Scale dependence in species turnover reflects variance in species occupancy

Daniel J. McGlinn[1] and Allen H. Hurlbert

Department of Biology, University of North Carolina, Chapel Hill, North Carolina 27599 USA

Abstract. Patterns of species turnover may reflect the processes driving community dynamics across scales. While the majority of studies on species turnover have examined pairwise comparison metrics (e.g., the average Jaccard dissimilarity), it has been proposed that the species–area relationship (SAR) also offers insight into patterns of species turnover because these two patterns may be analytically linked. However, these previous links only apply in a special case where turnover is scale invariant, and we demonstrate across three different plant communities that over 90% of the pairwise turnover values are larger than expected based on scale-invariant predictions from the SAR. Furthermore, the degree of scale dependence in turnover was negatively related to the degree of variance in the occupancy frequency distribution (OFD). These findings suggest that species turnover diverges from scale invariance, and as such pairwise turnover and the slope of the SAR are not redundant. Furthermore, models developed to explain the OFD should be linked with those developed to explain species turnover to achieve a more unified understanding of community structure.

Key words: distance decay; occupancy frequency distribution; scaling; spatial ecology; species–area relationship; species turnover.

Introduction

A key problem in community ecology is to understand species turnover or why species composition varies from place to place. Ecological processes such as environmental filtering, dispersal limitation, and competition, as well as processes such as speciation and extinction, all may leave their signature on patterns of species turnover (Shmida and Wilson 1985, Rosenzweig 1995, Hubbell 2001). Early work focused on understanding how species turnover is shaped by environmental gradients (Whittaker 1967, Peet and Loucks 1977). In a foundational study, Whittaker (1960) suggested that a variety of metrics, including the pairwise Jaccard index and the regional-to-local richness ratio ($\gamma/\bar{\alpha}$), can be used to quantify variation in species composition or beta-diversity. Thus Whittaker (1960) appeared to embrace a pluralistic approach to the study of species turnover that included examinations of different facets of variability in species composition; however, he did not describe how such a joint consideration would be undertaken in a rigorous way (Tuomisto 2010a).

Although more recent analyses of species turnover often differ from the early work of gradient analysis by not examining turnover along explicitly stated environmental gradients, the basic modes of analysis have not changed. Many studies have continued to focus on understanding species turnover via pairwise metrics of compositional similarity or dissimilarity (e.g., Jaccard and Sørensen indices; Koleff et al. 2003, Jurasinski et al. 2009, Anderson et al. 2011). Tuomisto (2010a) pointed out the Jaccard and Sørensen indices are derivatives of the regional-to-local richness ratio (what she referred to as $^0\beta_{Mt}$) when only two samples are considered. Additionally, recent work has developed multisite generalizations of the classic pairwise metrics of species turnover that can be applied to any number of quadrats simultaneously (Diserud and Ødegaard 2007). The regional-to-local richness ratio can also be linked to the study of the log–log species–area relationship (SAR; Arita and Rodriguez 2002, Scheiner et al. 2011). The slope of a spatially explicit SAR, commonly referred to as the z value, is defined as $\log_2[\gamma/\bar{\alpha}]/2$ for a system of four quadrats and is typically interpreted as the relative rate new species are added to the community as area is increased (White 2004, but see Tuomisto 2010b for an entropy-based interpretation of z). However, recent work has also demonstrated that z is directly related to mean relative species occupancy (Šizling and Storch 2004, Storch et al. 2007). The link between occupancy and the SAR is important because it demonstrates that z primarily reflects ecological and evolutionary drivers that shape how broadly distributed species are rather than how species covary spatially with one another, and it provides the opportunity to test hypotheses about how patterns of occupancy change as a function of scale using SARs (McGeoch and Gaston 2002, Storch et al. 2003, Hurlbert and Jetz 2010). This work has set the

[1] Present address: BNR 132, Department of Biology, Utah State University, Logan, Utah 84322-5305 USA. E-mail: danmcglinn@gmail.com

stage for developing a more general linkage between species occupancy and species turnover that also includes pairwise turnover metrics.

Despite the mathematical connection between pairwise metrics of similarity and the slope of the SAR via the regional-to-local richness ratio, these metrics are typically not jointly considered (Tuomisto 2010b). However, some researchers have suggested that they capture similar information and that these two metrics should be unified (Harte and Kinzig 1997, Morlon et al. 2008, Tjørve and Tjørve 2008). It should be intuitively clear that if there is zero pairwise turnover between any of the samples in a region then the slope SAR for that region must be also be zero (i.e., no new species are accumulated if species composition does not vary across the samples). Similarly, if each pairwise comparison yields complete turnover, then the slope of the SAR must be maximized because every species encountered is a new species. Therefore, pairwise metrics of turnover and the rate of change in the SAR clearly overlap in the information they convey. Several authors have demonstrated that it is possible to go beyond these simple intuitive links between the SAR and turnover when considering the special case of two samples. Specifically, Tjørve and Tjørve (2008) demonstrated that the proportion of species shared between two samples (g), a pairwise metric of compositional similarity, may be expressed as a function of the slope of the log–log SAR, z:

$$g = 2 - 2^z. \tag{1}$$

Eq. 1 provides an exact transformation between the slope of the SAR and a metric of compositional variability between two samples (in this case, Sørensen's index). This finding implies that the wealth of empirical information on SARs may be brought to bear on questions concerning the nature of species turnover. However, it has been less appreciated that unless species accumulation follows the strict model of scale invariance (i.e., a constant fraction of new species are gained between two samples of equal area irrespective of spatial scale), Eq. 1 cannot be extended from the two-sample case to a collection of N samples. In other words, if more than two samples are considered at a time, then for a given z a range of pairwise turnover values are actually possible (Fig. 1A). Given that almost all nested species–area relationships are evaluated by aggregating four or more quadrats at a fine scale to some coarser scale and the fact that strict scale invariance is rarely exhibited by SARs (e.g., Palmer and White 1994, Plotkin et al. 2000), Eq. 1 will rarely apply in its exact form.

The purpose of our study is to examine the scaling of species turnover by relating a pairwise metric of species turnover and the slope of the SAR when considering more than two samples. Within this framework, the assumption of scale invariance implied by Eq. 1 may be treated simply as a special case in which pairwise turnover and the slope of the SAR provide completely redundant information. If the assumption of scale-invariant species turnover is invalid, then by considering both pairwise turnover metrics and the SAR slope together we may gain more information than considering either alone. We develop a more general understanding of the aspects of community structure and distribution that contribute to this lack of redundancy between pairwise turnover and the SAR using analytically derived formulas for a four-quadrat case and simulations that generalize our primary results to an arbitrary number of quadrats. Our work is relevant to gridded species distribution data that are increasingly common in macroecological investigations. In our analytical treatment, we considered a four-quadrat case because quadrats are typically aggregated in sets of four (Hui 2009).

We find that species turnover in three plant communities systematically deviates from Eq. 1, and that this deviation depends on the degree of variance in the species' occupancy distribution. Thus, this study provides the first formal linkages between the SAR, turnover, and variance in occupancy. While we investigate this problem from the context of the Jaccard index of turnover, and the slope of the nested log–log SAR, our framework could be applied to any pairwise or multisite index of turnover based on presence–absence data (Tuomisto 2010b). The emphasis of this study is on attempting to relate measurements of turnover at different spatial grains (e.g., two vs. four units of area). In this light, our study could be equally well be viewed as an examination of how multisite metrics of turnover (e.g., Diserud and Ødegaard 2007) change as a function of the number of sites considered (Tuomisto 2010a, b).

ANALYTICAL LINKS BETWEEN SAR AND PAIRWISE TURNOVER

Here we develop a mathematical framework that defines the extreme values of average pairwise turnover, as measured by the Jaccard turnover index (one minus the Jaccard similarity index [Legendre and Legendre 1998:256]), given the slope of the log–log SAR, z, for a set of four, equally spaced quadrats. In this sampling context, a spatially explicit SAR (Type IIA or IIA) and spatially implicit SAR (Type IIB or IIIB) are equivalent (Scheiner et al. 2011). We chose to use the Jaccard turnover index as the pairwise metric of turnover because of it is popularity and simplicity; however, our proofs and findings apply to all presence–absence turnover metrics that may be expressed in terms of $\bar{\alpha}$, mean quadrat richness, and γ, the total number of unique species present over all quadrats (Jost 2006, Tuomisto 2010a, b). Note that the terms $\bar{\alpha}$ and γ are not tied to an absolute spatial scale and simply refer to richness at the grain and extent, respectively, of any particular comparison. For our purposes, we will reformulate Eq. 1 as derived by Tjørve and Tjørve (2008) in terms of Jaccard's index of turnover (Appendix A):
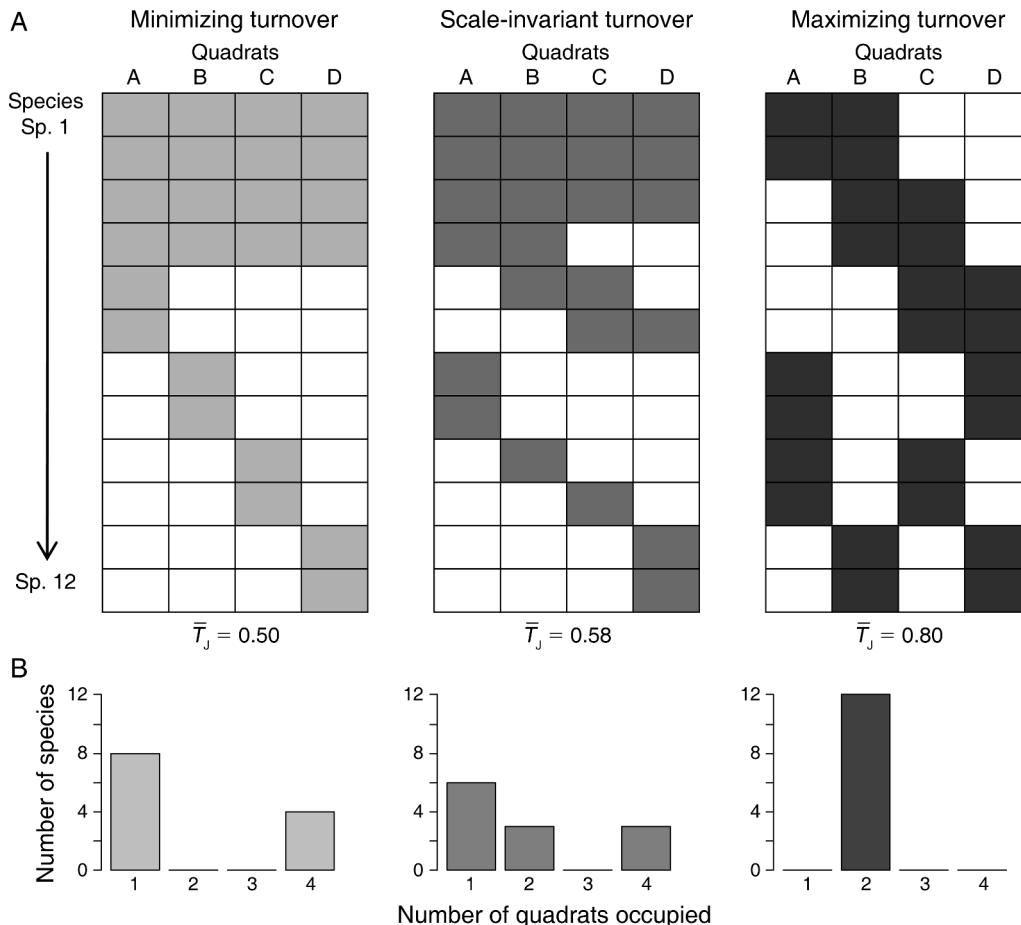
FIG. 1. (A) The species-by-site, presence–absence matrices and (B) accompanying histograms of the occupancy distribution for three hypothetical communities in which mean quadrat richness ($\bar{\alpha}$) = 6, total richness across all quadrats (γ) = 12, and the slope of the log–log species–area relationship (z) = 0.5. Shaded cells represent species presences, and the white regions represent species absences. The communities vary from one that minimizes turnover [average Jaccard turnover ($\bar{T}_J$) = 0.5] to one that maximizes turnover [$\bar{T}_J$ = 0.8]. The three communities only differ in their occupancy distributions, which are simply the row sums of the species-by-site matrices. Note that the community that minimizes turnover has the highest variance in occupancy and that the opposite is true for the community that maximizes turnover. The expected $\bar{T}_J$ based on Eq. 2 assuming scale invariance is 0.59.

$$T_J = 2 - 2^{1-z}. \qquad (2)$$

Now we will demonstrate that given $\bar{\alpha}$ and γ, and therefore the species–area slope, a range of average pairwise turnover values are possible. This can be seen clearly in three hypothetical communities displayed in Fig. 1. In all three communities, $\bar{\alpha} = 6$ and $\gamma = 12$ and therefore $z = 0.5$; however, the average level of pairwise turnover, as measured by Jaccard's index of dissimilarity ($\bar{T}_J$), varied from 0.5 to 0.8. Each species-by-site matrix contains 24 total species occurrences (i.e., $4\bar{\alpha}$); however, the matrices differ in how these occurrences are distributed across the γ species.

The distribution of the total number of occurrences of each species (i.e., the row sums of the species-by-site matrix), is one simple spatially implicit expression of the spatial distribution of the species (Fig. 1B). Because species identities are not of interest in this framework, we are only interested in the total number of species that

occurred in a given number of quadrats. This can be expressed by the occupancy frequency distribution (OFD), which we express as the vector $\mathbf{x} = [x_1, x_2, x_3, x_4]$ where x_i is the number of species that occur in exactly i quadrats (McGeoch and Gaston 2002, Hui and McGeoch 2007b).

The values of $\mathbf{x}$ must meet the following set of constraints to maintain $\bar{\alpha}$ and γ. The first constraint is that the sum of the elements of $\mathbf{x}$ must equal the total number of species in the assemblage: $\sum_{i=1}^{4} x_i = \gamma$. The second constraint is that the total number of occupied cells in the species-by-site matrix must be equal to $\sum_{i=1}^{4} i x_i = \sum_{i=1}^{4} \alpha_i = 4\bar{\alpha}$ (i.e., row sums = column sums) that links the elements of $\mathbf{x}$ to $\bar{\alpha}$. Last, it should be obvious that γ is constrained to a range of possible values defined by $\bar{\alpha}$: when all species are shared across all four quadrats (i.e., all species are x_4 species), then $\gamma = \bar{\alpha}$, and when all species are only present in a single quadrat

TABLE 1. Minimum and maximum possible values of pairwise turnover ($\bar{T}_J$) as a function of the slope of the species–area relationship, z.

Parameter	Domain of z	$\bar{T}_J$
Minimum $\bar{T}_J$	$[0, 1]$	$(2^{2z} - 1)/2^{2z-1} + 1)$
Maximum $\bar{T}_J$	$[0, \log_2(4/3)/2)$	$2 - 2^{1-2z}$
	$[\log_2(4/3)/2, 0.5)$	$\left(2^{2z} - \dfrac{2}{3}\right)\Big/\left(2^{2z-1} + \dfrac{2}{3}\right)$
	$[0.5, 1]$	$(2^{2z-1} + 1)/(2^{2z-2} + 2)$

Note: These functions were derived using an assumption of constant quadrat richness (Appendices B and C).

(x_1 species), then $\gamma = 4\bar{\alpha}$. Therefore we may state the following inequality: $\bar{\alpha} \leq \gamma \leq 4\bar{\alpha}$. This set of constraints allowed us to solve for the OFDs that maximize and minimize pairwise turnover respectively (Appendix B). These theorems indicate that OFDs with minimal variance in occupancy have maximum species turnover and similarly that maximum variance in occupancy results in minimum species turnover.

Based on the theorems for the OFDs that maximized and minimized pairwise turnover and with the additional constraint that quadrat richness is constant (i.e., $\alpha_i = \bar{\alpha}$), we were also able to derive analytical expressions of average pairwise turnover using Jaccard's index of turnover ($\bar{T}_J$). These theorems express $\bar{T}_J$ as a function of $\bar{\alpha}$ and γ and separately as a function of z (Appendix B). To facilitate the extension of our analytical framework to other pairwise metrics of turnover for presence–absence data, we also provided the resulting matching/mismatching components (a = number of species shared and $b = c$ = number of unique species) for each pairwise comparison (Legendre and Legendre 1998:254). Table 1 displays the expressions that define the constraint space between $\bar{T}_J$ and z.

To examine if these analytical results generally hold beyond a system of four quadrats (i.e., for an arbitrary number of quadrats) we conducted a simple numerical experiment. We were specifically interested in comparing average pairwise turnover for two sets of n^2 quadrats that have similar levels of $\bar{\alpha}$ and γ but differ strongly in their variance in occupancy. While an analytical solution to this general question may be possible, here we simply provide the results of 500 runs of a simulation in which we randomly placed 50 species on a 16×16 grid ($n^2 = 256$).

Occupancy of the ith species, p_i, in the "high-variance" community was generated from a lognormal (LOGN) distribution ($p_i^{LOGN} \sim LOGN(\mu = 0, \sigma^2 = 1)$). To ensure that these values ranged from 0 to 1 we transformed the p_i values by dividing them by the largest p_i^{LOGN} value. Therefore the most common species occupied every quadrat in the grid, while other species had occupancy values as low as 1/256. Occupancy in the "no-variance" (or even) community was set at $p^{EVEN} = \sum_{i=1}^{50} p_i^{LOGN}/50$ for all species. This was done to ensure

that $\sum_{i=1}^{50} p_i^{EVEN} = \sum_{i=1}^{50} p_i^{LOGN} = E[\alpha]$ where $E[\alpha]$ represents the expected value of α.

Because we are only interested in calculating average pairwise turnover at the grain of a single grid cell, the spatial distribution of individuals was of no consequence; therefore, we randomly distributed presences across the grid. For each run of our simulation we calculated pairwise $\bar{T}_J$ and the global z considering $\bar{\alpha}$ at the scale of a single cell and γ across all 256 grid cells.

It is possible in empirical data sets that variation in quadrat richness (i.e., a violation of our only assumption) may result in $\bar{T}_J$ values outside of the derived constraint space. This problem may be compounded when richness is low in which case small amounts of variance in richness can have large impacts on turnover estimates. Therefore, in this study we restrict the application of our framework to spatial sample grains in which average richness is at least 10. Additionally we analyzed simulated data sets in which we allowed richness to vary among quadrats (Appendix C).

EMPIRICAL DATA SETS

We used our analytical framework to examine the relationship between the SAR and pairwise turnover in three multi-scale plant community data sets. The plant communities we examined differed greatly in their constituent species, community spatial patterns, environmental heterogeneity, and sampling designs. For each community, we identified sets of four quadrats of equal area arranged such that each quadrat was an equal distance from the spatial centroid of the set. For each set, we calculated z based upon average quadrat richness and total richness across the set, and we calculated $\bar{T}_J$ using the six unique pairwise comparisons between the four quadrats.

The first data set, forest data, was from a mixed hardwood forest located on the Oosting Natural Area, in the piedmont of central North Carolina, USA (data and sampling design fully described in Palmer et al. [2007]). This data set was a good candidate for examining spatial turnover because it provides an unprecedented range of spatial grains. For our purposes, we examined 754 sets of four spatially contiguous quadrats that were drawn from four spatial grains: 4, 16, 256, and 1024 m^2.

The second data set, grassland data, was located on the Tallgrass Prairie Preserve, in the Flint Hills of northeastern Oklahoma, USA (data and sampling design fully described in McGlinn et al. [2010]). The grassland data was composed of 20 100-m^2 quadrats located semi-randomly on a 1-km grid. A nested series of spatially non-contiguous subplots were sampled in each of the corners of the quadrats. The quadrats were annually resampled for 12 years. For our purposes, we examined two spatial grains: 1 and 10 m^2 at a temporal grain of a single year for the entire 12-year period, which resulted in 480 sets of four quadrats.
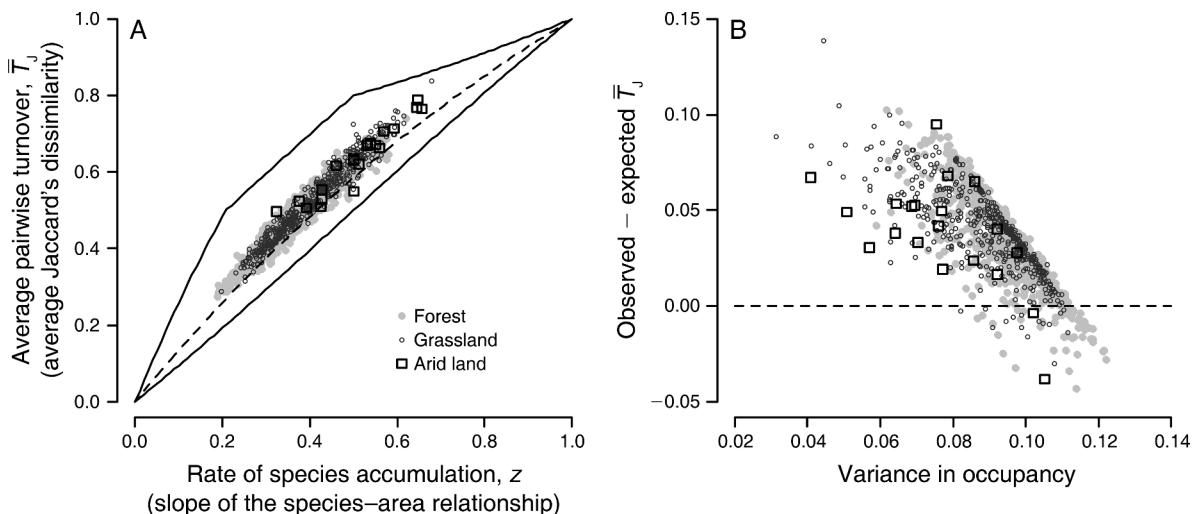
FIG. 2. (A) The relationship between the slope of the species–area relationship and average pairwise turnover for the three empirical data sets and (B) the difference in the empirical values from the model of scale-invariant turnover vs. variance in occupancy. In panel A, the constraint envelope is defined by the solid lines and the scale-invariant prediction is the dashed line. The forest data represent the following spatially contiguous windows: 4–16, 16–64, 256–1024, and 1024–4096 m^2. The grassland data comprised scales of 1–4 and 10–40 m^2 in non-contiguous spatial windows. The arid-land data were only from spatially non-contiguous windows ranging from 4 m^2 to 16 m^2 (the plot-to-web scale).

The last data set, arid-land data, is from the Sevilleta LTER in central New Mexico, USA. These data were collected from four different arid to semiarid habitats: (1) Chihuahuan/plains grassland, (2) creosote scrubland, (3) juniper–oak savannah, and (4) pinyon–juniper woodland (Muldavin 2008, 2009). In contrast to the forest and grassland data sets, the arid-land data set was not designed explicitly for evaluating spatial scale; however, it still has a natural hierarchical structure (see Appendix D for diagrams of sampling design). Within each of the four sites, data collection revolved around five rodent-trapping webs arrayed roughly linearly and with centers 260 to 460 m apart. A group of four 1-m^2 quadrats separated by approximately 1 m and arrayed in a square (hereafter referred to as a plot) were sampled 100 m in each cardinal direction from the center of each web. Thus, within each site there are five webs, 20 plots, and 80 quadrats. For our purposes, we only examined the plot-to-web scale (4 to 16 m^2).

For each data set, we calculated the residual deviance from the expectation of scale invariance (Eq. 2). We estimated the independent fractions of variance explained in the residuals by variance in occupancy and the coefficient of variation (CV) in species richness ($\sigma_\alpha / \mu_\alpha$) using ordinary least squares multiple regression. We were interested in the CV of richness because our framework assumes that richness is constant and therefore we wished to assess if violations of this assumption contributed strongly to our observed results. We used the CV of richness because the standard deviation of richness is generally not independent of average richness and interpretations across spatial scales would thus be less straightforward.

RESULTS

The constraint envelope for the maximum and minimum possible average pairwise turnover values ($\bar{T}_J$) as a function of the slope of the SAR (z) is displayed in Fig. 2A. The maximum $\bar{T}_J$ values ($\bar{T}_{Jmax}$) displayed three distinct phases of increase that were defined by three nonlinear equations (Table 1). Each change in phase occurs due to shifts in the OFD that are necessary to maximize pairwise turnover. The first phase ($z \in [0, \log_2(4/3)/2)$) is composed of only x_4 and x_3 species, the second phase ($z \in [\log_2(4/3)/2, 0.5)$) is composed of x_3 and x_2 species, the third phase ($z \in [0.5, 1]$) has only x_2 and x_1 species (see Appendix B). The minimum $\bar{T}_J$ values ($\bar{T}_{Jmin}$) fell tightly along the 1:1 line between $\bar{T}_J$ and z (Fig. 2A), although this function is also intrinsically nonlinear (Table 1). The constraint space was widest ($\bar{T}_{Jmax} - \bar{T}_{Jmin} \approx 0.3$) for intermediate values of z (i.e., $0.2 < z < 0.5$). The results of the simulation were in agreement with our analytical results (Fig. 3). Specifically, the community in which there was no variance in occupancy resulted in higher average pairwise turnover values than the community in which occupancy was variable (in this case right skewed).

The relationship between $\bar{T}_J$ and z was similar across the three empirical data sets (Fig. 2A). The majority of the empirical values (90–97%), irrespective of ecosystem, lay above the expectation for scale-invariant species turnover provided by Eq. 2. The multiple regressions indicated that most of the variation in the deviance from the expected values was negatively correlated with variance in species' occupancy values (partial $R^2 = 0.25$–0.76; Fig. 2B). In contrast, the CV of richness explained a negligible amount of the deviance in
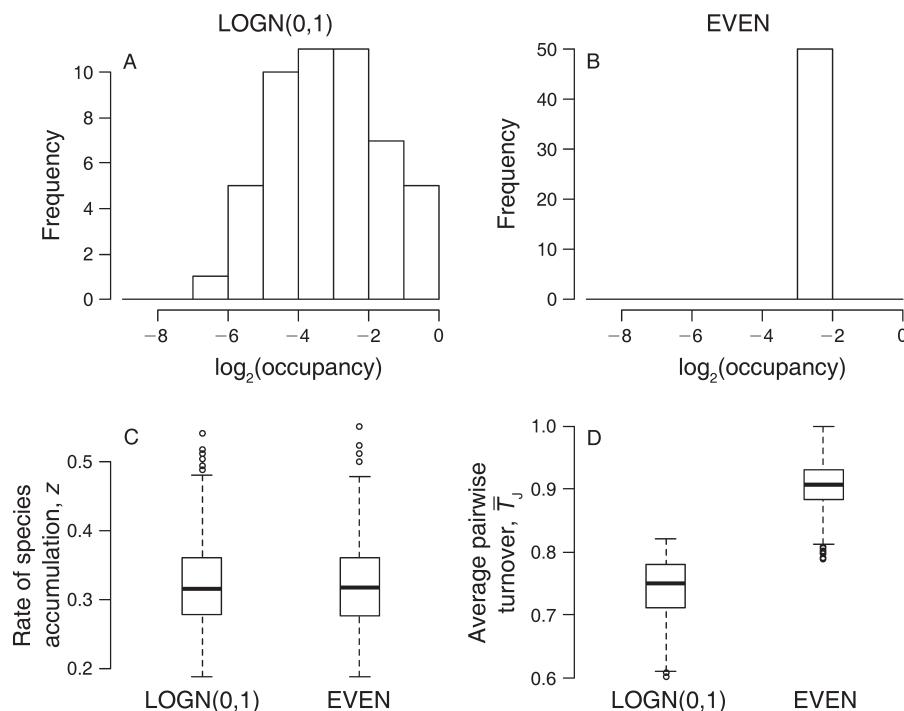
FIG. 3. The results of the numerical experiment conducted on the 16×16 grid. In panels A and B, the lognormal [LOGN(0, 1)] and the completely even [EVEN] occupancy frequency distributions are displayed for a single run of the simulation. Panels C and D display box-and-whisker plots that visually summarize the distribution of the rate of species accumulation and average pairwise turnover values, respectively. Plots show the minimum and maximum (bottom and top of box), interquartile range (whiskers), and outliers (circles). We designed our simulations such that there would be little to no difference in z values, but we expected that the community with lower variance in occupancy to result in higher values of $\bar{T}_J$, which is what we observed.

turnover values (partial $R^2 = 0.03–0.09$). Simulations in which we directly manipulated the CV of richness indicated that increasing the CV of richness increased the intercept but not the slope of the relationship between deviance from scale invariance and variance in occupancy (Appendix C). As in the example in Fig. 1, the empirical assemblages in which variance in occupancy was low had pairwise turnover values much higher than expected, while those with high variance in occupancy were much closer to or even less than the expected value.

Discussion

Our analytical framework links average pairwise turnover ($\bar{T}_J$) and the slope of the SAR (z) by defining the mathematical constraints imposed by a fixed value of z on the occupancy frequency distribution (OFD) and ultimately $\bar{T}_J$. Prior studies have established the expected value of pairwise turnover due to scale-invariant turnover, which suggests complete redundancy between the species–area relationship (SAR) and pairwise turnover (Harte and Kinzig 1997, Harte et al. 1999, Tjørve and Tjørve 2008), but data from three distinct plant communities suggest that empirical observations often deviate systematically from scale invariance. Furthermore, our framework and empirical results suggest that deviations from scale invariance are strongly related to

the degree of variance in the OFD. The novel links we have developed between turnover and the OFD should result in stronger tests of ecological theory and more unified understanding of biodiversity patterns.

The constraint space

The potential range of variation in $\bar{T}_J$ varies with the magnitude of z. The nature of this dependence reflects the constraints placed on the OFD by a given rate of species accumulation. To see this clearly consider that as γ approaches $\bar{\alpha}$ (i.e., $z \to 0$), both pairwise turnover and the species accumulation rate must converge at 0 given that the assemblage must be completely composed of x_4 species. Similarly, as γ approaches $4\bar{\alpha}$ (i.e., $z \to 1$), pairwise turnover and the species accumulation rate must converge at 1 because the OFD must be increasingly composed of only x_1 species. Because of the necessary convergence of z and $\bar{T}_J$, the greatest possible range of pairwise turnover values occurred at intermediate rates of species accumulation (i.e., z values between 0.2 and 0.5).

Scale-dependence of species turnover

Developing links between the SAR and pairwise turnover would be straightforward if ecological communities displayed scale-invariant patterns of turnover (Tjørve and Tjørve 2008); however, as our study and

many others have demonstrated, patterns of spatial turnover do not strictly adhere to models of scale invariance (Palmer and White 1994, Plotkin et al. 2000, Crawley and Harral 2001, Fridley et al. 2005). Specifically, in our study we found that the empirical data generally had a higher average level of pairwise turnover than that predicted by scale invariance. This appeared to be true across ecosystems and for different spatial grains, and it suggests that OFDs in the empirical communities were typically characterized by communities somewhere in between the hypothetical "scale-invariant" and "maximizing turnover" communities in Fig. 1.

The majority of studies addressing scale-invariant patterns of turnover have examined the degree of linearity in the log–log SAR. Therefore, it is useful to note that our observation of systematically larger values of average pairwise turnover than predicted under scale invariance is consistent with a concave-down log–log SAR (Plotkin et al. 2000, Crawley and Harral 2001, McGlinn and Palmer 2009), with more positive deviations from expected reflecting a greater degree of concavity. Although there is still a lively debate about the functional form of the SAR (e.g., Tjørve 2009), our framework suggests there are clear mathematical limits on the degree of concavity and thus the range of possible nested SAR shapes.

Interpreting the importance and potential drivers of the occupancy distribution

A key finding of our work is that the shape of the OFD influences the scale-dependence of pairwise turnover and thus the relationship between turnover and the SAR. Specifically, both our analytical derivations and empirical patterns suggest that there is a strong negative relationship between variance in the OFD and average pairwise turnover (Fig. 2B). This finding provides for the possibility of linking hypotheses developed for patterns of occupancy with patterns of turnover and vice versa. To demonstrate this potential here we use our theoretical framework to show that the core-satellite model (Hanski 1982), a process-based model designed initially to explain occupancy patterns, makes predictions about patterns of species turnover that generally do not agree with our empirical findings. Additionally, we describe how the scaling of environmental heterogeneity, which is typically linked to patterns of turnover (Triantis et al. 2003, Lundholm 2009), is expected to influence patterns of occupancy. Both of these extensions emphasize how our framework has established a firm link between patterns of occupancy and patterns of turnover, which result in a more unified understanding of community structure.

There are many studies outside the context of species turnover that have explored the drivers of the OFD (reviewed in McGeoch and Gaston 2002). For example, the core-satellite hypothesis suggests that an assemblage composed of core (x_4) species and satellite (x_1) species

reflects the equilibrium outcome of interspecific variation in local colonization and extinction rates in the absence of environmental heterogeneity (Hanski 1982). Our framework demonstrates that a core-satellite occupancy distribution is one end of the spectrum, and that it should result in minimum levels of pairwise species turnover at a given z value. However, our empirical results suggested that the plant communities we examined typically did not display bimodal, core-satellite OFDs and exhibited greater pairwise turnover than predicted by that hypothesis. This finding may indicate that the sites and spatial grains we examined are not characterized by equilibrium patterns of colonization and extinction as envisioned by Hanski (1982), and that other factors such as environmental heterogeneity are influencing the OFD (e.g., Brown 1984).

Environmental heterogeneity is frequently invoked as an explanation for the shape of the SAR. If species differ in their environmental requirements, then the more habitats or environments that are encountered the larger the z value will be (Triantis et al. 2003, Hurlbert and Jetz 2010). This line of reasoning has also been successfully applied to patterns of mean occupancy in which case we expect that mean occupancy will decrease as the degree of environmental heterogeneity increases because more rare or satellite species are represented in the samples (McGeoch and Gaston 2002, Storch et al. 2003). More recently it has also been suggested that the spatial geometry of environmental heterogeneity (e.g., the way in which variance in the environment scales with area) should be mirrored by the shape of the SAR (Palmer 2007, McGlinn and Palmer 2011). If environmental heterogeneity scales in a concave-up pattern (i.e., new habitats are encountered slowly as a function of area because of strong local autocorrelation) then the OFD will have high variance characteristic of a core-satellite pattern. In contrast, if the environment scales in a concave-down pattern (i.e., new habitats are encountered rapidly with an increase in area) then the OFD will have low variance with all species occurring at a similar level of occupancy. Thus our study provides a link between the scaling of environmental heterogeneity and the shape of the OFD. These expectations rest upon the assumptions that heterogeneity in the relevant environmental variable(s) can be quantified and that species respond to the environment at the spatial scale of the sample.

McGeoch and Gaston (2002) suggested that there are many potential mechanisms that may influence OFDs, and that these explanations are strongly scale dependent. As such, they urged that assemblage wide patterns of occupancy should be examined as a function of spatial grain. Our framework provides one potential way to accomplish this goal, as well as to understand how these patterns of occupancy contribute to the observed patterns of spatial species turnover.

Although our study represents a first attempt to explicitly examine the role of variance in occupancy in

explaining patterns of species turnover, previous studies have recognized that there is a direct link between mean relative species occupancy ($\bar{p}$) and z:$\bar{p} = \bar{\alpha}/\gamma$ and therefore in our scenario of four quadrats $z = -\log_2(\bar{p})/2$ (Šizling and Storch 2004, Storch et al. 2007). It is clear that the primary axis of variation in pairwise turnover is due to a negative correlation with $\bar{p}$ (i.e., the redundancy between the SAR and pairwise turnover); however, the secondary axis of variation in pairwise turnover (i.e., the novel information gained by considering both the rate of species accumulation and pairwise turnover simultaneously) reflects variance in the OFD. Additionally, Hui and McGeoch (2008) demonstrated that it is important to consider how species' occupancy changes with scale if we wish to link individual species distributions with patterns of community structure.

Several studies that have examined the response of the SAR to variance in the rank abundance distribution (RAD) have found somewhat analogous results. Specifically, it has been shown that low variance in the RAD (e.g., complete evenness) results in SARs with more spatial turnover at intermediate scales than expected under scale invariance (i.e., concave-down log–log SARs; He and Legendre 2002, McGlinn and Palmer 2009). It seems reasonable to draw parallels between these studies linking the RAD and the SAR with our study because (1) our empirical results are in-line with a concave-down log–log SAR, (2) the OFD should be monotonically related to the RAD of a community (Magurran 2007, Borregaard and Rahbek 2010), and (3) a hypothetical limiting case exists in which the OFD is equivalent to the RAD: when only a single individual occupies each quadrat. Furthermore, there has been some encouraging progress in predicting patterns of abundance in species simply from knowledge of species occupancy patterns (He and Gaston 2000, Hui and McGeoch 2007a, Conlisk et al. 2009, Borregaard and Rahbek 2010). However, the OFD and RAD are not completely analogous to one another, and the links between these two patterns deserve further analytical and empirical development.

Conclusions

We have developed a framework for examining the relationship between pairwise turnover and the slope of the SAR. We have found that there is substantial variability in the relationship between these two facets of turnover, and that previous attempts to link them assuming a simple model of scale-invariant species turnover are inadequate. Pairwise turnover and the SAR slope provide unique pieces of information about species distribution and co-occurrence, and by considering both in conjunction we may be better able to understand how these patterns are potentially shaped by colonization–extinction dynamics and variation in the environmental template.

Literature Cited

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

Arita, H. T., and P. Rodriguez. 2002. Geographic range, turnover rate and the scaling of species diversity. Ecography 25:541–550.

Borregaard, M. K., and C. Rahbek. 2010. Causality of the relationship between geographic distribution and species abundance. Quarterly Review of Biology 85:3–25.

Brown, J. H. 1984. On the relationship between abundance and distribution of species. American Naturalist 124:255–279.

Conlisk, E., J. Conlisk, B. Enquist, J. Thompson, and J. Harte. 2009. Improved abundance prediction from presence-absence data. Global Ecology and Biogeography 18:1–10.

Crawley, M. J., and J. E. Harral. 2001. Scale dependence in plant biodiversity. Science 291:864–868.

Diserud, O. H., and F. Ødegaard. 2007. A multiple-site similarity measure. Biology Letters 3:20–22.

Fridley, J. D., R. K. Peet, T. R. Wentworth, and P. S. White. 2005. Connecting fine- and broad-scale species–area relationships of southeastern U.S. flora. Ecology 86:1172–1177.

Hanski, I. 1982. Dynamics of regional distribution: the core and satellite species hypothesis. Oikos 38:210–221.

Harte, J., and A. P. Kinzig. 1997. On the implications of species–area relationships for endemism, spatial turnover, and food web patterns. Oikos 80:417–427.

Harte, J., S. McCarthy, K. Taylor, A. Kinzig, and M. L. Fischer. 1999. Estimating species–area relationships from plot to landscape scale using species spatial-turnover data. Oikos 86:45–54.

He, F. L., and K. J. Gaston. 2000. Estimating species abundance from occurrence. American Naturalist 156:553–559.

He, F. L., and P. Legendre. 2002. Species diversity patterns derived from species–area models. Ecology 83:1185–1198.

Hubbell, S. P. 2001. The unified neutral theory of biodiversity and biogeography. Princeton University Press, Princeton, New Jersey, USA.

Hui, C. 2009. On the scaling patterns of species spatial distribution and association. Journal of Theoretical Biology 261:481–487.

Hui, C., and M. A. McGeoch. 2007a. Capturing the "droopy-tail" in the occupancy–abundance relationship. Ecoscience 14:103–108.

Hui, C., and M. A. McGeoch. 2007b. Modeling species distributions by breaking the assumption of self-similarity. Oikos 116:2097–2107.

Hui, C., and M. A. McGeoch. 2008. Does the self-similar species distribution model lead to unrealistic predictions. Ecology 89:2946–2952.

Hurlbert, A. H., and W. Jetz. 2010. More than "more individuals": the nonequivalence of area and energy in the scaling of species richness. American Naturalist 176:E50–E65.

Jost, L. 2006. Entropy and diversity. Oikos 113:363–375.

Jurasinski, G., V. Retzer, and C. Beierkuhnlein. 2009. Inventory, differentiation, and proportional diversity: a consistent terminology for quantifying species diversity. Oecologia 159:15–26.

Koleff, P., K. J. Gaston, and J. J. Lennon. 2003. Measuring beta diversity for presence–absence data. Journal of Animal Ecology 72:367–382.

Legendre, P., and L. Legendre. 1998. Numerical ecology. Elsevier, Boston, Massachusetts, USA.

Lundholm, J. T. 2009. Plant species diversity and environmental heterogeneity: spatial scale and competing hypotheses. Journal of Vegetation Science 20:377–391.

Magurran, A. E. 2007. Species abundance distributions over time. Ecology Letters 10:347–354.

McGeoch, M. A., and K. J. Gaston. 2002. Occupancy frequency distributions: patterns, artefacts and mechanisms. Biological Reviews 77:311–331.

McGlinn, D. J., P. G. Earls, and M. W. Palmer. 2010. A 12-year study on the scaling of vascular plant composition in an Oklahoma tallgrass prairie. Ecology 91:1872.

McGlinn, D. J., and M. W. Palmer. 2009. Modeling the sampling effect in the species–time–area relationship. Ecology 90:836–846.

McGlinn, D. J., and M. W. Palmer. 2011. Quantifying the influence of environmental texture on the rate of species turnover: evidence from two habitats. Plant Ecology 212:495–506.

Morlon, H., G. Chuyong, R. Condit, S. Hubbell, D. Kenfack, D. Thomas, R. Valencia, and J. L. Green. 2008. A general framework for the distance-decay of similarity in ecological communities. Ecology Letters 11:904–917.

Muldavin, E. 2008. ANPP quadrat sampling data for pinyon–juniper site 1999–2001. Albuquerque, NM: Sevilleta Long Term Ecological Research Site Database: SEV187. http://sev.lternet.edu/project_details.php?id=SEV187

Muldavin, E. 2009. A long-term study to monitor net primary production across four distinct ecosystems: creosote dominant shrubland, black grama dominant grassland, blue grama dominant grassland, pinon/juniper woodland. Albuquerque, NM: Sevilleta Long Term Ecological Research Site Database: SEV129. http://sev.lternet.edu/project_details.php?id=SEV129

Palmer, M. W. 2007. Species–area curves and the geometry of nature. Pages 15–31 in D. Storch, P. L. Marquet, and J. H. Brown, editors. Scaling biodiversity. Cambridge University Press, Cambridge, UK.

Palmer, M. W., R. K. Peet, R. A. Reed, W. Xi, and P. S. White. 2007. A multiscale study of vascular plants in a North Carolina piedmont forest. Ecology 88:2674.

Palmer, M. W., and P. S. White. 1994. Scale dependence and the species–area relationship. American Naturalist 144:717–740.

Peet, R. K., and O. L. Loucks. 1977. Gradient analysis of southern Wisconsin forests. Ecology 58:485–499.

Plotkin, J. B., et al.. 2000. Predicting species diversity in tropical forests. Proceedings of the National Academy of Sciences USA 97:10850–10854.

Rosenzweig, M. L. 1995. Species diversity in space and time. Cambridge University Press, Cambridge, UK.

Scheiner, S. M., A. Chiarucci, G. A. Fox, M. R. Helmus, D. J. McGlinn, and M. R. Willig. 2011. The underpinnings of the relationship between space, time, and species richness. Ecological Monographs 81:195–213.

Shmida, A., and M. V. Wilson. 1985. Biological determinants of species diversity. Journal of Biogeography 12:1–20.

Šizling, A. L., and D. Storch. 2004. Power-law species–area relationships and self-similar species distributions within finite areas. Ecology Letters 7:60–68.

Storch, D., A. L. Šizling, and K. J. Gaston. 2003. Geometry of the species–area relationship in central European birds: testing the mechanism. Journal of Animal Ecology 72:509–519.

Storch, D., A. L. Šizling, and K. J. Gaston. 2007. Scaling species richness and distribution: uniting the species–area and species–energy relationship. Pages 300–321 in D. Storch, P. A. Marquet, and J. H. Brown, editors. Scaling biodiversity. Cambridge University Press, Cambridge, UK.

Tjørve, E. 2009. Shapes and functions of species–area curves (II): a review of new models and parameterizations. Journal of Biogeography 36:1435–1445.

Tjørve, E., and K. M. C. Tjørve. 2008. The species–area relationship, self-similarity, and the true meaning of the z-value. Ecology 89:3528–3533.

Triantis, K. A., M. Mylonas, K. Lika, and K. Vardinoyannis. 2003. A model for the species–area–habitat relationship. Journal of Biogeography 30:19–27.

Tuomisto, H. 2010a. A diversity of beta diversities: straightening up a concept gone awry. Part 1. Defining beta diversity as a function of alpha and gamma diversity. Ecography 33:2–22.

Tuomisto, H. 2010b. A diversity of beta diversities: straightening up a concept gone awry. Part 2. Quantifying beta diversity and related phenomena. Ecography 33:23–45.

White, E. P. 2004. Two-phase species–time relationships in North American land birds. Ecology Letters 7:329–336.

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

Whittaker, R. H. 1967. Gradient analysis of vegetation. Biological Reviews of the Cambridge Philosophical Society 42:207–264.

SUPPLEMENTAL MATERIAL

Appendix A

Jaccard's turnover index in terms of the slope of the species–area relationship (*Ecological Archives* E093-028-A1).

Appendix B

Derivation of the occupancy distributions, matching/mismatching components, and average Jaccard turnover index for communities that minimize and maximize species turnover (*Ecological Archives* E093-028-A2).

Appendix C

The influence of variation in species richness on simulated turnover patterns (*Ecological Archives* E093-028-A3).

Appendix D

The Sevilleta sampling design (*Ecological Archives* E093-028-A4).