Received: 22 September 2020 | Revised: 26 October 2021 | Accepted: 17 November 2021
DOI: 10.1111/geb.13437

RESEARCH ARTICLE

Global Ecology and Biogeography A Journal of Macroecology
WILEY

A null model for quantifying the geometric effect of habitat subdivision on species diversity

David C. Deane[1,2] | Dingliang Xing[1,3] | Cang Hui[4,5] | Melodie McGeoch[2] | Fangliang He[1]

[1]Department of Renewable Resources, University of Alberta, Edmonton, Alberta, Canada

[2]Department of Ecology, Evolution and Environment, La Trobe University, Bundoora, Victoria, Australia

[3]Zhejiang Tiantong Forest Ecosystem National Observation and Research Station, School of Ecological and Environmental Sciences, East China Normal University, Shanghai, China

[4]Centre for Invasion Biology, Department of Mathematical Sciences, Stellenbosch University, Matieland, South Africa

[5]Biodiversity Informatics Unit, African Institute for Mathematical Sciences, Cape Town, South Africa

Correspondence
David C. Deane, Department of Ecology, Evolution and Environment, La Trobe University, Bundoora, VIC 3086, Australia.
Email: d.deane@latrobe.edu.au

Editor: Volker Bahn

Abstract

Aim: To derive null models for the expected number of species shared among multiple samples or habitat patches, allowing exploration of the geometric effects of subdivision on species diversity.

Location: Global.

Major taxa studied: Predominantly sessile organisms.

Methods: The occurrence probability of a species in a subdivided area depends on its abundance and spatial pattern over a known habitat extent. The joint probability that two subareas share a species is the product of the probability of species occurrence in each subarea provided that the latter probability is independent. The sum of this probability over all species is the number of species the two subareas share, or zeta diversity of order 2. Generalizing from 2 to *m* subareas yields a null model for zeta diversity of order *m*. From zeta diversity, many metrics (e.g., beta and gamma diversity) for the *m* habitat patches can be calculated, revealing the effects of increasing habitat fragmentation.

Results: The null models show the geometric effects of subdivision depend on patterns of spatial distribution of species within a landscape and evenness of species abundance distribution. For aggregated assemblages, increasing subdivision decreases shared species, increases beta diversity and results in higher total species richness in subdivided habitat than an equal contiguous area.

Main conclusions: To correctly interpret diversity patterns in fragmented habitat the geometric effects of subdivision must be considered. Our models explain why fragmented habitat could have higher diversity than continuous habitat of equal area but predict a threshold patch-size above which this will not occur (here *c*. 100 ha). Apparently positive diversity effects of subdivision, including more species in groups of small patches, are probable outcomes of spatial aggregation of assemblages. The shared species null models offer an analytical tool for exploring the geometric effects of subdivision on diversity while controlling for total habitat area.

KEYWORDS
fragmentation per se, geometric effect of habitat subdivision, intraspecific spatial pattern, random placement model, sampling theory, shared species, species richness, zeta diversity

1 | INTRODUCTION

What difference in species richness should be expected between a single contiguous patch of habitat and an equal area of habitat subdivided among several smaller patches? This question came to prominence in the 1970s SLOSS debate, in which researchers questioned whether a Single Large Or Several Small reserves would protect more species (Diamond, 1975; Simberloff & Abele, 1976). Related questions have recently contributed to conflicting opinion on the biodiversity impacts of the breaking apart of habitat versus habitat loss from land clearance (i.e., the fragmentation 'per se' debate; Fahrig, 2017; Fahrig et al., 2019; Fletcher et al., 2018). Recently, May et al. (2019) showed via simulation that increasing habitat subdivision can, counterintuitively, increase the survival probability of any species whose individuals have aggregated spatial distributions; referred to as the geometric effect of habitat loss. Given this finding, if the consequences of fragmentation for diversity are to be comprehensively understood, this geometric effect from habitat subdivision needs to be distinguished from changes to diversity that arise subsequently from impacts of habitat loss and isolation on species demographic rates (Chisholm et al., 2018; Hanski et al., 2013; Kobayashi, 1985; May et al., 2019). To resolve this poorly understood phenomenon for multispecies assemblages, we develop an analytical null model for shared species in subdivided habitat. With such a null model it is possible to infer the direction and magnitude of any likely geometric effects of habitat fragmentation to quantify their probable effects on observed species diversity patterns.

Asking how many species a group of discrete patches share is an intuitive way to appreciate the spatial distribution of species and to compare their composition. It has long been recognized that key to understanding the effects of habitat subdivision on total species number is the degree of overlap in species composition between the patches—the proportion of species they share (Simberloff & Abele, 1976; Tjørve & Tjørve, 2008). The number of species shared between two habitat patches is central to many measures of pairwise dissimilarity, or beta diversity (reviewed in Koleff et al., 2003). Extending this idea, zeta (ζ) diversity predicts the expected number of species shared among any number of samples or sites, where the number of samples is referred to as the order of zeta (Hui & McGeoch, 2014). Zeta diversity has been used to show how shared species are a common currency from which multiple diversity patterns can be derived (Arita, 2017; Hui & McGeoch, 2014; Lu et al., 2019). Analogous to the partitioning of gamma diversity into alpha and beta diversity components, zeta diversity partitions diversity of groups of samples into multiple components from which other diversity measures can be expressed. For example, Jaccard's pairwise similarity index is $\zeta_2 / \left(2\zeta_1 - \zeta_2\right)$, where ζ_i is zeta diversity of order i (Hui & McGeoch, 2014). Thus, a null model for shared species, or zeta diversity (these terms are used interchangeably), allows the calculation of many diversity metrics useful for understanding the effects of subdivision (see Methods), including the expected number of species confined to only a single patch, of interest from the perspective of extinction risk in a metapopulation framework (Hanski, 1999).

The number of species shared by multiple patches can be predicted from the probability that each species occurs in those patches. These occurrence probabilities for a given area can be derived using methods based on spatial sampling theory, which combine species abundance distributions (SADs) and species spatial distributions to predict areal diversity patterns and scaling relationships (e.g., Arrhenius, 1921; Coleman, 1981; Harte et al., 2008; He & Legendre, 2002; Kobayashi, 1985; Morlon et al., 2008; Plotkin & Muller-Landau, 2002). Such methods can be used to explore probable diversity outcomes by varying the properties of these distributions (e.g., evenness in abundance or intraspecific aggregation), for example, to analyse species–area relationships (Harte et al., 2008; He & Legendre, 2002; Wilber et al., 2015) and pairwise compositional similarity (Morlon et al., 2008; Plotkin & Muller-Landau, 2002). Here we use spatial sampling theory to derive a null model for zeta diversity as a first step in predicting the consequences for diversity when a given habitat area is subdivided into several discrete patches. With such a null model, the predicted zeta diversity can be used to quantify a range of metrics (Hui & McGeoch, 2014) beneficial for understanding the geometric diversity effects of subdivision.

Our aims in this paper are therefore to (a) derive an analytical model to predict the expected number of shared species (zeta diversity) in subdivided areas for random and non-random distribution of species, (b) validate model predictions for shared species and other related metrics in subdivided habitat, and (c) use the models to demonstrate how shared species, and some resulting patterns of diversity, will change under different constraints, both ecological (e.g., different relative species-abundance distribution) and geometrical (e.g., patch area and number).

2 | METHODS

2.1 | Model derivation

2.1.1 | Shared species under random placement

Assuming complete knowledge of the abundance of every species in a study extent, A, and random distribution of a species within it, the occurrence probability of the species in an area a is (Arrhenius, 1921; He & Gaston, 2000):

$$P_i = 1 - \left(1 - \frac{a}{A}\right)^{N_i} \tag{1}$$

where N_i is the abundance of species i in the extent of the study area A.

In two samples of size a from the same community, the expected number of shared species is the sum of the product of the occurrence probability (Equation 1) from the two areas, a, across all species S. Among m disjoint samples of equal area a ($m < A/a$), under random placement of individuals (the probability that species i occurs in both sample k and l equals the product of occurrence probabilities in each sample, $P_{k\cap l,i} = P_{k,i} \cdot P_{l,i}$), the expected number of shared species (i.e., zeta diversity of order m, denoted $\zeta_{m|a}$) is simply:

$$\zeta_{m|a} = \sum_{i=1}^{S} P_{1\cap 2\cap \ldots \cap m,i} = \sum_{i=1}^{S} P_{1,i} \cdot P_{2,i} \cdot \ldots \cdot P_{m,i} = \sum \left[1-\left(1-\frac{a}{A}\right)^{N_i}\right]^{m} \tag{2}$$

where the expected number of species shared in m samples depends only on the sample area a and abundance vector N_i. Note when $m = 1$, Equation (2) as a function of sample area a is simply the random placement species–area model (Arrhenius, 1921; Coleman, 1981).

2.1.2 | An approximate expectation for shared species under non-random placement

Most populations of species are aggregated in space (Condit et al., 2000; He et al., 1997). In this case the joint probability of occurrence for a species is dependent on the distance between sampling locations (Diggle, 2013). This means the derivation adopted above for random placement of individuals is not a general solution (i.e., $P_{k\cap l,i} \neq P_{k,i} \cdot P_{l,i}$ generically). However, it is possible to use a similar approach to approximate the case where samples are adequately separated in space for their sampling probability to be independent of distance, or the area of interest (e.g., sampling grain) is large relative to the typical clustering pattern. That is, aggregation can be measured in different dimensions (Hui et al., 2010), where the number of individuals across samples is over-dispersed (variance > mean) but with samples randomly placed in space (i.e., $P_{k\cap l,i} \approx P_{k,i} \cdot P_{l,i}$).

To account for non-random (aggregated to regular) distributions of individuals among samples, models have been derived from the negative binomial (He & Legendre, 2002) and finite negative binomial distribution (Zillio & He, 2010). Notwithstanding the argument in the previous paragraph, under a negative binomial distribution of individuals among randomly placed samples ($P_{k\cap l,i} \approx P_{k,i} \cdot P_{l,i}$), the expected number of shared species in m samples of area a, or expressed as $\alpha = a/A$ (the proportion of total study extent in one sample) is:

$$\zeta_{m|\alpha} = \sum_{i=1}^{S} \left[1-(1-\alpha)\left(1+\frac{N_i\alpha}{k_i}\right)^{-k_i}\right]^{m} \tag{3}$$

where k_i is a parameter accounting for the spatial distribution of species i, with other notation following Equation (2). The term $(1 - \alpha)$ is required to ensure the number of species is zero when $a = 0$, and S when $a = A$, describing areal sampling without replacement (He & Legendre, 2002; Zillio & He, 2010).

Alternatively, based on the finite negative binomial distribution of individuals among randomly placed samples (Zillio & He, 2010), the equivalent model is:

$$\zeta_{m|\alpha} = \sum_{i=1}^{S} \left[1-\frac{\Gamma\left(N_i+k_i/\alpha-k_i\right)\Gamma\left(k_i/\alpha\right)}{\Gamma\left(N_i+k_i/\alpha\right)\Gamma\left(k_i/\alpha-k_i\right)}\right]^{m} \tag{4}$$

where $\Gamma(n) = (n - 1)!$ is the gamma function and other notation is as for Equation (3).

In both Equations (3) and (4), clustering parameter k_i adjusts the mean density in subarea a for species i, representing the effect of intraspecific spatial patterns. Aggregated, conspecific individuals have an attractive effect, meaning k_i takes positive values, while in regular distributions conspecifics are inhibitory, resulting in negative values of k_i (He & Gaston, 2000; Zillio & He, 2010). Parameter k_i can therefore be varied systematically to describe different conspecific spatial distributions, from aggregated to regular to random (e.g., He & Legendre, 2002; Wilber et al., 2015). As k approaches infinity (either negative or positive; He & Gaston, 2000), conspecific individuals follow random placement and Equations (3) and (4) converge to Equation (2).

In reality, every species follows a unique spatial pattern, that is, has a unique k_i value (Condit et al., 2000; He et al., 1997), and this value changes with area. If the value of k for every species at each scale of interest had to be estimated individually, it would result in an unfeasibly large number of parameters to fit Equations (3) and (4) (i.e., one for every species). Because the value of k_i changes with area, k_i needs to be estimated at each scale. Fortunately, there are two simplifications that make this unnecessary. First, aggregation tends to scale with area as a function of species abundance (see 'Scaling properties of the k parameter'). Second, a single community-level estimate of k is adequate to model emergent patterns of diversity (Plotkin & Muller-Landau, 2002). Here we achieve this using a scaling factor $1/c$, where $k_i = N_i\alpha/c$ and $N_i\alpha$ is the abundance of species i in subarea a (note, $\alpha = a/A$), and c is the estimated community level scaling parameter at this spatial scale. Substituting $c = N_i\alpha/k_i$ into Equations (3) and (4) we obtain the negative binomial (Equation 5) and finite negative binomial (Equation 6) shared species null models of individuals among randomly placed samples:

$$\zeta_{m|a} = \sum_{i=1}^{S} \left(1-(1-\alpha)(1+c)^{-\alpha N_i/c}\right)^{m} \tag{5}$$

and

$$\varsigma_{m|a} = \sum_{i=1}^{S} \left[1-\frac{\Gamma\left(N_i\left(1+\frac{1}{c}-\frac{\alpha}{c}\right)\right)\Gamma\left(\frac{N_i}{c}\right)}{\Gamma\left(N_i\left(1+\frac{1}{c}\right)\right)\Gamma\left(\frac{N_i}{c}(1-\alpha)\right)}\right]^{m} \tag{6}$$

where c is now a constant for a given α and other notation is the same as above. The use of a community-level scaling factor not only decreases complexity in the use of the model for prediction of species diversity at different scales, but also provides a mean to validate model predictions where it can be used as a model parameter to fit Equations (5) and (6) to observed data. As the use of Equations (5) and (6) should make little quantitative difference, the additional complexity of Equation (6) may not be warranted for numerical calculation and we validated both models for comparison (see Supporting Information Appendix S2). However, in the main text we choose to use Equation (6) because it is the exact model within a finite habitat extent (Zillio & He, 2010). Note that when $m = 1$, Equation (5) is the species–area model for the negative binomial distribution (He &

Legendre, 2002) and Equation (6) is the species–area model for the finite negative binomial distribution.

2.2 | Calculating diversity in subdivided habitat from shared species

Consider a formerly continuous landscape of area *A* that has been subjected to instantaneous habitat destruction, resulting in *m* randomly placed discrete remnant patches of subdivided habitat, each with area *a* (i.e., the total remaining habitat is $m \times a \le A$) and separated by a non-habitat matrix. We can use the random or non-random null models of individuals among randomly placed samples (Equations 2 or 6) to predict zeta diversity (orders 2 to *m*), from which the expected effects of subdivision on diversity for the *m* randomly placed patches can be calculated for any desired fraction of total area *A*. We use the zeta diversity partition to calculate three diversity metrics: (a) species accumulation across 1 to *m* patches, that is, species-accumulation curve; (b) pairwise Sørensen dissimilarity; and (c) the number of species found in only a single patch (single-patch endemics). We summarize these calculations below and refer interested readers to their original derivations (Hui & McGeoch, 2014; McGeoch et al., 2019).

The total number of species accumulated from 1 to *m* patches of size *a* is the species-accumulation curve. If samples are interpreted as remnant patches following land clearance, the number of species in *m* patches represents the total species diversity of the resulting subdivided landscape (i.e., gamma diversity). The difference in the number of expected species in a single, continuous patch of size $m \times a$ versus the total species diversity in *m* discrete patches each of area *a*, is essentially the difference in species number expected due to subdivision alone (as considered in the SLOSS debate). To explore this difference, we vary *m* from 2 to 32 samples, adjusting patch area *a* to maintain a constant total area. Species richness in individual patches will decrease from the shrinking patch area with increasing *m*, but the total habitat area remains constant. From the zeta diversity components (ζ_j), species accumulation is calculated recursively (Hui & McGeoch, 2014):

$$St_{m,a} = \sum_{j=1}^{m} (-1)^{j+1} \binom{m}{j} \zeta_j \qquad (7)$$

where $St_{m,a}$ = the total number of species contained in *m* patches of area *a* and ζ_j = the number of species shared in *j* patches ($j = 1, 2, \ldots, m$). Thus, zeta diversity can be used to calculate the number of species found in all *m* patches, or only a subset of these. Equation (7) offers a critical model for assessing the effect of landscape fragmentation on diversity. This is done by comparing Equation (7) and the species–area models ($m = 1$) defined by Equations (2) or (6) but replacing subarea *a* by $m \times a$. That is, the difference between $St_{m,a}$ and $\zeta_{1|ma}$ is attributable only to habitat subdivision as the total habitat area ($= m \times a$) is constant in the comparison; it thus defines the geometric effect of habitat subdivision.

Another important facet of diversity in subdivided habitat is the number of subpopulations each species maintains. For a constant patch size, extinction risk is a function of the number of occupied patches in a metapopulation (Hanski, 1999), species most at risk in a subdivided landscape are those that are endemic to a single patch (hereafter single-patch endemics or SPE). The number of SPE can be calculated from zeta diversity according to the formula (Hui & McGeoch, 2014): $St_{m,a} - St_{m-1,a}$, where $St_{m,a}$ and $St_{m-1,a}$ are calculated from Equation (7). Similarly, the number of species confined to a given number (*n*) of patches is $St_{m,a} - St_{m-n,a}$.

2.3 | Model validation

Our approach affords a spatially implicit expectation for the diversity of randomly placed patches of habitat arising from a formerly continuous landscape were all the habitat loss and subdivision to occur instantaneously, that is, the geometric fragmentation effects (May et al., 2019). The value of the null models is in establishing an expectation for the effects of subdivision (and this alone) on diversity for given abundances and spatial patterns of species. For the underlying non-random model to be valid we require inter-patch distances to be large enough to ensure the joint probability of occurrence of the species at any two locations is independent of their distance. Our simulations and empirical results have confirmed this assumption (Supporting Information Figures S3.1, S3.2, Appendix S3) and the non-random model (Equations 5 or 6) should provide a reasonable approximation for empirical species. This is not surprising for forest tree species, as Condit et al. (2000) showed that for stem-mapped forest plots distance decay in neighbourhood density mainly occurs at the fine scale, on the order of tens of metres. This suggests these data should provide a suitable test for the non-random model. Thus, we used both empirical data from stem-mapped plots and simulated forests with specific spatial patterns for model validation.

Empirical data were from the 50-ha Barro Colorado Island (BCI) stem-mapped forest plot (Condit et al., 2012), 2005 census, which recorded 211,845 live stems (with diameter at breast height ≥ 1 cm) distributed among 301 tree/shrub species. Stem-mapped forest data such as BCI are useful for testing spatial patterns of diversity, as they provide precise (*x*, *y*) coordinates and species identity for all individuals within a total plot area (typically 50 ha). It is also worth noting that BCI data show little change in shared species with distance, particularly with increasing zeta order (Supporting Information Figure S3.1). This is typical of landscapes of similar spatial extent, which offer useful data to test our models. For the following simulations, we retained the same total number of individuals and species as the empirical data but varied either the spatial distribution of individuals of each species (aggregated, random or regular) while retaining the empirical abundance distribution, or the species abundance distribution while retaining a constant spatial pattern (i.e., maintaining approximately the same level of aggregation as the empirical data).

In model validation we compared five intraspecific spatial patterns, those of the empirical data and four simulated forests with individuals positioned at random, at higher and lower strength of aggregation than BCI, and regularly. We report the results for the empirical and random plots in the main text and the remainder in Supporting Information Appendix S2. The locations of individuals in the random plot were simulated from a Poisson point process, while the aggregated and regular plots used Thomas and Strauss point pattern processes, respectively, simulated using R package 'spatstat' (Baddeley et al., 2015). The two sigma values used for the Thomas process (i.e., for aggregated plots), 35 and 50 m, respectively, represented slightly less and slightly more aggregated distributions than the empirical data (which had diversity patterns consistent with a sigma of approximately 40 m). In all simulations, the number of trees for each species was kept as the observed in BCI.

We also compared four different species abundance distributions, again using the empirical SAD and simulating three common SAD models with the same number of species and individuals as the empirical BCI data. Simulated SADs were based on the broken stick, log series and lognormal models, each simulated using R package 'sads' (Prado et al., 2017). For technical detail on simulations see Supporting Information Appendix S1.

To create data to validate the models, we sampled 20 randomly positioned, non-overlapping square quadrats. To test model performance when changing the size of subdivided area (or grain) and to model the scaling relationship for *k* (Equation (8), see below) we sampled at 10 different grains (quadrat size = 25, 100, 225, 400, 625, 900, 1,225, 1,600, 2,025 and 2,500 m^2). At each grain, sampling was repeated 200 times and the shared species and related diversity metrics (see 'Calculating diversity in subdivided habitat from shared species') were calculated for each sampling iteration. The mean and sampling variation (95% sampling intervals) were calculated for the 200 iterations and compared with model predictions.

2.3.1 | Model fitting and validation

The random placement model (Equation 2) has no free parameters, but the non-random model (Equation 6) required a single fitting step: estimating *c* at each sampling grain. All other models and diversity calculations using the zeta diversity partition were parameter free. We used numerical optimization (Brent's method) to find the value for *c* that minimized the absolute difference between mean observed richness in one quadrat for each grain and the value predicted by Equation (6) with $m = 1$ (i.e., the number of species in a subarea, *a*). R code to fit and apply the models for prediction is available at https://github.com/deaned01/SubdivHabitatDiversity/ and archived at https://doi.org/10.5281/zenodo.5767025.

The estimate for *k* was then used in the finite negative binomial model (Equation 6) to predict shared species ζ_m of *m* order at sampling grain *a* and the predicted shared species were used to calculate species accumulation curves, total species diversity, pairwise Sørensen dissimilarity and single-patch endemic species. Goodness-of-fit of these metrics was assessed using the coefficient of determination (R^2) and relative root-mean-square error (RMSE), defined as: $\mathrm{RMSE} = \sqrt{1/m\sum_{i=1}^{m}\left(x_i - o_i/o_i\right)^2}$, where: o_i = observed, x_i = predicted and m = the number of patches (from 5 to 20). Model fit was also illustrated graphically by showing predicted versus observed (mean ± 95% confidence limits) values. We present the validation results for Equations (2) and (6) and corresponding diversity metrics applied to the empirical and simulated data, respectively, in the main text and present all other validation results (e.g., for simulated SADs and spatial patterns) in Supporting Information Appendix S1.

2.3.2 | Scaling properties of the *k* parameter

For the non-random models, community-level parameter *k* scales according to a power law (Plotkin & Muller-Landau, 2002) and can be estimated at any scale once a base scale value is known (He & Hubbell, 2003). Quantifying this scaling relationship is critical to predicting shared species for different areas (and therefore levels of subdivision). Using scaling factor *c* as a fitting parameter allows us to test this relationship, which follows:

$$c_a = c_0\left(\frac{a}{a_o}\right)^z \tag{8}$$

where c_a is the value of the scaling parameter at the spatial grain *a*, c_0 is the value of the parameter quantified at base scale a_0 and *z* is a parameter describing the scaling relationship. We tested the fit of Equation (8) using the estimate for *c* at the 10 sampling grains identified during model validation (25–2,500 m^2), adopting 400 m^2 as the base scale (a_0) to estimate c_0. Choice of base scale is effectively arbitrary (He & Hubbell, 2003) but for this study extent, 400 m^2 represents a suitable trade-off between capturing aggregation at species level while avoiding saturated distributions of species (where the species occurs in every cell) at large spatial grains. Other base scales gave similar results.

2.4 | Model predictions for diversity in subdivided habitat

Having validated the model, we did two sets of simulations to explore how (a) shared species and the total species number in subdivided landscape were affected by changes in sampling grain, species abundance and spatial distributions, and (b) how subdivision alters observed diversity for a constant total habitat area. In simulations we retained the same constraints on area and species abundance as for validations.

2.4.1 | Predicting the expected change in the number of shared species for different sampling grain, SAD and spatial pattern

For shared species, we predicted the outcome for $m = 20$ using Equations (2) and (6) by varying (a) patch area using six sampling grains: 0.05, 0.1, 0.25, 0.5, 0.75 and 1 ha; (b) relative abundance under three models of contrasting evenness (broken stick, log-normal and log series, see Supporting Information Appendix S1); and (c) spatial distribution using aggregated, random and regular patterns at a given grain (0.25 ha) and for a log normal SAD. Parameter *c* was estimated at 0.25 ha from the non-random model (Equation 6) for the empirical and aggregated data as described in the previous section and a value of −0.05 was set to represent regular patterns. The sensitivity of shared species to these variations was illustrated using normalized zeta decline (normalizing shared species by *S*, the total number of species in the global extent rather than the total richness predicted across all samples) and gamma diversity (Equation 7).

2.4.2 | Predicting the expected effects of subdivision on diversity for constant total area and contrasting SAD and spatial pattern

Finally, we used the above scenarios of varying spatial and abundance distribution (see Model validation; Supporting Information Appendix S1) to explore the effects of increasing subdivision while retaining a constant total habitat area. The impacts of fragmentation for diversity are thought to become most pronounced when less than approximately 30% of habitat remains (Andrén, 1994; Hanski et al., 2013). We assumed 90% destruction of the habitat and divided the remaining area into 2, 4, 8, 16 and 32 equal sized patches retaining the remaining proportion of habitat at 10% (note retaining 20 and 30% of habitat produced qualitatively similar results). For each level of subdivision (represented by the number of patches into which the area was subdivided), we used the non-random shared species model with *c* scaled according to Equation (8) fit to the empirical or simulated data (Supporting Information Appendix S2) to calculate zeta diversity, from which we calculated total species richness, pairwise Sørensen dissimilarity and the proportion of single-patch endemics. We present total species richness in the subdivided habitat as proportional change from the richness predicted for a single contiguous patch. Thus, positive values indicate more species were present within subdivided habitat, negative values the reverse. All simulations and modelling were done in R 3.5.1 (R Core Team, 2014). The complete range of model validation and predictions presented in the Results and their underlying motivations are summarized in Table 1.

3 | RESULTS

3.1 | Shared species model validation

At 400-m^2 sampling grain, predicted shared species for the non-random and random models were all within sampling error in the observed mean for empirical and simulated random plots (Figure 1; see Supporting Information Figures S2.1–S2.6 in Appendix S2 for additional model validations under other non-random spatial patterns and abundance distributions). For the empirical plot, the non-random model (Equation 6) tended to overestimate (Figure 1a), particularly with increasing sampling grain (Supporting Information Figure S2.7, Appendix S2). Mean RMSE for shared species predictions ranged from 0.03 to 0.35 and was largest for empirical and regular spatial patterns, while R^2 exceeded .99 for all models and all spatial patterns (Supporting Information Table S2.1). For regular spatial patterns, models under-predicted shared species (Supporting Information Figure S2.1).

The estimated *c* parameter for the finite negative binomial model closely followed the scaling relationship of Equation (8) (Figure 2) for the BCI data, although for larger sampling areas the estimated value

TABLE 1 Summary of modelling stages undertaken and the underlying question/motivation for each. Analyses are divided into model validation and model application, with the relevant display item presenting the results in either the main text or Supporting Information Appendix S2

Topic	Question addressed	Result
Model validation		
1. Shared species	Do the models predict the number of shared species for simulated and empirical data?	Figure 1, Supporting Information Figures S2.1, S2.5, Table S2.1
2. Area-scaling in *c* parameter	Does Equation (8) describe area-scaling in the *c*-parameter of the non-random shared species model (Equation 5/6)?	Figure 2, Supporting Information Figures S2.8, S2.10
3. Diversity patterns	Do the predicted shared species provide reliable predictions of the expected diversity?	Figure 3, Supporting Information Figures S2.2–2.4
Model predictive application		
4. Shared species	How does the number of shared species change with changes in sampling grain, relative abundance and intraspecific spatial patterns?	Figure 4, Supporting Information Figure S2.9
5. Diversity patterns	What are the expected geometric effects of subdivision for diversity?	Figure 5, Supporting Information Figures S2.11, S2.12

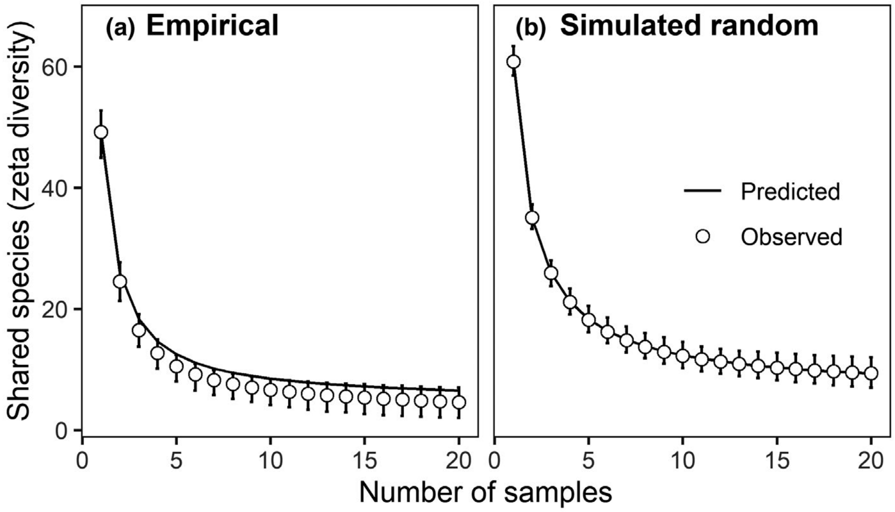

FIGURE 1 Predicted and observed shared species for twenty 400-m^2 non-overlapping quadrat samples randomly located in observed and simulated 50-ha forest plots. Both plots use the empirical Barro Colorado Island (BCI) species abundance distribution (SAD) but species vary in spatial distribution: (a) empirical BCI data, and (b) simulated random BCI plot. Each panel shows the mean and sampling error (95% confidence interval) in observed shared species from 200 repeat samples. Solid lines are model predictions for (a) using the best estimated c (=1.06) for Equation (6), and (b) parameter-free random placement model (Equation 2)

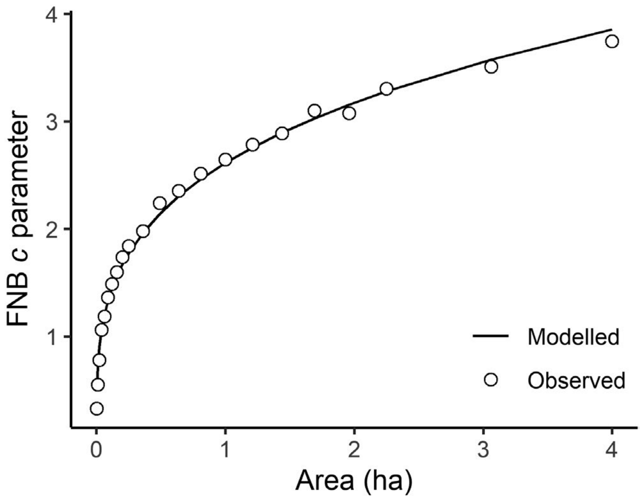

FIGURE 2 Scaling behaviour of the Finite negative binomial (FNB) *c* parameter as a function of sampling grain. Points show the estimated value of *c* in Equation (6) fit to mean diversity at the sampling grain indicated. The line is the fit of Equation (8) to a base scale of 400 m^2, with estimated $c_{400} = 1.06$ and the exponent value $z = 0.28$

was more prone to deviate from model predictions (Figure 2). The estimated model was $c_a = 1.06 \cdot (a/400)^{0.28}$, with $R^2 = .99$.

3.2 | Diversity prediction validation

Using estimated zeta diversity from the shared species null models (Equations 2 and 6), total species richness for multiple samples and single-patch endemic species were accurately predicted for all sampling grains (25–2,500 m^2; Supporting Information Figure S2.2) and spatial patterns (Figure 3, Supporting Information Figures S2.3–S2.6; all $R^2 > .99$; all RMSE < 0.07; Supporting Information Table S2.1). Species accumulation curves for empirical data were accurately predicted (Figure 3, Supporting Information Figures S2.2, S2.3 and S2.6), despite declining accuracy for the non-random shared species model (Equation 6) with increasing sampling grain (Supporting Information Figure S2.7). There was little difference in the predictive performance of the two non-random models (Equations 5 and 6) for any of the predictions, although where differences were observed, the finite negative binomial model was superior (Supporting Information Table S2.1).

3.3 | Changes in shared species for contrasting ecological and sampling geometry

The expected proportion of total plot diversity shared among samples (the decrease in zeta diversity with increasing numbers of samples) was a nonlinear decreasing function of sample size under any of the constraints analysed (Figure 4a–c). With decreasing sampling area, less even SAD, or as spatial pattern changes from regular through random toward aggregation, the proportion of shared species decreased (Figure 4a–c). The accumulation of species across all samples had the opposite trend to decline in shared species, increasing for greater decreases in zeta. For a given sampling area, the proportion of species accumulated was most sensitive to the evenness of the SAD (Figure 4d–f). For example, the exponent of a power law species–area curve (i.e., $s \propto area^z$) fitted to the data in Figure 4e changed by almost a factor of three between the broken stick ($z = 0.07$) and log series SAD ($z = 0.20$), while for a log normal SAD the different spatial patterns resulted in an increase of *c*. 50% ($z = 0.10$ to 0.15). In contrast, zeta retention was most strongly affected by area and spatial structure, with rare species rapidly lost among small samples and under increasing

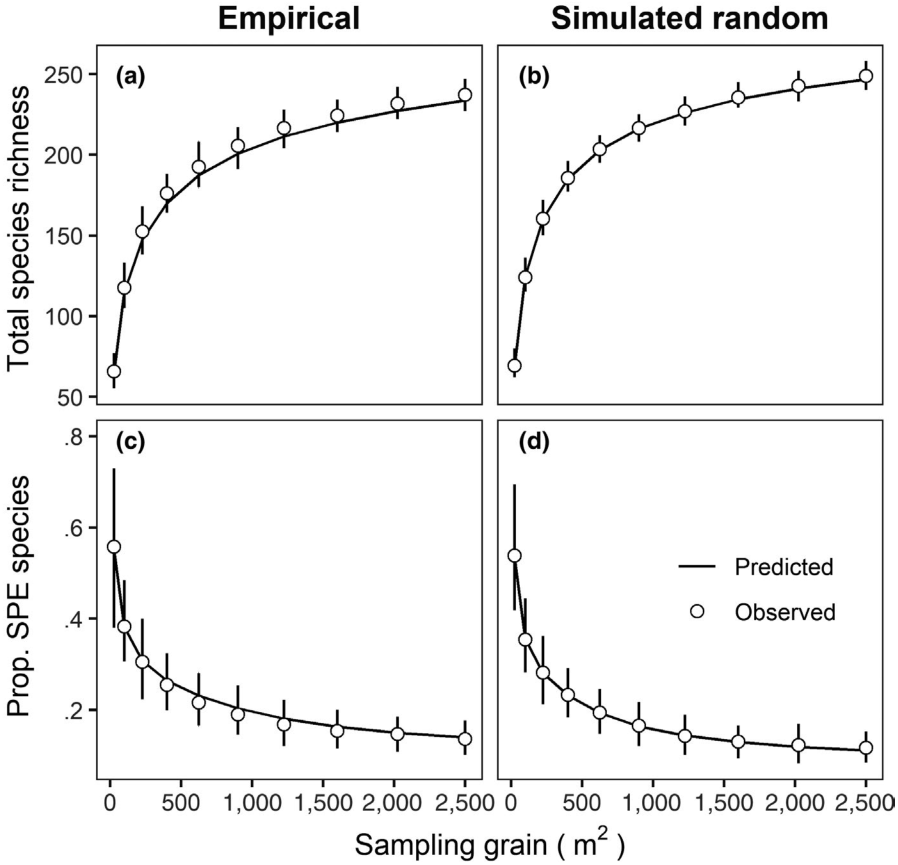

FIGURE 3 Observed and predicted total species richness (top row) and proportion of species found in only a single patch (bottom row) for 20 samples from empirical (left column) and simulated random (right column) 50-ha plots at 10 different sampling grains. (a) Total species in all samples for empirical Barro Colorado Island (BCI) data, (b) total species in all samples under random placement, (c) single-patch endemics (SPE) for BCI data (number of SPE ranged from 32–47 species), and (d) single-patch endemics under random placement (range: 29–45 species). Lines show the prediction for Equation (7), using shared species predictions from Equation (6) (a, c) and Equation (2) (b, d). Points show the mean (error bars = 95% confidence interval) observed species from 200 repeat samples from the BCI data

intraspecific aggregation (Supporting Information Figure S2.5). When individuals were randomly positioned (Supporting Information Figure S2.5) there was little difference between retention of species in the assemblage with increasing order of zeta.

3.4 | Effects of subdivision on diversity patterns

For aggregated assemblages, increasing subdivision of habitat area results in an increasing number of species being sampled relative to a single contiguous area, and more aggregated assemblages produce a greater proportional increase (Figure 5a). For a given level of aggregation, less even distribution of abundance results in a larger subdivision effect, increasing the number of species in subdivided relative to contiguous habitat (Figure 5b). Under random placement, the total species number does not change with subdivision, while for regular assemblages the total number of species sampled decreased with increasing subdivision (Figure 5a). For aggregated assemblages, the increase in species number slows down with subdivision and the most rapid changes occur as the area is broken up into a small number of patches (e.g., subdividing from 2–4 patches predicts 6 additional species, but between 16 and 32 fewer than 3; Figure 5a, aggregated scenario). For the simulated scenarios, the proportional increase in species between a single patch and subdivided habitat was modest, remaining below 10% even for the simulated highly aggregated assemblage.

Pairwise Sørensen dissimilarity had a qualitatively similar increase with subdivision but was more linear than changes in relative species number (Figure 5b). Pairwise Sørensen dissimilarity was also apparently more sensitive to changes in evenness of abundance than to changes in spatial pattern (Figure 5e). The proportion of species found in only a single patch decreased rapidly when total area was subdivided among few patches but differed little with increasing subdivision. For example, in the empirical data, doubling the number of patches from 2 to 4 reduced the predicted number of species in only one patch from 48 to 37, while doubling from 16 to 32 patches predicted 30 and 29 species, respectively. As with the other metrics, this proportion also increased with increasing aggregation (Figure 5c). However, the proportion of single-patch endemics was more sensitive to changes in relative abundance than to aggregation for a given number of subdivided areas (Figure 5f).

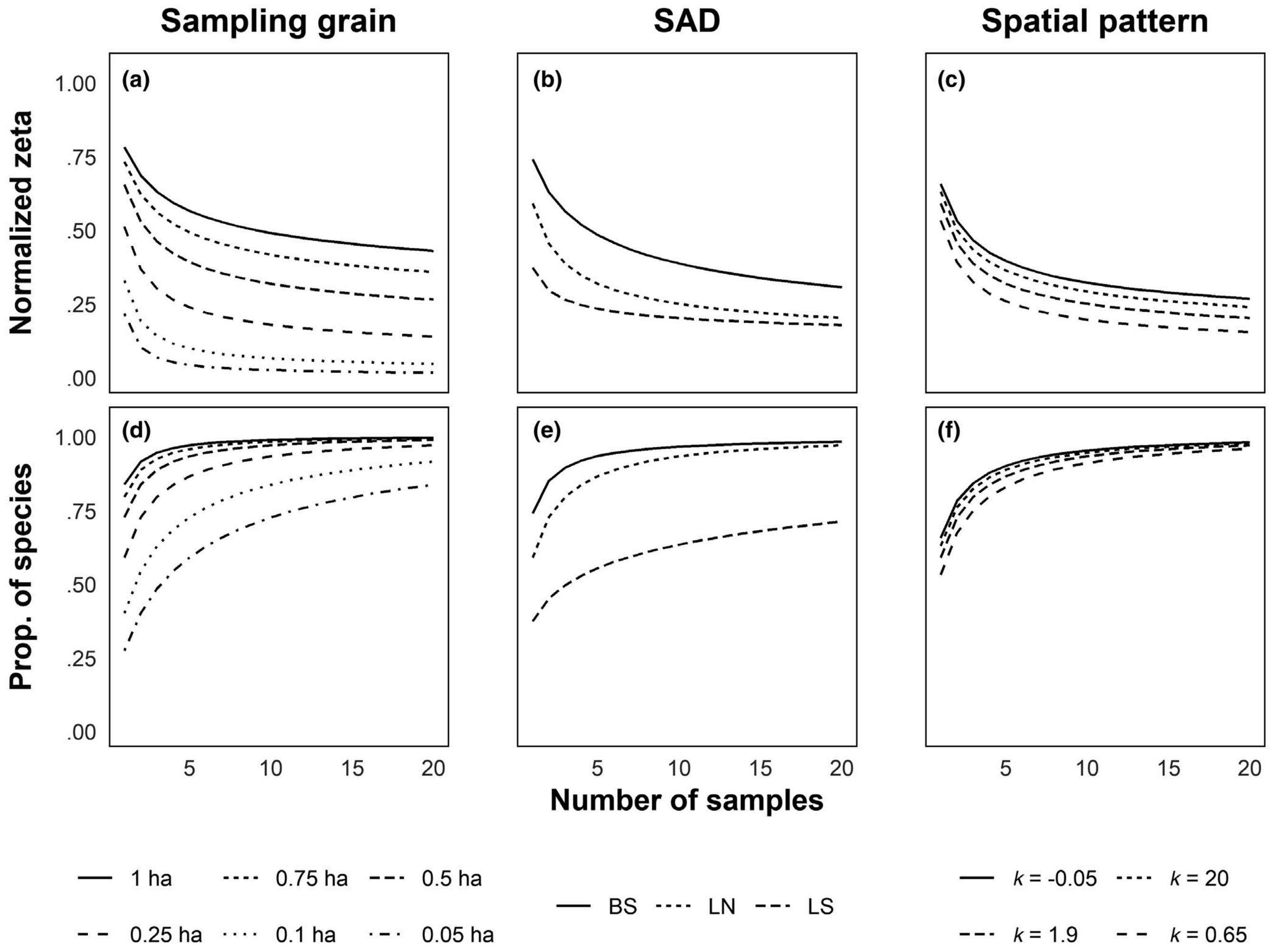

FIGURE 4 Changes in normalized zeta diversity decline (ratio of the number of shared species and total species number; top row) and the corresponding proportion of total species sampled (bottom row) of 20 samples, under varying: (a, d) sampling grain, (b, e) species abundance distributions (SADs), and (c, f) intraspecific spatial pattern. Panels (b, c) and (e, f) were calculated using a 0.25-ha sampling grain. Values on the *y* axis represent the proportion of species shared or accumulated from a species pool of 301. Panels (a), (c), (d) and (f) used the lognormal (LN) SAD. Other SAD used were: BS = broken stick; LS = log series). Parameter *k* in panels (c, f) simulated different intraspecific spatial patterns: −0.05 = regular, 20 = random, 1.9 = observed aggregation in Barro Colorado Island (BCI) at 0.25 ha, and 0.65 = aggregation predicted from the simulated landscape with the aggregation parameter sigma of the Thomas point process = 0.25

4 | DISCUSSION

4.1 | Intraspecific patterns determine the diversity of subdivided habitat

A higher total species richness in subdivided, than contiguous, habitat of equal total area is a widely recognized empirical pattern (reviewed in Fahrig, 2020) but claims of any positive diversity effects from subdivision remain controversial (e.g., Fahrig et al., 2019; Fletcher et al., 2018). The null models derived and validated here show the presence of more species in subdivided, than contiguous, habitat is an inevitable geometric effect of landscape fragmentation in case of aggregated spatial distributions. As such, this geometric expectation is an appropriate null hypothesis for testing the effect of landscape fragmentation on diversity (May et al., 2019). A similar idea appeared in Kobayashi (1985) almost 40 years ago but it remains widely unrecognized and its implications little explored (although see May et al., 2019). While this does not preclude the possibility of some positive effect of subdivision on species diversity in actual landscapes (e.g., Fahrig, 2017), the simplest explanation for greater richness of small patches in SLOSS-type patterns is that these arise from intraspecific aggregation with the strength of differences dependent upon species-abundance distribution patterns (Figure 5, Supporting Information Figures S2.3, S2.6).

Of course, the eventual post-fragmentation landscape diversity of formerly contiguous habitat will unfold over time and depend on many interacting factors relating to changes in area and isolation of patches (reviewed in Haddad et al., 2015). Most fragmentation modelling studies focusing on these long-term outcomes of biotic relaxation (where species' populations and demographic rates adjust to the new carrying capacity of a landscape following clearance) have shown well-established negative effects for diversity (e.g., Chase et al., 2020; Fletcher et al., 2018; Haddad et al., 2015) but few consider geometric effects (but see: Kobayashi, 1985; May et al., 2019).

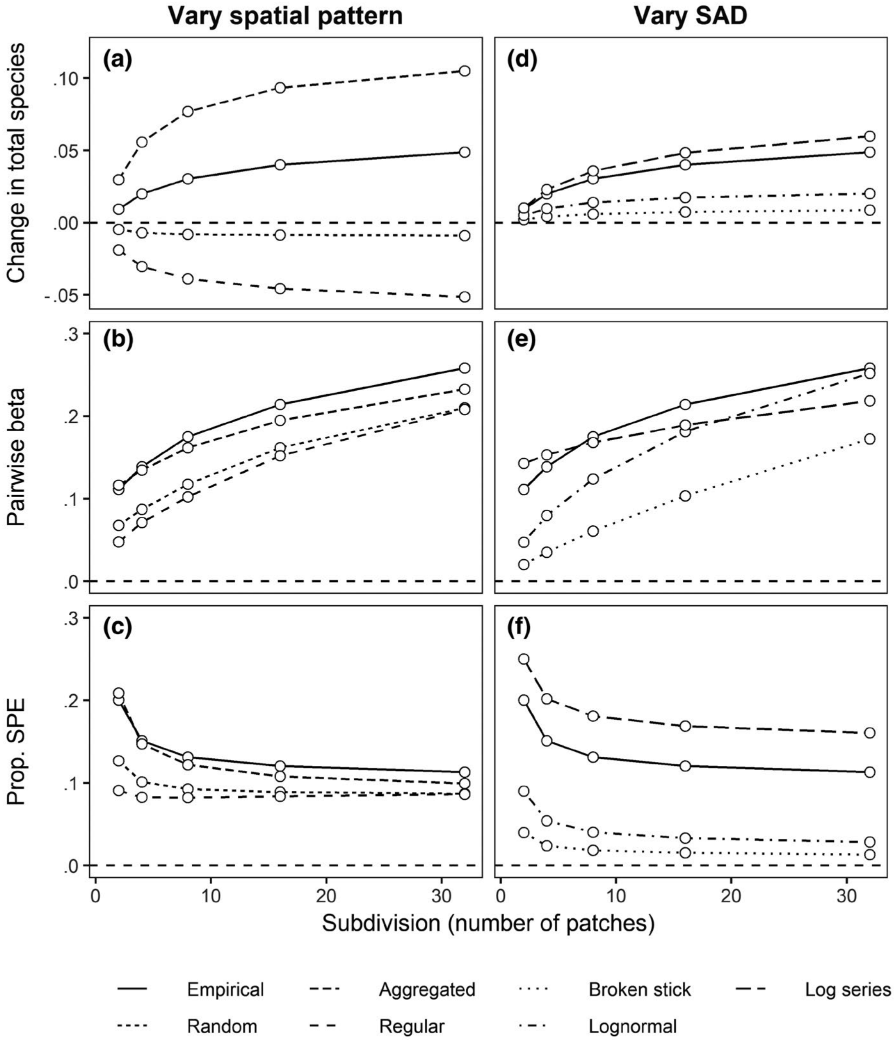

FIGURE 5 Predicted geometric effects of subdivision on diversity as a function of spatial pattern (left column) and evenness in abundance (right column). Diversity is compared for the total number of species relative to that predicted in a single contiguous area (top row), pairwise Sørensen dissimilarity (centre row) and the proportion of species found in only a single patch, that is, single-patch endemics (SPE; bottom row). Each point represents model predictions for the relative diversity metric as 10% of habitat area sampled from the original extent is increasingly subdivided ($m = 2, 4, 8, 16, 32$). Each simulation is based on the empirical data of Barro Colorado Island (BCI) and the curves corresponding to predictions for the observed data are shown in all panels

Although different fragmentation models predict wide variation in the long-term outcomes of fragmentation for diversity, they tend to show strong negative effects on total species number with increasing fragmentation, particularly as total habitat amount declines (Arnillas et al., 2017; Chisholm et al., 2018; Claudino et al., 2015; Hanski et al., 2013; Rybicki et al., 2020; Thompson et al., 2017). Both Hanski et al. (2013) and Rybicki et al. (2020) used simulations to show fragmentation matters little when a high proportion of habitat remains in the landscape but becomes important as total habitat area in the landscape falls below about 30%. However, even at such a small total habitat area, our models show the geometric effects of subdivision could have the reverse effect, resulting in positive fragmentation effects on landscape species richness, while potentially spreading the risk of landscape-level extinction. For example, by reducing the proportion of species in the landscape confined to a single patch (Figure 5), the risk of local extinction for some species could decrease, relative to a single large patch, depending on the amount of dispersal among patches and how local extinction risk changes with patch size (McCarthy et al., 2005).

Because of the static nature of our models, any predictions would more closely approach reality in fragmented landscapes for long-lived sessile organisms than for highly mobile taxa, and for fragmentation that occurs over a short period of time. While our method does not consider any long-term implications of biotic relaxation on diversity, it is conceivable these initial patterns could become stronger over time, because post-fragmentation processes would drive species to be more aggregated, a weaker form of biotic relaxation (He & Hubbell, 2013). This would likely further increase the relative number of species in groups of smaller patches, consistent with much empirical evidence (e.g., Deane et al., 2020; Fahrig, 2017, 2020; Quinn & Harrison, 1988; Simberloff & Gotelli, 1984). Similarly, our null models predict only total species number, not species composition, nor abundance. It is possible the greater richness in subdivided habitat might be limited to the most common species in the pre-clearance landscape. However, within our validation data, there was a difference in the mean abundance of species sampled in a single patch from those sampled in subdivided habitat, the latter tending to sample relatively rare species from the original landscape (Supporting Information Figure S2.11).

4.2 | The effect of spatial pattern and abundance distributions on diversity in subdivided habitat

The roles of species abundances and spatial distributions for scaling species richness and pairwise compositional similarity have been established from prior applications of spatial sampling theory (e.g., He & Legendre, 2002; Kobayashi, 1985; Morlon et al., 2008; Plotkin & Muller-Landau, 2002). Greater levels of intraspecific aggregation, or less even abundance distributions, both reduce species density in small areas. For a constant total number of species, this means fewer species are shared among pairs of samples (i.e., higher species turnover; Morlon et al., 2008; Plotkin & Muller-Landau, 2002) and species–area curves have steeper slopes (He & Legendre, 2002). Here we show the consequences of this for subdivided habitat, where smaller patches, less even abundance distributions or more aggregated assemblages share fewer species between subareas and therefore accumulate a greater fraction of total original species number than a single continuous area (Figures 4 and 5).

Unlike species richness, the proportion of single-patch endemics varies little with contrasting spatial distributions. This confirms the mathematical proof in He and Hubbell (2011), that endemics–area curves follow random placement, irrespective of spatial aggregation. In contrast, the evenness of abundance distributions has a strong influence on the proportion of single-patch endemic species, where more even distributions result in a smaller proportion of such species. This suggests species abundance distributions have implications for the relative impacts of immediate and long-term extinction risk (Kitzes & Harte, 2015), which our shared species models can be used to explore for subdivided habitat.

4.3 | The role of spatial scale

The value of the community-level clustering parameter in the non-random shared species model scales with sample area according to a power function. This not only simplifies the use of the model in simulation (Supporting Information Appendix S2) but also suggests a general prediction. As k increases, the shared species—and therefore any diversity metric calculated from this—will begin to resemble that for random placement even for aggregated assemblages. Typically, this behaviour occurs at a value of k exceeding *c*. 10 (Zillio & He, 2010). As we confirm here, subdivision makes no difference to our expectation for diversity if individuals are randomly positioned (He & Hubbell, 2011; Kobayashi, 1985; May et al., 2019). Thus, it is reasonable to expect that if all the patches under consideration were large enough that the corresponding k parameter for the original assemblage exceeded 10, then geometric effects of subdivision would not be evident. Extrapolating the area-scaling relationship for c shown in Figure 2, a minimum expected patch size to produce diversity patterns consistent with random placement would exceed 100 ha. Of 68 studies of fragmented landscapes analysed in Deane and He (2018), over 80% of patches were less than this total area confirming that geometric effects of fragmentation for diversity should not be overlooked (May et al., 2019). Interestingly, in a review of the fragmentation literature for birds and mammals, Andrén (1994) proposed random placement as a reasonable expectation for diversity of patches, particularly during the early stages of fragmentation when remnant patches are largest and geometric effects of subdivision would be most likely small.

4.4 | Model contribution

While a quantitative understanding of shared species in subdivided habitat is known to be essential for calculating total species diversity (Higgs & Usher, 1980; Simberloff & Abele, 1976), prior approaches have typically been derived on the basis of the species–area relationship (e.g., Harte & Kinzig, 1997; Higgs & Usher, 1980; Tjorve & Tjorve, 2008). Our analytical solution for shared species is derived from probability of species occurrence, a more basic quantity for describing species distribution, which is directly linked to species spatial pattern and abundance. As such, the diversity expectation for subdivided habitat depends only on spatial distribution and abundance of species. This expectation for the number of shared species unlocks the use of the zeta diversity partition (Hui & McGeoch, 2014) to calculate various diversity metrics in subdivided habitat. The estimation of single-patch endemics is one example and an important contribution of this study. This value can be predicted for any number of patches, offering useful context for understanding geometric effects on potential future extinction risk from a metapopulation perspective. New measures expressed in terms of zeta diversity are also emerging, along with methods to relate changes in zeta diversity to environmental gradients (Hui et al., 2018; Latombe et al., 2017, 2018; Lu et al., 2019; McGeoch et al., 2019). With our approach to estimating shared species, such measures become available for predicting the expected geometric effects of subdivision for species diversity.

For application of the non-random model it is important to be aware the approximate joint probabilities (i.e., $P_{k\cap l,i} \approx P_{k,i} \cdot P_{l,i}$ for species i in patches k and l) apply only for sampling grain and inter-sample distances at which abundances are independent, which will vary for each community. Model validations suggest they are reasonable for forest plots of tens to hundreds of hectares, an extent smaller than species' distributional ranges. This might not be a limitation when simulating the geometric effects of late-stage fragmentation, where mean nearest-neighbour interpatch distances increase (Andrén, 1994) and the assumption that patches would be adequately spaced to be independent is reasonable. Another assumption is that spatial distribution of species should follow Poisson or negative binomial distributions. Although those patterns are widely observed for empirical species (Condit et al., 2000; He et al., 1997; Zillio & He, 2010), departure from those patterns, for example, the presence of strong directional and spatially autocorrelated environmental gradients, or variation in spatial patterns among species, could contribute to the deviation between the expected and observed diversity. Future development of the models could consider how to include heterogeneous environments and explicit incorporation of spatial location of fragments to provide a more general model.

5 | CONCLUSION

Our shared species null models of individuals among randomly placed samples provide an analytical framework for exploring the geometric diversity effects of subdivision, and our resulting characterizations of these effects yield some important insights. First, they confirm that more species are expected for a given amount of subdivided habitat whenever intraspecific spatial patterns are predominately aggregated (Kobayashi, 1985; May et al., 2019). As a starting hypothesis, the observation of more species in fragmented habitat is perhaps best interpreted as an artefact of subdivision rather than a positive effect. Likely geometric effects of subdivision for diversity should be ruled out before any biological explanation is proposed, particularly for long-lived, sessile taxa. Second, we show that any additional 'benefits' from such geometric effects decrease in magnitude with increasing subdivision. This could be interpreted as support for a strategy of maintaining a few large habitat patches as the optimal compromise between maintaining large populations and spreading risk among multiple reserves (e.g., Soule & Simberloff, 1986). Finally, the models suggest that increasing subdivision also decreases the proportion of species found within a single patch. This raises some interesting questions regarding the rates of species loss that would be likely to occur from a single large fragment relative to a few smaller patches when only a small fraction of habitat remains.

ACKNOWLEDGMENTS

We thank C. Frost, S. Nielsen, J. Pither and C. St Clair for helpful comments on an early draft and V. Bahn and two anonymous referees for suggestions that improved the paper. This work was supported by a Discovery Grant of the Natural Science and Engineering Research Council of Canada. C.H. and F.H. acknowledge support from the National Research Foundation of South Africa (NRF) and the International Development Research Centre (IDRC, Canada) through the South Africa-Canada Research Chairs Mobility Initiative (ref no. CANM190215418471). C.H. is supported by the National Research Foundation of South Africa (grant 89967).

AUTHOR CONTRIBUTIONS

D.D., F.H. D.X. and C.H. conceived the methods. D.D. performed all analyses with contributions from all authors on the methods and presentation of results D.D. wrote the first draft and revisions with input from all authors.

DATA AVAILABILITY STATEMENT

The BCI stem-mapped data used in this study are in the public domain.

ORCID

David C. Deane https://orcid.org/0000-0003-2144-624X
Cang Hui https://orcid.org/0000-0002-3660-8160

REFERENCES

Andrén, H., & Andren, H. (1994). Effects of habitat fragmentation on birds and mammals in landscapes with different proportions of suitable habitat- a review. *Oikos*, *71*, 355–366. https://doi.org/10.2307/3545823

Arita, H. T. (2017). Multisite and multispecies measures of overlap, co-occurrence, and co-diversity. *Ecography*, *40*, 709–718. https://doi.org/10.1111/ecog.01942

Arnillas, C. A., Tovar, C., Cadotte, M. W., & Buytaert, W. (2017). From patches to richness: Assessing the potential impact of landscape transformation on biodiversity. *Ecosphere*, *8*, e02004. https://doi.org/10.1002/ecs2.2004

Arrhenius, O. (1921). Species and area. *The Journal of Ecology*, *9*, 95–99. https://doi.org/10.2307/2255763

Baddeley, A., Rubak, E., & Turner, R. (2015). *Spatial point patterns: Methodology and applications with R*. Chapman and Hall/CRC Press.

Chase, J. M., Blowes, S. A., Knight, T. M., Gerstner, K., & May, F. (2020). Ecosystem decay exacerbates biodiversity loss with habitat loss. *Nature*, *584*(7820), 238–243. https://doi.org/10.1038/s41586-020-2531-2

Chisholm, R. A., Lim, F., Yeoh, Y. S., Seah, W. W., Condit, R., & Rosindell, J. (2018). Species-area relationships and biodiversity loss in fragmented landscapes. *Ecology Letters*, *21*, 804–813. https://doi.org/10.1111/ele.12943

Claudino, E. S., Gomes, M. A. F., & Campos, P. R. A. (2015). Extinction debt and the role of static and dynamical fragmentation on biodiversity. *Ecological Complexity*, *21*, 150–155. https://doi.org/10.1016/j.ecocom.2014.11.011

Coleman, B. D. (1981). On random placement and species-area relations. *Mathematical Biosciences*, *54*, 191–215. https://doi.org/10.1016/0025-5564(81)90086-9

Condit, R., Ashton, P. S., Baker, P., Bunyavejchewin, S., Gunatilleke, S., Gunatilleke, N., Hubbell, S. P., Foster, R. B., Itoh, A., LaFrankie, J. V., Lee, H. S., Losos, E., Manokaran, N., Sukumar, R., & Yamakura, T. (2000). Spatial patterns in the distribution of tropical tree species. *Science*, *288*, 1414–1418. https://doi.org/10.1126/science.288.5470.1414

Condit, R., Lao, S., Pérez, R., Dolins, S. B., Foster, R. & Hubbell, S. (2012). Barro Colorado Forest census plot data (Version 2012). https://repository.si.edu/handle/10088/20925 or https://doi.org/10.5479/data.bci.20130603

Deane, D. C., & He, F. (2018). Loss of only the smallest patches will reduce species diversity in most discrete habitat networks. *Global Change Biology*, *24*, 5802–5814. https://doi.org/10.1111/gcb.14452

Deane, D. C., Nozohourmehrabad, P., Boyce, S. S. D., & He, F. L. (2020). Quantifying factors for understanding why several small patches host more species than a single large patch. *Biological Conservation*, *249*, 108711. https://doi.org/10.1016/j.biocon.2020.108711

Diamond, J. M. (1975). The island dilemma: Lessons of modern biogeographic studies for the design of natural reserves. *Biological Conservation*, *7*, 129–145. https://doi.org/10.1016/0006-3207(75)90052-X

Diggle, P. J. (2013). *Statistical analysis of spatial and spatio-temporal point patterns* (3rd ed.). CRC Press.

Fahrig, L. (2017). Ecological responses to habitat fragmentation per se. *Annual Review of Ecology, Evolution, and Systematics*, *48*, 1–23.

Fahrig, L. (2020). Why do several small patches hold more species than few large patches? *Global Ecology and Biogeography*, *29*, 615–628. https://doi.org/10.1111/geb.13059

Fahrig, L., Arroyo-Rodríguez, V., Bennett, J. R., Boucher-Lalonde, V., Cazetta, E., Currie, D. J., Eigenbrod, F., Ford, A. T., Harrison, S. P., Jaeger, J. A. G., Koper, N., Martin, A. E., Martin, J.-L., Metzger, J. P., Morrison, P., Rhodes, J. R., Saunders, D. A., Simberloff, D., Smith, A. C., … Watling, J. I. (2019). Is habitat fragmentation bad for biodiversity? *Biological Conservation*, *230*, 179–186. https://doi.org/10.1016/j.biocon.2018.12.026

Fletcher, R. J., Didham, R. K., Banks-Leite, C., Barlow, J., Ewers, R. M., Rosindell, J., Holt, R. D., Gonzalez, A., Pardini, R., Damschen, E. I., Melo, F. P. L., Ries, L., Prevedello, J. A., Tscharntke, T., Laurance, W. F., Lovejoy, T., & Haddad, N. M. (2018). Is habitat fragmentation good for biodiversity? *Biological Conservation*, *226*, 9–15. https://doi.org/10.1016/j.biocon.2018.07.022

Haddad, N. M., Brudvig, L. A., Clobert, J., Davies, K. F., Gonzalez, A., Holt, R. D., Lovejoy, T. E., Sexton, J. O., Austin, M. P., Collins, C. D., Cook, W. M., Damschen, E. I., Ewers, R. M., Foster, B. L., Jenkins, C. N., King, A. J., Laurance, W. F., Levey, D. J., Margules, C. R., … Townshend, J. R. (2015). Habitat fragmentation and its lasting impact on Earth's ecosystems. *Science Advances*, *1*, e1500052. https://doi.org/10.1126/sciadv.1500052

Hanski, I. (1999). *Metapopulation ecology*. Oxford University Press.

Hanski, I., Zurita, G. A., Isabel Bellocq, M., & Rybicki, J. (2013). Species-fragmented area relationship. *Proceedings of the National Academy of Sciences USA*, *110*, 12715–12720. https://doi.org/10.1073/pnas.1311491110

Harte, J., & Kinzig, A. P. (1997). On the implications of species-area relationships for endemism, spatial turnover, and food web patterns. *Oikos*, *80*, 417–427. https://doi.org/10.2307/3546614

Harte, J., Zillio, T., Conlisk, E., & Smith, A. B. (2008). Maximum entropy and the state-variable approach to macroecology. *Ecology*, *89*, 2700–2711. https://doi.org/10.1890/07-1369.1

He, F. L., & Gaston, K. J. (2000). Estimating species abundance from occurrence. *The American Naturalist*, *156*, 553–559. https://doi.org/10.1086/303403

He, F. L., & Hubbell, S. P. (2003). Percolation theory for the distribution and abundance of species. *Physical Review Letters*, *91*, 198103. https://doi.org/10.1103/PhysRevLett.91.198103

He, F., & Hubbell, S. P. (2011). Species-area relationships always overestimate extinction rates from habitat loss. *Nature*, *473*, 368–371. https://doi.org/10.1038/nature09985

He, F. L., & Hubbell, S. (2013). Estimating extinction from species-area relationships: Why the numbers do not add up. *Ecology*, *94*, 1905–1912. https://doi.org/10.1890/12-1795.1

He, F. L., & Legendre, P. (2002). Species diversity patterns derived from species-area models. *Ecology*, *83*, 1185–1198.

He, F., Legendre, P., & LaFrankie, J. V. (1997). Distribution patterns of tree species in a Malaysian tropical rain forest. *Journal of Vegetation Science*, *8*, 105–114. https://doi.org/10.2307/3237248

Higgs, A. J., & Usher, M. B. (1980). Should nature reserves be large or small. *Nature*, *285*, 568–569. https://doi.org/10.1038/285568a0

Hui, C., & McGeoch, M. A. (2014). Zeta diversity as a concept and metric that unifies incidence-based biodiversity patterns. *The American Naturalist*, *184*, 684–694. https://doi.org/10.1086/678125

Hui, C., Veldtman, R., & McGeoch, M. A. (2010). Measures, perceptions and scaling patterns of aggregated species distributions. *Ecography*, *33*, 95–102. https://doi.org/10.1111/j.1600-0587.2009.05997.x

Hui, C., Vermeulen, W., & Durrheim, G. (2018). Quantifying multiple-site compositional turnover in an Afrotemperate forest, using zeta diversity. *Forest Ecosystems*, *5*, 15.

Kitzes, J., & Harte, J. (2015). Predicting extinction debt from community patterns. *Ecology*, *96*, 2127–2136. https://doi.org/10.1890/14-1594.1

Kobayashi, S. (1985). Species diversity preserved in different numbers of nature reserves of the same total area. *Researches on Population Ecology*, *27*, 137–143. https://doi.org/10.1007/BF02515486

Koleff, P., Gaston, K. J., & Lennon, J. J. (2003). Measuring beta diversity for presence-absence data. *Journal of Animal Ecology*, *72*, 367–382. https://doi.org/10.1046/j.1365-2656.2003.00710.x

Latombe, G., Hui, C., & McGeoch, M. A. (2017). Multi-site generalised dissimilarity modelling: Using zeta diversity to differentiate drivers of turnover in rare and widespread species. *Methods in Ecology and Evolution*, *8*, 431–442. https://doi.org/10.1111/2041-210X.12756

Latombe, G., Richardson, D. M., Pysek, P., Kucera, T., & Hui, C. (2018). Drivers of species turnover vary with species commonness for native and alien plants with different residence times. *Ecology*, *99*, 2763–2775. https://doi.org/10.1002/ecy.2528

Lu, M. Y., Vasseur, D., & Jetz, W. (2019). Beta diversity patterns derived from island biogeography theory. *The American Naturalist*, *194*, E52–E65. https://doi.org/10.1086/704181

May, F., Rosenbaum, B., Schurr, F. M., & Chase, J. M. (2019). The geometry of habitat fragmentation: Effects of species distribution patterns on extinction risk due to habitat conversion. *Ecology and Evolution*, *9*, 2775–2790. https://doi.org/10.1002/ece3.4951

McCarthy, M. A., Thompson, C. J., & Possingham, H. P. (2005). Theory for designing nature reserves for single species. *American Naturalist*, *165*, 250–257. https://doi.org/10.1086/427297

McGeoch, M. A., Latombe, G., Andrew, N. R., Nakagawa, S., Nipperess, D. A., Roige, M., Marzinelli, E. M., Campbell, A. H., Verges, A., Thomas, T., Steinberg, P. D., Selwood, K. E., Henriksen, M. V., & Hui, C. (2019). Measuring continuous compositional change using decline and decay in zeta diversity. *Ecology*, *100*, e02832. https://doi.org/10.1002/ecy.2832

Morlon, H., Chuyong, G., Condit, R., Hubbell, S., Kenfack, D., Thomas, D., Valencia, R., & Green, J. L. (2008). A general framework for the distance-decay of similarity in ecological communities. *Ecology Letters*, *11*, 904–917. https://doi.org/10.1111/j.1461-0248.2008.01202.x

Plotkin, J. B., & Muller-Landau, H. C. (2002). Sampling the species composition of a landscape. *Ecology*, *83*, 3344–3356.

Prado, P. I., Miranda, M. D., & Chalom, A. (2017). sads: Maximum likelihood models for species abundance distributions. R package version 0.4.0. https://CRAN.R-project.org/package=sads

Quinn, J. F., & Harrison, S. P. (1988). Effects of habitat fragmentation and isolation on species richness - evidence from biogeographic patterns. *Oecologia*, *75*, 132–140. https://doi.org/10.1007/BF00378826

R Core Team. (2014). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. http://www.R-project.org/

Rybicki, J., Abrego, N., & Ovaskainen, O. (2020). Habitat fragmentation and species diversity in competitive communities. *Ecology Letters*, *23*, 506–517. https://doi.org/10.1111/ele.13450

Simberloff, D. S., & Abele, L. G. (1976). Island biogeography theory and conservation practice. *Science*, *191*, 285–286. https://doi.org/10.1126/science.191.4224.285

Simberloff, D., & Gotelli, N. (1984). Effects of insularization on plant-species richness in the prairie-forest ecotone. *Biological Conservation*, *29*, 27–46.

Soule, M. E., & Simberloff, D. (1986). What do genetics and ecology tell us about the design of nature-reserves. *Biological Conservation*, *35*, 19–40. https://doi.org/10.1016/0006-3207(86)90025-X

Thompson, P. L., Rayfield, B., & Gonzalez, A. (2017). Loss of habitat and connectivity erodes species diversity, ecosystem functioning, and stability in metacommunity networks. *Ecography*, *40*, 98–108. https://doi.org/10.1111/ecog.02558

Tjorve, E., & Tjorve, K. M. C. (2008). The species-area relationship, self-similarity, and the true meaning of the z-value. *Ecology*, *89*, 3528–3533.

Wilber, M. Q., Kitzes, J., & Harte, J. (2015). Scale collapse and the emergence of the power law species-area relationship. *Global Ecology and Biogeography*, *24*, 883–895. https://doi.org/10.1111/geb.12309

Zillio, T., & He, F. L. (2010). Modeling spatial aggregation of finite populations. *Ecology*, *91*, 3698–3706. https://doi.org/10.1890/09-2233.1

BIOSKETCH

This manuscript arose as a collaboration of scientists with a shared interest in the practical application of macroecological methods to inform conservation and improve ecological understanding.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

How to cite this article: Deane, D. C., Xing D., Hui C., McGeoch M., & He F. (2021). A null model for quantifying the geometric effect of habitat subdivision on species diversity. *Global Ecology and Biogeography*, 00, 1–14. https://doi.org/10.1111/geb.13437