Research

EDITOR'S CHOICE

Ecography 37: 309–320, 2014
doi: 10.1111/j.1600-0587.2013.00407.x

Subject Editor: Jens-Christian Svenning. Accepted 11 October 2013

Scale dependency in the functional form of the distance decay relationship

Jeffrey C. Nekola and Brian J. McGill

J. C. Nekola (jnekola@unm.edu), Dept of Biology, Univ. of New Mexico, Albuquerque, NM 87131, USA. – B. J. McGill, School of Biology and Ecology, Univ. of Maine, 5751 Murray Hall, Orono, ME 04469, USA.

We examine a novel mathematical approach which posits that the decay of similarity in community composition with increasing distance (aka distance decay) can be modeled as the sum of individual species joint-probability vs distance relationships. Our model, supported by analyses of these curves from three datasets (North American breeding birds, North American taiga plants, and tropical forest trees), suggest that when sampling grain is large enough to avoid absences due to stochastic sampling effects, and/or sampling extent is large enough to generate species turnover through the deterministic crossing of environmental and/or geographical range limits, species joint-probability over increasing distance will generally exhibit exponential decay. However, at small scales where occurrence is driven more by stochastic sampling effects, species joint-probability curves exhibit a power-law decay form. Lacking a theoretical prediction of how individual species joint-probability relationships combine to generate community distance decay, we also performed a meta-analysis of 26 ecological and 4 human-system datasets, using non-linear regression to mean and quantile non-linear regression at tau = 0.95 for linear, exponential, and power-law decay forms. These analyses demonstrate that the functional form of community distance decay – as shown by comparison of AIC ranks – is largely determined by observational scale, with power law decay prevailing within domains where the species pool remains constant, while exponential decay prevails at larger scales over which the species pool varies, paralleling the patterns predicted in our mathematical approach.

The decay of compositional similarity with increasing inter-observation distance (aka distance decay) is a ubiquitous pattern found across a wide array of physical, biological, social and other complex systems (Tobler 1970, Nekola and Brown 2007). It has garnered considerable interest within the fields of community ecology (Condit et al. 2002, Green et al. 2004, Soininen et al. 2007), macroecology (McKnight et al. 2007, Blanchette et al. 2008, Qian et al. 2009) and conservation biology (Bell 2003, Steinitz et al. 2005), and has been suggested to shed important insights into underlying fundamental mechanisms (Nekola and White 1999, Hubbell 2001, Tuomisto et al. 2003, Gilbert and Lechowicz 2004).

In spite of this interest, controversy exists regarding its expected functional form. While a number of empirical analyses have firmly established exponential decay (Nekola and White 1999, Qian et al. 2005, 2009, Jobe 2007), linear (Blanchette et al. 2008, Perez-del-Olmo et al. 2009) and power law (Harte et al. 1999, Condit et al. 2002, Green et al. 2004) forms have also been documented. In spite of this, all theoretical derivations of the distance decay relationship (Harte and Kinzig 1997, Hubbell 2001, Chave and Leigh 2002, Houchmandzadeh 2008, Morlon et al. 2008) generate power-law or power-law-like forms.

Here we begin to develop analytical underpinnings for a theory of community distance decay based on joint-probabilities for individual species. We show that: 1) expectation of exponential or power-law forms for individual species joint-probability curves is a function of sampling scale, and 2) the community distance decay curve is a sum of single species joint-probability curves. Because a theoretical prediction for the shape of community distance decay curves from individual species joint-probability curves is not analytically tractable (except in a few specific cases), we consider this issue through empirical meta-analysis of 26 community datasets collected over both spatial and temporal extents, covering a large range of ecosystems (simulated data to abyssal plain to continental mountains, arctic tundra to tropics), taxa groups (including mammals, birds, lepidoptera, robber flies, land snails, vascular plants, bryophytes), sample grains (100 m^2 – subcontinental and daily-yearly) and total extent (3–9000 km and 1–25 yr). Four human system examples are also considered (1 spatial, 3 temporal; Nekola and Brown 2007) to

identify if general patterns in community distance decay functional forms exist.

Individualistic model of distance decay

Single species distance decay

We begin with consideration with the decay of joint-probability (Palmer 2005) for a single species/agent/event with increasing inter-observation distance. Specifically, if comparing two plots L and M of the same size A at distance d apart (Fig. 1) for species i, then:

$$p_i(d_{LM}) = \text{Prob}\{\text{species } i \text{ present at } M \mid \text{species } i \text{ present at } L\} \quad (1)$$

or more generally:

$$p_i(d) = \text{Prob}\{\text{species present at any site distance } d \text{ from a site where species } i \text{ is present}\} \quad (2)$$

While this definition has a clear relationship to the occupancy function used by numerous authors (Harte et al. 1999, He and Legendre 2002), in these previous formulations p has been a function of the plot area, A, rather than the distance, d, from reference plot L. Thus, our approach is the first to explicitly consider spatial structure.

It is clear that under biologically realistic conditions where species have finite ranges, $p_i(d)$ needs to fit the following criteria: 1) $p_i(0) = 1$; 2) $p_i(\infty) = 0$; 3) $p_i(d_{AB}) = p_i(d_{BA})$ (i.e. isotropy or direction does not matter); 4) $p_i(d_1) \geq p_i(d_2)$ if $d_1 < d_2$ (i.e. p_i monotonically decreasing); 5) $p_i(d_1 + d_2) = g(p(d_1), p(d_2))$ for some functional form g if d_1 and d_2 are linear and contiguous as shown in Fig. 1 (i.e. decomposability with distance). Note that 1, 2 and 4 combined can be used to show that $0 \leq p_i(d) \leq 1$ if $0 \leq d \leq \infty$.

There are an infinite number of functional forms that fit these criteria. Three that have been commonly used to analyze community distance decay (Nekola and White 1999, Soininen et al. 2007) include the linear [$p_i(d) = 1 - cd$], exponential [$p_i(d) = \exp(-cd)$], and power law [$p_i(d) = d^{-c}$]. Additionally, Harte (2011) suggests a logarithmic decay [$p_i(d) = 1 - c^*\log(d)$] form. All contain a single parameter, c, which denotes decay rate. Note that the power law form violates criterion 1, the linear form violates criterion 2, while the logarithmic form violates both. As a result, these functional forms can in practice be used only over limited intervals: [0, d_{max}] where $d_{max} = d$ at which $p_i(d) = 0$ for the linear, [$\in$, ∞] for the power law, and [$\in$, d_{max}] for the logarithmic. While Palmer (2005) calculated $p_i(d)$ for tropical trees at the La Selva Biological Station in Costa Rica, he did not propose that it take on any particular mathematical form. However, his plot of $p_i(d)$ across all species and sites appears to demonstrate an exponential trace.

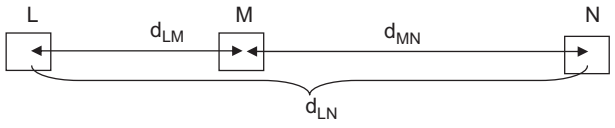

Figure 1. Heuristic diagram demonstrating the relationship between sample plots and distances used in the theoretical model.

What form should g in criterion 5 take? This function describes how p_i should behave when a long distance d is decomposed into two smaller distances d_1 and d_2 (where $d_1 + d_2 = d$ and lie along a straight line; Fig. 1). Clearly it should be commutative and associative. The most obvious, simple forms are $g(x,y) = x + y$ or $g(x,y) = x^*y$. However, the additive form violates criterion 4 [e.g. $p(6) = p(4 + 2) = p(4) + p(2) > p(4)$ but $p(6) < p(4)$] and can easily produce $p_i > 1$. The multiplicative form fits well with our intuitive sense of how similarity decays [i.e. $p_i(d_1 + d_2) = p_i(d_1)^*p_i(d_2)$], and is essentially the law of independent probabilistic events [P(A and B) = P(A)*P(B)]. Note that $p_i(d_1 + d_2) = p_i(d_1)p_i(d_2)$ can only be satisified by $p_i(d) = \exp(cd)$ with criteria 1 and 2, requiring that $c < 0$. A negative exponential form for the decay of joint occurrence probability with increasing distance thus conforms well to criteria 1–5.

It is instructive to proceed with a more rigorous examination of the form of g. Considering the straight line connecting L to M to N in Fig. 1, via the law of total probability (Karlin and Taylor 1975):

$$\begin{aligned}&\text{P(present at N}|\text{present at L)}\\ &\quad = \text{P(present at N}|\text{present at M)}\\ &\qquad \times \text{P(present at M}|\text{present at L)}\\ &\qquad + \text{P(present at N}|\text{NOT present at M)}\\ &\qquad \times \text{P(NOT present at M}|\text{ present at L)}\end{aligned} \quad (3)$$

or:

$$p_i(d) = p_i(d_2)p_i(d_1) + X(d_2)(1 - p_i(d_1)) \quad (4)$$

where: $X(d)$ is the probability of a species being present at one point conditioned on it NOT being present at another point [note this is not simply $1 - p_i(d)$].

More simply:

$$p_i(d) = p_i(d_1)p_i(d_2) + Y(d_1,d_2) \quad (5)$$

where $Y(d_1,d_2) = \text{Prob}\{\text{present at N, absent at M}|\text{present at L}\}$ or essentially the probability of encountering an 'occurrence hole' (sensu Rapoport 1982, Hurlbert and White 2007) within the range of species i. $Y = 0$ when species i is never absent at M if present at L and N. In such cases, we have the multiplicative form for g and $p_i(d) = \exp(-cd)$. If $Y \neq 0$, then g is more complex, and $p_i(d)$ is not a simple exponential.

Community distance decay

Assuming $p_i(d)$ is exponential for one species, then what distance decay form should be expected across a community assemblage? As is common in theoretical ecology, we will work with the Sørenson similarity measure due to its tractable mathematics (Plotkin and Muller-Landau 2002, Morlon et al. 2008). Let S_{LM} be the number of species found in both plots L and M, while S_L is the number of species found in plot L and similarly for S_M. We assume approximately $S = S_L = S_M$. because systematic changes in richness will lead to bias in parameter estimation (Jobe

2007). Then Sørenson (d_{LM}) = Sørenson (L,M) = S_{LM}/average (S_L,S_M). If we sum the probability that a species is present in M over all species present in L, this gives:

$$\begin{aligned} \text{Sørenson}\,(L,M) &= \sum\nolimits_{i=1}^{S_L} p_i(d_{LM}) \Big/ S \\ &= \frac{1}{S}\sum_{i=1}^{S} \exp(-c_i d) = E_i\ \exp(-c_i d) \end{aligned} \tag{6}$$

where E_i is the expectation (or averaging) operator taken with respect to i or over all species. We call this the community aggregated distance decay function. This makes the strong assumption that the joint probability decay is independent between all species.

A sum of exponentials as found in Eq. 6 can take on a vast variety of shapes, dependent entirely on the distribution of c_i values. If we treat c_i as a random variable, C, from which the c_i are drawn and using the tools for calculating the expected value of a function of a random variable (Lindgren 1976 p. 113) we have:

$$\text{Sør}(d) = E_i\ \exp(-c_i d) = \int \exp(-xd) f(x)\, dx \tag{7}$$

Where $f(x)$ is the pdf of the random variable describing the distribution of C. A uniform distribution for C across [0, c] yields $1 - [\exp(-c^* d))/(c^* d)]$. In the limiting case of c having a Dirac delta distribution (i.e. c_i is constant with no variation), then Eq. (6) and (7) give community Sørenson similarity as an exponentially decaying function of distance. Assuming C is distributed according to an exponential distribution yields $\text{Sør}(d) = c/(c + d)$ while assuming C is distributed according to a power law yields $\text{Sør}(d) = \Gamma(c)c/(b^{c*} d^c)$. However, as will be shown in the results, C is strongly right-truncated, making a Beta distribution the most likely. While this does not yield an analytically solvable integral, a partial understanding can realized by taking a Taylor-series expansion of $E_i \exp(-c_i d)$. Let C now be the mean value of c_i ($C = \bar{c_i}$) and $\Delta_i = c_i - C$, with the expansion around C:

$$E_i\ \exp(-c_i d) = E_i\ \exp(-(C + \Delta_i)d) = \exp(-Cd)E_i \left[1 + \frac{(\Delta_i d)^2}{2!} - \frac{(\Delta_i d)^3}{3!} + \frac{(\Delta_i d)^4}{4!} - \frac{(\Delta_i d)^5}{5!} + \ldots\right] \tag{8}$$

The infinite series inside the brackets is a function of d that alternates in sign, with $k^n/n!$ converging to zero for large enough n (where $k = \Delta_i d$). Thus, while the series does converge, it occurs in a way that depends on Δ_i and d. In particular if $\Delta_i d$ is > 1 over much of the range of interest for d, the possibility exists for very slow convergence and strong deviations from $\exp(-Cd)$. Nothing further can be said about this situation without a detailed knowledge of the distribution and magnitude of Δ_i relative to d. After examining empirical distributions in the Results, we will present further analysis in the Discussion as to when convergence on $\exp(-Cd)$ may be expected.

Equation 6 can be expanded for the case where the evaluated species pool is not just the species present at L but includes all taxa in the regional pool. To do this, we redefine S to be the richness of the regional species pool, with w_i being the occupancy rate for species i (i.e. the fraction of sites supporting species i) or equivalently the probability that species i is found at site L. Then we have:

$$\text{Sør}(d) = \frac{1}{S}\sum_{i=1}^{S} w_i \exp(-c_i d) = E_i w_i\ \exp(-c_i d) \tag{9}$$

We call this the weighted community aggregate distance decay function. This curve should closely approximate the traditional community-level decay of similarity curve which is generated by calculating pairwise similarity between samples (Nekola and White 1999).

Model assumptions

Two main assumptions underlie our approach. First, to produce an exponential decay of joint-probability over increasing distance for a single species we assumed that $Y(d_1,d_2) = 0$ or equivalently that a species must occur at M if present at L and N. In effect Y informs about the probability that a sample plot falls inside an occurrence hole within a species range. The value of Y should show strong scale dependence: if sample grain is small with significant sampling error and extent does not exceed the environmental and/or geographic range limits of species within the pool, then it is likely that the main reason for compositional turnover will be sampling effects (Nekola and White 1999). In such instances $Y \neq 0$, and because in such situations $p_i(d)$ is not exponential, this form will also not provide a good fit to the community data. The expected shape in such cases should initially show a more rapid decrease in joint-probability than the exponential expectation because of occurrence holes, but will then ultimately asymptote > 0 because the probability of a given species within the pool is also > 0. Although not a rigorous proof, at a minimum the above demonstrates that a more power-law like than exponential form would result from cases where $Y \neq 0$. Multiple theoretical distance decay models assuming such small grains and uniform species pools have all produced power law decay forms (Harte and Kinzig 1997, Chave and Leigh 2002, Houchmandzadeh 2008, Morlon et al. 2008).

When sample grain is large relative to the organism/agent/event (making for guaranteed detection), with extent being large enough to pass beyond species environmental and/or geographic range limits (allowing the community pool to vary across the sample), then the main cause of species turnover will be the exceeding of species distributional limits, minimizing the chance that a species will be absent at M but present at L and N. In these situations $Y = 0$, and the exponential form should represent a good approximation. Note that these conditions may be met at various absolute scales. For instance, if strong local spatial or temporal gradients exist, even over very limited extents the community pool will not be uniform and the exponential form should occur.

The second major assumption is that the decay of each species joint-probability is independent of that exhibited by others. This allows summing of individual decay curves to produce Eq. 6. Such an assumption of independence between co-occurring species is one of the foundational bases of modern community ecology (Gleason 1926, Curtis 1955, Whittaker 1975) that accurately reproduces a number of community patterns (McGill 2010, 2011, He and Legendre 2002, Green and Plotkin 2007). Although species pairs can be identified which clearly show non-independent distributions (e.g. red bellied woodpecker and red-headed woodpecker – Root 1989; American beach and its obligate parasite beachdrops – Voss 1985), these interactions represent such a small fraction of all possible pairwise interactions (Barker and Mayhill 1999) that the assumption of independence is a valid approximation.

Summary of model predictions

1) Individual species will exhibit exponential decay in their joint-probability with increasing distance if and only if observational scale is large enough to allow a varying species pool and to avoid quasi-stochastic occurrence holes.

2) The Sørenson decay of community similarity is equivalent to the occupancy-weighted community average of the individual $p_i(d)$ curves in the regional species pool.

3) The weighted community aggregate curve (Eq. 8) will fit better than the unweighted curve in cases where species are not equally likely across the species pool (or equivalently that the species pool changes over the extent studied), but will fit worse when the species pool is constant across sites.

4) The functional form for community distance decay [$Sør(d)$] depends on the distribution of c_i, the decay rates of for individual joint-probability curves. Unfortunately this distribution has never been studied and we have no empirical knowledge of its shape, making impossible the prediction of community distance decay functional form. Even when all single species joint-probability curves are exponential, community decay curves of innumerable forms can be produced.

Methods

Test of model predictions

Three different datasets were used to test the general accuracy of our approach. The first is the vascular plant data of LaRoi (1967), which was sampled from 34 nine-ha (300×300 m) mature, undisturbed upland white spruce forest plots at relatively equal distances across boreal North America. A total of 220 vascular plant species were encountered, with individual plot richness ranging from 19–56. The second is the North American Breeding Bird Survey (BBS), which is run annually at over 2000 routes across the continental US and southern Canada (Sauer et al. 2011). A 3-min point count is taken every half mile (approximately 0.8 km) for 50 stops along a transect of 25 miles (approximately 40 km). An individual route typically has 50–100 bird species detected and overall almost 600 species of birds have been observed at least once. To avoid issues with detection error and year-on-year variability, we took average abundance over the years 2003–2007, and used only the 1430 routes that had a quality code reflecting appropriate weather and observer ability for all five years. Thus, a bird was counted as being present at a given route if it appeared at least once in five years. The third dataset was the Barro Colorado Island tree plot containing 225 species (Condit et al. 2000). Every tree >10 cm DBH was recorded within each of the 50 one hectare plots. Distance was recorded between hectare midpoints.

Due to their continental extents, we a priori expect the LaRoi and BBS datasets to have individual site compositions drawn from varying species pools. Decay of individual species joint-probabilities in these data should thus generally take on an exponential form. However, because the BCI dataset covers only a 1 km extent, all samples are likely to be drawn from only two nested species pools (young within old forest; Svenning et al. 2004). In this situation, species absences are likely due to stochastic sampling effects and/or disturbance history, with a power-law like decay-form prevailing. In all three cases the community distance decay curve should be the sum of the individual species decay curves, although the fit should be worse for the BCI data because of the greater impact of stochastic sampling errors. While no explicit prediction can be made for the functional form of the community distance decay curve, the heuristic considerations given above suggest that an exponential form may be more likely at large grains and/or extents which sample across varying species pools. We also predict the weighted community aggregate curve should perform better for the La Roi and BBS data since they have large enough extents/strong enough gradients to exceed the limits of individual species, while conversely the BCI data should be better fit by an unweighted curve.

To test the individual species portion of the model (prediction 1), for each dataset we fit an exponential decay curve to the each species joint-probability by calculating similarity (1 if the species was present in both plots; 0 of not) and distance between each occupied site and all others. Because calculations were only conducted from occupied sites, we constrained similarity to equal 1 at the origin and fit only a single parameter to the decay function. Other functional forms (e.g. linear, power law, logarithmic) were not fit because of their violations of criteria 1 and/or 2. We then calculated mean similarity across distance bins of 200 km (LaRoi and BBS) or 200 m (BCI). Due to the computational challenges in the 1403 site $\times 600$ species BBS data (reflecting a total potential number of $\sim 6 \times 10^8$ calculations), prior to model fit we randomly removed absent sites in order to generate occupancy rates of approximately 10%. We then sampled 10 000 individuals from all site by site pairs. We recorded the c and r^2 value for each species. To calculate the community aggregated curve (Eq. 6), we predicted values every 200 km or 200 m, took their average, and plotted a line through the data. We calculated the weighted community aggregate curve (Eq. 7) similarly except that each species was weighted by its observed occupancy. Because singleton and doubleton species do not possess enough degrees of freedom to accurately fit an exponential model (singletons are points in space while doubletons can only demonstrate linear patterns), we

also calculated a community aggregate curve in which each species was equally weighted ($w_i = 1/S$) after removal of singletons and doubletons. We term this the filtered community aggregate curve.

The traditional community distance decay curve was also calculated for each dataset. Because of its use in model formulation, Sørenson similarity was used. Since the total BBS data would represent almost 2 000 000 pairwise comparisons, we limited analysis to 10 000 randomly-drawn pairs. To these data we then fit a linear decay curve using ordinary least squares (OLS) regression and an exponential and power-law decay curves using non-linear least squares (NLS) regression. Note that under the assumption of normal errors, OLS and NLS give maximum likelihood estimates allowing the use of the Akiake information criteria (AIC), which we used to compare the three models. We also plotted a LOESS line on these plots to allow for easy visual comparison of the central tendency in the empirical data vs the community aggregated curves.

Meta-analysis of community similarity datasets

Lacking a theoretical prediction for the functional form of community distance decay, we also conducted a meta-analysis of 26 ecological and 4 human-system datasets. Two datasets represent simulated community turnover data using COMPAS (Minchin 1987), with composition lists being generated for 50 randomly-spaced sites along a single gradient containing 20 species. Two different turnover scenarios were used: constant turnover where all species possess the same modal abundance, standard deviation, and are equally spaced along the gradient, and lognormal turnover, where species modes, standard deviations, and gradient placement follow a lognormal distribution. The remaining data are empirical, represent examples of both spatial and temporal decay, and attempt to maximize ecosystem range, taxa groups, sample grains, and extents. Information regarding the habitat type, taxa group, maximum extent, sample grain, number of samples, and additional notes on these datasets is found in Table 2. These data include: Galapagos Islands vascular plant flora (Preston 1962); North America upland taiga vascular plants (LaRoi 1967) and bryophytes (LaRoi and Stringer 1976); fen-restricted vascular plants in northeastern Iowa (Nekola 1994); Martha's Head to Bermuda abyssal plain bivalves (Allen and Sanders 1996); Belize terra firma tropical forest (Bird 1998); entire vascular plant floras for all algific talus slope sites along Buck Creek, Clayton County, Iowa (Nekola 1999); Peruvian and Panamanian terra firma forest (Condit et al. 2002); Barro Colorado Island terra firma forest (Condit et al. 2000); New Zealand land snails (Barker 2005); eastern North American and Asian vascular plants (Qian et al. 2005); Iowa county butterfly faunas (Schlicht et al. 2007); small rodent and winter annuals plants at Portal, Arizona (Ernest et al. 2009); eastern North American rock outcrop, Atlantic coastal plain acidophile, and southwestern USA montane land snail faunas (Nekola 2010); North American Breeding Bird Survey (Sauer et al. 2011); southern Flint Hills tallgrass prairie vascular plants (McGlinn and Palmer 2011); seasonal butterfly and skipper emergences at the Rock Island Preserve, Linn County, Iowa (F. Olsen pers comm.); and seasonal robberfly emergences at the Sevilleta LTER (H. Lease pers. comm.). Representative human-system datasets include: concert setlists for Cowboy Junkies (<www.setlist.com>) and Phish (<www.ihoz.com/PhishStats.html>) performances; all commercially available garden vegetable varieties in the US and Canada as reported by the Garden Seed Inventory of the Seed Savers Exchange; all ingredients reported for the cuisines of Ethiopia, Hungary, India, Iran, Ireland, Korea, Mexico, Norway, Puerto Rico, and Thailand (Smith 1990). The Cowboy Junkies setlist, Garden Seed Inventory, and global cuisine distance decay analyses were previously reported in Nekola and Brown (2007).

For each of these datasets, Jaccard similarity was calculated between each pairwise combination of samples, while distance was calculated as the linear difference in space or time between observations. The Jaccard metric was employed because of its common use in distance decay analyses (Nekola and White 1999). The Jaccard and Sørenson indices differ only slightly in their denominator values, and perform in a highly similar (and for all intents, identical) fashion. These data were then subjected to non-linear regression to the mean using the 'nls' routine in R, and quantile regression using the 'nlrq' routine in R with tau = 0.95. Quantile regression was used because the central tendency may not be as important in the distance decay relationship as the shape of the upper bound (Rocchini and Cade 2008). For each of these regression approaches three different functional forms were fit to the data: $S_d = S_i - cd$ (linear decay); $S_d = S_i\ e^{-cd}$ (exponential decay) and $S_d = Ad^{-c}$ (power-law decay) where S_d = intersample similarity at distance d; S_i = initial similarity or nugget (Nekola and White 1999); A = arbitrary intercept; d = intersample distance; c = rate of similarity change. We did not fit logarithmic decay [$S_d = a - c(\log(d))$] because we found it to neither outperform power-law fits in the three intensively studied datasets (LaRoi, BBS, BCI) nor in those data which appeared by eye to most likely to express the logarithmic functional form (Tallgrass Prairie Preserve, Iowa Butterfly County occurrences, Portal rodents, and Cowboy Junkes setlists).

AIC values were calculated for each regression and ranked from 1 (least) to 3 (largest) across the three functional forms for each dataset. We then constructed a contingency table for each of the regression methods, showing 1st, 2nd, and 3rd AIC ranks vs linear, exponential, and power-law fits. A Fisher-exact test was conducted on these tables to determine if cell frequency violated null expectations.

Results

Model test

All three model predictions were confirmed across the three datasets. In general, the exponential decay form was found to be a good approximation for taiga vascular plant and continental bird species joint-probabilities with increasing observational distance (Fig. 2). While the vast majority of LaRoi and BBS species were fit by an exponential model

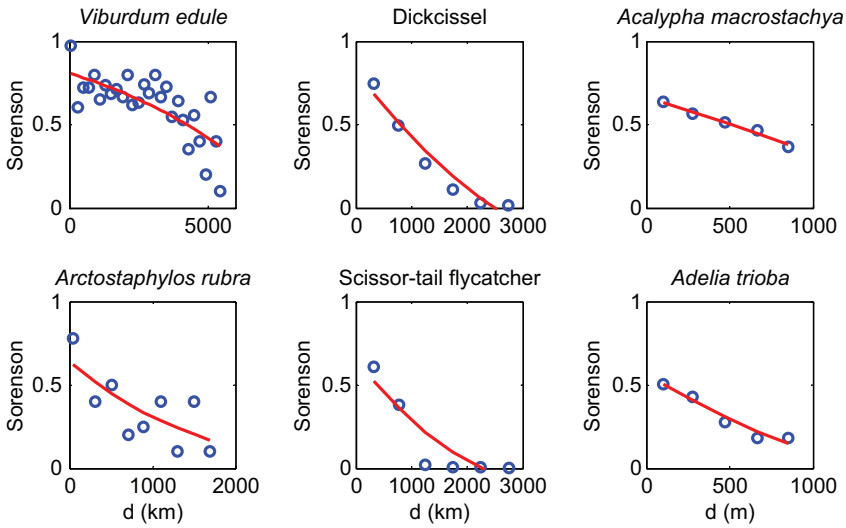

Figure 2. Individual species joint-probability decay with distance [$p_i(d)$] curves for arbitrary species. Blue dots are observed similarities. Red line is LOESS smoothed regression. Left column represents two species from the LaRoi dataset. Middle column represents two species from the Breeding Bird Survey dataset. Right column represents two species from the Barro Colorado Island dataset. Note that for a few species (like *Viburnum edule*) exponential decay is a poor fit. However, as seen in the bottom row of Fig. 3 there only a few such species.

with $r^2 > 0.9$, a noticeably higher frequency of poor fits were noted in the BCI data (Fig. 3). These results confirm prediction 1. It should be noted, however, that six BBS species (1% of total), present at only a handful of sites, demonstrated increasing similarity with distance. These were removed from subsequent analysis. The distribution of c_i values was unimodal with a strong right truncation, making exponential and power law distributions for C a poor description of the empirical data (Fig. 3). Visual comparison of the weighted community aggregate vs the LOESS line for the true Sørenson decay (Fig. 4) shows a remarkable convergence in the LaRoi and BBS data, confirming prediction 2. Table 1 also demonstrates that the community weighted aggregate produces exponential decay parameters that are very close to the observed true community decay values. This is quite remarkable given that no parameters in the weighted community aggregate curve were adjusted to make them fit the true Sørenson decay curve. Although no community aggregate curve provides a good fit to the empirical BCI community decay data, the unweighted curve is a better fit, confirming prediction 3. Linear, exponential and power law forms were then fit to the empirical Sørenson community distance decay (Fig. 4). Community distance decay in the LaRoi and BBS data are best fit by the exponential form, while the BCI data is best fit by a power law (Table 1).

Meta-analyses

All 26 ecological and 4 human-system datasets demonstrated a negative relationship between community similarity and

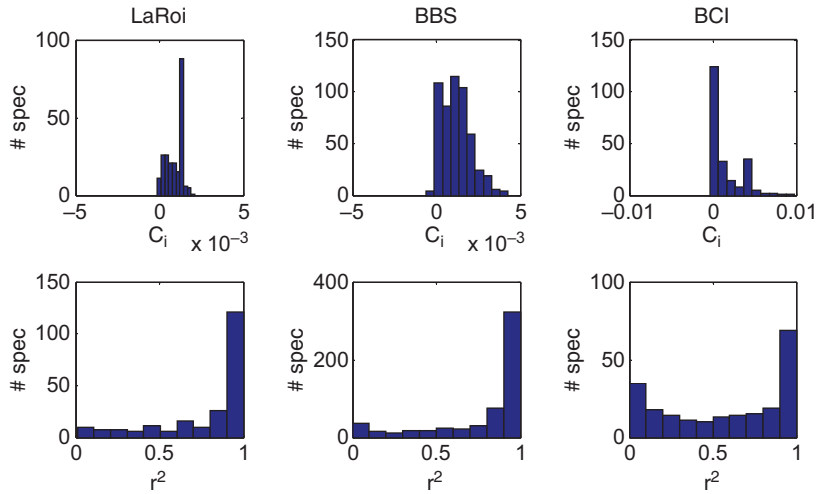

Figure 3. Results of exponential fits to individual species joint-probability occurrence decay. Top row shows distribution of estimated coefficents, c_i. Bottom row shows a histogram of r^2 values for these fits. Left column is for LaRoi data, middle column is for the Breeding Bird Survey, and the right column is for the Barro Colorado Island data.

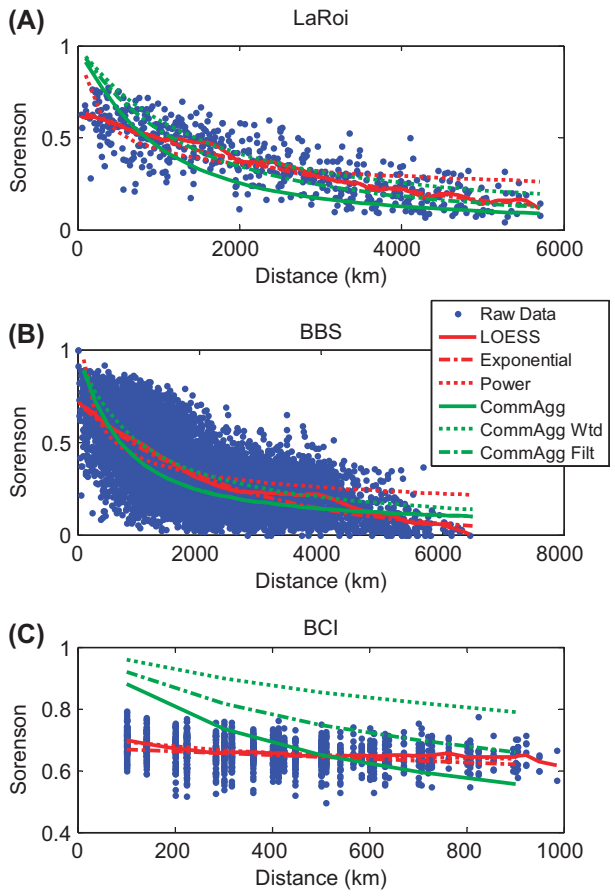

Figure 4. Empirical distance decay relationships, with each blue dot representing a plot pair. Red solid line shows LOESS fit. Red dash-dotted red line shows an exponential fit. Red dotted line shows a power law fit. The green lines show the community aggregated curves. The solid green line equally considers all species (i.e. Eq. 6). The dotted green line weights species by their occupancy (i.e. Eq. 7). The dash-dotted line includes all species equally but eliminates species found in 2 or fewer sites. (A) LaRoi boreal plant data. (B) North American Breeding Bird Survey (note only 10 000 random pairs were used out of the over 2 000 000 possible). (C) Barro Colorado Island data.

distance (Table 2). Based on AIC ranks, each of the three functional forms were found to provide a best-fit to at least some of the ecological datasets under both non-linear regression to mean and quantile regression to tau = 0.95. However, in all four human-system datasets the power-law model was best fit to the data, followed by the exponential model, with the linear model providing the poorest fit. For ecological systems exponential decay was best fit 14 times to the mean, and 13 times to tau = 0.95 (Table 3). Power-law forms were best fit in 8 and 6 cases, respectively, while linear was the best fit in 4 and 7 cases. While exponential equations never provided the poorest fits, in almost half of the datasets linear fits performed the most poorly (13 and 10, respectively), while power-law fits performed worst in at least 50% of datasets (13 and 16, respectively). Fisher-exact tests on these contingency tables indicate a highly significant ($p = 0.00002$ and $p = 0.000014$) deviation of cell frequencies from a uniform null model.

Discussion

These analyses clearly demonstrate that the distance decay of community similarity reflects the aggregate of individual species joint-probability decay functions across the species pool. As shown in both large-scale detailed datasets (BBS, LaRoi), when the grain size and/or extent are large enough to generate exponential individual joint-probability decay, the community decay curve is also best fit by this form. This outcome is replicated in the meta-analysis, where large scale samples crossing into multiple species pools also exhibited exponential community distance decay. Examples include: simulated communities along a strong environmental gradient; vascular plant flora of the Galapagos archipelago; vascular plant and bryophyte assemblages across boreal North America; vascular plants of North American and Asian temperate forests; North American Atlantic coastal plain land snails; BBS data; and the adult emergence of both butterflies/skippers and robberflies along a seasonal gradient.

Table 1. Summary of results for three datasets.

	True		Community aggregate		Weighted community aggregate	
	C	AIC	C	AIC	C	AIC
LaRoi						
Linear	−0.0001	−2578.8	−0.00012	−125.4	−0.0001	−142.4
Exponential	**−0.00027**	**−2590.4**	**−0.00056**	**−183.5**	**−0.00032**	**−185.1**
Power	−0.286	−2355.1	−0.482	−147.0	−0.353	−149.4
BBS						
Linear	−0.00011	−38294	−0.00007	−156.0	−0.00008	−167.5
Exponential	**−0.00039**	**−39287**	−0.00047	−202.3	**−0.00034**	**−215.1**
Power	−0.370	−37635	**−0.484**	**−211.5**	−0.409	−199.0
BCI						
Linear	−0.00006	−7646	−0.00039	−31.0	−0.00021	−43.2
Exponential	−0.00009	−7647	−0.00059	−34.0	**−0.00024**	**−45.4**
Power	**−0.0367**	**−7677**	**−0.202**	**−39.6**	−0.085	−40.9

Table 2. Source, system, sample scale, and regression information for the analyzed datasets. The upper AIC values and ranks represent statistics generated from non-linear regression to mean, while the lower values represent statistics generated from quantile non-linear regression for tau = 0.95.

Source	Location	Habitat	Taxa group	Max Extent	Grain	No. samples	Notes	AIC value/rank order		
								Linear	Exponential	Power
COMPAS	Simulated					50	Constant turnover	−3665.7/2 4730.4/2	−3956.1/1 3837.2/1	−1478.3/3 5430.9/3
COMPAS	Simulated					50	Lognormal turnover	−1317.6/2 7018.5/2	−1976.0/1 6937.4/1	−721.0/3 7372.3/3
Preston 1962	Galapagos Islands	Volcanic archipelago	Vascular plants	286.2 km	1.1–4640 km^2	18		−222.7/2 51.0/1	−224.6/1 51.6/2	−206.7/3 61.5/3
LaRoi 1967	North America	Upland taiga	Vascular plants	5687.4 km	9 ha	34		−1212.6/3 1770.0/2	−1244.1/1 1729.3/1	−1027.2/2 2116.4/3
LaRoi and Stringer 1976	North America	Upland taiga	Bryophytes	5687.4 km	9 ha	34		−1205.1/2 1810.7/2	−1221.7/1 1804.4/1	−1122.2/3 1919.9/3
Nekola 1994	Northeastern Iowa	Fen	Specialist vascular plants	254.0 km	0.1–50.1 ha	10		−91.1/1 −91.7/1	−91.0/2 −90.7/2	−89.7/3 −87.4/3
Allen and Sanders 1996	Atlantic abyssal plain	Total fauna	Protobranch bivalves	777.5 km	ca 0.1 ha	21	< 3000 m depth only	−286.8/1 498.3/1	−274.4/2 501.9/2	−208.7/3 522.6/3
Bird 1998	Belize	Tropical forest	Canopy trees	205.5 km	1 ha	30		−798.2/3 1290.1/3	−810.0/2 1282.6/2	−947.6/1 1068.2/1
Nekola 1999	Buck Creek valley, Iowa	Algific talus slope	Vascular plants	3.1 km	0.125–1 ha	7		−83.1/1 −109.2/1	−82.9/2 −109.1/2	−77.6/3 −107.5/3
Condit et al. 2000	Barro Colorado Island	Tropical forest	Trees > 10 cm dbh	1 km	1 ha	50		−4167.7/3 −13446/3	−4168.8/2 −13448/2	−4199.0/1 −13496/1
Condit et al. 2002	Panama	Tropical forest	Canopy trees	110.6 km	1 ha	100		−9045.8/3 38969.4/3	−11850/1 37667.5/1	−11125/2 37820.6/2
Condit et al. 2002	Peru	Tropical forest	Canopy trees	102.2 km	1 ha	15		−275.4/3 −105.6/3	−276.1/2 −107.8/2	−305.8/1 −136.2/1
Barker 2005	New Zealand	Total fauna	Land snails	1590.9 km		28		−460.6/2 1197.3/1	−463.0/1 1205.8/2	−416.5/3 1371.6/3
Qian et al. 2005	E North America	Total flora	Vascular plants	1895.3 km	67000–249 000 km^2	25	47 chosen comparisons	−64.6/2 −107.2/1	−67.0/1 −106.6/2	−64.3/3 −38.3/3
Qian et al. 2005	E Asia	Total flora	Vascular plants	1870.6 km	67000–249 000 km^2	18	40 chosen comparisons	−61.4/3 −93.0/2	−76.2/1 −116.9/1	−70.9/2 −10.2/3
Schlicht et al. 2007	Iowa	County faunas	Butterflies and skippers	502.1 km	1042–2524 km^2	60	Limited to well sampled (40 + taxa) counties	−3973.9/3 9840.0/3	−3975.8/2 9839.1/2	−3982.2/1 9838.4/1
Ernest et al. 2009	Portal, Arizona	Desert scrub and grassland	Winter annual plants	16 yr	Season summary for 250 m^2	17	Plot no. 22 from 1989–2005	10.4/3 264.7/3	10.3/2 265.3/1	9.7/1 265.4/2
Ernest et al. 2009	Portal, Arizona	Desert scrub and grassland	Rodents	25 yr	Season summary for 250 m^2	26	Plot no. 22 from 1976–2005	−301.9/3 1227.5/3	−310.9/2 1219.0/2	−326.0/1 1197.8/1

(Continued)

Table 2. (Continued).

Source	Location	Habitat	Taxa group	Max Extent	Grain	No. samples	Notes	AIC value/rank order		
								Linear	Exponential	Power
Nekola 2010	US Atlantic Coast	Acid habitats	Land snails	2013 km	10 000 m^2	15		−147.1/2 70.9/2	−148.5/1 69.9/1	−146.5/3 72.0/3
Nekola 2010	E North America	Rock outcrop	Land snails	3031.7	10 000 m^2	48		−1269.0/2 6446.9/1	−1297.1/1 6462.6/2	−1225.6/3 6553.0/3
Nekola 2010	SW North America	Montane forest	Land snails	841.1 km	10 000 m^2	14		−104.9/1 26.4/2	−102.9/2 26.3/1	−95.7/3 41.0/3
McGlinn and Palmer 2011	Oklahoma	Tallgrass prairie	Vascular plants	15.6 km	100 m^2	151	Comparisons with plot no. 5 only	−358.0/3 7.8/3	−359.8/2 5.6/2	−367.5/1 −11.0/1
McGlinn and Palmer 2011	Oklahoma	Tallgrass prairie	Vascular plants	9 yr	Season summary for 100 m^2	10	Plot no. 5 from 1997–2006	−149.5/3 −156.7/2	−151.0/2 −157.2/1	−155.6/1 −152.2/3
Sauer et al. 2011	North America	Total fauna 2002–2007	Breeding birds	6139.8 km	~16 ha	~3700	1000 draws from all pairwise comparisons	−968.9/2 −8687.5/2	−1084.8/1 −8698.1/1	−878.4/3 −8405.5/3
H. Lease pers. comm.	Sevielleta LTER	Desert scrub	Robber flies	193 d	1 d	26	2006 season	−173.9/3 1289.9/3	−231.7/1 1269.6/1	−227.5/2 1275.2/2
F. Olsen pers. comm.	Rock Island Preserve, Iowa	Tallgrass prairie	Butterflies and skippers	174 d	2–4 h	38	1990 season	−713.1/3 3301.1/3	−741.9/1 3244.2/1	−718.3/2 3277.2/2
Nekola and Brown 2007	Concert setlists	Cowboy Junkies		6856 d	1 show	33	1987 to 2006	−894.8/3 −4559.7/3	−921.3/2 −4657.3/2	−1047.1/1 −4910.7/1
<www.ihoz.com/PhishStats.html>	Concert setlists	Phish		6070 d	1 show	18	1983–2000	−485.5/3 −46.5/3	−511.5/2 −51.0/2	−525.5/1 −69.9/1
Nekola and Brown 2007	Commercially available varieties	Garden Seed Inventory		23 yr	1 Garden Seed Inventory	7	1981 to 2004	−40.4/3 −183.4/3	−50.4/2 −199.4/2	−54.4/1 −231.6/1
Nekola and Brown 2007	Ingredient lists	Global cusines		16500 km	1 ethnic region	10		−126.1/3 −453.2/3	−128.5/2 −454.7/2	−136.6/1 −456.3/1

Table 3. Contingency table of AIC rank vs functional form of the distance decay relationship. The smallest AIC value among the three functional forms was assigned a rank of '1', the second smallest a rank of '2', with the largest being assigned a rank of '3'. The first number in a cell represents the number of datasets achieving that given rank based on AIC values calculated from non-linear regression to mean, while the second represents the number achieving that rank based on quantile non-linear regression for tau = 0.95.

	Mathematical distance decay form		
AIC rank	Linear	Exponential	Power law
1	4/7	14/13	8/6
2	9/9	12/13	5/4
3	13/10	0/0	13/16

Fisher exact test for identical cell frequencies: p = 0.00001966 for non-linear regression to mean; p = 0.00001374 for quantile non-linear regression.

However when community data is sampled at small grains or within a mostly uniform species pool at limited extents (BCI), exponential decay is a poorer describer of joint-probability decay with community distance decay exhibiting a power-law like form. This result was also replicated in the meta-analysis. Examples include: vascular plants within 100 m^2 quadrats along a 15.6 km extent in the Tallgrass Prairie Preserve in northeastern Oklahoma; canopy trees within 1 ha tropical forest stands along a 1 km extent at BCI, a 100 km extent in lowland Peru, and a 200 km extent in Belize; seasonal summaries for rodents within a 250 m^2 sample over a 25 yr extent in Portal, Arizona. While Iowa county butterfly faunas would appear a violation, with power-law decay being favored even though ~25% of the fauna reaching its range limit within the state boundary, it is also important to realize that range-limit species also tend to be rare. Since rare species do not exhibit a distance decay signal (Nekola and White 1999), they will not contribute greatly to the overall shape of the decay function. As a result, the observed power law shape is largely due to the ~75% of the fauna which is cosmopolitan.

Power-law community distance decay has also previously been shown to describe the variation of vascular plant composition at <10 km extents within Colorado subalpine meadows (Harte et al. 1999) and for soil microbes over ~1000 km extents in Australia (Green et al. 2004). It is important to note in this latter example that because of their exemplary long-distance dispersal abilities (Green and Bohannan 2006) local microbe communities are likely drawn from a much more extensive and uniform species pool than would be the case for most macroscopic organism groups.

The predominance of power-law community distance decay in human-systems is also readily explained using this framework. Because of the immense human capacity to move ideas and goods across large spatial scales, and to store and retrieve data across long time periods, absolute range limits for commercial goods or ideas are likely not crossed at either planetary spatial extents or decadal time spans. For instance, global trade over the last 600 yr has allowed essentially all domesticated crops to be available world wide. As a result, many new world domesticates are now common aspects in old world cuisines – e.g. tomatoes in Italian, chiles in Indian/Ethiopian/Chinese/Thai, while old-world domesticates now feature prominently in new world cuisines – e.g. wheat, cinnamon, anise, chickens and pigs in Mexico. The existence of public and private seed banks has also made it possible for 'lost' commercial varieties to be made available after absences of a few decades. And, it is also highly unlikely that either the Cowboy Junkies or Phish would be unable to play any song in their repertoire given more than a few days' rehearsal. As a result, agents/events within the analyzed human systems tend to be drawn from a uniform pool, with absences being largely the result of occurrence holes or stochastic sampling effects. Such situations should – and do – lead to power-law community distance decay.

While we are unable to analytically prove that exponential community distance decay should be expected in situations where individual species joint-probability curves are also largely exponential, consideration of c_i values for LaRoi and BBS in conjunction with Eq. 8 (the Taylor series expansion), helps explain this correspondence. Recall that convergence to $\exp(-Cd)$ is dependent upon how fast $(\Delta_i d)^n/n!$ converges to zero as n increases. Our empirical data (Fig. 3) indicate that the distribution of Δ_i (or equivalently, c_i) is strongly truncated on both tails, making the values of c_i strongly centered around the mean. As a result c_i (or more specifically $\Delta_i = c_i - \bar{c}_i$) is never larger than a small multiple of $\bar{c}_i$ (e.g. 2–3 times). Second we note that $\bar{c}_i$ is order-of-magnitude close to but slightly larger than $1/d$ such that $c_i d$ is slightly greater than 1 (i.e. approximately in the interval [1,10] for most of the values of d in the decay curve graphs). Since, by tail truncation, Δ_i is never larger than $2\bar{c}_i$ or $3\bar{c}_i$, then $\Delta_i d$ also must fall in the range of roughly 1–10. Additionally note that $(\Delta_i d)^n/n!$ converges on zero quite quickly when $\Delta_i d$ is in the range [1,10]. Also note $\Delta_i d$ is less than 1 when d is close to zero, which also forces a fast convergence on zero. As a result, the Taylor series expansion necessitates a community distance decay function close to $\exp(-Cd)$ when species joint-probability occurrence is exponential and the distribution of c_i values is strongly truncated and order of magnitude close to $1/d$. While providing a numerically based understanding for the preponderance of observed exponential-like community distance decay, this analysis can not be construed as a general proof, as the end result is strongly dependent upon the form of the joint-probability decay curve and c_i distribution properties.

These analyses also document two additional important additional insights regarding the nature of ecological communities. First, it is striking how similar the parameterized and predicted coefficients are between the true and weighted aggregate community distance decay function for large-scale data (Fig. 4; Table 1). This correspondence provides strong evidence of individualistic species sorting (Gleason 1926). Second, the different expected forms for distance decay within (power law) vs between (exponential) communities allows for an empirical test of the observational scales under which communities exist. For instance, in Panamanian tropical forest, strong power-law decay apparent at distances of 1 km or less indicate that samples are being drawn from within the same community. However, the exponential decay observed at distances of 5–100 km indicates that at this larger scale different communities are being surveyed.

Since its popularization within ecological research over a decade ago, there has been a mismatch between theoretical models and empirical results of distance decay analysis. We hope the approach described within this paper helps resolve this conundrum. Specifically, sampling scales associated with small grain, limited extents, and absent/limited environmental gradients are expected to have power-law community distance decay, while sampling scales associated with large grain and strong environmental gradients are expected to have exponential community distance decay.

Acknowledgements – The following graciously provided us with their raw datasets for analysis: Gary Barker (New Zealand land snail data); Matt Barthel (Garden Seed Inventory data); Neil Bird (Belize tropical forest data); Jim Brown, Morgan Ernst and Tom Valone (Portal rodent and vascular plant data); Richard Condit (Panama and Peru tropical forest data); Hilary Lease (Sevilleta LTER robberfly data); Frank Olsen (Rock Island Preserve lepidoptera data); Mike Palmer and Dan McGlinn (Tallgrass Prairie Preserve data); Peter White (eastern Asian and North American regional flora data). Peter Minchin provided the COMPAS code for simulating community turnover along environmental gradients. Initial work was funded through a Santa Fe Inst. workshop grant.

References

Allen, J. A. and Sanders, H. L. 1996. The zoogeography, diversity and origin of the deep-sea protobranch bivalves of the Atlantic: the epilogue. – Prog. Oceanogr. 38: 95–153.

Barker, G. M. 2005. The character of the New Zealand land snail fauna and communities: some evolutionary and ecological perspectives. – Rec. West. Aust. Mus. Suppl. 63: 52–102.

Barker, G. M. and Mayhill, P. C. 1999. Patterns of diversity and habitat relationships in terrestrial mollusc communities of the Pukeamaru Ecological District, northeastern New Zealand. – J. Biogeogr. 26: 215–238.

Bell, G. 2003. The interpretation of biological surveys. – Proc. R. Soc. B 270: 2531–2542.

Bird, N. M. 1998. Sustaining the yield: improved timber harvesting practices in Belize, 1992–1998. – Natural Resources Inst., Univ. of Greenwich, NRI Catalog Services, CAB International, Wallingford, UK.

Blanchette, C. A. et al. 2008. Biogeographical patterns of rocky intertidal communities along the Pacific coast of North America. – J. Biogeogr. 35: 1593–1607.

Chave, J. and Leigh, E. G. Jr 2002. A spatially explicit neutral model of β-diversity in tropical forests. – Theor. Popul. Biol. 62: 153–168.

Condit, R. et al. 2000. Spatial patterns in the distribution of tropical tree species. – Science 288: 1414–1418.

Condit, R. et al. 2002. Beta-diversity in tropical forests. – Science 295: 666–669.

Curtis, J. T. 1955. A prairie continuum in Wisconsin. – Ecology 36: 558–566.

Ernest, S. K. M. et al. 2009. Long-term monitoring and experimental manipulation of a Chihuahuan Desert ecosystem near Portal, Arizona, USA. – Ecology 90: 1708.

Gilbert, B. and Lechowicz, M. J. 2004. Neutrality, niches, and dispersal in a temperate forest understory. – Proc. Natl Acad. Sci. USA 101: 7651–7656.

Gleason, H. A. 1926. The individualistic concept of plant association. – Bull. Torrey Bot. Club 53: 7–26.

Green, J. L. and Bohannan, B. J. M. 2006. Spatial scaling of microbial biodiversity. – Trends Ecol. Evol. 21: 501–507.

Green, J. L. and Plotkin, J. B. 2007. A statistical theory for sampling species abundances. – Ecol. Lett. 10: 1037–1045.

Green, J. L. et al. 2004. Spatial scaling of microbial eukaryote diversity. – Nature 432: 747–750.

Harte, J. 2011. Maximum entropy and ecology. – Oxford Univ. Press.

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

Harte, J. et al. 1999. Estimating species–area relationships from plot to landscape scale using species-turnover data. – Oikos 86: 45–54.

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

Houchmandzadeh, B. 2008. Neutral clustering in a simple experimental ecological community. – Phys. Rev. Lett. 101: 078103.

Hubbell, S. P. 2001. The unified neutral theory of biodiversity and biogeography. Monographs in population biology no. 32. – Princeton Univ. Press.

Hurlbert, A. H. and White, E. P. 2007. Ecological correlates of geographical range occupancy in North American birds. – Global Ecol. Biogeogr. 16: 764–773.

Jobe, R. T. 2007. Estimating landscape-scale species richness: reconciling frequency- and turnover-based approaches. – Ecology 89: 174–182.

Karlin, S. and Taylor, H. M. 1975. A first course in stochastic processes, 2nd ed. – Academic Press.

LaRoi, G. H. 1967. Ecological studies in the boreal spruce-fir forests of the North American taiga. I. Analysis of the vascular flora. – Ecol. Monogr. 37: 229–253.

LaRoi, G. H. and Stringer, M. H. 1976. Ecological studies in the boreal spruce-fir forests of the North American taiga. II. Analysis of the bryophyte flora. – Can. J. Bot. 54: 619–643.

Lindgren, B. W. 1976. Statistical theory. – Macmillan Publishing.

McGill, B. J. 2010. Towards a unification of unified theories of biodiversity. – Ecol. Lett. 13: 627–642.

McGill, B. J. 2011. Linking biodiversity patterns by autocorrelated random sampling. – Am. J. Bot. 98: 481–502.

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

McKnight, M. W. et al. 2007. Putting beta-diversity on the map: broad scale congruence and coincidence in the extremes. – PLoS Biol. 5: 2424–2432.

Minchin, P. R. 1987. Simulation of multidimensional community patterns: towards a comprehensive model. – Plant Ecol. 71: 145–156.

Morlon, H. et al. 2008. A general framework for distance-decay of similarity in ecological communities. – Ecol. Lett. 11: 904–917.

Nekola, J. C. 1994. Environment and vascular flora of northeastern Iowa fen communities. – Rhodora 96: 121–169.

Nekola, J. C. 1999. Paleorefugia and neorefugia: the influence of colonization history on community pattern and process. – Ecology 80: 2459–2473.

Nekola, J. C. 2010. Acidophilic terrestrial gastropod communities of North America. – J. Molluscan Stud. 76: 144–156.

Nekola, J. C. and White, P. S. 1999. Distance decay of similarity in biogeography and ecology. – J. Biogeogr. 26: 867–878.

Nekola, J. C. and Brown, J. H. 2007. The wealth of species: ecological communities, complex systems, and the legacy of Frank Preston. – Ecol. Lett. 10: 188–196.

Palmer, M. W. 2005. Distance decay in an old-growth neotropical forest. – J. Veg. Sci. 16: 161–166.

Perez-del-Olmo, A. et al. 2009. Not everything is everywhere: the distance decay of similarity in a marine host–parasite system. – J. Biogeogr. 36: 200–209.

Plotkin, J. B. and Muller-Landau, H. C. 2002. Sampling the species composition of a landscape. – Ecology 83: 3344–3356.
Preston, F. W. 1962. The canonical distribution of commonness and rarity. – Ecology 43: 185–215, 410–432.
Qian, H. et al. 2005. Beta diversity of angiosperms in temperate floras of eastern Asia and eastern North America. – Ecol. Lett. 8: 15–22.
Qian, H. et al. 2009. The latitudinal gradient of beta diversity in relation to climate and topography for mammals in North America. – Global Ecol. Biogeogr. 18: 111–122.
Rapoport, E. H. 1982. Areography: geographical strategies of species. – Pergamon Press.
Rocchini, D. and Cade, B. S. 2008. Quantile regression applied to spectral distance decay. – Geosci. Remote Sens. Lett. 5: 640–643.
Root, T. 1989. Atlas of wintering North American birds. – Univ. of Chicago Press.
Sauer, J. R. et al. 2011. The North American Breeding Bird Survey, results and analysis 1966–2009. Version 3.23.2011. – USGS Patuxent Wildlife Research Center, Laurel, MD.
Schlicht, D. W. et al. 2007. The butterflies of Iowa. – Univ. of Iowa Press.
Soininen, J. et al. 2007. The distance decay of similarity in ecological communities. – Ecography 30: 3–12.
Smith, J. 1990. The frugal gourmet on our immigrant ancestors. – Harper-Collins.
Steinitz, O. et al. 2005. Predicting regional patterns of similarity in species composition for conservation planning. – Conserv. Biol. 19: 1978–1988.
Svenning, J.-C. et al. 2004. Ecological determinism in plant community structure across a tropical forest landscape. – Ecology 85: 2526–2538.
Tobler, W. R. 1970. A computer movie simulating urban growth in the Detroit region. – Econ. Geogr. 46: 234–240.
Tuomisto, H. et al. 2003. Dispersal, environment, and floristic variation of western Amazonian forests. – Science 299: 241–244.
Voss, E. G. 1985. Michigan flora. – Cranbrook Inst. of Science, Bloomfield Hills, MI, Bulletin 59.
Whittaker, R. H. 1975. Communities and ecosystems. – MacMillan Publishing.