scale.

### *Step 1: Create a rarefaction curve for the sum of all samples: γ-scale*

If we pool the two (or more) samples, we can calculate the γ-scale rarefaction curve (solid, thick black line in Figure 2). This curve shows the number of species for a random sample of individuals from the whole metacommunity or time series; for any sample of $n$ individuals, $S_n$ is the expected number of species in that sample. This type of rarefaction curve is sometimes referred to as an individual-based rarefaction curve or random sampling model. The curve and its variance has been derived analytically for sampling without replacement (Hurlbert 1971, see Coleman et al. 1982 for formulation for sampling with replacement). Here, because we calculate rarefaction using all samples, the γ-scale curve represents a 'null expectation' of the number of species for $n$ individuals, when all individuals of all species occur randomly across the samples (in space or time).

### *Step 2: Create a rarefaction curve for each individual sample and average them: α-scale*

Next, we calculate the rarefaction curves for the individual samples (Fig 2, thin solid black lines) and average them to obtain the α-scale rarefaction curve up to the number of individuals ($n$) that provides the target level of coverage (Fig. 2, thick dashed gray line; Engel et al. 2021). Here coverage refers to how close the γ curve has come to a hypothetical asymptote (i.e., it is an estimate of sample completeness, Chao and Jost 2012).

### *Step 3: Compare the α and γ- scale curves to estimate the β-scale patterns*

The classical Whittaker's $\beta_S$ metric is calculated as $\gamma / \bar{\alpha}$ where $\bar{\alpha}$ is average sample richness (Whittaker 1960). Within our framework, these values are represented by the ends of the rarefaction curves (i.e., the average number of species per site [$\bar{\alpha}$], and all species observed across all sites in a region or time points [$\gamma$], Fig. 2).

To estimate non-random spatial structure, we must compare the (average) α-scale curve (dashed line) to the γ-scale curve, after standardizing for sampling effects. We control for the numbers of individuals sampled (i.e., sampling effects) by comparing the γ- and α-scale rarefaction curves at the same value of number of individuals (*n*).

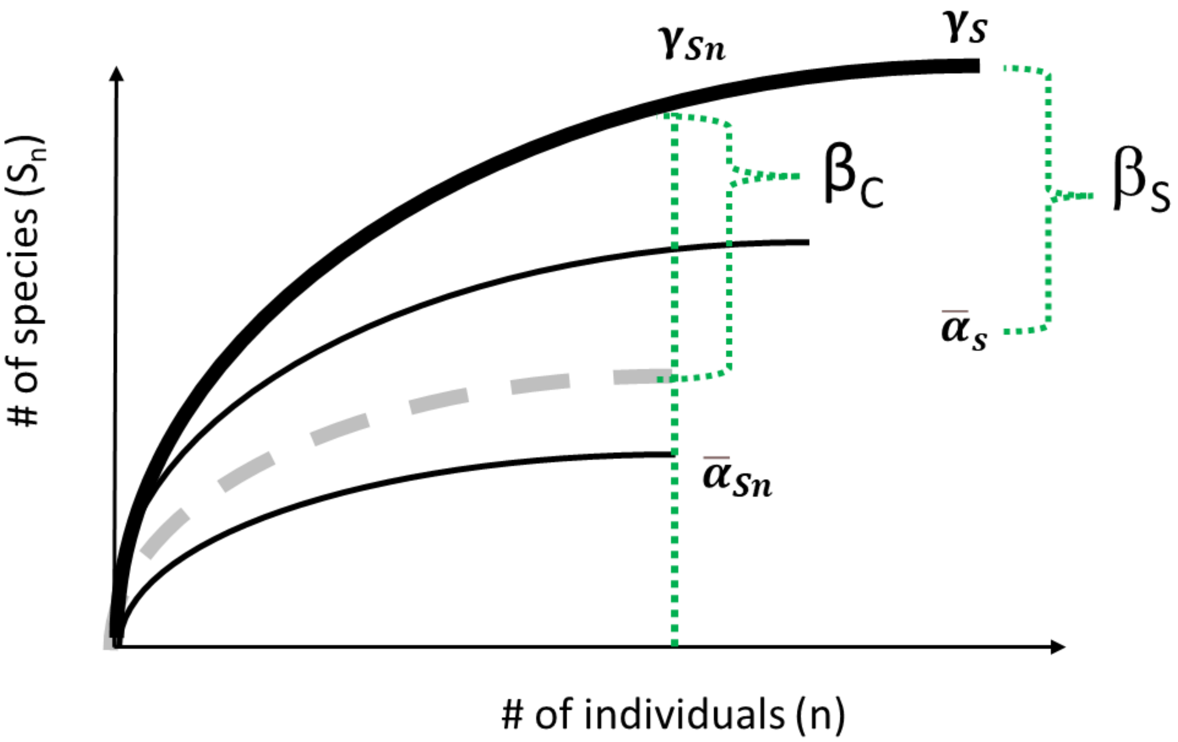

Figure 2: Individual based rarefaction curves at the $\alpha$- and $\gamma$-scales. Thin solid black lines depict two α-scale curves, with their average shown by the thick dashed gray line; solid thick black line shows the $\gamma$-scale curve, which is calculated by pooling the two $\alpha$-scale samples. Whittaker's $\beta_S$ is calculated as the ratio of $\gamma_S$ and $\bar{\boldsymbol{\alpha}}_{\boldsymbol{S}}$. In contrast, coverage-based $\beta_C$ controls for the number of individuals by using the ratio of $\gamma_{Sn}$ and $\bar{\boldsymbol{\alpha}}_{\boldsymbol{Sn}}$, where *n* provides a target degree of coverage that

adjusts for variation in the regional SAD (species pool) when comparing between meta-communities (eqn 5 from Engel et al. 2021).

To illustrate the behavior of these metrics, we simulated four scenarios similar to those shown on Figure 1, and calculated $\beta_S$ and $\beta_C$. All scenarios have 50 species in the regional pool, but they vary from the starting community in either their evenness, total number of individuals, or conspecific aggregation. When individuals of all species are distributed randomly and only the evenness of the SAD decreases (Fig. 3A: high evenness 3B: low evenness) or the total number of individuals decreases (Figure 3A: $N = 1000$; 3C: $N = 400$), we see that the average $\alpha$-scale curve (dashed gray line) falls directly on top of the $\gamma$-scale curve (solid black line, Figure 3 E-G), and low evenness and fewer individuals are associated with an increases in $\beta_S$, but $\beta_C$ is equal to one in both cases (insets on Figure 3F and G). However, when we add non-random structure via species aggregation (Figure 3D), the $\alpha$- and $\gamma$-scale IBR curves diverge (Figure 3H), and both metrics are greater than one (compare inset Figure 3H to 3E).

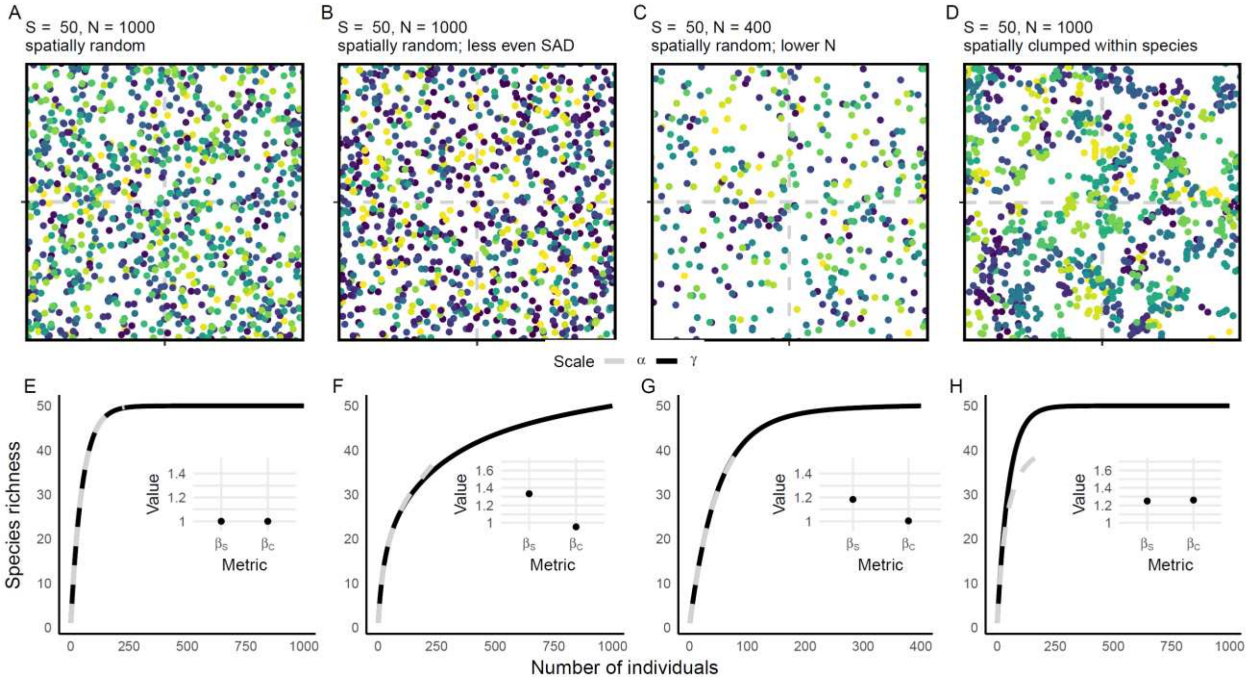

Figure 3: Quantitative illustration that $\beta_S$ responds to changes to changes in evenness and the total number of individuals, whereas $\beta_C$ only responds to changes in within-species aggregation. Simulated communities are shown in panels (A-D) in which different colored dots represent individuals of different species. The landscape ($\gamma$ scale) is divided into four quadrats ($\alpha$ scale, dashed gray lines). Panels E-H show the corresponding rarefaction curves associated with each artificial community for the $\gamma$- (solid black line) and average α-scales (dashed gray line). Inset on these panels is the value of each $\beta$ diversity metric described. Note that $\beta = 1$ means species composition does not vary among samples.

If species are randomly distributed among sites (or time points; Fig. 3A-C), then species will likely be sampled at all sites, and the α- and γ-scale curves will fall on top of each other (Fig. 3E-G). However, if species display conspecific aggregation (i.e., individuals within a species are clumped, Fig. 3D) such that they are non-randomly distributed in space or time, then the α-scale curve will fall below the γ-scale curve (Fig. 3H), because new species will be encountered across different sites or time points due to the within species aggregation, pulling the γ-scale curve up relative to the α-scale curve. The resulting ratio of $\gamma_{Sn}$ and $\bar{\alpha}_{Sn}$, which we call coverage-based β diversity ($\beta_C$, Engel et al. 2021), is indicative of the degree to which species show (non-random) intraspecific aggregation among sites or time periods.

Thus, $\beta_C$ reflects the degree of non-randomness in the spatial or temporal distribution of species within the domain of $(0, \infty)$. The example in Figure 3D shows a case where there is a non-random distribution of species composition among samples, and $\beta_C>1$. However, it is also plausible that the α- and γ-scale curves could completely overlap (see e.g., Fig. 3E-G), in which case we would conclude that even though there is β-diversity (i.e., $\beta_S>1$), this is simply due to random placement effects ($\beta_C=1$). Finally, species can also show conspecific segregation (i.e., individuals within a species are overdispersed more than random), where the α-scale curve falls above the γ-scale curve, and $\beta_C < 1$ (not shown).

One additional benefit of βc (that is not illustrated in Figure 3 but described in detail in Engel et al. 2021) is that this metric is unbiased when comparing β-diversity across meta-communities that differ in the size of their species pools (e.g., in temperate vs. tropical environments, or across strong environmental gradients). This is accomplished by computing $\beta_C$ within each meta-community at the same level of sample coverage or completeness (Chao and Jost 2012). In effect, ensuring that γ-scale sample coverage is the same for all metacommunities means that the value of $n$ (the number of individuals for which $\gamma_{Sn}$ and $\bar{\alpha}_{Sn}$ are calculated) varies among metacommunities.

To summarize, traditional measures of variation in species composition across area or time ($\beta_S$) are shaped by both random and non-random sampling processes, and we can isolate the non-random structure in space or time in determining that scaling by calculating $\beta_C$ (Table 1). Furthermore, we can evaluate these β-diversity measures for a wide variety of questions concerning species compositional shifts in space and time. We provide R code to calculate classical $\beta_S$ and $\beta_C$ (as well as several other β metrics which we do not show here for simplicity) in mobr::calc_beta_div (McGlinn et al. 2022).

Table 1. Multiplicative $\beta$ diversity metrics and the effects that they capture. Species abundance distribution (SAD) effects are due to changes in species evenness and/or the size of the species

pool, N effects refer to changes in richness due to variation in the number of individuals sampled, while aggregation effects refer to changes in richness due to variation in how individuals are spatially or temporally distributed (clumped, random, or overdispersed).

| **Metric** | **Number of individuals sampled (*n*)** | **Effect controlled for** | **Effects captured** |
|---|---|---|---|
| $\beta_S$ (Whittaker's) | Regional *n* vs average local *n* | none | SAD, N, and aggregation |
| $\beta_C$ (coverage) | Regional *n* equals local *n* and corresponds to a target level of coverage. Across meta-communities, coverage is fixed but *n* may vary. | N and SAD (evenness and size of pool) | aggregation |

## One approach, many questions: Some example applications.

There are several benefits to our approach. Rarefaction curves provide an intuitive visualization of $\alpha$- and $\gamma$-diversity patterns, the shape of the SAD, and the degree of variation in species composition that exists between samples. Moreover, the same family of measures can be used to estimate $\beta$-diversity, and to differentiate between random placement and non-random structure leading to biodiversity scaling for multiple related questions. We illustrate some of this potential using a case study. We examined compositional variation in bird diversity between natural and engineered riparian habitats using a subset of data from the Central Arizona-Phoenix Long-Term Ecological Research site (Warren et al. 2022). We focus on riparian habitats where water permanence was perennial, and contrast sites in engineered settings (including a landscaped riparian preserve, a constructed wetland, and a water retention area along the Salt River, each surrounded by urban or agricultural areas) with those in more natural environments (located along perennial river reaches and surrounded by desert). Point count surveys with a 40-m fixed radius were conducted by trained observers that recorded all birds seen and heard; we analyzed samples collected in spring between 2001 and 2016. Before calculating our metrics, we ensured that sample effort was consistent across all sites and years; this meant three sites were retained from each habitat (engineered and natural), and data from 2003 and 2009 were discarded due to missing samples.

Using these effort-standardized data, we address four questions examining how random and non-random components contribute to patterns of β-diversity through space and time: (1) does the total spatiotemporal variation in community composition differ between engineered and natural habitats? (2) How does spatial variation in community composition change through time in each of the two habitats? (3) Does the temporal variation in community composition differ between engineered and natural habitats? (4) Are there compositional differences between (rather than within) engineered and natural habitats, and do any differences change through time?

### *Q1) does the total spatiotemporal variation in community composition differ between engineered and natural habitats?*

We used all site-year combinations within each habitat to examine total spatiotemporal variation in community composition. γ-scale rarefaction curves combine all the samples across space and time within habitats, and show that the engineered habitat had more individuals, but fewer species than the natural habitat (Figure 4i). To examine spatiotemporal variation, we defined the α-scale as a single site-year combination within a habitat (Figure 4i inset shows α- and γ-scale curves). The greater number of species in the natural habitat compared to the engineered habitat resulted in higher $\beta_S$ in natural habitats. However, this pattern was reversed for $\beta_C$ when the influence of sampling effects were removed from the calculations (Figure 4ii), meaning that aggregation in time and space was similar in the engineered and natural habitats.

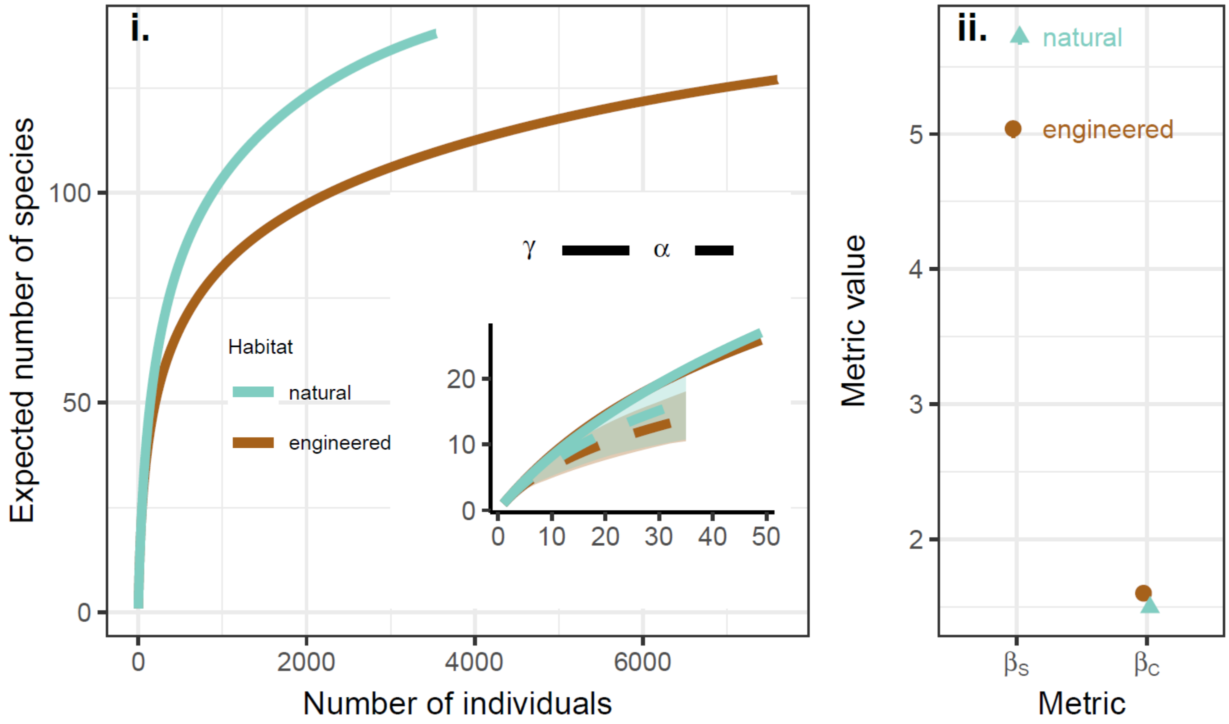

Figure 4. Total spatiotemporal $\beta$-diversity: (i) $\gamma$-scale rarefaction curves for each habitat, with inset showing $\gamma$- and average $\alpha$-scales (note an individual $\alpha$-scale curve was a single site in a single year); (ii) mean $\beta_S$ and $\beta_C$ [point] (95% quantile whiskers not visible) of total spatiotemporal $\beta$-diversity jackknife resamples in each habitat type.

### ***Q2) How does spatial variation in community composition change through time in each of the two habitats?***

Figure 5 shows the pattern of spatial $\beta$-diversity in engineered and natural habitats through time. $\beta_S$ increased through time for the natural sites, indicating that those communities were becoming more different from one another through time (opposite to the oft expected pattern of biotic homogenization, where communities become more similar through time and spatial $\beta$ diversity declines). There was no similar trend in spatial $\beta_S$ of the engineered sites, and by the end of the time series (but not the beginning) the engineered sites had lower levels of $\beta_S$ than the natural sites. However, this pattern qualitatively changed when the influence of non-random patterns was explicitly considered. $\beta_C$ indicates that species became less aggregated within engineered sites through time, suggesting biotic homogenization after random-placement mechanisms were controlled, and no change in $\beta_C$ in the natural habitat. Combined, these results suggest the apparent pattern of increasing differentiation in the natural habitat was mostly driven by sampling effects (e.g., altered numbers of individuals, and/or rare species), and that there was a weak decrease of within species aggregation across sites in the engineered habitat.

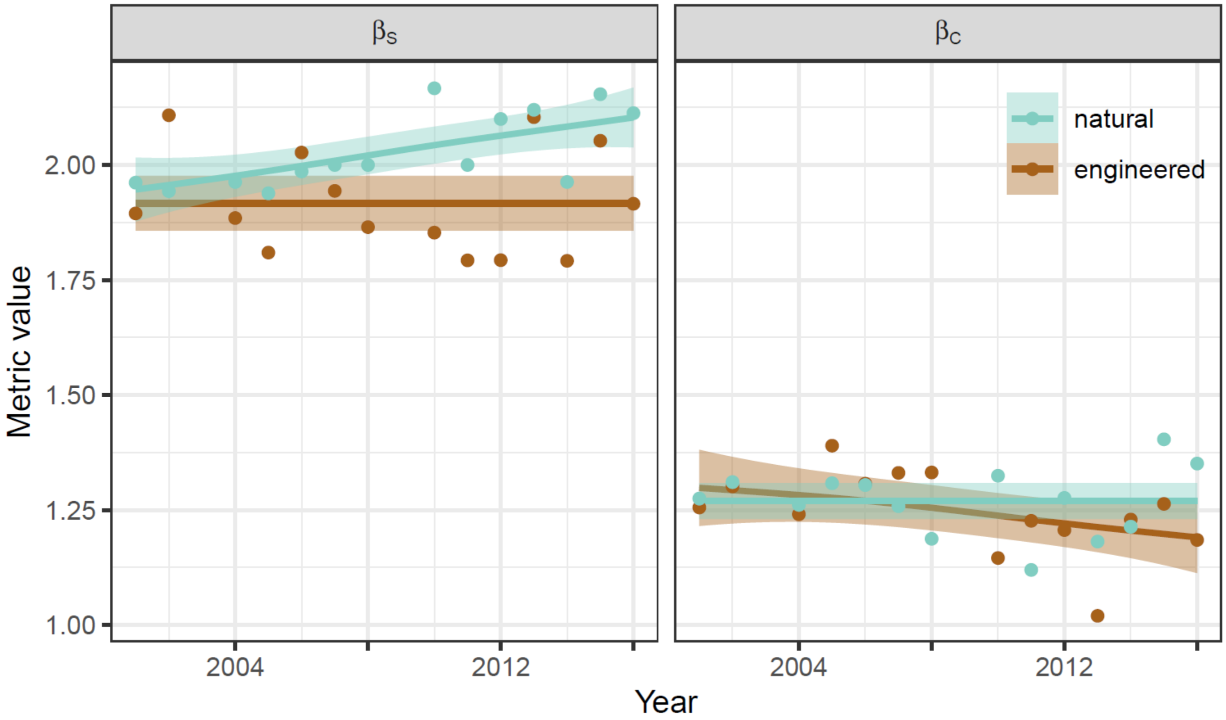

Figure 5. Spatial $\beta$-diversity as a function of time for $\beta_S$ and $\beta_C$ in the two habitats. Trend lines represent OLS linear models with their 95% CI.

### *Q3) Does the temporal variation in community composition differ between engineered and natural habitats?*

Across all years, sites in the engineered habitat had greater variation than sites in natural habitat in both the total number of individuals (i.e., the end points of the $\gamma$-scale curves on the x-axis, Fig. 6i), and shape of the SAD (reflected by greater variation in the curvature of the $\gamma$-scale rarefaction). On average, natural sites had slightly higher levels of temporal $\beta_S$ than the engineered sites, but the variation among sites (and only three replicates) meant there was no overall difference in temporal $\beta_S$ between habitats (Fig. 6ii). We conclude that the weak differences of temporal $\beta_S$ between habitats were primarily due to random sampling effects (Fig. 6i) because this pattern disappeared for $\beta_C$ (and $\beta_C$ was slightly higher in the engineered habitat). In both habitats, $\beta_S$ was also more than double the value of $\beta_C$, suggesting that more than 50% of year-to-year variation in community composition was due to changes in the number of individuals and/or rare species. The similar values $\beta_C$ in both habitats indicates that temporal autocorrelation of species presences did not differ much between natural and engineered habitats.

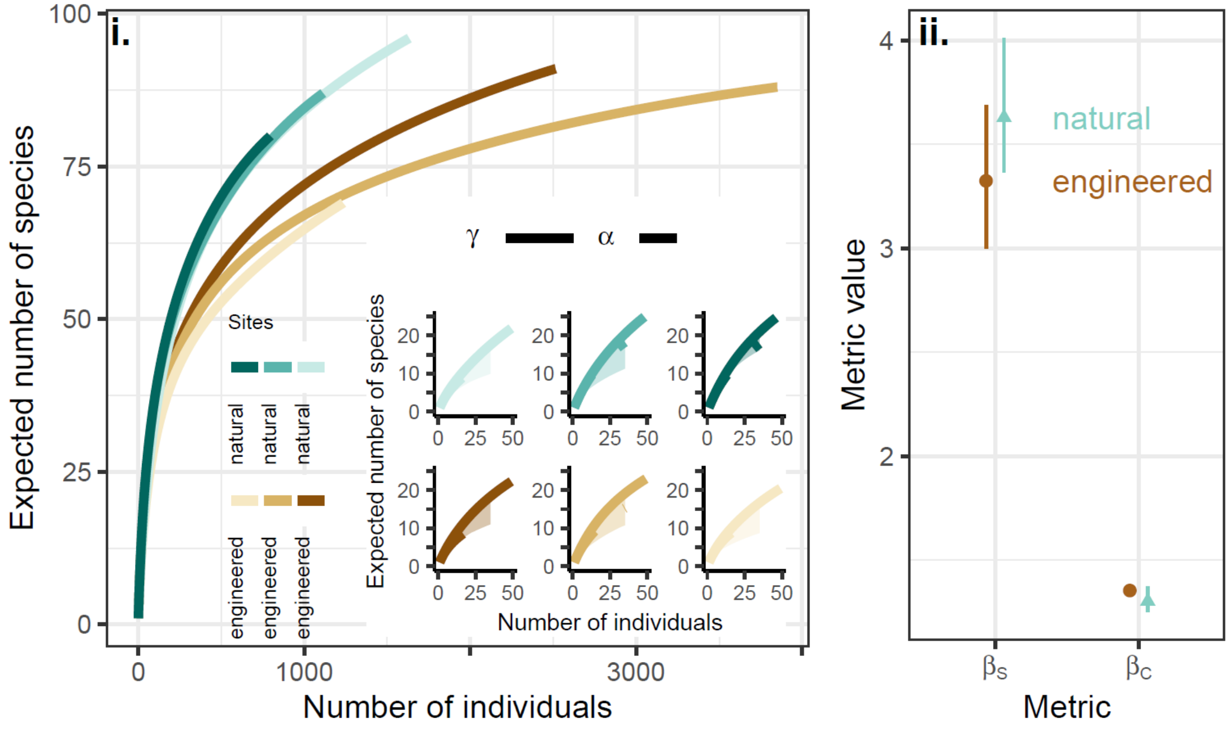

Figure 6. Temporal $\beta$-diversity: (i) γ-scale rarefaction curves for each site (i.e., all years combined), with inset showing $\gamma$- and the average α-scales for each site; (ii) temporal $\beta_S$ and $\beta_C$ (mean [point] and 95% quartiles [whiskers] of jackknife resamples) of total temporal $\beta$-diversity.

***Q4) Are there compositional differences between engineered and natural habitats, and do any differences change through time?***

Finally, the same concepts and tools that we used to examine variation in species composition within treatments can also be used to compare species composition between treatments through time. Essentially, this asks whether bird communities in natural and engineered sites are random subsets of a common larger species pool? Do non-random spatial patterns contribute to any differentiation? And do these patterns change over time? (Fig. 7). Here, the overall difference between treatments ($\beta_S$) was larger than one (there is some species turnover between habitats) and slightly declined through time (homogenization). However, when only non-random patterns were considered, $\beta_C$ was closer to, though still greater than one, and only slightly declined through time. This suggests that once we control for sampling effects that compositional differences between the habitats were relatively small but still detectable and not changing through time.

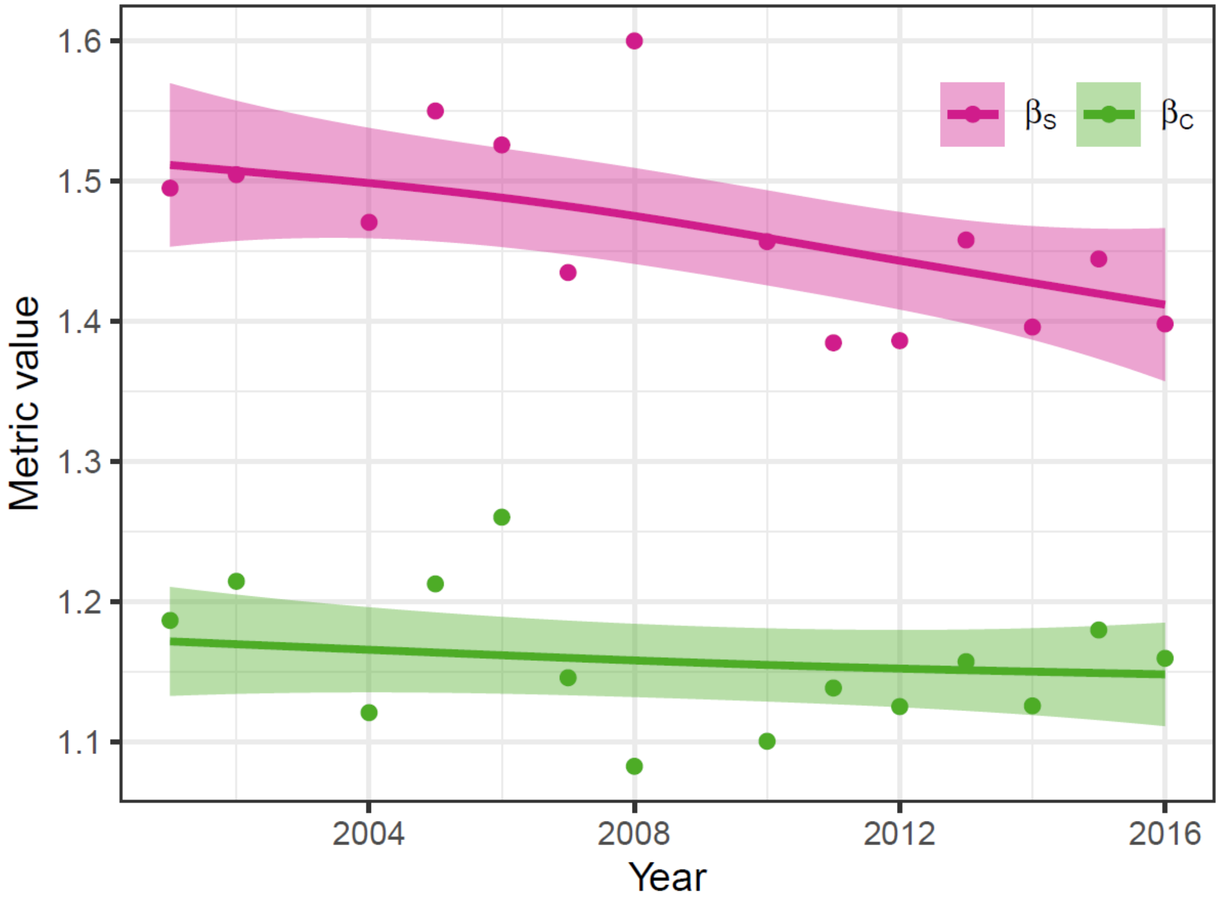

Figure 7. Compare species composition between the engineered and natural riparian habitat types (i.e., the $\beta$-diversity) as a function of time.

## Discussion

We have described and demonstrated an integrated framework for quantifying the underlying causes of β-diversity and namely if those causes are due to random sampling effects or aggregating mechanisms such as environmental filtering and dispersal limitation. The approach that we have described provides a generalized framework for comparing patterns of total β-diversity, which we call $\beta_S$, to those that specifically partial out the non-random patterns of β-diversity ($\beta_C$). We have demonstrated that any question relating to how species composition changes across samples, whether they be taken through space or time, can be subject to the same approach and metrics. For many cases, this greatly simplifies what can seem a complex endeavor of finding the 'right' β-metric for the question at hand.

Often, researchers switch β-diversity metrics and concepts when measuring compositional shifts within a metacommunity or among heterogeneous sites along an environmental gradient (e.g., Anderson et al. 2011). For example, within a metacommunity, estimates of β-diversity are often based on measures of dispersion in community composition among sites (e.g., Anderson et al. 2006). These measures, however, can be strongly influenced by both the relative abundances of species, and the size of the regional species pool. This means that randomization-based null models) are needed if one wants to compare levels of dispersion among different metacommunities, and/or make inferences regarding potential driving mechanisms (e.g., Chase et al. 2011, Kraft et al. 2011, Myers et al. 2013). However, the appropriate form of randomization for the null model remains contentious (Kraft et al. 2012, Qian et al. 2012, Mori et al. 2015, Tucker et al. 2016, Xing and He 2021). Our rarefaction-based approach can also be considered a type of null sampling model. However, comparing rarefaction curves has a number of benefits over other null model approaches: by calculating α- and γ-scale curves, β-diversity can be put back into the context of scale-dependent multicomponent changes in diversity (Chase et al. 2018, Blowes et al. 2022, Rolls et al. 2023); rarefaction curves can be based on analytical solutions improving efficiency; and, rarefaction curves can be visualized, making them more intuitive and easier to communicate than other null model approaches. Nevertheless, some of the concerns arising from the use of null models also apply to the approach overviewed here. For example, there is a strong 'Narcissus' effect (i.e., the outcome reflects the inputs) in developing null models to evaluate whether differences among samples deviate from a null expectation; the samples that are used to calculate γ-diversity influence the likelihood that they will deviate from a null expectation (Ulrich et al. 2017). The same is certainly true for the use of individual-based rarefaction curves in which deviations are mathematically constrained by the two end points of the rarefaction curve (McGlinn et al. 2021).

Baselga (2010) has advocated an approach that partitions measures of dissimilarity among samples (e.g., Jaccard's or Sorensen's index or an abundance-based equivalent) into measures that capture species turnover between samples, and those that account for the nestedness of species difference between samples (but see Šizling et al. 2022). In essence, the nestedness part of this partition is the same as our 'random-placement' effect, while turnover captures the essence of our β-diversity measures that capture non-random variation among samples. For example, in our case study we asked whether bird species in engineered and natural riparian habitats were a random subset of the same regional species pool (Figure 7). We found that $\beta_S$ values were quite high compared to the $\beta_C$ values, which indicates the turnover component is small relative to the nested component in Baselga's approach.

As with spatial β-diversity comparisons, there have been variable approaches to capture β-diversity through time (Legendre 2019, Magurran et al. 2019, Tatsumi et al. 2022). Most measures of temporal turnover calculate turnover as a metric of community dissimilarity through time. Often rates of change between an initial and subsequent samples, or the rate of decay in dissimilarity as a function of the time elapsed between samples being compared are estimated, which can then be compared across systems or taxa (e.g., Korhonen et al. 2010, Blowes et al. 2019). However, as with spatial β-diversity, these measures cannot discern whether observed rates of turnover are different from what would be expected from a random placement model through time. Authors have used different approaches to remedy this problem. For example, Dornelas et al. (2014) compared rates of temporal β-diversity to those expected from a neutral model (Hubbell 2001) to discern whether turnover rates were faster than expected under the assumption of neutral dynamics, while Stegen et al. (2013) used a null model to determine whether temporal turnover patterns were greater than expected from sampling effects. Temporal turnover can also be decomposed into changes due to abundances (similar to our ‘sampling’ effects) and changes due to species turnover (Shimadzu et al. 2015, Lamy et al. 2015). As with spatial β-diversity measures, our approach is similar, but simplifies the problem by asking whether temporal changes are non-random in a time series.

Recently, authors have developed approaches to partition the influence of species gains and losses to changes in spatial β-diversity through time (Rosenblad and Sax 2017, Tatsumi et al. 2021), and these have been expanded to incorporate changes in relative abundances (Tatsumi et al. 2022). These methods are useful for examining “winning” and “losing” species that underlie changes in spatial β-diversity through time. However, these methods risk isolating beta-diversity changes from local ($\alpha$) and regional ($\gamma$) scale changes, and are unable to disentangle random versus non-random structure associated with these changes. Thus, our approach can provide a complementary, and more complete picture into scale-dependent changes driving variation of spatial composition through time.

Finally, for simplicity we have focused here on two related metrics: $\beta_S$ and $\beta_C$. Other measures of β diversity with different weights on common and rare species (i.e., Hill numbers) (Jost 2007, Tuomisto 2010, Chao et al. 2012, 2023, but see Lande 1996) can also be calculated at different points along the rarefaction curves. For example, the metric based on Simpson's entropy, also known as the probability of interspecific interaction (PIE) (or Gini-Simpson index)(Hurlbert 1971) (where $q$=2 in the Hill number continuum; Jost 2007, Chao et al. 2014), can be visualized as the slope at the base of the rarefaction curve (Chase et al. 2018, McGlinn et al. 2019). These Hill numbers or numbers equivalents can also be used with the multiplicative diversity partition used here (i.e., $\gamma = \alpha*\beta$; Jost 2007), and result in an effective number of distinct communities, with the tuning parameter (i.e., order $q$) determining the sensitivity to rare and common species. Recently, Chao et al. (2023) also proposed a framework for standardizing beta-diversity that consider the joint influence of sampling effects and spatial/temporal aggregation. Their framework and ours both standardize biodiversity data to the same level of sample coverage when comparing β between meta-communities. However, an important difference between the approaches is that Chao et al. (2023) assume that individuals are independently sampled (i.e., randomly encountered), whereas we assume that individuals within a sample are not independent of one another due to aggregation. In fact, our primary intention here is to explicitly quantify the important contributions of aggregation to β-diversity, which cannot be directly measured with the Chao et al. (2023) approach.

## Conclusions

Ecologists are often interested in examining the role of metacommunity-level mechanisms such as dispersal limitation and environmental filtering for patterns of β-diversity (Vellend 2016, Leibold and Chase 2018). The generalized approach that we have described relies on a set of intuitive metrics from sampling theory to quantify total β-diversity ($\beta_S$), and $\beta$-diversity due to non-random aggregation ($\beta_C$), which will allow for stronger tests of hypotheses related to mechanisms expected to influence patterns of aggregation. In addition, the framework provides an integrated way to examine how changes at finer ($\alpha$) and coarser ($\gamma$) scales combine to determine variation in species composition ($\beta$). This places a central focus on scale-dependent diversity changes, with the potential to uncover deeper insights into scale-dependence by varying the focal spatial or temporal grain of the analysis. It remains an open question as to how much variation in β-diversity reflects random sampling effects vs non-random aggregation effects. Our framework provides a means of addressing this question across space and time.

## Acknowledgements:

DJM, SAB, TE, JMC gratefully acknowledge the support of the German Centre of Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig (funded by the German Research Foundation; FZT 118). ISM received funding from the European Union Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie grant agreement no. 894644. TE was supported by the German Research Foundation (DFG) within the project “Establishment of the National Research Data Infrastructure (NFDI)” in the consortium NFDI4Biodiversity (project number 442032008).

## Literature Cited

Adler, P. B. et al. 2005. Evidence for a general species-time-area relationship. Ecology 86:2032–2039.

Anderson, M. J. et al. 2011. Navigating the multiple meanings of β diversity: a roadmap for the practicing ecologist. Ecology Letters 14:19–28.

Anderson, M. J., K. E. Ellingsen, and B. H. McArdle. 2006. Multivariate dispersion as a measure of beta diversity. Ecology Letters 9:683–693.

Baselga, A. 2010. Partitioning the turnover and nestedness components of beta diversity. Global Ecology and Biogeography 19:134–143.

Blowes, S. A. et al. 2019. The geography of biodiversity change in marine and terrestrial assemblages. Science 366:339–345.

Blowes, S. A. et al. 2022. Local biodiversity change reflects interactions among changing abundance, evenness, and richness. Ecology:e3820.

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

Chao, A. et al. 2014. Rarefaction and extrapolation with Hill numbers: a framework for sampling and estimation in species diversity studies. Ecological Monographs 84:45–67.

Chao, A. et al. 2023. Rarefaction and extrapolation with beta diversity under a framework of Hill numbers: The iNEXT.beta3D standardization. Ecological Monographs n/a:e1588.

Chao, A., C.-H. Chiu, and T. C. Hsieh. 2012. Proposing a resolution to debates on diversity partitioning. Ecology 93:2037–2051.

Chao, A., and L. Jost. 2012. Coverage-based rarefaction and extrapolation: standardizing samples by completeness rather than size. Ecology 93:2533–2547.

Chase, J. M. et al. 2018. Embracing scale-dependence to achieve a deeper understanding of biodiversity and its change across communities. Ecology Letters 21:1737–1751.

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.

Chiarucci, A. et al. 2009. Spatially constrained rarefaction: incorporating the autocorrelated structure of biological communities into sample-based rarefaction. Community Ecology 10:209–214.

Coleman, B. D., M. A. Mares, M. R. Willig, and Y. H. Hsieh. 1982. Randomness, area, and species richness. Ecology 63:1121–1133.

Crist, T. O., and J. A. Veech. 2006. Additive partitioning of rarefaction curves and species-area relationships: unifying alpha-, beta- and gamma-diversity with sample size and habitat area. Ecology Letters 9:923–932.

Daskalova, G. N. et al. 2020. Landscape-scale forest loss as a catalyst of population and biodiversity change. Science 368:1341–1347.

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.

Dornelas, M. et al. 2014. Assemblage Time Series Reveal Biodiversity Change but Not Systematic Loss. Science 344:296–299.

Engel, T. et al. 2021. Using coverage-based rarefaction to infer non-random species distributions. Ecosphere 12:e03745.

Frank, S. A. 2014. Generative models versus underlying symmetries to explain biological pattern. Journal of Evolutionary Biology 27:1172–1178.

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.

Hubbell, S. P. 2001. The unified neutral theory of biodiversity and biogeography. Princeton Univ. Press, Princeton, NJ.

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.

Kobayashi, S. 1982. The rarefaction diversity measurement and the spatial distribution of individuals. Japanese Journal of Ecology 32:255–258.

Kobayashi, S. 1983. Another calculation for the rarefaction diversity measurement for different spatial distributions. Japanese Journal of Ecology 33:101–102.

Korhonen, J. J., J. Soininen, and H. Hillebrand. 2010. A quantitative analysis of temporal turnover in aquatic species assemblages across ecosystems. Ecology 91:508–517.

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 β Diversity Along Latitudinal and Elevational Gradients.” Science 335:1573–1573.

Lamy, T., P. Legendre, Y. Chancerelle, G. Siu, and J. Claudet. 2015. Understanding the Spatio-Temporal Response of Coral Reef Fish Communities to Natural Disturbances: Insights from Beta-Diversity Decomposition. PLOS ONE 10:e0138696.

Lande, R. 1996. Statistics and partitioning of species diversity, and similarity among multiple communities. Oikos 76:5–13.

Legendre, P. 2019. A temporal beta-diversity index to identify sites that have changed in exceptional ways in space–time surveys. Ecology and Evolution 9:3500–3514.

Legendre, P., D. Borcard, and P. R. Peres-Neto. 2005. Analyzing beta diversity: Partitioning the spatial variation of community composition data. Ecological Monographs 75:435–450.

Leibold, M. A., and J. M. Chase. 2018. Metacommunity Ecology. Princeton University Press, Princeton, New Jersey.

Magurran, A. E., M. Dornelas, F. Moyes, and P. A. Henderson. 2019. Temporal β diversity—A macroecological perspective. Global Ecology and Biogeography 28:1949–1960.

McGlinn, D. J. et al. 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.

McGlinn, D. J. et al. 2021. A multiscale framework for disentangling the roles of evenness, density, and aggregation on diversity gradients. Ecology:e03233.

Mori, A. S., S. Fujii, R. Kitagawa, and D. Koide. 2015. Null model approaches to evaluating the relative role of different assembly processes in shaping ecological communities. Oecologia 178:261–273.

Myers, J. A. et al. 2013. Beta-diversity in temperate and tropical forests reflects dissimilar mechanisms of community assembly. Ecology Letters:151–157.

Myers, J. A., J. M. Chase, R. M. Crandall, and I. Jiménez. 2015. Disturbance alters beta-diversity but not the relative importance of community assembly mechanisms. Journal of Ecology 103:1291–1299.

Olszewski, T. D. 2004. A unified mathematical framework for the measurement of richness and evenness within and among multiple communities. Oikos 104:377–387.

Palmer, M. W., D. J. McGlinn, and J. F. Fridley. 2008. Artifacts and artifictions in biodiversity research. Folia Geobotanica 43:245–257.

Preston, F. W. 1960. Time and space and the variation of species. Ecology 41:611–627.

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., X. Wang, and Y. Zhang. 2012. Comment on "Disentangling the Drivers of β Diversity Along Latitudinal and Elevational Gradients." Science 335:1573–1573.

Rolls, R. J. et al. 2023. Biotic homogenisation and differentiation as directional change in beta diversity: synthesising driver–response relationships to develop conceptual models across ecosystems. Biological Reviews 98:1388–1423.

Rosenblad, K. C., and D. F. Sax. 2017. A new framework for investigating biotic homogenization and exploring future trajectories: oceanic island plant and bird assemblages as a case study. Ecograph 40:1040–1049.

Sanders, H. L. 1968. Marine benthic diversity: a comparative study. American Naturalist 102:243–282.

Scheiner, S. M. et al. 2011. The underpinnings of the relationship between space, time, and species richness. Ecological Monographs 81:195–213.

Scholes, R. J. 2017. Taking the Mumbo Out of the Jumbo: Progress Towards a Robust Basis for Ecological Scaling. Ecosystems 20:4–13.

Shimadzu, H., M. Dornelas, and A. E. Magurran. 2015. Measuring temporal turnover in ecological communities. Methods in Ecology and Evolution 6:1384–1394.

Šizling, A. L. et al. 2022. Mathematically and biologically consistent framework for presence-absence pairwise indices. bioRxiv.

Stegen, J. C. et al. 2013. Stochastic and deterministic drivers of spatial and temporal turnover in breeding bird communities. Global Ecology and Biogeography 22:202–212.

Tatsumi, S., R. Iritani, and M. W. Cadotte. 2021. Temporal changes in spatial variation: partitioning the extinction and colonisation components of beta diversity. Ecology Letters 24:1063–1072.

Tatsumi, S., R. Iritani, and M. W. Cadotte. 2022. Partitioning the temporal changes in abundance-based beta diversity into loss and gain components. Methods in Ecology and Evolution 13:2042–2048.

Tucker, C. M., L. G. Shoemaker, K. F. Davies, D. R. Nemergut, and B. A. Melbourne. 2016. Differentiating between niche and neutral assembly in metacommunities using null models of β-diversity. Oikos 125:778–789.

Tuomisto, H. 2010. 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.

Ulrich, W. et al. 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.

Vellend, M. 2016. The Theory of Ecological Communities. Princeton University Press, Princeton, NJ, USA.

Warren, P., S. B. Lerman, H. Bateman, M. Katti, and E. Shochat. 2022. Point-count bird censusing: long-term monitoring of bird abundance and diversity in central Arizona-Phoenix, ongoing since 2000 ver 20. Environmental Data Initiative.

Whittaker, R. H. 1960. Vegetation of the Siskiyou Mountains, Oregon and California. Ecological Monographs 30:279–338.

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.