

REVIEW AND SYNTHESIS

Scale-dependent effect sizes of ecological drivers on biodiversity: why standardised sampling is not enough

Jonathan M. Chase^{1*} and Tiffany M. Knight²

¹Biodiversity Synthesis Laboratory, Saint Louis, MO, 63105, USA

²Department of Biology, Washington University in St. Louis, St Louis, MO, 63130, USA

*Correspondence: E-mail: jmchase88@gmail.com

Abstract

There is little consensus about how natural (e.g. productivity, disturbance) and anthropogenic (e.g. invasive species, habitat destruction) ecological drivers influence biodiversity. Here, we show that when sampling is standardised by area (species density) or individuals (rarefied species richness), the measured effect sizes depend critically on the spatial grain and extent of sampling, as well as the size of the species pool. This compromises comparisons of effects sizes within studies using standard statistics, as well as among studies using meta-analysis. To derive an unambiguous effect size, we advocate that comparisons need to be made on a scale-independent metric, such as Hurlbert's Probability of Interspecific Encounter. Analyses of this metric can be used to disentangle the relative influence of changes in the absolute and relative abundances of individuals, as well as their intraspecific aggregations, in driving differences in biodiversity among communities. This and related approaches are necessary to achieve generality in understanding how biodiversity responds to ecological drivers and will necessitate a change in the way many ecologists collect and analyse their data.

Keywords

Effect size, intraspecific aggregation, probability of interspecific encounter, rarefaction, species abundance distribution, species-accumulation curve.

Ecology Letters (2013) 16: 17–26

INTRODUCTION

Perhaps more so than any other subfield of ecology, studies on the ecological drivers that influence patterns of biodiversity are wrought with ambiguity and debate. For example, despite decades of intense empirical study and meta-analyses, there remains little consensus on the magnitude and direction of the influence of the most important natural drivers of biodiversity, including habitat productivity (Whittaker 2010; Adler *et al.* 2011), disturbance (Svensson *et al.* 2012; Fox 2013), heterogeneity (Hortal *et al.* 2009; Allouche *et al.* 2012) and habitat area (Scheiner *et al.* 2011; Triantis *et al.* 2012; Proença & Pereira 2013). Likewise, the magnitudes of the effects of anthropogenic drivers of biodiversity loss are highly variable (Murphy & Romanuk (2012) and often controversial, including effects of invasive species (Davis *et al.* 2011; Powell *et al.* 2011; Simberloff *et al.* 2013) and habitat destruction (He & Hubbell 2011; Rybicki & Hanski 2013). Because of the high degree of variability across studies, ecologists have not been able to conclusively answer some of the field's most important questions. For example, is biodiversity in certain types of ecosystems (e.g. biodiversity hotspots, islands) more susceptible to anthropogenic effects than others? Are certain taxa or functional groups of species more influenced by ecological drivers than others? Are some types of natural or anthropogenic ecological drivers more important to biodiversity than others?

The majority of empirical studies aimed at quantifying the magnitudes of effects of ecological drivers on biodiversity quantify differences in species richness, or related metrics that vary with species richness (e.g. Shannon's diversity entropy). However, the values of these biodiversity metrics are not constant, and are well known to be influenced both by *spatial grain* (the size of sampling unit) and *spatial extent* (the total area encompassed) of the study (See the

Glossary for definitions of italicised words) (e.g. Palmer & White 1994; Lande 1996; Scheiner *et al.* 2000; Gotelli & Colwell 2001). To remedy the scale-dependence in biodiversity metrics, the typical protocol among empirical ecologists is to standardise sampling effort between communities (e.g. Gotelli & Colwell 2001; Magurran & McGill 2011; Colwell *et al.* 2012). *Species density*, for example, refers the numbers of species in a given standardised sampling area, and *rarefied richness* refers to the numbers of species in a sample once the sampling effect is controlled by standardising the numbers species for a given number of individuals sampled. If sampling effort is not standardised, *richness extrapolations* are often used to estimate the numbers of species if sampling were more complete. These standardised values, then, are assumed to give an unbiased estimate of the differences (effect sizes) among communities and are used for statistical analyses and meta-analyses.

Although there is undoubtedly idiosyncrasy in how biodiversity responds to ecological drivers, here, we argue that much of the confusion and debate about the direction and magnitude of biodiversity responses might result from the fact that effect sizes measured using standardised analytical procedures are themselves scale-dependent. Two factors undermine the utility of effect sizes calculated from area- and individual-controlled estimates of biodiversity. First, in any given community, the number of species increases with increasing sampling grain and extent, known as the *species accumulation curve* (SAC), but it does so in a nonlinear, decelerating way. Except for the numerically implausible case where the SACs between two communities are exactly parallel, the difference between communities when measured at one grain or extent will differ from that at a different grain or extent, creating bias and ambiguity in the measured effect size (e.g. Cao *et al.* 2007; Sandel & Smith 2009; Giladi *et al.* 2011; Powell *et al.* 2013). For example,

Powell *et al.* (2013) showed that invasive plant species have large effects on native species biodiversity when data were measured at small spatial grains (i.e. 1–10 m²; which is the most common sampling grain used in invasion studies; Powell *et al.* 2011), but that this effect dissipated when data were measured at larger grains (i.e. 250–500 m²). Likewise, Dumbrell *et al.* (2008) showed that butterfly species richness was not strongly influenced by forest disturbance due to logging at the smallest extents, but the effect sizes increased at larger sampling extents. Second, while it has not been explicitly quantified, there are reasons to expect that the size of species pool size might influence the measured effect sizes at a given spatial grain/extent regardless of the true magnitude of the effects. For example, a standardised sampling protocol will sample a much smaller proportion of a community with a large species pool compared to a community with a smaller species pool, underestimating the true difference between the communities (Chao & Jost 2012). Likewise, localities from communities with larger species pools are probabilistically more likely to deviate from null expectations than those from communities with smaller species pools (e.g. Swenson *et al.* 2006; Lessard *et al.* 2012). Thus, differences in effect sizes among communities where the species pool varies may be confounded, such as comparisons among taxa (e.g. animals vs. plants), ecosystem types (e.g. aquatic vs. terrestrial) or biogeographic regions (e.g. temperate vs. tropics).

The implications of scale-dependence and species pool dependence on the effect sizes of ecological drivers on biodiversity are twofold. First, it implies that comparing effect sizes in standard analyses (e.g. *t*-tests, ANOVA), as well as among studies using meta-analyses, are limited or even misleading because effect sizes depend critically on choices of sampling grain and extent (e.g. Sandel & Smith 2009; Whittaker 2010; Powell *et al.* 2011, 2013). Second, understanding the nature of scale-dependence in response to ecological treatments can help to elucidate some of the possible mechanisms by which these treatments alter patterns of species' absolute and relative abundances, distributions and co-occurrences. The shape of the SAC in a community is determined by the underlying spatial distribution (occupancy and aggregation) and abundance (both absolute and relative) of each species (Preston 1962; He & Legendre 2002; Powell *et al.* 2011). As a result, any shifts in the abundances and distributions of species due to ecological drivers will be reflected in the shape of the SAC, which can allow some inference into the nature by which a particular treatment alters patterns of biodiversity (e.g. Powell *et al.* 2013).

In this article, we first show how the measured effect size (magnitude of difference between communities) of an ecological driver depends critically on both sampling grain/extent and the size of the species pool for most biodiversity metrics. This severely limits the utility of inferences made by statistical analyses comparing effect sizes within empirical studies, as well as meta-analyses across studies. For instance, in a single study of the influence of an ecological driver on biodiversity, researchers could reach very different conclusions about the magnitude (and even direction) of effect size simply by choosing a different grain or extent of sampling, or by comparing among taxa with differently sized species pools. Likewise, comparisons of the magnitude and direction of effect sizes in meta-analyses will be confounded by ambiguous effect size measurements due to variation in grain and extent of sampling in each study, as well as differences in the size of the species pools among studies. While many meta-analyses find systematic variation in effect

sizes among taxa (e.g. animals vs. plants), ecosystem types (e.g. aquatic vs. terrestrial) or biogeographic regions (e.g. temperate vs. tropical), these could just as likely be due to differences in the size of the species pool or systematic variation in sampling scale among these comparisons rather than any inherent ecological differences. Next, we show how quantifying multiple biodiversity metrics across several sampling scales can be used to better understand the scale-dependent effect sizes of ecological drivers and the underlying changes in species abundances and distributions that create these patterns. We emphasise the utility of one metric in particular, the *Effective Number of Species* (ENS) calculated from Hurlbert's (1971) *Probability of Interspecific Encounter* (PIE). This metric can provide an unambiguous measure of the effect size of an ecological driver on biodiversity regardless of spatial scale and species pool size, and can be used to disentangle the relative influence of changes in the total and relative abundances, as well as their spatial distributions, on the response of biodiversity. Finally, we conclude with an exposition on the relevance of this approach to several open questions in biodiversity research, as well as point towards future directions needed to fully elucidate the scale-dependent responses of biodiversity to natural and anthropogenic ecological drivers.

ECOLOGICAL DRIVERS AND THE SPECIES ACCUMULATION CURVE

There are multiple ways that investigators measure how biodiversity increases with spatial scale (e.g. Scheiner *et al.* 2011). Here, we use a *nested SAC*, where the numbers of species are counted in successively larger areas because it allows us to examine spatial grain and extent in the same framework, and avoids confusion with other types of species-area relationships, such as among islands or patches of different sizes. Importantly, the SAC is simply a way to distil information about the relative abundances and distributions of the populations of species in the community of interest and place it in an explicit spatial context; all other biodiversity metrics of interest (e.g. α -, β -, γ -diversity, diversity entropies, evenness) are numerical derivatives of this relationship.

The shape of an SAC (and by extension, all other biodiversity metrics) is determined by four fundamental properties of a given community (e.g. Preston 1962; May 1975; He & Legendre 2002): (1) The size of the species pool; this is determined both by regional/historical factors such as meta-community size and speciation rates as well as local biotic, environmental and dispersal limitation filters, (2) The total number of individuals that can live in a defined area; all else being equal, an area with more individuals will have more species, (3) The distribution of commonness and rarity of species in the community, known as the *Species Abundance Distribution (SAD)*; all else being equal, communities that have species that are more equitable in their relative abundances (i.e. more even) have higher diversity at smaller scales, but rise more slowly towards the maximum number of species in a given community (i.e. the species pool), whereas more uneven communities have lower local diversity but rise more rapidly towards the maximum level and (4) the degree of intraspecific *aggregation* (i.e. 'clumpiness') in the spatial distributions of species; communities where individuals of a species are distributed at random (low aggregation) have higher diversity at smaller scales but rise more slowly towards the maximum level, while communities with higher levels of intraspecific aggregation, where individuals of a species tend to co-occur more closely than

expected at random (high aggregation), have lower local diversity but rise more quickly towards the maximum. Thus, an ecological driver that influences any one of these factors (or combinations thereof) will change not just the position, but the overall shape of the SAC, leading to scale-dependent effect sizes when comparing two or more non-parallel SACs.

Although the vast majority of studies examining biodiversity's response to ecological drivers only compare values at one spatial grain/extent, some empirical analyses have explicitly quantified the scale-dependence of ecological drivers. One way that this has been done is by quantifying and comparing the intercepts and slopes of SACs among communities, (e.g. Benítez-Malvido & Martínez-Ramos 2003; Carey *et al.* 2006; Dumbrell *et al.* 2008; Sandel and Corbin 2012, Powell *et al.* 2013). While useful, there are at least two limitations to this approach. First, the SAC is often equated with the generalised species-area relationship and depicted as a power-law ($S = cA^z$ where S is species, A is area and c and z are constants); when logged, this power law is linear [$\log(S) = z\log(A) + \log(c)$], where the slopes (z) and intercepts (c) are compared among sites. However, because the power-law assumes there is no asymptote or change in slope with increasing area, parameter estimates from it may be biased on the spatial scales empiricists usually examine biodiversity (e.g. He & Legendre 1996, 2002). Second, changes in several different underlying patterns, including difference in total and relative abundances, as well as intraspecific aggregation, can lead to similar changes in the shape of the SAC. Thus, quantifying differences in intercepts and slopes across treatments or communities provide little insight into underlying processes by which ecological drivers alter biodiversity without deeper analyses (e.g. Powell *et al.* 2013).

A second way in which the effects of ecological drivers have been examined in a spatially explicit context is by partitioning diversity into local (α -diversity) and regional (γ -diversity) spatial scales, with a particular focus on the scaling factor between the two, known as β -diversity. This approach has also provided important insights into scale-dependent effect sizes of several ecological drivers (e.g. Chase & Leibold 2002; Tylianakis *et al.* 2006; Balata *et al.* 2007; Passy & Blanchet 2007; Gardezi & Gonzalez 2008; Kraft *et al.* 2011). However, there is a great deal of confusion and debate about how to estimate β -diversity and what it means and this confusion has mired progress in this area (e.g. Jost 2007; Tuomisto 2010a,b; Veech & Crist 2010; Anderson *et al.* 2011). Further, the values of any metric of β -diversity, just like changes in the shape of the SAC, cannot unambiguously indicate the underlying causes of change, such as changes in total and relative abundances or aggregations, without appropriate null models (e.g. Chase *et al.* 2011; Kraft *et al.* 2011). Finally, comparisons of β -diversity in response to ecological drivers cannot unambiguously identify the direction or magnitude of changes because its values critically depend on the how local and regional scales are defined (i.e. the grain and extent of sampling effort; Loreau 2000; Scheiner 2003). For these reasons, we suggest that analyses of β -diversity are limited in their ability to examine the scale-dependent effects of ecological drivers on biodiversity, and more continuous measures of biodiversity scaling are preferable.

For brevity, we do not explicitly consider influence of temporal scale on biodiversity (i.e. with increasing time spent sampling a location, more species will be found). As with spatial scaling, temporal scaling can have large influence on estimates of biodiversity metrics (e.g. White *et al.* 2010). However, the effects of increasing sampling

grain and extent in the temporal dimension will by-and-large mirror the results we show below for spatial grain and extent, and thus similar issues and solutions should be achievable.

COMPARISONS OF BIODIVERSITY METRICS ACROSS SCALES

Despite the recognised importance of sampling and spatial scale for estimates of biodiversity (e.g. Lande 1996; Gotelli & Colwell 2001; Olszewski 2004; Dauby & Hardy 2012), the majority of empirical studies still compare and analyse the effect of ecological drivers on biodiversity at a single sampling scale. The justification for doing so rests on the fact that if sampling effort is maintained constant between communities, the influence of scale is controlled by estimating the density of species for a given area and/or can be controlled richness rarefaction or extrapolation. It is not our intention to compare each biodiversity metric across a range of sampling grains and extents, or to explore all of the various ways by which these metrics change when the factors underlying the SAC (e.g. abundances, aggregations) are altered by ecological drivers. Instead, our goal is to illustrate how the most commonly used approaches for estimating and comparing effect sizes due to ecological drivers, by standardising to species density or rarefied richness, provide ambiguous results and misleading interpretations when considered at a single spatial scale, but can be used to more definitively estimate and compare effect sizes when spatial scale and abundance data are explicitly considered.

To illustrate the scale-dependence of effect sizes of an ecological driver, Fig. 1 presents an idealised comparison of two communities, each with differently shaped, non-parallel SACs or rarefaction curves as is typical in such comparisons. In this case, the difference between the two communities is due to changes in the SAD (rare species are disproportionately rarer in one community than the other). The number of species is shown as a function of either sampling area (species density) or the number of individuals (rarefied richness) along the x-axis, and three sampling grains/extents are shown; sample A, with a relatively small area or few individuals sampled; sample B with an intermediate area/individuals sampled; and, Sample C, with a larger area/individuals sampled. By comparing these two curves, it is clear why effect sizes of either species density or rarefied richness depend critically on the choice of grain/extent of sampling and that standard statistical analyses and meta-analyses comparing effect sizes will provide ambiguous answers. In this case, effect sizes are small in sample A, highest in sample B, and small again in sample C. Although not illustrated here, there are also many scenarios where the effect size would just decrease with sampling scale, increase with sampling scale, or even crisscross, where the effect of an ecological driver shifts from positive to negative, or vice versa. The only case where effect sizes would not vary with scale would be when two SACs are exactly parallel, which is an improbable scenario; even when SACs are parallel in log-log space, the measured effect size decreases with sampling area.

In addition to illustrating the problems of scale-dependence when comparing effect sizes of species density and rarefied richness, Fig. 1 also illustrates a metric that is much less scale-dependent. Olszewski (2004) has shown that a metric devised to indicate the degree of evenness in a community, Hurlbert's (1971) PIE, represents the slope at the base of the rarefaction curve in a community (depicted by arrows in the figure). PIE is a metric that essentially asks 'if two individuals are pulled from a community at random,

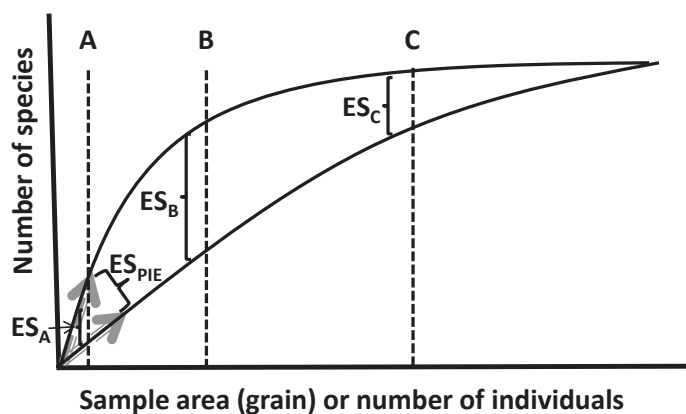


Figure 1 Hypothetical depiction of two communities that differ in their relative abundance, but not total number, of species in a community; community 1 is more even than community 2. The curves represent either the SAC (area sampled) or the rarefaction curve (individuals sampled). Three sampling grains are shown (a–c); Grain a is small compared to the extent of the community, but is typical of many empirical studies (e.g. 1 m² quadrats for vegetation), while b and c represent progressively larger sampling grains. The standardised effect size (labelled ES) for each sampling grain differs, showing that even when standardised for sampling area and/or individuals measured effect sizes are ambiguous. The slopes of the grey arrows emanating from the base of the rarefaction curve are quantified using the PIE, and the effect size of this difference (ES_{PIE}) provides a scale-independent effect size unless communities are differentially aggregated.

what is the probability that they are of different species? A community where the relative abundances of species are more even will have a higher PIE than one where a few species dominate and this is then reflected in the overall shape of their rarefaction curves. Importantly, PIE is the complement of Simpson's diversity index (D), which measures the probability that two individuals are the same species ($PIE = 1 - D$), and is sometimes called the Gini-Simpson index (e.g. Jost 2006).

Because PIE represents the slope of the rarefaction curve at its base, it is generally insensitive to sample grain/extent, just like Simpson's index (Lande 1996; Lande *et al.* 2000), and thus the difference in PIE between two communities can provide a scale-independent metric of effect size. However, as emphasised by Jost (2006), comparisons of metrics such as PIE and other diversity entropies are not meaningful in-and-of-themselves and must be converted into the ENS. ENS represents the number of equally abundant species (i.e. a perfectly even community) there would need to be to achieve the same diversity value as the one obtained. If all species in a community had exactly the same number of individuals, ENS would simply be the total number of species in that community; as the level of equitability decreases (estimated by PIE in this case), ENS becomes increasingly less than richness (i.e. rare species count as only a fraction of an 'effective' species). Although ENS based on Shannon's entropy has the most desirable mathematical properties (Jost 2006, 2007), it is more strongly influenced by sample size than Simpson's index and its derivatives (i.e. PIE) (Lande 1996; Dauby & Hardy 2012). Here, we use the ENS derived from Hurlbert's PIE, $ENS_{PIE} = 1 / \sum_{i=1}^S p_i^2$, where S is the number of species and p_i is the proportion of the community represented by species i (Jost 2006; Dauby & Hardy 2012). While the ENS_{PIE} effect size is generally insensitive to sample grain and extent when communities are distributed randomly, we will show below that sample grain and extent can

influence the values of PIE, leading to scale-dependent ENS_{PIE} effect sizes, when individuals are not spatially random (i.e. when individuals are aggregated) (see also Olszewski 2004).

Finally, we note that an often used and important group of methods that ecologists frequently use to compare biodiversity among two or more communities is to extrapolate species richness from subsamples. This approach assumes that there is some 'true' richness that each community has, and there are several parametric and nonparametric methods available to estimate this true richness based on the subsets of area sampled (reviewed in Colwell *et al.* 2012). While this can often be useful, especially when comparing among communities that have unequal sample size, we note that richness extrapolations can often obscure several important differences among communities that lead to scale-dependent patterns (i.e. the intercept and curvature of the SAC), and are limited when the SAC is not asymptotic, and thus are less useful for discerning the mechanisms by which ecological drivers alter biodiversity.

EFFECTS OF VARIATION IN ABSOLUTE ABUNDANCES, RELATIVE ABUNDANCES AND AGGREGATION ON EFFECT SIZES

Figure 2 presents the results from simulations examining the three primary mechanism by which an ecological driver can alter a species' SAC – absolute density, relative abundance and aggregations – for two differently sized species pools. Each simulation [performed in MATLAB (2011)] began with a 'control' community with an SAC that asymptotes at either the lower (30) or higher (45) level of richness in the species pool and whose slope is determined by a log-series SAD of randomly distributed individuals (i.e. not aggregated). We use Fisher's *et al.* (1943) log-series, $Y = (-1/\log(1-c)) \times c^X/X$, where Y is the relative abundance of each species, c is a coefficient ($c = 0.9$ for the control community) and X is the rank of species (the qualitative conclusions do not hinge on the specific equation underlying the SAD). Individuals were sampled from a spatial extent of 360 quadrats (with 80 individuals/quadrat initially) and quadrats were sequentially nested to increase the grain of sampling. In each simulation run except for those manipulating species aggregation, 80 individuals were allocated to each quadrat, and the species identity of each individual was chosen randomly based on its relative abundance; this information was used to create a SAC. The curves in Fig. 2 (left panels) represent the average SAC across 1000 replicate simulations.

Figure 2a shows the effect of simply reducing the density of individuals by 50% (i.e. to 40 individuals/quadrat) in response to a neutral ecological driver (i.e. equal effects on all species). Figure 2b shows the effect of an ecological driver that changes the SAD (but not density of individuals) towards a less even community (e.g. rare species more affected by the ecological driver, $c = 0.85$ in the log-series equation). Figure 2c is the same as Fig. 2b, but examines a more extreme change in the SAD ($c = 0.70$ in the log-series equation); in this case, some of the rarest species are not sampled in spatial extent of the community (360 quadrats) and are thus considered locally extirpated. Figure 2d shows the effect of an ecological driver that leads to aggregation of species in the community, but does not change the density of individuals or the SAD. Aggregation was achieved by choosing the species identity of each individual in a quadrat based on both its relative abundance and by the identity of the last species chosen (we increase the probability that this species will be chosen again by multiplying its relative abundance by

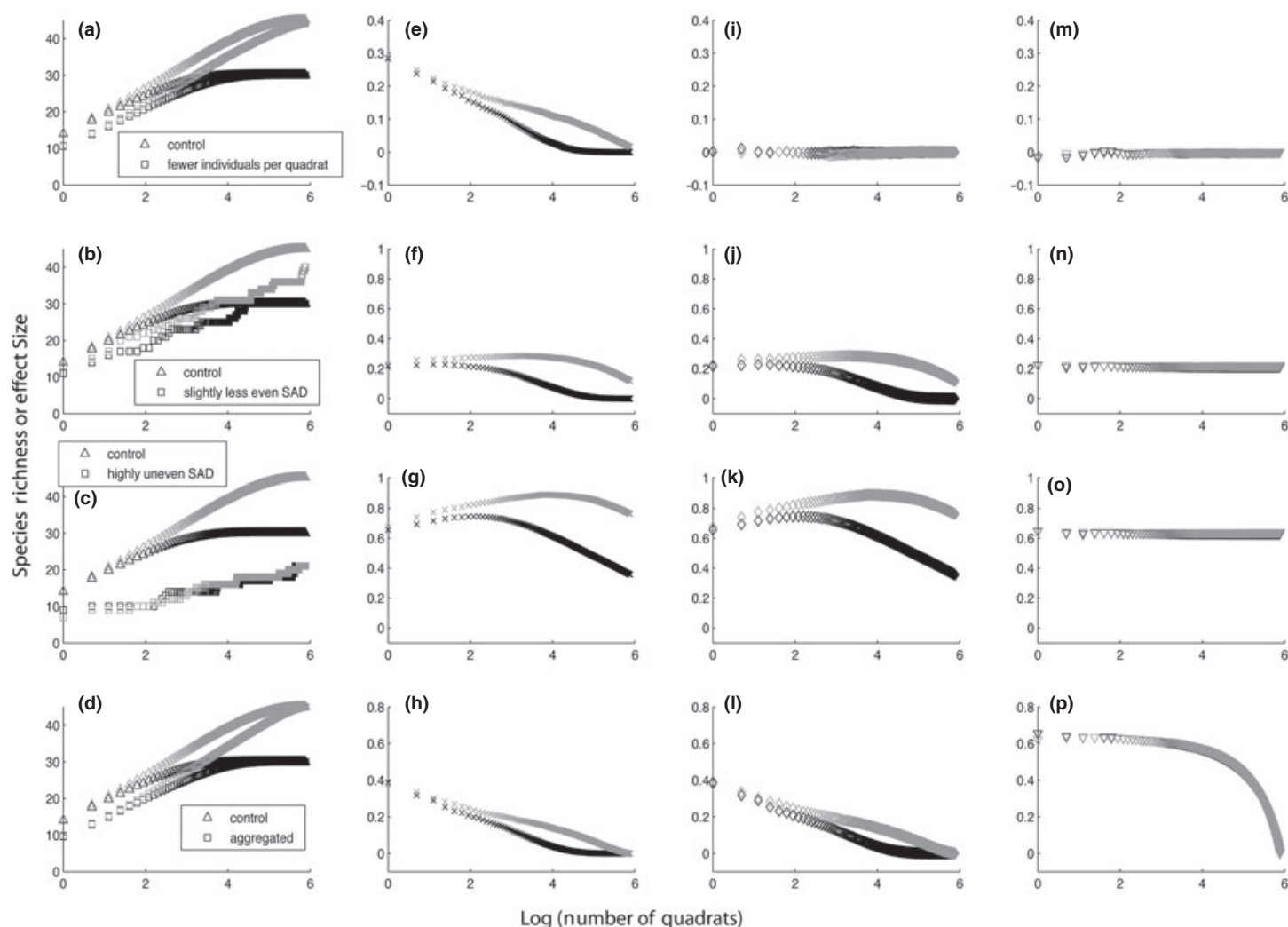


Figure 2 Relationship between species richness or biodiversity effect size and log sampling scale (number of quadrats sampled). The species pool of the communities represented is either 45 (grey symbols) or 30 (black symbols). This first column (a–d) shows species richness and sampling scale under a variety of treatments: control (80 individuals/quadrat, c in log series equation = 0.9, no aggregation), fewer individuals per quadrat (40 individuals/quadrat, c in log series equation = 0.9, no aggregation), slightly less even SAD (80 individuals/quadrat, c in log series equation = 0.85, no aggregation), highly uneven SAD (80 individuals/quadrat, c in log series equation = 0.7, no aggregation), and aggregation (80 individuals/quadrat, c in log series equation = 0.9, aggregation). The second column (e–h) shows the log ratio effect size for species richness and sampling scale. Effect size is $\ln(\text{control}) - \ln(\text{treatment})$. The third column (i–l) shows the log ratio effect size for rarefied species richness and sampling scale. The fourth column (m, n, o, p) shows the log ratio effect size for ENS_{PIE} and sampling scale.

eight). For brevity, each of these simulations only represents one direction by which ecological drivers alter the SAC, though the opposite effects (e.g. increases in the density of individuals, increases in evenness or decreases in aggregation) largely mirror those presented here.

In Fig. 2a (changing the density of individuals), Fig. 2b (moderate change to the SAD), and Fig. 2d (changing aggregation), the qualitative changes in the shape of the SACs in response to the ecological driver are largely similar; fewer species coexist at the smallest grains, but as sampling increased, the SACs converged. Figure 2c (dramatic change to the SAD) showed that species richness differed between the control and treatment even at the largest grain.

We present three measures of biodiversity and how the effect size varies with increasing sampling grain in the right three columns of Fig. 2. First, we present the effect size of species density at each sampling scale as the log response ratio $[\ln(\text{control density}) - \ln$

(treatment density)]. For each mechanism in isolation, effect size on species density was large at the smallest sampling grains and lowest at the largest grain sizes. However, when the SAD was strongly altered (Fig. 2g), effect sizes increased with grain in intermediate extents, until they decreased at the largest extents and this effect was pronounced with larger species pools. Importantly, for each mechanism, the effect sizes at a single sampling grain were always higher for the larger species pool than for the smaller species pool, even though the underlying mechanisms by which an ecological driver alters a community was exactly the same. This emphasises that the common practice of comparing and contrasting effect sizes in single studies or in meta-analyses between communities (e.g. among biogeographic regions or ecosystem types) or between different groups of species in a community (e.g. among different taxonomic groupings) that differ in the size of their species pool are misleading even when sampling grain is kept constant. Thus, while species den-

sity is the most commonly used metric to compare the influence of ecological drivers on biodiversity in empirical studies, it is clear that it is the most variable and least informative for estimating and comparing the magnitude of effect sizes or for elucidating potential mechanisms leading to the observed patterns.

Second, we present the log-ratio effect size of rarefied species richness in the third column of Fig. 2. Species richness was rarefied by sampling equal numbers of individuals per quadrat from both the control and treatment plots (Gotelli & Colwell 2001). When the ecological driver influences a community only by reducing the density of individuals, then the rarefied species richness effect size is zero and does not change with sampling grain regardless of the size of the species pool (Fig. 2, top row). For all other mechanisms examined in Fig. 2, the ecological driver did not influence individual density, and thus the rarefied species richness effect size relationship with sample grain is identical to the species density effect size relationship with sample grain. Thus, although differences in effect sizes of rarefied richness can also depend on sampling grain (e.g. if the SAD or aggregation changes), comparing this measure of effect size to the species density effect size allows one mechanism (the role of changes in the density of individuals) behind shifts in the SAC to be discerned.

Third, we present the log-ratio effect sizes of the ENS_{PIE} . Because PIE represents the base of the rarefaction curve, it controls for differences in the total density of individuals. As such, when an ecological driver influences a community only by reducing the density of individuals, the effect size of ENS_{PIE} is zero and does not change with sampling grain (Fig. 2m). When an ecological driver decreases the evenness of a community, it reduces the ENS_{PIE} , resulting in an effect size of ENS_{PIE} that is positive, but does not change with sampling grain. The magnitude of the ENS_{PIE} effect size depends on the magnitude of changes in the SAD, as can be seen by comparing Fig. 2n,o. Importantly, the magnitude ENS_{PIE} effect sizes were the same between communities with disparately sized species pools when the underlying strengths of effects on the SAD were the same. The constancy of ENS_{PIE} effect sizes across sample grain and species pool size was expected given its scale-independence (e.g. Dauby & Hardy 2012). Finally, when an ecological driver leads to aggregation in a community, this reduced the ENS_{PIE} at small, but not larger spatial scales, resulting in an ENS_{PIE} effect size relationship that decreases with sampling scale (Fig. 2p). This is because aggregation causes species to appear absent or rare at smaller sampling grains even though they have not changed their relative abundance across the extent of the sampled community. Once sampling grain is large enough, these last species are sampled and the difference in ENS_{PIE} between the two communities disappears.

Figure 3 shows the relationship between species richness or effect sizes with sampling scale when combinations of the factors that influence the shape of the SAC – the density of individuals, relative abundance, or aggregation – are altered pairwise combinations or all three together. This illustrates that comparing how several different metrics of biodiversity, and their effect sizes, vary with sampling scale can provide information about the mechanisms that underlie how an ecological driver influences communities. Comparing species density effect size relationship to the rarefied richness effect size curve (second and third columns in Fig. 3) allows one to determine whether ecological drivers alter the shape of the SAC by changing the density of individuals. If the density of individuals is not changed by an ecological driver, then these curves are identical

(e.g. third row of Fig. 3). The intercept and the shape of the ENS_{PIE} effect size relationship with sample scale provide information about how the ecological driver alters the SAD and aggregation of the community. When the effect size at the largest sampling scale deviates from zero, the ecological driver altered the SAD and thus altered the ENS_{PIE} in the community. If the effect size does not change with sampling grain or extent, then the ecological driver has not changed the propensity for species to aggregate (e.g. due to habitat heterogeneity or frequency-dependent interactions). However, when the shape of the relationship is not flat, then the ecological driver altered the aggregation of the community. In this case, the overall effect size of the driver can still be determined by calculating the ENS_{PIE} randomised across the extent of all sampled units (i.e. the effect size measured at the furthest point along the x-axis of Fig. 3, see also Olszewski 2004).

TOWARDS AN EMPIRICAL PROTOCOL FOR MEASURING AND ANALYZING IMPACTS OF ECOLOGICAL DRIVERS ON BIODIVERSITY

It is well known that the values obtained with most biodiversity metrics depend on spatial grain and extent (e.g. Lande 1996; Scheiner *et al.* 2000; Gotelli & Colwell 2001; Lande *et al.* 2000; Olszewski 2004; Jost 2006, 2007; Dauby & Hardy 2012). However, most empirical studies that measure multiple ecological treatments, as well as meta-analyses among those experiments, still assume that reporting biodiversity metrics that are standardised for sampling effort (i.e. through area- or individual-standardised analyses) is sufficient to provide a robust estimate of the effect size of the treatment on biodiversity. Our results emphasise that this is not the case; effect sizes vary with sampling grain and extent for most of the commonly reported biodiversity metrics. Thus, the majority of studies that have been performed cannot be used to unambiguously define the magnitude (and sometimes even direction) of the effect size of an ecological driver on biodiversity. This means that researchers studying the same exact system and point in time can get a different answer on the effect of an ecological driver depending on the (often arbitrary) chosen sampling grain and/or extent using standard statistical tools and corrections (see also Cao *et al.* 2007; Sandel & Smith 2009; Giladi *et al.* 2011; Powell *et al.* 2011, 2013).

In addition, it is usually assumed that standardising sampling effort is sufficient to provide a robust and comparable effect sizes when species pools differ, such as across biogeographic zones (e.g. latitude, island size), environmental gradients (e.g. productivity), ecosystem types (e.g. aquatic vs. terrestrial) or taxonomic groupings (e.g. animals vs. plants). However, our results emphasise that comparisons of biodiversity responses to ecological drivers are strongly confounded by the size of the species pool. Thus, a study could conclude that an ecological driver has a greater effect on biodiversity of one taxon compared to another, or one ecosystem compared to another, when in reality, the underlying mechanisms by which the driver influences the community are the same. Or it could conclude that the effects of the ecological driver are the same, when in reality, they are quite different. Thus, the standard practice of comparing effect sizes among different groups of species within a site, or among sites with differently sized species pools in a single study with a common sampling grain using standard statistical approaches for measuring and comparing effect sizes (e.g. t-tests, ANOVA), and of comparing the magnitude of effect sizes among taxonomic

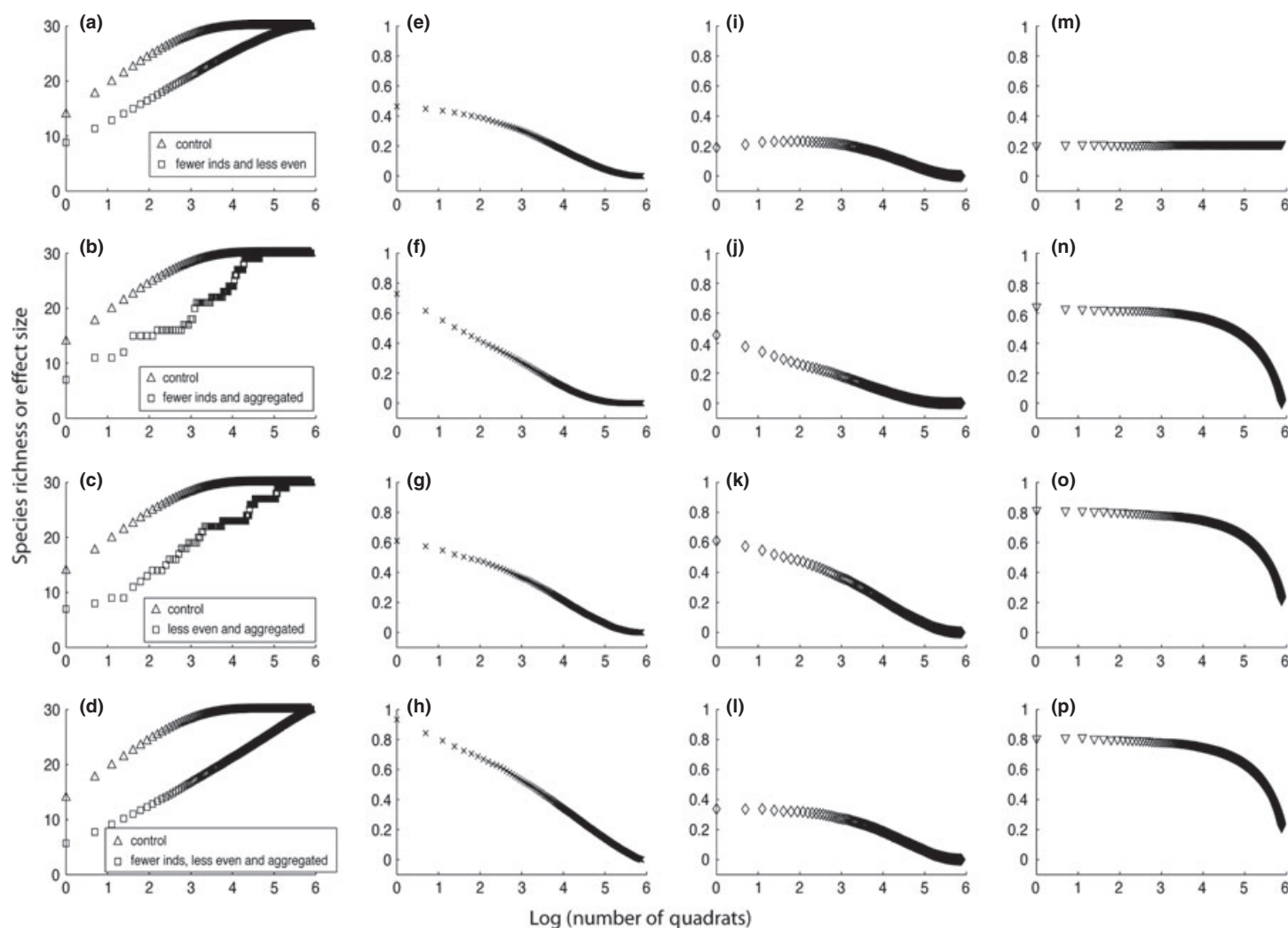


Figure 3 Relationship between species richness or biodiversity effect size and log sampling scale (number of quadrats sampled). The species pool of the communities represented is 30. This first column (a–d) shows species richness and sampling scale under a variety of treatment combinations: control (80 individuals/quadrat, c in log series equation = 0.9, no aggregation), fewer individuals per quadrat (40 individuals/quadrat), less even SAD (c in log series equation = 0.85), and aggregation. The effect sizes shown in columns 2–4 are identical to those described in Fig. 1 and always represent $\ln(\text{control}) - \ln(\text{treatment combination})$.

grouping, ecosystem types or biogeographic zones in meta-analyses when species pool sizes differ, can in fact, be quite misleading.

Next, we outline the necessary steps to gain less ambiguous information so that we can better understand how ecological drivers influence biodiversity across scales.

Collect abundance, coverage and/or biomass data and keep it separated by spatial location

Here, we advocate that it is essential, not just preferable, to measure the total density and relative abundance of individuals (or some estimate thereof) of species in the community, as well as their spatial locations, to obtain unbiased and interpretable effect sizes of ecological drivers on biodiversity. This has been emphasised by many mathematical ecologists who have recognised the severe sample-size bias that emerge when values of species richness are used for comparisons (e.g. Lande 1996; Jost 2006, 2007; Dauby & Hardy 2012). However, most empiricists still collect and compare differ-

ences in the numbers of species between samples (i.e. species density) because such comparisons seem to make the most intuitive sense and the data are typically easier to collect and analyse. Unfortunately, presence–absence data at a single spatial scale do not allow us to unambiguously quantify the magnitude (and sometimes even direction) of effect sizes of an ecological driver, nor can they be used to make comparisons of those effect sizes within or among communities. Presence–absence data at multiple spatial scales across communities, such as those used to construct and compare slopes and intercepts of SACs across communities, provide more information, but still do not allow an explicit consideration of the possible mechanisms that might lead to shifts in the SAC (i.e. changes in density, relative abundance or aggregation). Although comparisons of rarefied species richness, controlling for the numbers of individuals, provides information on how changes in the number of individuals alters species numbers, this does not provide an unbiased estimate of differences between communities because the value depends critically on sampling grain and extent. Comparisons of

the slope of the rarefaction curves at the base (estimated by PIE) are much less scale-dependent. Finally, spatially explicit data on the densities and relative abundances of individuals are most useful, because these data allow us to discern the importance of aggregation as a mechanism by which an ecological driver influences biodiversity patterns. However, in the absence of spatially explicit data, randomisations of ENS_{PIE} calculated from relative abundance data (or related metrics) across sample plots within a community can still provide unambiguous effect size measurements when randomised across the level of the entire extent of sampling.

Standardising effect size measurements

In order to fully understand how an ecological driver alters the shape of the SAC, three unambiguous effect sizes can be measured; changes in the density of individuals, changes in the relative abundance of individuals and changes in the aggregations among individuals.

(1) The effect size on the density of individuals. If there are differences in the shape of the SAC between communities, but no differences in the relationship between rarefied richness and sample grain, this implies that the entire SAC difference is simply due to differences in the density of individuals. Thus, this effect size measure gives an estimate of the contribution of changes in density to the shift in SACs and measured effect sizes

(2) The effect size of ENS_{PIE} . PIE represents the slope at the base of the rarefaction curve (Olszewski 2004), and thus controls for differences in the numbers of individuals, indicating the degree to which the relative abundances of species (i.e. SAD) have changed across treatments. In cases where ecological drivers do not alter the level of aggregation among species, effect sizes of ENS_{PIE} are largely scale-independent, and are also independent of species pool size (Figs 2 and 3), providing an unambiguous effect size for single studies and meta-analyses. However, when an ecological driver alters the aggregations among species, the effect size of ENS_{PIE} is also biased by sample grain (Figs 2p and 3n-p). In such cases, an unambiguous effect size of an ecological driver on the relative abundances of species in the entire community can be estimated by randomising samples among sites, or simply by lumping all of the individuals from every sample taken into a single analysis of ENS_{PIE} (i.e. the effect size measured at the furthest point along the x-axis in Figs 2 and 3). We also note that the process of extinction debt (i.e. the slow decline of a species' abundance towards local extinction) will influence the magnitude by which an ecological driver influences the relative abundances of species in a community and thus the measured ENS_{PIE} effect size. An ecological driver might appear to have a minimal influence on the relative abundance of species shortly after environmental conditions are changed, however, the effect on relative abundance will be much more dramatic after enough time has passed for local extinctions to appear.

(3) The effect size of an ecological driver on the degree of aggregation among species. The degree of aggregation or heterogeneity within a community can strongly influence the shape of the SAC and differences among communities in their degree of aggregation is one of the main mechanisms by which SAC curves can intersect (i.e. an aggregated community can have fewer species at smaller scales but more species at larger scales than an unaggregated community) (Lande *et al.* 2000). The aggregation among species within a community can be measured by comparing the ENS_{PIE} estimated

from within one or a few samples arranged close together to the ENS_{PIE} randomised across the extent of the community (Figs 2 and 3, see also Olszewski 2004). If the values of ENS_{PIE} within samples are similar to those across samples, individuals in the community are dispersed reasonably randomly, whereas if the within sample values of ENS_{PIE} are much higher (i.e. more uneven) than the across sample values, then aggregation is playing an important role in the shape of the SAC. We caution, however, that this effect size is the most tenuous of those we have advocated because it is contingent on the choice of sampling grain and extent, and in particular their ratios (just as measures of β -diversity depend on the often arbitrary choices of the scale at which α - and γ -diversity are measured; Loreau 2000). Nevertheless, the effect size measured from ENS_{PIE} randomised from the total sample of individuals from across the community extent provides the best understanding of the overall effect of ecological drivers. If only one effect size can be measured, we advocate this as the most important for use in statistical analyses and meta-analyses because it is the least ambiguous across sample grains and species pools.

CONCLUSIONS

Our study builds on two important findings from previous literature: (1) Changes in absolute abundances, relative abundances and aggregations of species will alter the shape of the SAC (e.g. He & Legendre 2002) and (2) most metrics of biodiversity are influenced by the sample grain and spatial extent of the study (e.g. Gotelli & Colwell 2001; Olszewski 2004; Dauby & Hardy 2012). We highlight the implications of these findings for the measurement and interpretation of biodiversity effect sizes. Any factor that influences the total abundance, relative abundance and/or aggregation of species within a community will shift the overall shape of its SAC (i.e. its intercepts and curvature), such that SACs of different communities will not be parallel. As a result, estimates of the difference in biodiversity between two or more communities (i.e. its effect size) are usually scale-dependent, even when sampling is standardised. Values of effect size from statistical tests and syntheses of these effect sizes in meta-analyses will be ambiguous and even misleading. Further, differences in the size of the species pool between groups compared within a study (e.g. animals vs. plants), across ecosystem types (e.g. aquatic vs. terrestrial) or across biogeographic zones (e.g. temperate vs. tropical communities) compromise interpretation of effect size similarities and differences. In all, despite decades of intense investigation, the limitations of current analytical techniques through standardised sampling at single spatial scales have left us with very limited insight on the nature by which natural and anthropogenic factors influence patterns of biodiversity.

Despite these limitations, we are optimistic that new syntheses will be forthcoming. For example, debates about the relative intensity of species extinctions that result from habitat destruction centre around whether rare species are disproportionately affected by habitat loss (e.g. Rybicki & Hanski 2013) or SADs remain constant (e.g. He & Hubbell 2011). Empirical results on this issue have been equivocal, but this could be due to differences across studies in the spatial grain measured and the size of the species pools being examined. More information could be gleaned from the empirical literature by comparing the effect sizes of ENS_{PIE} due to changes in habitat size to determine whether the overall shape of the SAD changes. Likewise, although patterns of how species richness varies

with increasing productivity at small spatial grains are highly idiosyncratic (e.g. Whittaker 2010; Adler *et al.* 2011), several empirical studies have suggested that more productive systems have higher levels of intraspecific aggregation, leading to larger effect sizes of productivity at larger relative to smaller spatial scales (e.g. Chase & Leibold 2002; Chalcraft *et al.* 2008; Gardezi & Gonzalez 2008). Comparing the effect sizes of ENS_{PIE} within and among spatial scales in many different systems could allow an exploration of generality in this pattern. Once we gain a better understanding of the patterns of how biodiversity responds to both natural and anthropogenic variation in the environment, through its influence on total abundances, relative abundances and spatial distributions of species, we can move forward to forge understanding of the mechanisms that influence these patterns, and how they vary through space and time.

ACKNOWLEDGEMENTS

We thank the participants 'Ecological Effects of Global Change' conference held in Paris, France on June 22, 2012 for stimulating discussion regarding the ideas presented here, and M. Hochberg and M. Holyoak for the invitation to participate in that conference and this special issue. Comments by several reviewers challenged us to improve our presentation and framework substantially. Our work in this area has been enhanced by conversations and collaborations with many colleagues, most notably K. Powell and the participants of the 'Gradients of β -diversity' Working Group supported by the National Center for Ecological Analysis and Synthesis (NCEAS), a centre funded by NSF (grant EF-0553768); the University of California, Santa Barbara; and the state of California. We also acknowledge the National Science Foundation (DEB-0241080, 0816113, 0949984) for research support that stimulated many of the concepts presented herein.

AUTHORSHIP

Both authors conceived the study. TK did the simulation modelling. JC wrote the first draft of much of the manuscript, and TK revised the draft.

Glossary

Species density

The number of species in a given standardised amount of area. Note, however, density needs to be compared on the scale at which it was measured; converting density to smaller areas (e.g. species/m²) will give misleading answers.

Nested Species Accumulation Curve (SAC)

The curve that depicts how species accumulate as successively more area is sampled. Smaller areas are subsumed within larger areas, and thus it is a nested measurement.

Species Abundance Distribution (SAD)

The proportional abundances of species in a community, which is usually dominated by a few common species, and then populated by many rare species. The shape of the SAD changes as the relative evenness of the community changes.

Intraspecific Aggregation

The degree to which individuals of a species are 'clumped' or distributed randomly in a community.

Rarefied Richness

The number of species in a community once any differences in the numbers of individuals among communities is controlled.

Richness Extrapolations

Parametric and nonparametric techniques developed to take observations of species numbers and relative abundance with limited sampling effort and projecting how many species are expected in the entire community.

Effective Number of Species (ENS)

An index of the numbers of species in a community that explicitly accounts for the fact that rare species have a disproportionate effect on measures of species richness, but not other diversity-based analyses. The ENS represents the number of equally abundant species (i.e. a perfectly even community) there would need to be to achieve the same diversity value as the one obtained. If the abundances of all species in a community were exactly identical, ENS would simply be the total number of species in that community; as the level of evenness decreases, ENS becomes increasingly less than richness (i.e. rare species count as only a fraction of an 'effective' species)

REFERENCES

- Adler, P.B., Seabloom, E.W., Borer, E.T., Hillebrand, H., Hautier, Y., Hector, A. *et al.* (2011). Productivity is a poor predictor of plant species richness. *Science*, 333, 1750–1753.
- Allouche, O., Kalyuzhny, M., Moreno-Rueda, G., Pizarro, M. & Kadmon, R. (2012). Area-heterogeneity tradeoff and the diversity of ecological communities. *Proc. Natl Acad. Sci.*, 109, 17495–17500.
- Anderson, M.J., Crist, T.O., Chase, J.M., Vellend, M., Inouye, B.D., Freestone, A.L. *et al.* (2011). Navigating the multiple meanings of β -diversity: a roadmap for the practicing ecologist. *Ecol. Lett.*, 14, 19–28.
- Balata, D., Piazzini, L. & Benedetti-Cecchi, L. (2007). Sediment disturbance and loss of beta diversity on subtropical rocky reefs. *Ecology*, 88, 2455–2461.
- Benítez-Malvido, J. & Martínez-Ramos, M. (2003). Impact of forest fragmentation on understory plant species richness in Amazonia. *Conserv. Biol.*, 17, 389–400.
- Cao, Y., Hawkins, C.P., Larsen, D.P. & Van Sickle, J. (2007). Effects of sample standardization on mean species detectabilities and estimates of relative differences in species richness among assemblages. *Am. Nat.*, 170, 381–395.
- Carey, S., Harte, J. & Del Moral, R. (2006). Effect of community assembly and primary succession on the species-area relationship in disturbed ecosystems. *Ecography*, 29, 866–872.
- Chalcraft, D.R., Cox, S.B., Clark, C., Cleland, E.E., Suding, K.N., Weiher, E. *et al.* (2008). Scale-dependent responses of plant biodiversity to nitrogen enrichment. *Ecology*, 89, 2165–2171.
- Chao, A. & Jost, L. (2012). Coverage-based rarefaction and extrapolation: standardizing samples by completeness rather than size. *Ecology*, 93, 2533–2547.
- Chase, J.M. & Leibold, M.A. (2002). Spatial scale dictates the productivity-diversity relationship. *Nature*, 415, 427–430.
- Chase, J.M., Kraft, N.J., Smith, K.G., Vellend, M. & Inouye, B.D. (2011). Using null models to disentangle variation in community dissimilarity from variation in α -diversity. *Ecosphere*, 2, art. 24.
- Colwell, R.K., Chao, A., Gotelli, N.J., Lin, S.Y., Mao, C.X., Chazdon, R.L. *et al.* (2012). Models and estimators linking individual-based and sample-based rarefaction, extrapolation and comparison of assemblages. *J. Plant Ecol.*, 5, 3–21.
- Dauby, G. & Hardy, O.J. (2012). Sampled-based estimation of diversity *sensu stricto* by transforming Hurlbert diversities into effective number of species. *Ecography*, 35, 661–672.
- Davis, M.A., Chew, M.K., Hobbs, R.J., Lugo, A.E., Ewel, J.J., Vermeij, G.J. *et al.* (2011). Don't judge species on their origins. *Nature*, 474, 153–154.

- Dumbrell, A.J., Clark, E.J., Frost, G.A., Randell, T.E., Pitchford, J.W. & Hill, J.K. (2008). Changes in species diversity following habitat disturbance are dependent on spatial scale: theoretical and empirical evidence. *J. Appl. Ecol.*, 45, 1531–1539.
- Fisher, R.A., Corbet, A.S. & Williams, C.B. (1943). The relation between the number of species and the number of individuals in a random sample of an animal population. *J. Anim. Ecol.*, 12, 42–58.
- Fox, J.W. (2013). The intermediate disturbance hypothesis should be abandoned. *Trends Ecol. Evol.*, 28, 86–92.
- Gardezi, T. & Gonzalez, A. (2008). Scale dependence of the species-energy relationship: evidence from fishes in thousands of lakes. *Am. Nat.*, 171, 800–815.
- Giladi, I., Ziv, Y., May, F. & Jeltsch, F. (2011). Scale-dependent determinants of plant species richness in a semi-arid fragmented agro-ecosystem. *J. Veg. Sci.*, 22, 983–996.
- Gotelli, N.J. & Colwell, R.K. (2001). Quantifying biodiversity: procedures and pitfalls in the measurement and comparison of species richness. *Ecol. Lett.*, 4, 379–391.
- He, F. & Hubbell, S.P. (2011). Species–area relationships always overestimate extinction rates from habitat loss. *Nature*, 473, 368–371.
- He, F. & Legendre, P. (1996). On species–area relations. *Am. Nat.*, 148, 719–737.
- He, F. & Legendre, P. (2002). Species diversity patterns derived from species–area models. *Ecology*, 83, 1185–1198.
- Hortal, J., Triantis, K.A., Meiri, S., Thébaud, E. & Sfenthourakis, S. (2009). Island species richness increases with habitat diversity. *Am. Nat.*, 174, E205–E217.
- Hurlbert, S.H. (1971). The nonconcept of species diversity: a critique and alternative parameters. *Ecology*, 52, 577–586.
- Jost, L. (2006). Entropy and diversity. *Oikos*, 113, 363–375.
- Jost, L. (2007). Partitioning diversity into independent alpha and beta components. *Ecology*, 88, 2427–2439.
- Kraft, N.J.B., Comita, L.S., Chase, J.M., Sanders, N.J., Swenson, N.G., Crist, T.O. *et al.* (2011). Disentangling the drivers of β -diversity along latitudinal and elevational gradients. *Science*, 333, 1755–1758.
- Lande, R. (1996). Statistics and partitioning of species diversity, and similarity among multiple communities. *Oikos*, 76, 5–13.
- Lande, R., DeVries, P.J. & Walla, T.R. (2000). When species accumulation curves intersect: implications for ranking diversity using small samples. *Oikos*, 89, 601–605.
- Lessard, J.-P., Belmaker, J., Myers, J.A., Chase, J.M. & Rahbek, C. (2012). Inferring local ecological processes amid species pool influences. *Trends Ecol. Evol.*, 27, 600–607.
- Loreau, M. (2000). Are communities saturated? On the relationship between α , β and γ diversity. *Ecol. Lett.*, 3, 73–76.
- Magurran, A.E. & McGill, B.J. (eds) (2011). *Biological Diversity: Frontiers in Measurement and Assessment*. Oxford University Press, New York, USA.
- May, R. (1975). Patterns of species abundance and diversity. In: *Ecology and Evolution of Communities* (eds Cody, M. & Diamond, J.). Belknap Press of Harvard University, Cambridge, USA, pp. 81–120.
- Murphy, G.E.P. & Romanuk, T.N. (2012). A meta-analysis of community response predictability to anthropogenic disturbances. *Am. Nat.*, 180, 316–327.
- 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. & White, P.S. (1994). Scale dependence and the species-area relationship. *Am. Nat.*, 144, 717–740.
- Passy, S.I. & Blanchet, F.G. (2007). Algal communities in human-impacted stream ecosystems suffer beta-diversity decline. *Divers. Distrib.*, 13, 670–679.
- Powell, K.I., Chase, J.M. & Knight, T.M. (2011). A synthesis on plant invasion effects on biodiversity across spatial scales. *Am. J. Bot.*, 98, 539–548.
- Powell, K.P., Chase, J.M. & Knight, T.M. (2013). Invasive plants have scale-dependent effects on native biodiversity by altering the species–area relationship. *Science*, 339, 316–318.
- Preston, F.W. (1962). The canonical distribution of commonness and rarity: Part I. *Ecology*, 43, 185–215.
- Proença, V. & Pereira, H.M. (2013). Species–area models to assess biodiversity change in multi-habitat landscapes: the importance of species habitat affinity. *Basic Appl. Ecol.*, 14, 102–114.
- Rybicki, J. & Hanski, I. (2013). Species–area relationships and extinctions caused by habitat loss and fragmentation. *Ecol. Lett.*, 16 (S1), 27–38.
- Sandel, B. & Corbin, J.D. (2012). Scale and diversity following manipulation of productivity and disturbance in Californian coastal grasslands. *J. Veg. Sci.*, 23, 906–918.
- Sandel, B. & Smith, A.B. (2009). Scale as a lurking factor: incorporating scale-dependence in experimental ecology. *Oikos*, 118, 1284–1291.
- Scheiner, S.M. (2003). Six types of species–area curves. *Glob. Ecol. Biogeogr.*, 12, 441–447.
- Scheiner, S.M., Cox, S.B., Willig, M., Mittelbach, G.G., Osenberg, C. & Kaspari, M. (2000). Species richness, species–area curves and Simpson's paradox. *Evol. Ecol. Res.*, 2, 791–802.
- Scheiner, S.M., Chiarucci, A., Fox, G.A., Helmus, M.R., McGlenn, D.J. & Willig, M.R. (2011). The underpinnings of the relationship of species richness with space and time. *Ecol. Monogr.*, 81, 195–213.
- Simberloff, D., Martin, J.-L., Genovesi, P., Maris, V., Wardle, D.A., Aronson, J. *et al.* (2013). Impacts of biological invasions: what's what and the way forward. *Trends Ecol. Evol.*, 28, 58–66.
- Svensson, J.R., Lindegarth, M., Jonsson, P.R. & Pavia, H. (2012). Disturbance–diversity models: what do they really predict and how are they tested? *Proc. R. Soc. B*, 279, 2163–2170.
- Swenson, N.G., Enquist, B.J., Pither, J., Thompson, J. & Zimmerman, J.K. (2006). The problem and promise of scale dependency in community phylogenetics. *Ecology*, 87, 2418–2424.
- Triantis, K.A., Guilhaumon, F. & Whittaker, R.J. (2012). The island species–area relationship: biology and statistics. *J. Biogeogr.*, 39, 215–231.
- 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.
- Tylianakis, J.M., Klein, A.-M., Lozada, T. & Tscharntke, T. (2006). Spatial scale of observation affects α , β and γ diversity of cavity-nesting bees and wasps across a tropical land-use gradient. *J. Biogeogr.*, 33, 1295–1304.
- Veech, J.A. & Crist, T.O. (2010). Toward a unified view of diversity partitioning. *Ecology*, 91, 1988–1992.
- White, E.P., Ernest, S.M., Adler, P.B., Hurlbert, A.H. & Lyons, S.K. (2010). Integrating spatial and temporal approaches to understanding species richness. *Philos. Trans. R. Soc. B. Biol. Sci.*, 365, 3633–3643.
- Whittaker, R.J. (2010). Meta-analyses and mega-mistakes: calling time on meta-analysis of the species richness–productivity relationship. *Ecology*, 91, 2522–2533.

Editor, Howard Cornell

Manuscript received 3 October 2012

First decision made 4 November 2012

Second decision made 22 February 2013

Manuscript accepted 12 March 2013