Oikos 000: 001–011, 2017
doi: 10.1111/oik.04620

Subject Editor: Werner Ulrich. Editor-in-Chief: Dries Bonte, Accepted 24 September 2017

Checkerboard score–area relationships reveal spatial scales of plant community structure

Gordon G. McNickle, Eric G. Lamb, Mike Lavender, James F. Cahill Jr., Brandon S. Schamp, Steven D. Siciliano, Richard Condit, Stephen P. Hubbell and Jennifer L. Baltzer

G. G. McNickle, Dept of Botany and Plant Pathology and Center for Plant Biology, Purdue Univ., West Lafayette, IN, USA. – E. G. Lamb (http://orcid.org/0000-0001-5201-4541) (eric.lamb@usask.ca), Dept of Plant Sciences, Univ. of Saskatchewan, Saskatoon, SK, Canada. – M. Lavender, Dept of Biological Sciences, Queen's Univ., Kingston, ON, Canada. – J. F. Cahill Jr., Dept of Biological Sciences, Univ. of Alberta, Edmonton, AB, Canada. – B. S. Schamp, Algoma Univ., Marie, ON, Canada. – S. D. Siciliano, Dept of Soil Science, Univ. of Saskatchewan, Saskatoon, SK, Canada. – R. Condit, Field Museum of Natural History, Chicago, IL, USA. – S. P. Hubbell, Dept of Ecology and Evolutionary Biology, Univ. of California Los Angeles, Los Angeles, CA, USA. – J. L. Baltzer, Dept of Biology, Wilfrid Laurier Univ., Waterloo, ON, Canada.

Identifying the spatial scale at which particular mechanisms influence plant community assembly is crucial to understanding the mechanisms structuring communities. It has long been recognized that many elements of community structure are sensitive to area; however the majority of studies examining patterns of community structure use a single relatively small sampling area. As different assembly mechanisms likely cause patterns at different scales we investigate how plant species co-occurrence patterns change with sampling unit scale. We use the checkerboard score as an index of species segregation, and examine species C-score1–sampling area patterns in two ways. First, we show via numerical simulation that the C-score–area relationship is necessarily hump shaped with respect to sample plot area. Second we examine empirical C-score–area relationships in arctic tundra, grassland, boreal forest and tropical forest communities. The minimum sampling scale where species co-occurrence patterns were significantly different from the null model expectation was at 0.1 m^2 in the tundra, 0.2 m^2 in grassland, and 0.2 ha in both the boreal and tropical forests. Species were most segregated in their co-occurrence (maximum C-score) at 0.3 m^2 in the tundra (0.54 × 0.54 m quadrats), 1.5 m^2 in the grassland (1.2 × 1.2 m quadrats), 0.26 ha in the tropical forest (71 × 71 m quadrats), and a maximum was not reached at the largest sampling scale of 1.4 ha in the boreal forest. The most important finding is that the dominant scales of community structure in these systems are large relative to plant body size, and hence we infer that the dominant mechanisms structuring these communities must be at similarly large scales. This provides a method for identifying the spatial scales at which communities are maximally structured; ecologists can use this information to develop hypotheses and experiments to test scale-specific mechanisms that structure communities.

Studies of community diversity and species co-occurrence patterns are frequently used to make inferences about the mechanisms structuring those communities (Gotelli and McCabe 2002, Götzenberger et al. 2012). While it is well known that estimates of diversity and other elements of community structure are sensitive to sample size and sample plot area (Gotelli and Colwell 2001, Chase and Knight 2013), experience and practical constraints mean that the majority of studies of species co-occurrence patterns use relatively standard sampling units (e.g. 0.25–1 m^2 in herbaceous communities or 20–100 m^2 in forests) (Kent and Coker 1992). Co-occurrence patterns can be sensitive to plot size (Reitalu et al. 2008, Maestre et al. 2009, Long et al. 2015), however no study has systematically examined co-occurrence scale relationships within communities. This sensitivity means that different studies which might use different sized sampling units, could find different patterns of species co-occurrence for somewhat trivial reasons. Here, we are interested in exploring how a commonly used index of species segregation – the checkerboard score (C-score; Stone and Roberts 1990) – varies across spatial scale. We will argue that there is a C-score–area relationship that follows from the species–area relationship, and that this relationship can potentially inform studies of diversity and community ecology.

Segregation of species may occur for many reasons (e.g. dispersal limitation, habitat filtering, competitive exclusion) (Diamond 1975, Gilpin and Diamond 1982, Drake 1990, Weiher and Keddy 2001, Blois et al. 2014). Importantly, significant negative co-occurrence does not imply a particular mechanism is operating, but simply quantifies the degree of species segregation in a set of data. A comprehensive exploration of how plant species co-occurrence patterns change with the scale of the sampling unit may improve our ability to distinguish between possible mechanisms underlying the structure of plant communities. By identifying the spatial scales where segregation of species is either high or low we can identify the spatial scales at which to conduct experiments designed to investigate the mechanisms that shape negative species co-occurrence.

Species co-occurrence data take the form of matrices, where the data are presence or absence, the rows represent species occurrences, and the columns represent sample site occurrences. For somewhat trivial reasons, sample unit scale, and body size are tightly linked because the species–area relationship affects the number of species found in a sample plot, while the size of the plants in the communities (e.g. trees versus herbs) influences the number of individuals that can occur inside a plot of a given area. For example, measurements taken at a spatial scale equivalent to the size of a single organism will necessarily have very low alpha diversity, and therefore will have few species co-occurrences within a single quadrat, resulting in a community (quadrat by species) matrix containing mostly absences. At the other end of the spectrum, samples taken at very large spatial scales relative to the size of an individual will contain most of the species in the community (regional species pool) in each quadrat, and thus the community matrix will contain mostly presences. Thus, at small and large spatial scales low co-occurrence levels will be common because species appear to either never co-occur (small scale estimates) or because species appear to always co-occur (large scale estimates) (Ulrich et al. 2017). This control that sample scale exerts on the degree of matrix fill is a somewhat trivial consequence of matrix algebra, however as we will argue, when explored in an ecological context, we believe this can have non-trivial implications for our understanding of community structure.

Indeed, it has been shown that the degree of matrix fill (i.e. the ratio of presences and absences), and the raw checkerboard score (C-score) produces a hump-shaped relationship (Ulrich and Gotelli 2013, Ulrich et al. 2017). If the species–area relationship within a community controls the degree of matrix fill in any observed community data set, then this hump-shaped relationship between C-score and area might help guide the study of ecological communities by identifying the scales at which there is the largest signal of species segregation. Here we argue that the point(s) of maximum negative co-occurrence can provide guidance about the scale(s) at which communities are maximally structured and the optimal spatial scales at which to investigate mechanisms structuring those communities.

Our objectives in this paper are twofold. We first use simulations to: 1) demonstrate that the hump-shaped matrix fill-C-score relationship previously described by (Ulrich and Gotelli 2013) can be driven by the species richness area relationship which produces a co-occurrence–area relationship independent from any other assumptions about community structure and; 2) explore how the C-score–area relationship varies as a function of the shape of the species–area curve. Second, we explore the observed C-score–area relationships in four different plant communities differing radically in plant body size and presumed dominant environmental drivers: arctic tundra, boreal forest, grassland and tropical forest. In each of these four case studies, we identify the spatial scales where there is significant species segregation. Answers to this question do not provide direct information about the mechanisms that structure these communities, but do provide important information about the scales of the dominant mechanism(s), and may provide guidance for the design of future experiments to test hypothesized mechanisms. We conclude by discussing some of the implications of the results obtained from the analysis of the four case studies.

Methods

Negative co-occurrence scores and species area curves

We used the checkerboard score (C-score) to quantify species segregation. The C-score is based on the number of species co-occurrences at shared sites along pairwise sub-matrices extracted from within the community matrix. Thus, the total number of checkerboard units (C_{ij}) for each species pair ij, was calculated as:

$$C_{ij} = \left(r_i - X_{ij}\right)\left(r_j - X_{ij}\right) \qquad (1a)$$

where r_i is the number of times species i occurs without species j, r_j is the number of times species j occurs without species i and X_{ij} is the number of sites where species i and j co-occur (Stone and Roberts 1990). However, C_{ij} described in Eq. 1a is well known to be sensitive to the dimensions of the matrix, altered by both the number of species (i.e. the number of rows) and the number of sites that are sampled (i.e. the number of columns). To normalize the score based on the number of rows, Stone and Roberts (1990) rescaled C_{ij} to an average calculated across all possible species pairs (P) for a given level of gamma diversity, S. For S species, P is given by:

$$P = S\left(S-1\right)/2 \qquad (1b)$$

Ulrich and Gotelli (2013) refined this normalization to eliminate bias caused by the number of sites sampled, n, by rescaling according to the number of possible site pairs, N:

$$N = (n(n-1)/2) \qquad (1c)$$

where n is the total number of replicate plots. Thus, at the community level, the normalized C-score that we used throughout our analyses is given by:

$$C_n = \sum\nolimits_{j=0}^{S} \sum\nolimits_{i<j} \frac{C_{ij}}{PN} \qquad (1d)$$

We use Eq. 1d to calculate a normalized C_n-score in every case in our subsequent analyses making our estimates independent of sample size and community matrix size, thus allowing us to focus exclusively on how negative co-occurrence changes with sample plot scale. We refer to it as the C_n-score through the remainder of this manuscript to emphasize that this metric is subtly different from the C-score of Stone and Roberts (1990), and from other formulations of the C-score.

Simulated C_n-score–area relationships

We first examined the C_n-score–area relationship using a numerical simulation where we generated random presence–absence community matrices at a variety of sampling scales and then calculated C_n-scores for those matrices. Spatial scale ranged from 0.1 to 100 (this is a numerical simulation so the scale here is unitless). Sample species richness was defined using the species area curve (Chase and Knight 2013):

$$S = BA^z \tag{2}$$

where S is the cumulative number of species detected in area A. The parameters B and z are constants that scale the minimum number of species and the rate of accumulation respectively.

For each spatial scale from 0.1 to 100, we generated an occurrence matrix with 50 sample plots (i.e. 50 matrix rows), and S_{max} columns defined by the maximum spatial scale and the species area curve (e.g. $S_{max} = B \times 100^z$). This simulates a series of alternative experiments where sampling effort was held constant, but the scale of the sampling unit used in each experiment differs. Each point on the curve represents one of these experiments that differs only in the scale of the sampling unit, which allows us to generate a C_n-score–area relationship that is related to the species–area relationship. Thus, the number of species occurrences per plot (row sum) was determined by entering plot size into Eq. 2; those occurrences were randomly distributed across the S_{max} columns in the row. Normalized C_n-scores (Eq. 1d) were calculated for the random occurrence matrix at each scale, allowing us to demonstrate the expected C_n-score and sampling scale relationship. We explored how the C_n-score–area relationship was influenced by the shape of the species–area curve by running these simulations for a range of species–area parameter values $B = (10, 20, 30, 50)$ with z held constant at $z = 0.25$, and $z = (0.1, 0.2, 0.3, 0.4)$ with B held constant at $B = 25$.

Empirical case studies

To examine empirical area–C_n-score relationships, we examined four case studies selected to represent a potentially wide range of spatial scales of species interactions based on the natural history of the resident species, overall diversity and environmental conditions. We examined a low shrub community (Arctic tundra), a perennial herbaceous community (grassland) with individual plant bodies at a scale of centimetres to metres, and two forest communities (boreal and tropical forest) with plant bodies at a scale of tens of meters or more. Each dataset is a map of stem locations, allowing us to repeatedly sample the community at a variety of sample unit scales via virtual quadrats. In each case, we placed 50 square virtual quadrats of some constant size A within the map by randomly selecting a plot corner coordinate without torus wrapping and extracted the presence–absence community matrix (species as columns and rows) with replacement. We then increased the quadrat size, A, and sampled another 50 quadrats to simulate the same experiment but with a different choice of quadrat size. We repeated this process at many quadrat areas until we ran out of space in the stem map. We compared the observed community matrices to two distinct null models. First, we used a fixed–fixed null model where row and column sums were held constant, but the pattern of co-occurrence was randomized using the permatswap function in the vegan library in R (Hardy 2008, Oksanen et al. 2016). This model is robust against type I errors (Lavender et al. 2016), but is also very conservative because it is designed for studies where the sample unit might vary in size (e.g. comparisons among islands; Connor and Simberloff 1979). In order to confirm that our results were robust against choice of null model, we also used a fixed-equiprobable model where only row sums (i.e. species occurrences) are held constant, but the column sums are allowed to vary with equiprobability. This method is less conservative, but more commonly applied to studies where the sample unit is always the same fixed size (Gotelli 2000). Briefly, both methods randomize matrices by shuffling 2×2 submatrices within the observed matrix 30 000 times (Lehsten and Harmand 2006) while maintaining row and/or column totals as described above. As the results of the fixed–fixed and fixed-equiprobable models were equivalent (Supplementary material Appendix 1 Fig. A1), we only present full results for the fixed–fixed models. For each observed matrix 1000 null matrices were sampled with 30 000 swaps between each sample and with a burn-in of 30 000 swaps prior to the first sample. We used the quasiswap method for generating null matrices, which does not produce sequential null matrices, but instead generates a matrix at each time step that is fully independent of previous matrices (Miklós and Podani 2004). The standardized effect size of the C_n-score (z) can be calculated based on the observed C_n-score (C_n) and the mean (μ) and standard deviation (σ) of the expected C_n-scores produced from the 1000 null matrices (Ulrich and Gotelli 2007, 2010):

$$z = \frac{C_n - \mu}{\sigma} \tag{3}$$

The z statistic is significant at the two-tailed $p = 0.05$ level for $-1.96 < z > 1.96$. We then repeat the process at the next spatial scale to generate a continuous C_n-score–area and effect-size–area relationship. A significant C_n-score effect size indicates that species are more or less segregated than would be expected by chance regardless of the magnitude of the C_n-score.

From Eq. 3, we expect the relationship between C_n and area, as well as μ and area to be hump-shaped for all empirical case studies. However, we have no a priori expectations about the relationship between σ and area, which likely depends on habitat heterogeneity across ecological scales. Thus, there we also have no a priori expectations about the relationship between z and area, which very likely can take on almost any shape. Therefore, for each case study we show the relationship between observed and expected C_n-score and area, and also the relationship between effect size and area. However, while we show the raw C_n-score, we caution readers that in the empirical case studies effect size is really the only meaningful way of interpreting the significance of the patterns in species segregation. We included plots of the observed and expected raw C_n-scores for several reasons. First, based on a priori understanding of matrix fill, we expect that the underlying patterns of significance would be produced from the hump-shaped relationship between both observed and expected C_n-score versus area relationships, and that it was important to show that this was present in the empirical data sets. Second, effect size is undefined in our null community simulations both because the mean of a null minus the mean of a null is zero, and because it is inappropriate to apply frequentist statistics to simulation results

(White et al. 2014). Thus, while the raw C-score is not useful for empirical understanding, it does provide a point of comparison between the empirical studies and the simulation results that is not possible with the effect size.

For each case study, species area curves were generated from the stem maps via custom code that randomly placed 50 quadrats at each scale sampled (see site descriptions below), and extracting the cumulative species richness versus area data. Equation 2 was then fit to these subsampled data using the nlm library in the base package of R. This method generated the precise species accumulation curves observed within the stem maps used to calculate the C_n-scores. This is equivalent to existing software methods for generating species-area curves, but works with the stem map data that we had for each community. Since R^2 is undefined for non-linear regression, we show the fit of the model as the standard error of each non-linear regression which represents the average distance between the data and the regression line in the units of the y-axis.

Case study 1: Arctic tundra

The High-Arctic tundra site was on the Truelove Lowland (75°40′N; 84°35′W), a 43 km² polar oasis on Devon Island, Nunavut. Sample plots were located on dry beach ridges and had plant communities dominated by low growing perennials such as *Salix arctica, Dryas integrifolia* and *Carex* spp. (Bliss et al. 1994, Lamb et al. 2016). Banerjee et al. (2011b) provides a full description of edaphic conditions at the site.

Five rectangular plots (90 × 200–220 cm) were surveyed in July 2008. Each plot was divided into contiguous 10 × 10 cm quadrats and species presence in each quadrat was recorded to produce a stem map precise to the nearest 10 cm. This produced 5 maps with 180 or 200 quadrats per plot, for a total of 940 quadrats containing data from 15 species. For the simulations we sampled from all of the five plots with replacement using randomly placed square virtual quadrats at nine scales ranging from 0.01 m² (10 × 10 cm) to 0.81 (90 × 90 cm) in size. Fifty randomly placed virtual quadrats were sampled among all five plots at each scale to generate the C_n-score by area relationship, and the standardized effect size.

Case study 2: boreal forest

The boreal forest site was located at Scotty Creek, Northwest Territories, Canada (61°18′N, 121°18′W). The site is located in a region of discontinuous permafrost, defined by a mosaic of permafrost plateaux, bogs and channel fens. Contiguous forest occurs only on the elevated permafrost plateaux, but trees do occur in the bogs and fens. The nutrient poor organic soils produce a forest dominated by black spruce *Picea mariana*. White spruce *P. glauca*, jack pine *Pinus banksiana* and tamarack *Larix laricina* are present but rare in this species poor community. A shrub layer dominated by *Betula* and *Rhododendron* species is found below the trees. This northern site is characterized by short cool growing seasons, long cold winters and relatively low annual precipitation (Quinton et al. 2009).

The dataset was a 2013–2014 census of all vegetation >1 cm diameter at breast height (DBH; measured at 1.3 m). The Scotty Creek Forest Dynamics Plot is part of the Smithsonian Centre for Tropical Forest Science – ForestGEO programme which follows a standardized sampling protocol (Condit 1998). The Scotty Creek Forest Dynamics Plot is 9.6 ha in area (120 × 800 m) and divided into 20 × 20 m grid cells. Every live stem was identified to species, and mapped to the nearest cm providing a detailed individual level map of a mature forest community composed of 38 053 individual trees, from 11 species. For the simulations, we sampled the map with replacement using randomly placed square virtual quadrats at scales ranging from 0.0025 ha (5 × 5 m) to 1 ha (100 × 100 m). Fifty randomly placed virtual quadrats were sampled at each scale as above.

Case study 3: grassland

The grassland site was a rough fescue short grass prairie located at the Roy Berg Kinsella Research Ranch, in Kinsella, Alberta, Canada (53°50′N, 111°33′W). The site is in the Aspen Parkland ecoregion and is comprised of a mosaic of trembling aspen *Populus tremuloide*) stands and short-grass prairie. The plant community is dominated by C_3-grasses including *Festuca hallii*, *Hesperostipa curtiseta* and *Poa pratensis* and includes more than 40 forb species of varying abundance (Lamb 2008). The site was historically winter-grazed by bison, and had been last grazed by cattle five years prior to sampling.

In the summer of 2005, a single haphazardly placed 3 × 3 metre plot was divided into 36 0.5 × 0.5 m grid cells. Every live stem was identified to species and mapped to the nearest centimetre providing a detailed map of the distribution of 1889 stems from 29 species. For the simulations, we sampled the map with replacement using randomly placed square virtual quadrats at scales ranging from 0.1 m² (10 × 10 cm) to 2.25 m² (1.5 × 1.5 m). Fifty randomly placed virtual quadrats were sampled at each scale as above.

Case study 4: tropical forest

Our tropical forest case study system was the Barro Colorado Island (BCI) forest dynamics plot in Panama (9°15′N, –79°85′W) (Condit 1998, Hubbell et al. 1999). The system is a species-rich old-growth tropical forest in a highly seasonal environment with 2500 mm of precipitation annually, a four month dry season, and modest topographic variation.

The data set was the publically available 2010 census (census 7) of trees >1 cm diameter at breast height (DBH) (Condit et al. 2012). BCI is in the same network of forest plots as Scotty Creek and the details of the sampling design and protocol are the same (Condit 1998) except that BCI is a 50 ha plot (500 × 1000 m). As in the boreal forest, every live stem >1 cm DBH and taller than 1.3 m, was identified to species and mapped to the nearest cm within the 50-ha plot. As above, this provided us with a detailed individual level map of the mature forest community from 223 176 individual trees, for 301 species. For the simulations, we sampled the map with replacement using randomly placed square virtual quadrats at scales ranging from 0.0025 ha

(5 × 5 m) to 25 ha (500 × 500 m). Fifty randomly placed virtual quadrats were sampled at each scale as above.

Data deposition

Data available from the Dryad Digital Repository: < http://dx.doi.org/10.5061/dryad.5f876 > (McNickle et al. 2017) and Smithsonian Tropical Forest Network (Scotty and BCI data). All code used in this manuscript is archived at: < https://github.com/ggmcnickle/CScore >.

Results

Numerical simulations

Not surprisingly, in all simulations, there was a hump-shaped C_n-score–area relationship, though at z = 0.1 the peak was shifted nearly to zero (Fig. 1). The exponent, *z*, changes the slope of the SAR in log–log space and thus the rate at which the matrix fills with increasing sample unit scale. Consequently, increasing *z* moves the peak of the C_n-score–area relationship to larger and larger scales. The constant, *B*, changes the y-intercept of the SAR in log–log space, and consequently has no effect on the C_n-score–area relationship. The constant *B* does, however, change the step-like appearance of the C_n-score–area relationship that arises because species richness is an integer, and so lower gamma diversity reduces the smoothness of the relationship. Effect size is undefined for the numerical simulations since the simulated communities are already null communities.

Empirical – species by area relationships

The exponent *z* of the species area curves (rate of species accumulation) was lowest in the species rich tropical forest (z = 0.18), and was similar in the other three communities (z approximately 0.3) (Fig. 2). The BCI z-value is higher than the previously published value of 0.146 (Condit et al. 2005), likely due to differences in the plot selection approaches. As expected, the intercept *B* increased for the more species rich communities (Table 1). The species area curve sampling for the tundra often included plots with no plants even at larger

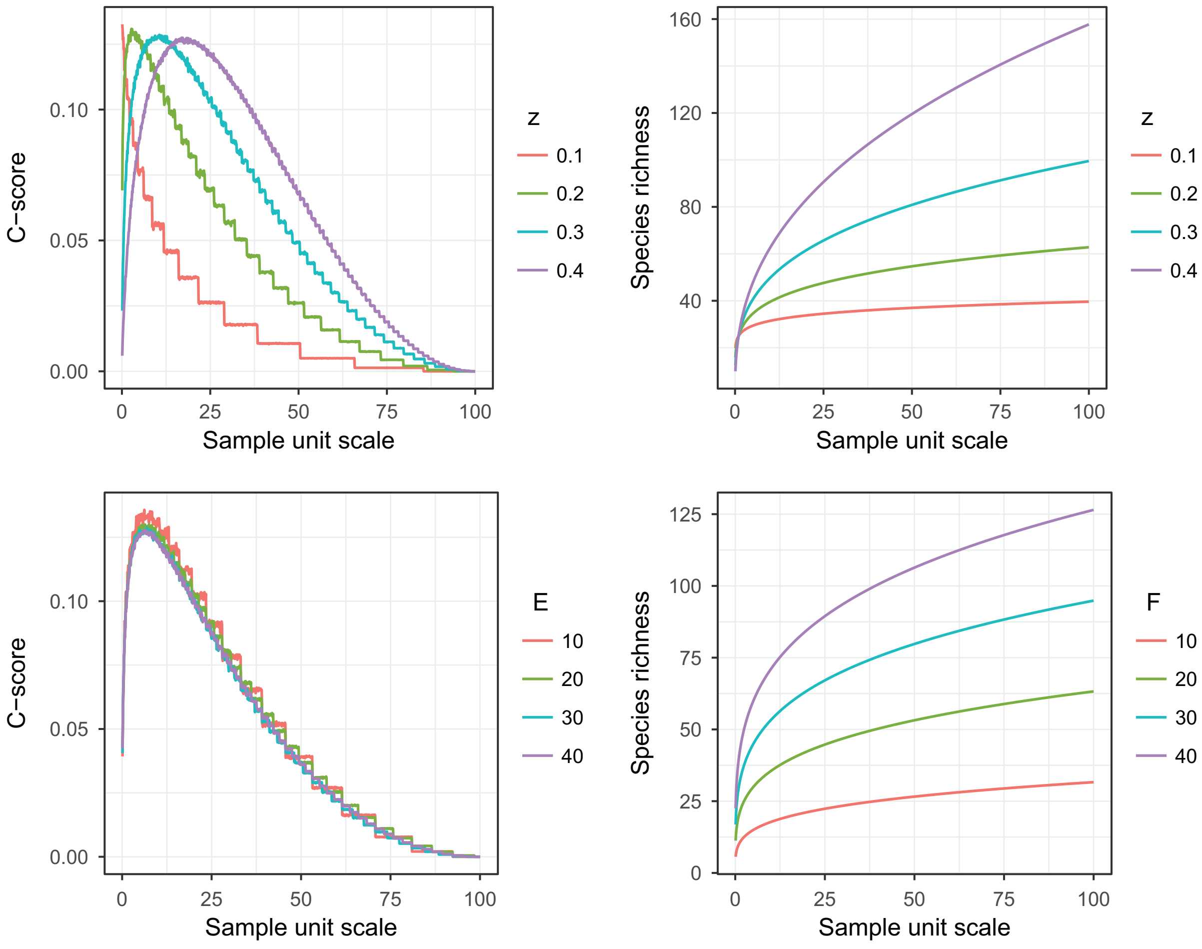

Figure 1. Effects of changing species–area relationship parameters on C-score–area relationships (left panels), the species area–relationships are also shown (right side). In panels (a) and (b) z was varied as indicated in the figure legend and B was held constant at B = 25. In panels (c) and (d) B was varied as indicated in the figure legend and z was held constant at z = 0.25.

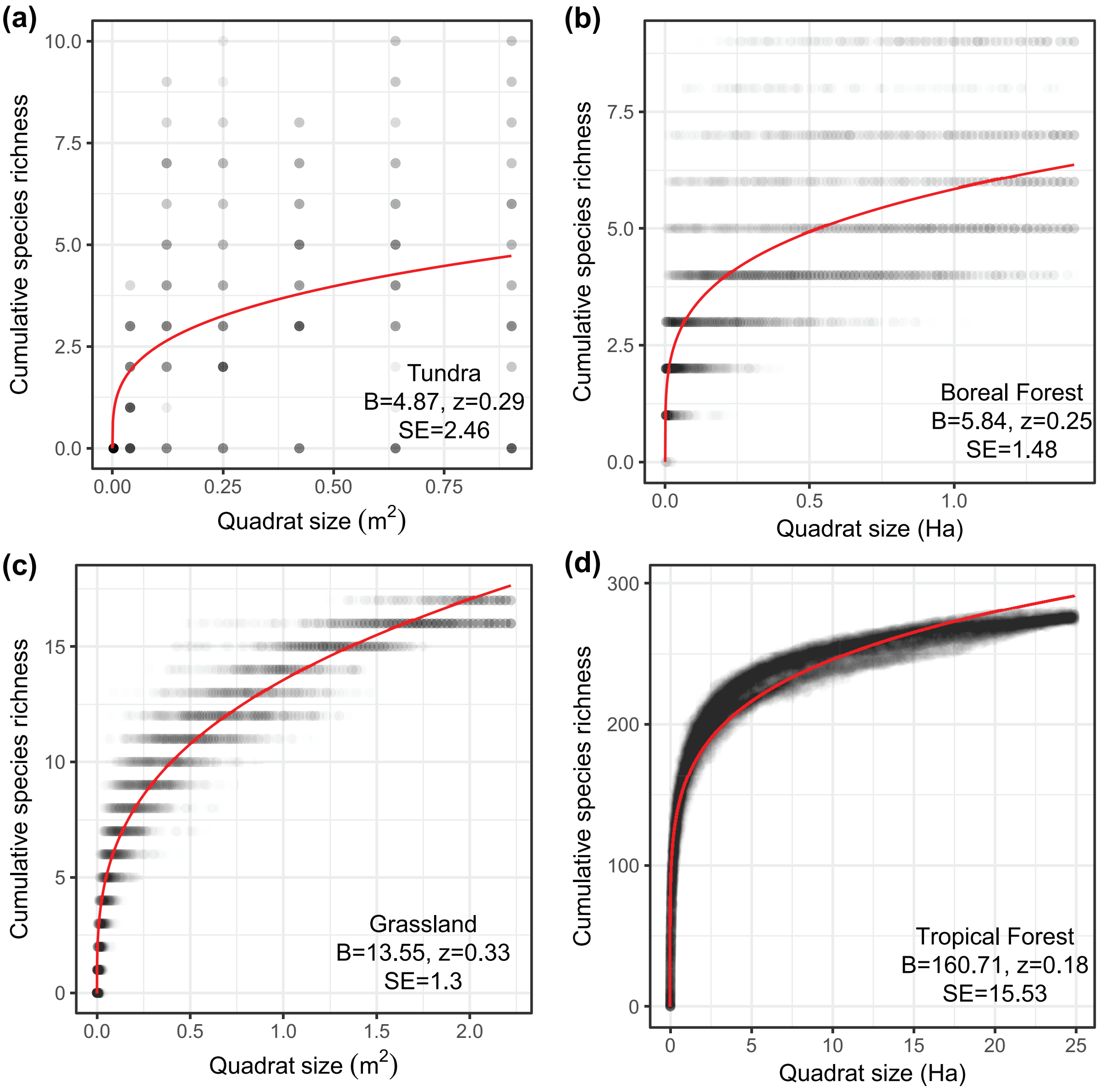

Figure 2. Species area curves for (a) tundra, (b) boreal forest, (c) grassland and (d) tropical forest sites. Darker symbols indicate larger numbers of overlapping datapoints. SE is the standard error of the fitted non-linear model (since R^2 is undefined for non-linear regression): SE represents the average distance between the data and the regression line in the units of the y-axis.

sampling areas, highlighting the patchy nature of that community, and resulting in high error around the fitted curve.

Empirical – C_n-score by area relationships

The fixed-fixed and fixed-equiprobable null models produced similar results (Supplementary material Appendix Fig. A1). Since there were no major differences in conclusions from either null model, in what follows we present and discuss the more conservative fixed-fixed null model only.

We expect the raw C_n-score by area relationships to be hump shaped. Indeed, all case studies, except the boreal forest, had hump shaped C_n-score–area relationships for both observed and expected null C_n-scores (Fig. 3). The maximum C_n-score (the scale at which species are most segregated in their co-occurrence) was reached in the tundra at 0.3 m^2 (0.54 × 0.54 m quadrats) and in the grassland community at 1.5 m^2 (1.2 × 1.2 m quadrats). The maximum C_n-score in the boreal forest was not reached at the maximum sampling scale of 1.4 ha, while in the tropical forest, maximum C_n-score was observed at 0.26 ha (71 × 71 m quadrats) (Fig. 3). In all cases the maximum observed C_n-score was significantly different from the expected null C_n-score.

This significance of the observed C_n-scores are most clearly seen from, the standardized effect size, which indicates that the observed C_n-score is more segregated than would be expected by chance regardless of the absolute value of the C_n-score. Unlike the C-score–area relationship which we expect to be hump shaped, we have no expectation about the shape of the effect size–area relationship. Indeed, these curves took on a variety of complex shapes (Fig. 4). The minimum scale that the C_n-score effect size was significantly different from random was at 0.1 m^2 in the tundra, 0.2 m^2 in grassland, and 0.2 ha in both the boreal and tropical forest (Fig. 4).

Table 1. Summary of empirical descriptions of the communities.

Parameter	Communities			
	Tundra	Boreal forest	Grassland	Tropical forest
Gamma diversity	15	11	29	311
Minimum community richness (*B*)	4.87	5.84	13.55	160.71
rate of species accumulation (*z*)	0.29	0.33	0.25	0.18
Area of max. C-score	0.3 m^2	–	1.5 m^2	0.26 ha
Min area significant effect size	0.1 m^2	0.2 ha	0.2 m^2	0.2 ha
Area of max. effect size	0.2 m^2	–	1.5 m^2	–

Only the tundra and grassland data sets were sampled at a sufficient spatial scale to estimate the maximum effect size which occurred at 0.2 m^2 and 1.5 m^2 respectively (Fig. 4a, c). Neither forest community reached a maximum effect size, including the tropical forest where we were able to sample 25 ha quadrats, which strikes us as remarkable (Fig. 4b, d).

Discussion

It is well known that estimates of diversity vary with the spatial scale of plots and the sampling effort of a survey (Gotelli and Colwell 2001, Chase and Knight 2013). We have shown here via numerical simulation that the relation-

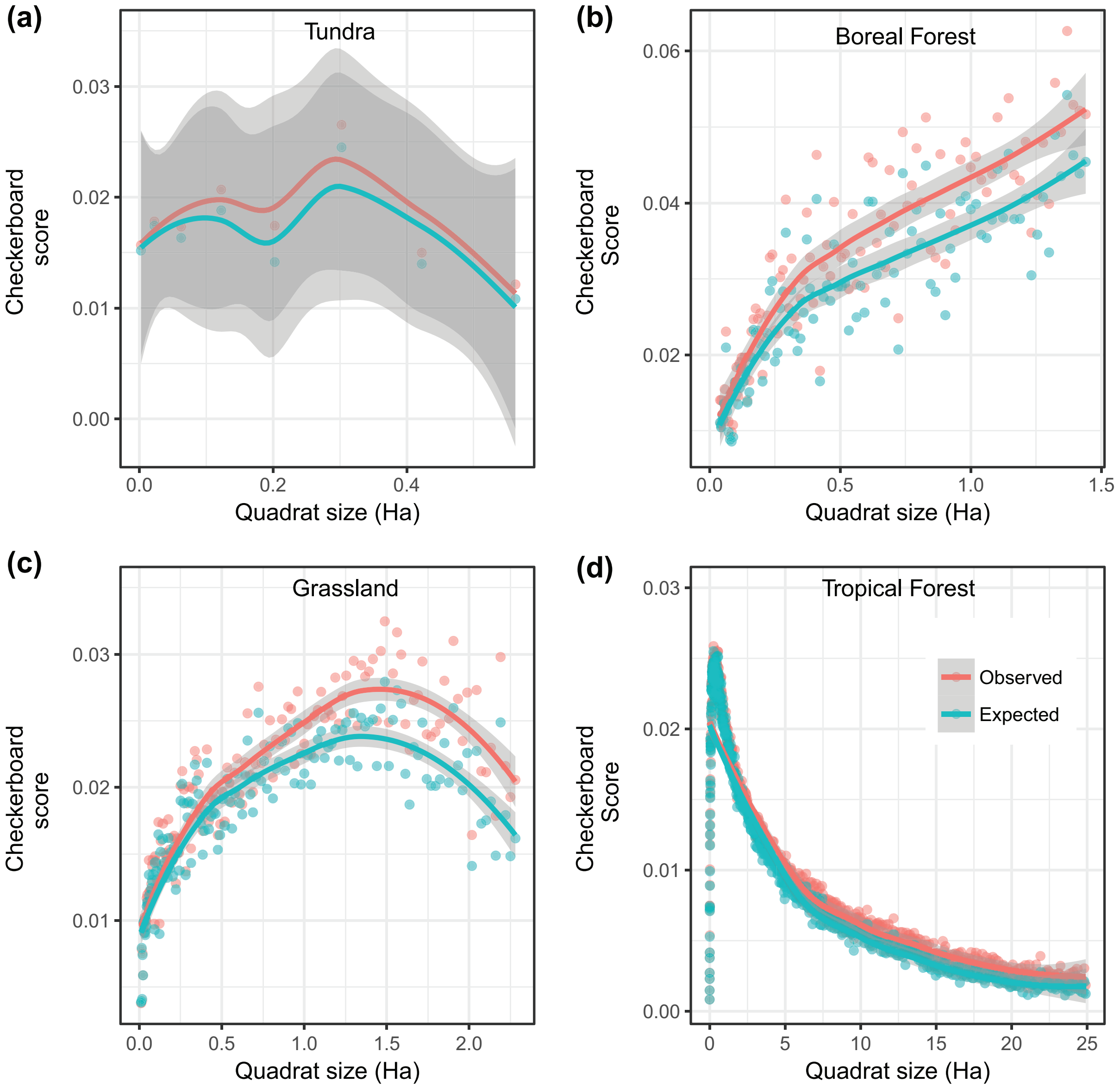

Figure 3. Observed and expected C_n score–area relationships for (a) tundra, (b) boreal forest, (c) grassland and (d) tropical forest sites. Lines are fit using a non-parametric loess function, and grey zones indicate 1 SE around the line. These results are from the fixed-fixed null model only.

ship between species richness and area sample unit scale can influence measures of species co-occurrence such as the C_n-score (Fig. 1). This happens because the accumulation of species with increasing area necessarily produces a C_n-score–area relationship which is hump shaped. This is neither novel nor surprising as it has previously been shown that varying matrix fill causes such a hump-shaped relationship with C-score (Ulrich and Gotelli 2013, Ulrich et al. 2017). Many mechanisms can lead to varying matrix fill; Ulrich et al. (2017) for example, have shown that changing levels of beta diversity (species turnover) across gradients of gamma diversity (regional species pool size) can produce similar hump-shaped C-score patterns. Here we have shown that within a single community varying matrix fill tied to the species–area relationship can also generate a general C_n-score–area relationship. Specifically, we show that intermediate plot sizes with intermediate levels of matrix fill (proportion of 1s in the matrix) will produce maximum C_n-scores, even in randomly assembled communities (Fig. 1). Importantly, our four case studies show that the scale of the intermediate plot size in real ecological communities cannot be predicted a priori, but represents the scale at which there is the largest signal of negative species co-occurrence. The identification of the scale with the maximum statistically significant level of negative species co-occurrence is the scale at which community ecologists are most likely to detect the mechanisms that structure communities. While we illustrate this phenomenon for the C_n-score metric, this observation is likely to apply more broadly across many approaches to the study of species co-occurrence.

Based on this linkage between the species-area relationship and co-occurrence, we argue that the measurement of negative co-occurrence at any one arbitrary spatial scale does not provide complete evidence about how communities are

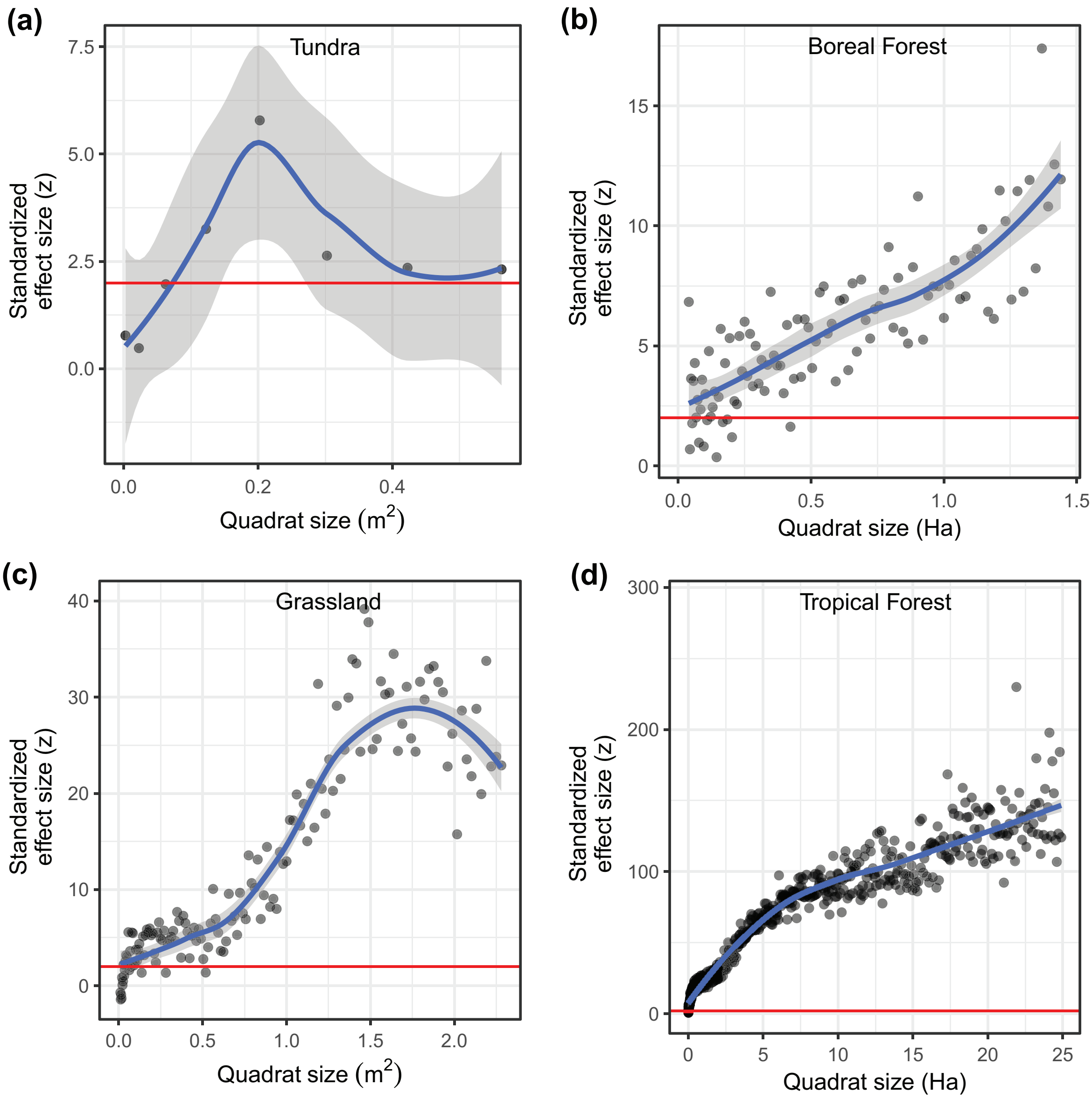

Figure 4. Standardized effect size–area relationships for (a) tundra, (b) boreal forest, (c) grassland and (d) tropical forest sites. Values above the horizontal red line are statistically significant. Lines are fit using a non-parametric loess function, and grey zones indicate 1 SE around the line. These results are from the fixed-fixed null model only.

structured. Many have pointed out that studies of biodiversity are highly scale dependent and that traditional standardised sampling protocols are often not enough to fully study patterns of biodiversity (Gotelli and Colwell 2001, Chase and Knight 2013). Indeed, in all of our empirical case studies, we found scales at which C_n-scores were not different from what was expected by chance under the null model (implying communities that are not strongly structured with respect to negative co-occurrence) and also scales with scores greater than expected (implying that scale-dependent mechanisms are driving co-occurrence patterns) (Fig. 4). The scales at which significant negative co-occurrence was observed appear to depend, in part, on the scale of the organisms in the system, and potentially on the relationship between abiotic variability and mean organism size. The smaller size of grassland and tundra plants corresponded to a smaller scale of structure compared to the large scale for boreal and tropical trees. Importantly, the plot size where we began to see evidence of significant negative co-occurrence was typically quite large compared to standard sampling units (Table 1). Having identified the scales at which significant structure is encountered, it becomes much more feasible to undertake empirical studies to determine whether mechanisms such as species interactions, abiotic drivers, or habitat filtering are the key drivers of patterns at these scales. A first step in such studies may be to use a pairwise analysis approach such as Blois et al. (2014) to identify species pairs with significant co-occurrence patterns at a given scale, and to partition geographic distance effects into those attributable to environmental drivers and other factors (e.g. dispersal distance).

In the grassland and tundra communities, the minimum scale of significant C_n-scores and the scale of the maximum observed C-score could conceivably match the scale at which root systems are able to directly interact (Schenk et al. 1999, Frank et al. 2010, Lamb et al. 2016). Further, many species in the grassland and arctic tundra systems are clonal. Interconnected clones may spread over areas equivalent to the scales of 0.2–1.5 m^2 detected by our analyses, suggesting that shoot and root competition at the scale of clones might also be an important factor structuring these communities (de Kroon et al. 1992, Wildová 2004). Thus the scales of community structure detected in the grassland are consistent with biotic mechanisms of species segregation such as direct plant–plant competition for soil resources among and between individuals. Alternatively, the geospatial patterns of important environmental factors drive small-scale species segregation. At the tundra site, for example, soil moisture has a 1.6 m spatial range (Banerjee et al. 2011a), tundra microbial communities often have a range of ~2 m (Banerjee et al. 2011b), and nutrient cycling a range between 2 and 4 m (Banerjee and Siciliano 2012a, b). The potential role of environmental or competitive mechanisms structuring co-occurrence patterns should not be surprising, as there is evidence from a variety of ecosystems that both biotic and abiotic factors can drive species co-occurrence patterns (Dullinger et al. 2007, Reitalu et al. 2008).

In contrast to the grassland and tundra systems, community structure at the scale of 0.26 to >1.5 hectares in the boreal and tropical forests is unlikely to be driven by resource competition. At these scales, individual root systems would require a radius of ~35 m (tropical) and more than 150 m (boreal) for the outer root tips of two trees to directly interact. Common mycorrhizal networks (CMN) could extend further than roots, as trees in a temperate coniferous forest 20 m apart can be connected (Beiler et al. 2015), but there is no evidence that CMNs routinely connect individuals across 70 m or more. Such large-scale patterns of co-occurrence could be driven through a variety of ecological mechanisms including the dispersal of seeds and distribution of different regeneration niches, herbivory at the scales at which large animals move such as the Janzen–Connell effect, or pathogen impacts on aggregated populations (Janzen 1970, Connell 1971, Grubb 1977, Condit et al. 2000). Given that species richness of the tropical forest is not fully captured in the 50 ha BCI plot (Condit et al. 1996, 2005), there is also the possibility larger-scale biotic driven patterns could exist. Finally, as in the tundra and grassland systems, these forest community patterns could be driven by the scale of abiotic effects (Condit et al. 2000). At the boreal forest site, the landscape is composed of large repeating plant community units driven largely by hydrology and permafrost dynamics with large areas (peat plateaux) dominated by black spruce and wetter and more nutrient-rich fens dominated by larch (Camill 1999). The maximum sampling scale of 1.4 ha was likely not large enough to capture these large scale hydrological features that drive species distributions.

One implication of these results is that the failure of many co-occurrence studies to find evidence for significant co-occurrence patterns (Götzenberger et al. 2012, García-Baquero and Crujeiras 2015) may simply be due to sampling at an arbitrary spatial scale which is different from the scale at which the community is structured. The analysis presented here suggests that previous studies finding weakly or unstructured communities may not have used a sample spatial scale suitable to the community and its structuring processes. Because we may not have been looking for evidence at the spatial scales where processes such competitive exclusion occurs we could erroneously conclude that such mechanisms are not important. For example, our own past study of competition in the grassland system used as a case study here found limited evidence for competition structuring community diversity (Lamb and Cahill 2008). That study, however, measured the community at 0.25 m^2, and the analysis here shows that we might not always expect to find evidence of competitive exclusion at such a small scale (Fig. 4b).

Our results suggest that a multi-scale sampling approach will be an important first step in studies examining species co-occurrence. Our results also suggest that negative co-occurrence is not a characteristic of communities, implying that sufficient sampling does not bring one to a fixed community measure of negative co-occurrence, as is the case with other community metrics like richness and evenness. Estimating the C_n-score–area relationship can therefore provide a method to identify the scales at which communities are most strongly structured, and hence the scales at which the ecological or abiotic processes structuring those communities are the strongest. Identifying the scale provides clues as to the mechanisms that are most important in that system and can guide subsequent experimentation. Our simulations suggest that this hump shaped C_n-score–area relationship will be universal, and our four community analyses support this view. It remains to be seen if this is a general feature of the C_n-score–area relationship, or if this was particular to the four case studies included in this study. A multi-modal C_n-score–area relationship, for example, would imply multiple mechanisms operating at multiple scales.

In conclusion, we have shown here that the species-area relationship naturally produces a C-score–area relationship (Fig. 1). By analyzing four ecologically different case studies from arctic tundra to tropical forest, we show that the dominant scales of structure in these systems were surprisingly large relative to plant body system size in these systems. It has long been recognized that estimates of diversity are influenced by field sampling choices, but it has been difficult to make general suggestions about how one should design studies to account for this. The C_n-score–area relationship can provide guidance for identifying the spatial scales where species interactions influence community structure, it can also help develop hypotheses about the mechanisms that structure communities, and we believe provides a next step towards ecological understanding and unifying a wide array of otherwise contradictory studies.

Acknowledgements – Funding – Field data collection was funded by the International Polar Year, NSERC Discovery and NSERC PDF to SDS and EGL (Devon Island), CFCAS, CFI and Ontario ERA to JLB (Scotty), NSERC Discovery and ACA Biodiversity to JFC and GGM (Kinsella) The BCI forest dynamics research project was founded by S. P. Hubbell and R. B. Foster and is now managed by R. Condit, S. Lao and R. Perez under the Center for Tropical Forest Science and the Smithsonian Tropical Research Institute in Panama. Numerous organizations have provided funding, principally the US National Science Foundation, and hundreds of field workers have contributed. BCI data were obtained from < http://ctfs.si.edu/webatlas/datasets/bci/ >. Analysis and writing were funded in part by NSERC Discovery Grants to EGL and BSS.
Statement of authorship – GGM and EGL contributed equally to this paper.

References

Banerjee, S. and Siciliano, S. D. 2012a. Factors driving potential ammonia oxidation in Canadian Arctic ecosystems: does spatial scale matter? – Appl. Environ. Microbiol. 78: 346–353.

Banerjee, S. and Siciliano, S. D. 2012b. Spatially tripartite interactions of denitrifiers in arctic ecosystems: activities, functional groups and soil resources. – Environ. Microbiol. 14: 2601–2613.

Banerjee, S. et al. 2011a. Soil spatial dependence in three Arctic ecosystems. – Soil Sci. Soc. Am. J. 75: 591–594.

Banerjee, S. et al. 2011b. Evidence of high microbial abundance and spatial dependency in three Arctic soil ecosystems. – Soil Sci. Soc. Am. J. 75: 2227–2232.

Beiler, K. J. et al. 2015. Topology of tree–mycorrhizal fungus interaction networks in xeric and mesic Douglas-fir forests. – J. Ecol. 103: 616–628.

Bliss, L. C. et al. 1994. Patterns of plant distribution within two polar desert landscapes. – Arct. Alpine Res. 26: 46–55.

Blois, J. L. et al. 2014. A framework for evaluating the influence of climate, dispersal limitation, and biotic interactions using fossil pollen associations across the late Quaternary. – Ecography 37: 1095–1108.

Camill, P. 1999. Patterns of boreal permafrost peatland vegetation across environmental gradients sensitive to climate warming. – Can. J. Bot. 77: 721–733.

Chase, J. M. and Knight, T. M. 2013. Scale-dependent effect sizes of ecological drivers on biodiversity: why standardised sampling is not enough. – Ecol. Lett. 16: 17–26.

Condit, R. 1998. Tropical forest census plots: methods and results from Barro Colorado Island, Panama and a Comparison with other plot. – Springer Verlag and RG Landes Company.

Condit, R. et al. 1996. Species–area and species–individual relationships for tropical trees: a comparison of three 50-ha plots. – J. Ecol. 84: 549–562.

Condit, R. et al. 2000. Spatial patterns in the distribution of tropical tree species. – Science 288: 1414.

Condit, R. et al. 2005. Geographic ranges and b-diversity: discovering how many tree species there are where. – Biologiske Skrifter 55: 57–71.

Condit, R. et al. 2012. Barro Colorado Forest census plot data, 2012 version.

Connell, J. H. 1971. On the role of natural enemies in preventing competitive exclusion in some marine animals and in rain forest trees. – In: den Boer, P. J. and Gradwell, G. R. (eds), Dynamics of populations. Proc. Adv. Study Inst. on Dynamics of numbers in populations, Oosterbeek, the Netherlands, pp. 298–312.

Connor, E. F. and Simberloff, D. 1979. The assembly of species communities: chance or competition? – Ecology 60: 1132–1140.

de Kroon, H. et al. 1992. Size hierarchies of shoots and clones in clonal herb monocultures: do clonal and non-clonal plants compete differently? – Oikos 63: 410–419.

Diamond, J. M. 1975. Assembly of species communities. – In: Cody, M. L. and Diamond, J. M. (eds), Ecology and evolution of communities. Harvard Univ. Press, pp. 342–444.

Drake, J. A. 1990. The mechanics of community assembly and succession. – J. Theor. Biol. 147: 213–233.

Dullinger, S. et al. 2007. Weak and variable relationships between environmental severity and small-scale co-occurrence in alpine plant communities. – J. Ecol. 95: 1284–1295.

Frank, D. A. et al. 2010. Grassland root communities: species distributions and how they are linked to aboveground abundance. – Ecology 91: 3201–3209.

García-Baquero, G. and Crujeiras, R. M. 2015. Can environmental constraints determine random patterns of plant species co-occurrence? – Ecol. Evol. 5: 1088–1099.

Gilpin, M. E. and Diamond, J. M. 1982. Factors contributing to non-randomness in species co-occurrences on Islands. – Oecologia 52: 75–84.

Gotelli, N. J. 2000. Null model analysis of species co-occurrence patterns. – Ecology 81: 2606–2621.

Gotelli, N. J. and Colwell, R. K. 2001. Quantifying biodiversity: procedures and pitfalls in the measurement and comparison of species richness. – Ecol. Lett. 4: 379–391.

Gotelli, N. J. and McCabe, D. J. 2002. Species co-occurrence: a meta-analysis of J. M. Diamond's assembly rules model. – Ecology 83: 2091–2096.

Götzenberger, L. et al. 2012. Ecological assembly rules in plant communities – approaches, patterns and prospects. – Biol. Rev. 87: 111–127.

Grubb, P. J. 1977. The maintenance of species-richness in plant communities: the importance of the regeneration niche. – Biol. Rev. Camb. Phil. Soc. 52: 107–145.

Hardy, O. J. 2008. Testing the spatial phylogenetic structure of local communities: statistical performances of different null models and test statistics on a locally neutral community. – J. Ecol. 96: 914–926.

Hubbell, S. P. et al. 1999. Light-gap disturbances, recruitment limitation, and tree diversity in a neotropical forest. – Science 283: 554–557.

Janzen, D. H. 1970. Herbivores and the number of tree species in tropical forests. – Am. Nat. 104: 501–528.

Kent, M. and Coker, P. 1992. Vegetation description and analysis. – Wiley.

Lamb, E. G. 2008. Direct and indirect control of grassland community structure by litter, resources, and biomass. – Ecology 89: 216–225.

Lamb, E. G. and Cahill, J. F. 2008. When competition does not matter: grassland diversity and community composition. – Am. Nat. 171: 777–787.

Lamb, E. G. et al. 2016. A high-throughput belowground plant diversity assay using next-generation sequencing of the trnL intron. – Plant Soil 404: 361–372.

Lavender, T. M. et al. 2016. The influence of matrix size on statistical properties of co-occurrence and limiting similarity null models. – PLoS One 11: e0151146.

Lehsten, V. and Harmand, P. 2006. Null models for species co-occurrence patterns: assessing bias and minimum iteration number for the sequential swap. – Ecography 29: 786–792.

Long, W. et al. 2015. Changes in patterns of species co-occurrence across two tropical cloud forests differing in soil nutrients and air temperature. – Biotropica 47: 416–423.

Maestre, F. T. et al. 2009. On the relationship between abiotic stress and co-occurrence patterns: an assessment at the community level using soil lichen communities and multiple stress gradients. – Oikos 118: 1015–1022.

McNickle, G. G. et al. Data from: Checkerboard score-area relationships reveal spatial scales of plant community structure. Dryad Digital Repository, < https://doi.org/10.5061/dryad.5f876 >.

Miklós, I. and Podani, J. 2004. Randomization of presence–absence matrices: comments and new algorithms. – Ecology 85: 86–92.

Oksanen, J. F. et al. 2016. vegan: community ecology package ver. 2.4-1. – < https://cran.r-project.org >, < https://github.com/vegandevs/vegan >.

Quinton, W. L. et al. 2009. Peatland hydrology of discontinuous permafrost in the Northwest Territories: overview and synthesis. – Can. Water Resour. J. 34: 311–328.

Reitalu, T. et al. 2008. Plant species segregation on different spatial scales in semi-natural grasslands. – J. Veg. Sci. 19: 407–416.

Schenk, H. J. et al. 1999. Spatial root segregation: are plants territorial? – Adv. Ecol. Res. 28: 145–180.

Stone, L. and Roberts, A. 1990. The checkerboard score and species distributions. – Oecologia 85: 74–79.

Ulrich, W. and Gotelli, N. J. 2007. Disentangling community patterns of nestedness and species co-occurrence. – Oikos 116: 2053–2061.

Ulrich, W. and Gotelli, N. J. 2010. Null model analysis of species associations using abundance data. – Ecology 91: 3384–3397.

Ulrich, W. and Gotelli, N. J. 2013. Pattern detection in null model analysis. – Oikos 122: 2–18.

Ulrich, W. et al. 2017. The tangled link between β- and γ-diversity: a Narcissus effect weakens statistical inferences in null model analyses of diversity patterns. – Global Ecol. Biogeogr. 26: 1–5.

Weiher, E. and Keddy, P. 2001. Ecological assembly rules: perspectives, advances, retreats. – Cambridge Univ. Press.

White, J. W. et al. 2014. Ecologists should not use statistical significance tests to interpret simulation model results. – Oikos 123: 385–388.

Wildová, R. 2004. Below-ground spatial pattern of rhizomes in a grassland community and its relevance to above-ground spatial pattern. – Plant Ecol. 174: 321–338.

Supplementary material (available online as Appendix oik-04620 at < www.oikosjournal.org/appendix/oik-04620 >). Appendix 1.