COMMUNITY ECOLOGY

Toward ecologically explicit null models of nestedness

Jeffrey E. Moore · Robert K. Swihart

Abstract A community is ''nested'' when species assemblages in less rich sites form nonrandom subsets of those at richer sites. Conventional null models used to test for statistically nonrandom nestedness are under- or over-restrictive because they do not sufficiently isolate ecological processes of interest, which hinders ecological inference. We propose a class of null models that are ecologically explicit and interpretable. Expected values of species richness and incidence, rather than observed values, are used to create random presence–absence matrices for hypothesis testing. In our examples, based on six datasets, expected values were derived either by using an individually based random placement model or by fitting empirical models to richness data as a function of environmental covariates. We describe an algorithm for constructing unbiased null matrices, which permitted valid testing of our null models. Our approach avoids the problem of building too much structure into the null model, and enabled us to explicitly test whether observed communities were more nested than would be expected for a system structured solely by species–abundance and species–area or similar relationships. We argue that this test or similar tests are better determinants of whether a system is truly nested; a nested system should contain unique pattern not already predicted by more fundamental ecological principles such as species–area relationships. Most species assemblages we studied were not nested under these null models. Our results suggest that nestedness, beyond that which is explained by passive sampling processes, may not be as widespread as currently believed. These findings may help to improve the utility of nestedness as an ecological concept and conservation tool.

Keywords Forest songbirds · Random placement model · Site–species matrix · Species–abundance · Species–area

Introduction

An important goal in community ecology is to explain species assemblage patterns across a set of sites that are spatially distinct but subject to the same regional species pool, such as islands in an archipelago or habitat fragments in a disturbed landscape. A set of spatially distinct communities is considered to be ''nested'' when the assemblages in less species-rich sites form nonrandom subsets of those at progressively richer sites (Patterson and Atmar 1986; Patterson 1987). In a perfectly nested system, any species present at a given site will be found at all sites with equal or greater species richness, and any species absent from a particular site will be absent from all less rich sites as well. In reality, natural systems deviate from perfect order; some species that are absent from one site occur at some species-poorer sites, whereas some species present at a given site are absent from a handful of richer sites. Thus, statistical evaluation is required to determine whether observed species distributions are significantly more nested than would be expected under some null process.

J. E. Moore · R. K. Swihart
Department of Forestry and Natural Resources,
Purdue University, West Lafayette,
IN 47907-2061, USA

J. E. Moore (✉)
Duke Center for Marine Conservation,
135 Duke Marine Lab Road, Beaufort,
NC 28516-9721, USA
e-mail: jemoore@duke.edu

A fundamental challenge of nestedness analysis is to determine what constitutes *nonrandom* pattern. We emphasize that there is some expected amount of species subsetting that will occur randomly as a result of fundamental and well-documented ecological principles, such as variation in species commonness, or relationships between species richness and factors such as island area or habitat diversity. We believe that for nested-subsets theory to be ecologically informative and useful as a conservation tool, it should strive to identify unique patterns that are not merely artifacts of these basic principles. Such revelations would help detect the presence of additional mechanisms working to structure a community, such as strong inter-specific interactions (e.g., competition, predation), strong ordering of extinction or colonization risks, etc.

Unfortunately, researchers have long been hindered in their ability to appropriately test for nonrandom community nestedness, because conventionally available null models do not sufficiently isolate the ecological processes that we wish to test. Most null models used in the nestedness literature are relatively uninformative because they only allow comparison of observed pattern to those expected to arise under unrealistically simple processes (Wright et al. 1998). In contrast, unduly restrictive null models are available that may mask the detection of non-random pattern because too much of that pattern may be built into the null model itself (Diamond and Gilpin 1982; Wright and Reeves 1992; Cook and Quinn 1998; Jonsson 2001). The goal then, is to develop a null model that is neither overly simplistic nor overly restrictive, but that is suitable for testing specific ecological hypotheses arising from nested-subsets theory.

In this paper, we review null models that have been proposed for nestedness analyses and other site-species matrix analyses in community ecology, particularly focusing on the ecological interpretation of each null model, and the susceptibility of each to rejection under different circumstances. We then describe a new class of null models that generates pattern expected to occur under an ecologically explicit set of conditions that are realistic but not unduly restrictive. Our null models allow one to ask specific questions, e.g., are the observed communities more nested than would be expected for a system structured solely:

(a) by random placement of individuals across sites of different sizes?
(b) by empirical relationships between richness and habitat diversity?
(c) by empirical relationships between richness and area + isolation?

We used these and conventional models to test for nestedness of several datasets that encompassed a range of taxa.

Data matrices

The data for nestedness analyses consist of binary indicators of species occurrence (incidence) at a set of sampled sites, arranged in a matrix with rows representing sites, and columns representing species (Fig. 1). Each matrix element a_{ij} contains either a 1 or a 0, indicating whether species j was present or absent at site i, respectively. Depending on the metric used to measure nestedness, columns (species) of the matrix are typically sorted in descending order of species incidence (i.e., the species that occurred at the most sites is in the left-most column), whereas rows (sites) may be sorted in descending order of either species richness or of the value of some site variable that the researcher believes may be driving the nestedness pattern. One metric that does not require the matrix to be presorted is that of the nestedness temperature calculator (Atmar and Patterson 1993, 1995), which algorithmically sorts the matrix to a configuration that minimizes its "temperature" (T). Most nestedness metrics and their calculations have been reviewed elsewhere (Wright et al. 1998), so we do not duplicate that effort here (also see Brualdi and Sanderson (1999) for a more recent metric). Our point in describing the data matrix, however, is that when the rows and columns are sorted, either by hand or by computer, 1s tend to cluster toward the upper left triangle of the matrix, and it is the degree to which they do so that dictates the value of the various metrics. Importantly, this tendency for upper left triangular fill is expected for a set of communities in which species vary in incidence rates and in which sites vary in richness as a function of covariates such as area, isolation, or productivity. This expectation has important implications for interpretation of different null models.

Conventional null models

Regardless of how the nestedness metric is calculated, its statistical significance is usually evaluated by Monte Carlo methods. The elements of the observed site–species matrix are randomly shuffled a large number of times according to a set of constraints that define the null model. If the metric value for the observed data falls above or below a certain cutoff percentile (α) of the simulated metric values, the system is deemed significantly nested. Following the working definition of Gotelli and Graves (1996, pp 3–4), and for the purposes of analyzing site–species matrices, we view a null model as one that should randomize the data to generate a stochastic pattern with respect to a particular ecological process of interest. It is critical that certain elements of the data be contrained, so that the randomization produces "a pattern that would be expected in the absence of a particular ecological mechanism." Thus, the null model should incorporate a certain amount of

Fig. 1a–e Examples of random site × species matrices (30 × 15) generated under four different null models, based on data for 15 forest songbird species in fragmented landscapes of northern Indiana, collected May–June 2001–2003. Model **a** *R00* fills all matrix elements equiprobably; **b** *R0* and **c** *R0'* constrain only row *or* column totals to equal those of observed data, respectively; **d** R_{ff} constrains both row and column totals to equal those of **e** observed data. Matrices are sorted in order of decreasing species richness (top-down) and species incidence (left–right). *Solid diagonal line* is to help visualize matrix packing. *Shaded* 1s indicate species occurrence at a site. *T* is the matrix temperature

a) *R00*   *T* = 65.33

b) *R0*   *T* = 58.3

c) *R0'*   *T* = 22.5

d) R_{ff} *T* = 14.3

e) Observed data   *T* = 14.7
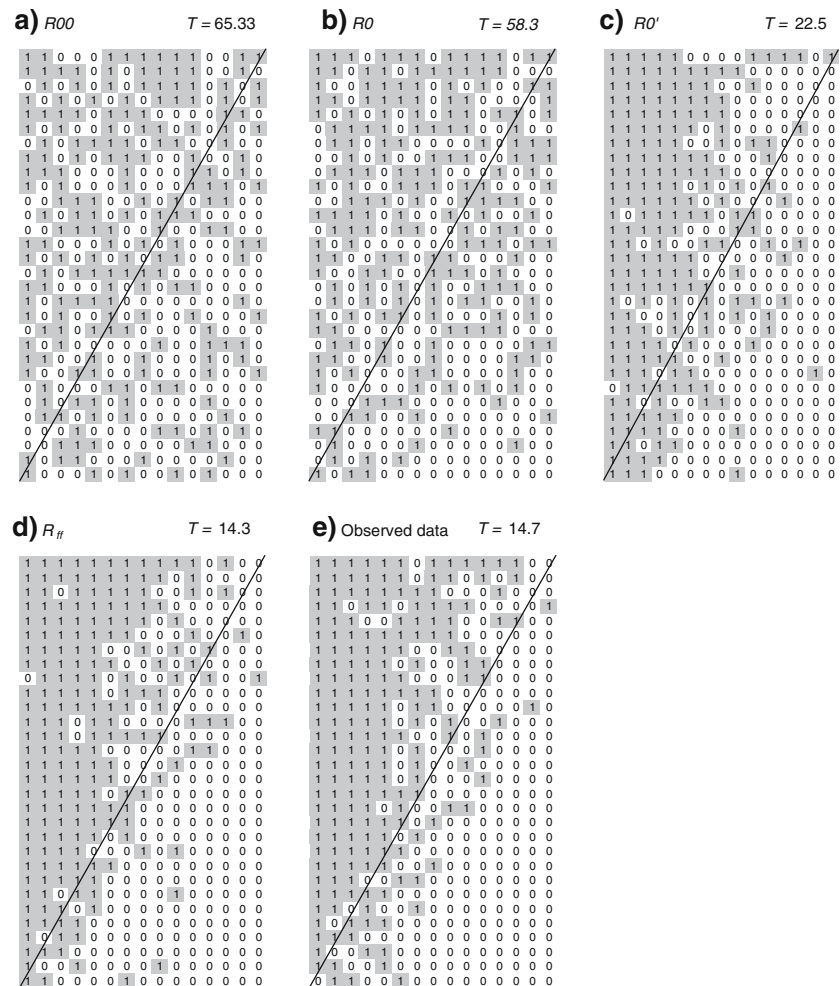
biological structure that can confidently be assumed to be operating in a system, so as to isolate the pattern or process of true interest. This is analogous to conducting a controlled lab experiment in which all elements of the treatment and control groups are held constant except for the single factor being tested. Models with insufficient structure will be easily rejected without having tested predictions of the ecological theory (Gotelli and Graves 1996, p. xii).

The least restrictive null model in traditional nestedness analysis constrains the total number of 1s in the randomized matrix to equal that of the observed data matrix, but otherwise all matrix elements have an equal probability of being filled. This is the null model of the nestedness temperature calculator (Atmar and Patterson 1995), denoted *R00* by Wright et al. (1998; Table 1). Most randomized matrices under this null model, even after row and column sorting, will not show any strong fill tendency toward either the left or top of the matrix (Fig. 1a), and will have nestedness metric values that reflect this. If the observed data show sufficient variation in either richness across sites or

incidence across species, the fill will be greater toward the top or left of the sorted matrix, respectively, and will thus yield a nestedness metric value far from the null expectation. Thus, *R00* simply tests for whether communities are more nested than would be expected for a system with uniformly distributed species richness and occurrence frequencies. Clearly, *R00* is an under-restrictive test for most systems, as it lacks the biological structure required to generate a realistic pattern that might be expected for a system in which sites vary in richness, or in which some species are more widespread than others. This problem seems widely recognized in the nestedness literature (Brualdi and Sanderson 1999; Jonsson 2001; Fisher and Lindenmayer 2002), but in spite of this, results based on this null hypothesis are still reported (e.g., Fernández-Juricic 2000; Feeley 2003; Azeria 2004; Martínez-Morales 2005).

A slightly more restrictive null model is one that constrains either row or column totals in the randomizations to equal those in the observed data matrix. If row totals are fixed, the null model is based on the ecological assumption

Table 1 Summary of null models used to test for community nestedness

Null model	Source	Site richness (row) constraints[a]	Species incidence (column) constraints[a]	Matrix fill constraints	Interpretation: ''significant'' nestedness explained by variation[b]...
$R00$	Atmar and Patterson (1995)	None	None	Same as obs	...in observed richness or incidence
RANDNEST	Jonsson (2001)	None	∝ Obs	$E(\text{fill})^c$ = obs	...in observed richness
RANDOM0, $R0$	Patterson and Atmar (1986)	Same as obs	None	Same as obs	...in observed incidence
$R0'$		None	Same as obs	Same as obs	...in observed richness
$R_{ij\text{mean}}$	Bascompte et al. (2003)	∝ Obs	∝ Obs	$E(\text{fill})^c$ = obs	...beyond observed richness or incidence
RANDOM1, $R1$	Patterson and Atmar (1986)	Same as obs	∝ Obs	Same as obs	...beyond observed richness or incidence
RANDOM2, $R2$	Wright et al. (1998)	Same as obs	∝ Obs	Same as obs	...beyond observed richness or incidence
R_{ff}	Connor and Simberloff (1979)	Same as obs	Same as obs	Same as obs	...beyond observed richness of incidence
R_{ecol}	This study	∝ Obs	∝ Est or obs	$E(\text{fill})^c$ = est	See text

Apart from R_{ecol}, null models are generally listed from least to most constrained

[a] Model constraints in which row or column totals to fill proportionally to observed values (∝ obs) are theoretical, as some are not able to successfully implement this constraint in practice. See text

[b] Interpretations, particularly for models with both row and column constraints (apart from R_{ecol}), are theoretical. In practice, some of these null models are prone to Type I or II errors. See text

[c] ''$E(\text{fill})$ = obs'' and ''$E(\text{fill})$ = est'' denote that the expected fill value (i.e., mean for many randomizations) should equal that of the observed data matrix (obs) or that of the estimated fill (est) under an ecologically explicit null model, respectively

that different sites vary in richness (identically to the observed variation), but that expected species occurrence rates are fairly uniform (i.e., all species will be observed with similar frequency). This null model has been denoted RANDOM0 (Patterson and Atmar 1986) or $R0$ (Wright et al. 1998; Table 1), and generates randomized matrices that fill more prominently toward the top on average, but do not show any strong fill tendency toward the left, and have nestedness metric values that reflect this (Fig. 1b). If observed occurrence frequencies vary sufficiently from uniform expectation, the observed data matrix will have top-down fill similar to the randomized matrices, but also will show stronger fill toward the left. Thus, the observed pattern and its nestedness metric will differ significantly from those expected under the null model. In this case, the null hypothesis will be rejected, implying that communities are more structured than expected for a system with uniform species incidence. Thus, RANDOM0 or $R0$ is effectively just a test for community structure due to variation in species commonness.

Analogous to $R0$ is a model in which total matrix fill and column totals are fixed, but rows are filled equiprobably. We denote this model $R0'$ because of its structural similarity to $R0$ (Table 1). This model makes the ecological assumption that species vary in incidence (identically to the observed data) but that richness across sites is from a uniform distribution. A conceptually similar model to $R0'$ is RANDNEST (Jonsson 2001), in which elements within columns are simply filled with probability equal to the occurrence rate for species j. Rows are thus filled equiprobably, and total matrix fill is not constrained to equal

the observed fill. These null models thus generate a matrix pattern expected if some species occur more frequently than others but sites do not vary in richness. Randomized matrices will fill strongly toward the left, but not toward the top (Fig. 1c). If richness does vary significantly across sites, the observed matrix will have similar left–right fill to randomized matrices but will fill more strongly toward the top of the matrix than most randomized matrices, and hence will have a nestedness value substantially different from the null expectation. These null models effectively just test for variation in species richness, i.e., for whether communities are more nested than would be expected for a system with uniform richness across sites. In spite of the fact that the null models $R0$, $R0'$, and RANDNEST principally test for community structure that occurs solely by virtue of nonuniform variation in richness or occurrence, and in spite of similar criticism from other researchers in the nestedness literature (e.g., Simberloff and Martin 1991; Brualdi and Sanderson 1999), these are the models most commonly used in nestedness analyses to this day (e.g., Fleishman and MacNally 2002; Berglund and Jonsson 2003; Cook et al. 2004; Ficetola and Bernardi 2004; Wethered and Lawes 2005; Higgins et al. 2006).

In fact, it is well recognized that some species are more abundant and widespread than others and that sites vary in richness according to gradients of area, isolation, productivity, habitat diversity, and so on. Most studies of nestedness sample sites that vary in area, and the species–area relationship is one of the most fundamental rules in ecology (MacArthur and Wilson 1967; Rosenzweig 1995). In such systems, any of the above null models are likely to

be rejected, regardless of whether species compositions are structured by processes different from the fundamental principles just mentioned. Wright et al. (1998) anticipated that random sampling from a simple species–abundance distribution would yield strong nestedness under a null model like $R00$, $R0'$ or $R0$, and several studies have shown examples of this for simulated and real data (Cutler 1994; Andrén 1994; Higgins et al. 2006). Higgins et al. (2006) essentially viewed nestedness as a passive result of species–abundance relationships (SAR) and island-size distributions (ISD), and sought to identify characteristics of SAR and ISD that affected the degree of nestedness. In their study, and in all other studies that test for nestedness against the previously described null models, nestedness has thus been implicitly defined as the degree of community subsetting that occurs in systems as a result of SAR and ISD, compared to a system not described by one or both of these. In other words, under conventional approaches, nestedness is merely a property of SAR and ISD.

We believe that a more informative null model of nestedness is one that will generate an expected and realistic distribution of species assemblages, given some expected variation in species occurrence rates and richness governed by well-recognized processes, but in the absence of additional structuring mechanisms. This requires both row and column totals to be constrained in generating random matrices. Matrices created under such a model will fill toward both the top and left, similar to the observed data matrix, so rejection of the null hypothesis requires that the observed data show additional structure beyond that expected due to predictable variation in richness and occurrence frequencies alone. Here we provide an overview of previous null models that somehow constrain both row and column marginals, and then in the following section we provide an alternative approach.

One type of null model that constrains both row and column totals is the RANDOM1 model (Patterson and Atmar 1986), also denoted $R1$ (Wright et al 1998; Table 1). This model constrains row totals (species richness) to equal the observed distribution, but randomly assigns 1s to each column with probability proportional to observed species incidence frequencies. Several studies revealed that the $R1$ model is highly prone to type I error because the randomized marginal totals are systematically biased from the observed marginals (Simberloff and Martin 1991; Wright and Reeves 1992; Cook and Quinn 1998), so Wright et al. (1998) developed a similar null model ($R2$; called RANDOM2 by Jonsson 2001) that fixes row sums and assigns 1s to each column with probability proportional to squared incidence frequencies (Table 1). This null model reportedly produced random matrices with marginals more similar to the observed marginals. Bascompte et al. (2003) and Rodríguez-Gironés and Santamaría (2006) generated

random matrices by assigning 1s to each matrix element a_{ij} with probability equal to the average occurrence probabilities of row i and column j. We denote this $R_{ij\text{mean}}$ (Table 1). This approach should lead to problems similar those of $R1$, however, because the expected marginal totals for random matrices will be lower than the observed totals for large observed values, and higher than the observed totals for small observed values. The most constrained null model fixes both row and column sums in the randomized matrices to be equal to those of the observed data matrix. This model, which we denote R_{ff} (ff = "fixed–fixed"; Table 1), was proposed by Connor and Simberloff (1979) to evaluate Diamond's (1975) species "assembly rules," and has subsequently been applied many times in the analysis of species co-occurrence (Stone and Roberts 1990; Manly 1995; Sanderson et al. 1998; Gotelli 2000; Gotelli and McCabe 2002), but only rarely in the nestedness literature (Brualdi and Sanderson 1999; McAbendroth et al. 2005).

An ecologically explicit null model

Proponents of null models that constrain both row and column totals assert what we have argued above: not constraining both row and column marginals fails to test an ecological prediction of interest. The criticism of these constrained nulls, however, is that observed marginal totals may themselves reflect important pattern, in which case constraining both row and column totals to reflect the observed marginals would build too much structure from the data into the null model, thus making it too difficult to reject (e.g., Diamond and Gilpin 1982; Wright and Reeves 1992; Cook and Quinn 1998; Jonsson 2001). This seems likely to be the case. Consider a 10×10 matrix with 50% fill (e.g., 50 of the 100 elements have 1s). If row and column marginals are $\mathbf{r} = \mathbf{c} = [10, 9, 8, ..., 1]$, the matrix is nonpermutable and will describe perfect nestedness (Jonsson 2001). Similarly, if $\mathbf{r} = \mathbf{c} = [5, 5, 5, ..., 5]$, the matrix will be completely non-nested. Jonsson (2001) noted a relationship between the distribution of species richness (row) marginals and matrix temperature (T), and we have noted that for a fixed set of row and column marginals, the range of possible matrix temperatures is rather small (see "Results"). Clearly then, system nestedness is at least partially reflected in the distributions of the marginals themselves.

For this reason, we propose a different type of null model: one that constrains both row and column totals to reflect not the observed marginals but rather a priori expectations for richness or occurrence frequencies (both if possible), based on a particular ecological hypothesis. The test for nestedness is therefore, in part, a test for differences between the observed and expected marginal distributions

themselves. Here we present several case study analyses. In the first, we extended a random placement model (Coleman et al. 1982) to derive expected values for species richness (rows) and incidence (columns) of forest songbirds across a set of fragmented Midwestern landscapes, based on the number of individuals of each species detected during surveys and the amount of habitat observed in each landscape. This approach, which we also applied to the dataset presented in Coleman et al. (1982), obviously requires abundance data, not just presence–absence data. Next, for these two datasets and four others available in the primary literature, we derived expected values of richness by fitting species–area models to the data and constraining incidence according to observed values. These latter cases are examples of what may be a ''best available option'' for constructing a null model when no abundance data are available to construct a priori expectations of occurrence frequencies. Finally, in two of the latter datasets, we derived expected richness values by fitting relationships between species richness and habitat diversity or richness versus area + isolation. In all of these examples, we explicitly tested whether real communities were more (or less) nested than would be expected for a system in which communities are structured according to, say, species–abundance and species–area relationships or species–diversity relationships, but are otherwise assembled randomly. We denote this general class of ecologically explicit models as R_{ecol} (Table 1).

Materials and methods

Forest songbirds in Indiana, USA

Study area and sampling design

Data used in our first analysis were collected as part of a regional study to examine effects of habitat fragmentation on biodiversity for several taxonomic groups in the upper Wabash River basin of north-central Indiana (e.g., Swihart and Moore 2004; Moore and Swihart 2005; Rizkalla and Swihart 2006). This basin is >20,000 km^2 and is predominantly in agricultural use. Forest habitats are patchily distributed and cover only about 8% of the total study area. During May–June 2001–2003, we sampled for occurrence of songbird species at 950 forest point locations, distributed across 35 sample ''landscapes'' that were each 4.8 × 4.8 km (23 km^2) in size. Point counts lasted 10 min and had a fixed 50-m radius. Each point was visited once. We used a stratified random sampling algorithm to select landscapes representative of the range of variation in land-cover types in the region, and to select forest patches representative of the size distribution in each landscape

(Moore and Swihart 2005). For the purposes of the current analysis, landscapes represent sites in the site–species matrix, and we defined the bird community of interest a priori as those 15 species detected during our study that are native to the study area and breed obligately in closed canopy forest. The commonnesses of these species ranged from very rare in our study area (e.g., prothonotary warbler *Protonotaria citrea*, detected only at two point counts) to very common (e.g., red-eyed vireo *Vireo olivaceus*, detected at 297 counts and in 34 of 35 landscapes). Though landscapes were all the same size, they varied in amount of forest habitat from 134 to 2,505 ha, and landscapes with more forest contained greater observed species richness ($R^2 = 0.59$, $P < 0.001$). Our interest was in whether communities observed in each landscape would conform to a nonrandomly nested pattern. Sampling effort was not identical in each landscape, so to guard against effects of effort on observed richness, we removed the five least-sampled landscapes from analysis. After controlling for forest area in the 30 retained landscapes, effort and observed richness were not correlated ($P = 0.22$). Effort in these landscapes varied from 15 to 87 counts (<40 for all but four landscapes).

Species–area relationship and random placement

Coleman et al. (1982) proposed a theory of random placement to estimate expected richness across K sites of differing area when information on species abundance is available, and when individuals are assumed to be distributed randomly. Let N be the total number of individuals observed from S species, with n_i equaling the number of individuals of species i. If the area of the kth region is a_k, then the relative area of region k is

$$\alpha_k = a_k \Big/ \sum\nolimits_{k=1}^{K} a_k. \tag{1}$$

The random placement hypothesis says that if individuals are distributed randomly, the probability of a given individual occurring at site k is proportional to the area of k, i.e., α_k. The probability of the individual not occurring on k is thus $1 - \alpha_k$. If we assume that conspecifics distribute themselves independently of each other, the probability that no individual from the ith species occurs on k is $(1 - \alpha_k)^{n_i}$. This quantity summed for all species is the expected number of species not occurring on site k. Thus, the expected number of species occurring on site k is

$$\bar{s}(\alpha_k) = S - \sum_{i=1}^{S} (1 - \alpha)^{n_i}. \tag{2}$$

For our analysis, we redefine N as the total number of species ''point occurrences'' across our set of C point

counts (e.g., if five species were each detected at two point counts, $N = 10$ point occurrences). Thus, n_i is the number of point occurrences for species i, i.e., the number of point count locations where species i was detected. To generate an expectation of observed richness across K landscapes, we assume that the probability that a given point occurrence will be in landscape k is proportional to the relative amount of forest in k and also to the number of counts conducted in k. Therefore, let a_k in equation 1 be the amount of forest in landscape k. Let the relative effort in landscape k be $E_k = c_k/C$, where c_k is the number of point counts conducted in k. Then, the probability of a given point occurrence occurring in k is

$$B_k = \alpha_k E_k \Big/ \sum\nolimits_{k=1}^{K} \alpha_k E_k \qquad (3)$$

Replacing α_k in Eq. 2 with B_k yields an expected richness in landscape k of

$$\bar{s}(B_k) = S - \sum_{i=1}^{S} (1 - B_k)^{n_i}. \qquad (4)$$

The expected variance in richness for each site (following Coleman et al. 1982) is

$$\sigma^2(B_k) = \sum_{i=1}^{S} (1 - B_k)^{n_i} - \sum_{i=1}^{S} (1 - B_k)^{2n_i}. \qquad (5)$$

We can also estimate the expected value and variance for each species incidence (number of sites where each species occurs) simply by replacing S in Eqs. 4 and 5 with the number of landscapes K and summing across different values of B_k instead of different values of n_i. Thus, the expected incidence for species i is

$$\bar{I}(B_i) = K - \sum_{k=1}^{K} (1 - B_k)^{n_i} \qquad (6)$$

with the variance for $\sigma^2(B_i)$ similar to Eq. 5.

If the observed richness is consistent with values expected under a random placement hypothesis, then approximately two-thirds of the observed values will fall within one SD of the expected values when both are plotted against B_k (Coleman et al. 1982). Similarly, we expect at least two-thirds of the observed occurrence values to fall with the expected value ± 1 SD for each species.

Species–area relationship without random placement

We also estimated expected richness at each site by fitting the classical power function, $S = cA^z$, to the data, where S is richness at a site, A is site area, and c and z are

parameters to be estimated. Fitted values of S were used as expected row marginals. Observed values of occurrence frequencies were used as expected column marginals.

Datasets from primary literature

We used the random placement model to generate expected richness and occurrence marginals (27×38) from data on birds nesting on islands in Pymatuning Lake, Pennsylvania–Ohio, USA (Coleman et al. 1982, 1979 data only). For this dataset plus four others, we also fit a power function (above) or exponential function ($S = C + z\log A$) to data on richness and site area to generate expected richness at each site. We used observed occurrence frequencies as expected column marginals. The datasets were: boreal mammals (matrix size $= 19$ sites $\times 14$ species) and birds (13×12) in the Great Basin, USA (Brown 1978; Skaggs and Boecklen 1996), IUCN Red List vascular plant and fungal communities (46×35) in boreal forest patches in Sweden (Berglund and Jonsson 2003), and tropical forest birds (10×31) in the Usambara Mountains, Tanzania (Newmark 1991).

In addition, we calculated alternative richness (row) expectations for bird datasets of Brown (1978) and Newmark (1991), based on relationships identified in those studies. For the Brown (1978) dataset, we estimated richness as a linear function of the habitat diversity score reported at their sites (Johnson 1975). For the Newmark (1991) dataset, we estimated richness as an exponential function of both forest-fragment area and isolation distance, i.e., $S = C + z\log A + k\log D$.

Calculating nestedness

We evaluated nestedness using the temperature (T) metric of Atmar and Patterson (1993). Generally, this metric reflects the number of ''unexpected'' species occurrences and absences (idiosyncrasies) in the observed matrix, relative to a perfectly nested matrix. Idiosyncrasies close to the upper left or lower right corner of the ''packed'' matrix contribute greater unexpectedness than those closer to the center of the matrix. Packing refers to the resorting of rows and columns in a manner that yields the lowest possible value of T for that matrix. Higher values of T suggest greater disorder; they are ''hotter'' with greater ''entropy,'' and thus less nested. We calculated T for each presence–absence matrix using the algorithm implemented in the program BINMATNEST (Rodríguez-Gironés and Santamaría 2006), which is based on that of the Nestedness Temperature Calculator (NTC; Atmar and Patterson 1995), but is more robust to problems that have been identified in the NTC.

Testing for nonrandomness

We used Monte Carlo methods to evaluate the statistical significance of T for our Indiana dataset and each of the published data matrices against several null models. These included our ecologically explicit null models (random placement, area function, area-isolation function, habitat diversity), plus two conventional models for comparison: $R0'$ (incidence values are fixed) and R_{ff} (both incidence and richness are fixed). We obtained significance results for all models by calculating the P value as the proportion of 100 random matrices with a smaller or equal T value. The dataset available in Coleman et al. (1982) did not include a presence–absence matrix, but did include marginal totals for rows and columns of the matrix. For these data, we therefore compared the mean T obtained from random matrices generated under model R_{ff} (using the observed marginals) against the distribution of random matrices of other null models. The mean T from R_{ff} should have provided a close approximation of T for the unknown real matrix, since the variance of T generated under model R_{ff} is very small (Table 2).

We generated random matrices for our ecologically explicit null models by estimating a probability of occurrence (p_{ij}) for each element a_{ij}, and then assigning 1s or 0s to each element according to independent Bernoulli trials. Therefore, neither row, nor column, nor total fill of a given random matrix was constrained to equal the observed or expected values. We feel this is important for two reasons. First, it treats the observed data matrix as a random variable drawn from a distribution of possibilities that could arise under the same expectation (i.e., observed data are viewed as a single realization of stochastic natural processes). Second, for practical purposes, it is necessary to allow marginals and total fill to vary if the observed matrix is to occur within the null space. In other words, given that the

expected marginals and fill under the null model may differ from the observed matrix, the only way to generate random matrices that are like the real data is to allow matrix fill and marginal sums to vary from the null expectation.

Calculating the p_{ij} in such a way that expected marginal sums equal those of the observed data (or in our case, equal the estimated sums from an empirical model) is nontrivial (Gilpin and Diamond 1982; Gotelli and Graves 1996; Cook and Quinn 1998), and is not successfully accomplished by filling each a_{ij} with probabilities proportional to the product of squared marginal sums r_i and c_j, as in the $R2$ model (JEM, unpublished data). However, we developed an algorithm that worked well (Fig. 2). We began by calculating $r_i c_j / \Sigma r_i$ for each element, where r_i and c_j are the sums for row i and column j, as calculated under the null model. We used Σr_i in the denominator of this calculation instead of another indicator of total fill, because the sum of row and column marginals are not always equal in our null models that generated expectations for the row marginals only ($\Sigma r_i = \Sigma c_j$ in the random placement model, however). This calculation can generate expectations >1 for some matrix elements, so we had to redistribute the excess probabilities across rows and columns. Gilpin and Diamond (1982) and Cook and Quinn (1998) also did this but were vague in their explanation of how the redistribution occurred or how well their algorithm performed. For each $p_{ij} > 1$, we distributed the full excess across elements in the same row and *again* across the same column in proportion to the contribution of those elements to their respective marginals (excluding other elements with $p_{ij} > 1$). This inflates the total expectation within the matrix because the probabilities are redistributed twice, but this approach is necessary to maintain row and column sums identical to those computed before the redistribution. The surplus was then removed from the matrix by subtracting from each element: $r_i^+ c_j^+ / \Sigma r_i^+ c_j^+$, where r_i^+ and c_j^+ are the amount of

Table 2. Nestedness analysis results for six datasets and five different null model tests (see text)

| Null model Dataset	T_{obs}	$R0'$ $\bar{T} \pm SD(P)$	Random placement $\bar{T} \pm SD(P)$	Species–area function[a] $\bar{T} \pm SD(P)$	Other R_{ecol}[b] $\bar{T} \pm SD(P)$	R_{ff} $\bar{T} \pm SD(P)$
Forest songbirds, IN	14.7	23.7 ± 2.1 (<0.01)	14.0 ± 2.3 (0.63)	17.3 ± 2.7 (0.17)		14.4 ± 0.4 (0.79)
Nesting birds, PA-OH	2.5 ± 0.4[c]	26.8 ± 2.5 (<0.01)	7.16 ± 1.6 (<0.01)	4.1 ± 1.3 (0.10)		
Plants and fungi, Sweden	7.6	26.9 ± 1.7 (<0.01)		10.2 ± 2.0 (0.01)		7.25 ± 0.5 (0.76)
Mammals, Great Basin	14.0	30.8 ± 3.0 (<0.01)		14.0 ± 3.2 (0.49)		12.6 ± 0.8 (0.98)
Songbirds, Great Basin	20.3	23.2 ± 4.3 (0.20)		17.8 ± 4.6 (0.68)	14.8 ± 4.0 (0.88)	21.3 ± 1.0 (0.14)
Songbirds, Tanzania	6.0	34.7 ± 4.0 (<0.01)		5.6 ± 2.1 (0.64)	7.2 ± 2.7 (0.39)	6.4 ± 0.6 (0.26)

T_{obs} is the nestedness temperature for the observed data. P values are the proportion of random matrices out of 100 with $T_{random} \leq T_{obs}$

[a] Species–area null model based on power function for IN songbirds, Great Basin songbirds and mammals; based on exponential function for PA-OH songbirds, plants and fungi in Sweden, and Tanzania songbirds

[b] For Great Basin songbirds, ''Other R_{ecol}'' was based on fitted linear relationship between species richness and habitat diversity ($R^2 = 0.81$); for Tanzania songbirds, it was based on the exponential model: Richness = log(Area) + log(Isolation) (Adj $R^2 = 0.96$)

[c] T_{obs} for this dataset was estimated as the mean T from 100 matrix randomizations, based on observed row and column sums, under model R_{ff}

Fig. 2a–f Evaluation of our random-matrix generation algorithm, in terms of its ability to generate null matrices with expected marginal sums equal to those estimated by our ecologically explicit null models. Each panel shows row (species richness) sums for the six datasets used in our nestedness analyses: **a** forest birds in Indiana, USA; **b** songbirds in Pennsylvania–Ohio, USA; **c** boreal mammals in Great Basin, USA; **d** boreal birds in Great Basin, USA; **e** fungi and plants in Sweden; **f** tropical mountain birds in Tanzania (see text for complete dataset citations). The x-axis depicts row-sum estimates from a fitted species–area (power or exponential) relationship. The y-axis depicts the mean row sums from 100 random matrices (*filled circle*) generated under that null model, ±1 SD (*dotted lines*). Points should lie along the line of unity (*solid line*) if null matrices are unbiased. Also shown are the observed row sums from the real datasets (×). Results for column sums (species incidence), not shown, are qualitatively identical
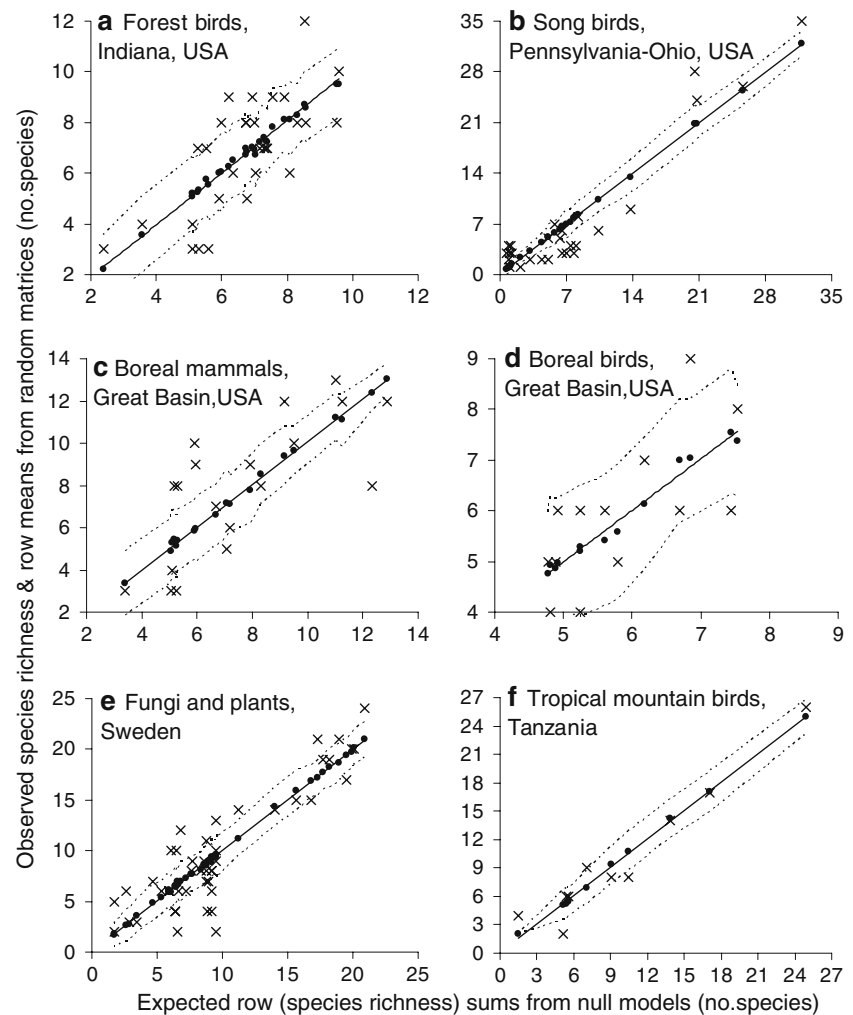
probability added to each row and column following the redistribution, compared to the previous r_i and c_j. This process was iterated several times (≤4) until all $p_{ij} \leq 1$.

We generated random matrices for $R0'$ and R_{ff} in Eco-Sim (Gotelli and Entsminger 2004). For these, column ($R0'$) or row and column (R_{ff}) totals equaled those of the observed matrix. For R_{ff}, we generated null matrices using the independent swap algorithm, which has good statistical properties for creating matrices with fixed-row and fixed-column constraints (Gotelli and Entsminger 2001, but see Miklós and Podani 2004).

Results

Forest songbirds in Indiana, USA

Variation in species richness and incidence

The distribution of species incidences for 15 forest-obligate songbirds in north-central Indiana was consistent with

expectations generated by the random placement model (Fig. 3a). For 11 of 15 species (73%), the number of landscapes in which they occurred was within one SD of the expectation under a stochastic process. For only one species (pileated woodpecker *Dryocopus pileatus*) was the observed species incidence beyond (lower than) two SD of the expectation. The distribution of species richness differed slightly from expectation under a random placement model, however. For only 16 of the 30 landscapes (53%) did the number of observed species fall within one SD of the expected value (Fig. 3b). Ninety percent of the landscapes had observed richness within two SD of the expected value. Departure from randomness was due to richness being lower than expected from the random placement model in 21 (70%) of the 30 landscapes. Total fill (number of 1s) for the observed data matrix was 204, whereas the fill of the expected matrix was 225. Thus, while richness increased with the amount of forest cover in a landscape, it did so at a slower rate than would be observed if individual birds from all species randomly settled in landscapes in proportion to their amount of forest cover.
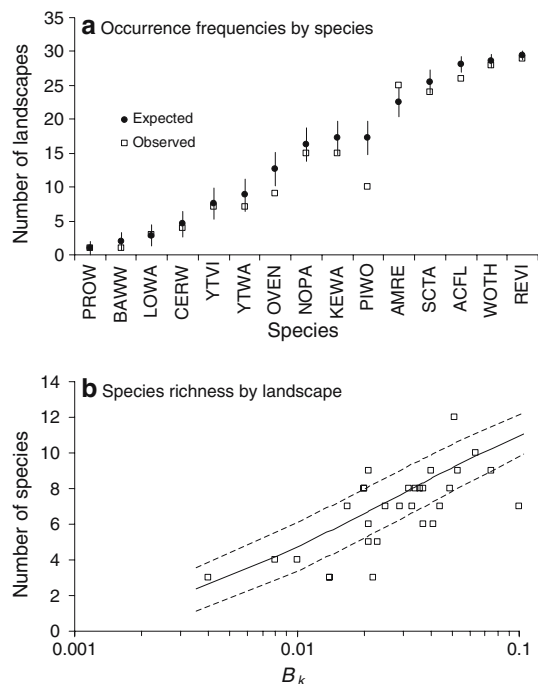
Fig. 3a–b Comparison of observed (*open square*) and expected values under a random placement model for **a** number of study landscapes in Indiana where each forest songbird species occurs, and **b** species richness in each landscape. *Error bars* and *dotted lines* depict 1 SD. In **b**, expected richness (*solid line*) increases with B_k (*x*-axis), which is the proportion of total forest cover (ha) and sampling effort (number of point counts) across K landscapes that occurs in each 23-km^2 landscape (k; see text). Bird species codes: *ACFL* acadian flycatcher *Empidonax virescens*, *AMRE* American redstart *Setophaga ruticilla*, *BAWW* black-and-white warbler *Mniotilta varia*, *CERW* cerulean warbler *Dendroica cerulea*, *KEWA* Kentucky warbler *Oporornis formosus*, *LOWA* Louisiana waterthrush *Seiurus motacilla*, *NOPA* Northern parula *Parula americana*, *OVEN* ovenbird *Seiurus aurocapillus*, *PIWO* pileated woodpecker *Dryocopus pileatus*, *PROW* prothonotary warbler *Protonotaria citrea*, *REVI* red-eyed vireo *Vireo olivaceus*, *SCTA* scarlet tanager *Piranga olivacea*, *WOTH* wood thrush *Hylocichla mustelina*, *YTVI* yellow-throated vireo *Vireo flavifrons*, *YTWA* yellow-throated warbler *Dendroica dominica*

This could be explained by clustering of conspecifics within landscapes (Martín and Goldenfeld 2006), rather than individuals settling independently, so variation in community richness was perhaps better described by a power function (Fig. 4a) than a random placement model.

Nestedness

Forest songbird communities in Indiana were significantly nested under null model *R0′* ($P < 0.01$, Table 2), but not under our null models based on random placement ($P = 0.63$) or a species–area power function ($P = 0.17$), or under model R_{ff} ($P = 0.79$). Thus, our study communities were much more nested than would be expected for a set of sites with uniformly distributed species richness, but not

more than a system structured by those agents that gave rise to the observed area–richness relationships.

Island-nesting birds in Pennsylvania–Ohio

Variation in species richness and incidence

As reported by Coleman et al. (1982), the distribution of species richness on islands in Pymatuning Lake conformed reasonably well to expectations under a random placement model (Fig. 5b). However, the distribution of species occurrence did not (Fig. 5a). Nearly half (44%) of observed occurrence frequencies were beyond one SD of the expectation. Thirteen percent were beyond two SD. Departures from expectation increased for more common species (Fig. 5b). For example, the three most common species (gray catbird *Dumetella carolinenis*, song sparrow *Melospiza melodia*, and yellow warbler *Dendroica petechia*) occurred on many more islands than expected under the random placement model, given their observed abundance. Total fill (number of 1s) for the observed data matrix was 203, whereas the fill of the expected matrix was 187. One likely explanation for these results is that *habitat area* for these species probably did not scale proportionally with island area. For example, the largest island in this study consisted mainly of pine and deciduous forest (Coleman et al. 1982), neither of which is likely to have harbored many song sparrows. Incorporating this factor into the model would have increased the expected incidence frequency of this species because a lower proportion of the observed song sparrow abundance would be assigned to that largest island.

Nestedness

Island bird communities on Pymatuning Lake were significantly nested under null model *R0′* ($P < 0.01$; Table 2) and under our null model based on a random placement ($P = <0.01$), but were not significantly nested under our null model based on a species–area exponential function ($P = 0.10$). Thus, the study communities of Coleman et al. (1982) were more nested than would be expected for a set of islands with uniformly distributed species richness or in which individuals settled randomly on islands in proportion to island area, but were not strongly nested compared to a system structured by processes underlying the observed species–area relationship (Fig. 4b).

Nestedness in other systems

Except for Great Basin boreal songbirds, communities of the remaining four systems we analyzed were significantly

Fig. 4a–f Relationships between patch area (*A*) and species richness (*S*) for six datasets (**a–f**) used in our nestedness analyses. Area data, reported as in original studies, are in hectares for all datasets except for **c** and **d**, which are in square miles (see text for dataset citations). Observed data (*circle*) were fit by exponential or power functions. Fitted estimates of species richness (*filled squares*) were used as expected row marginals (r_i) for null matrices in nestedness analyses
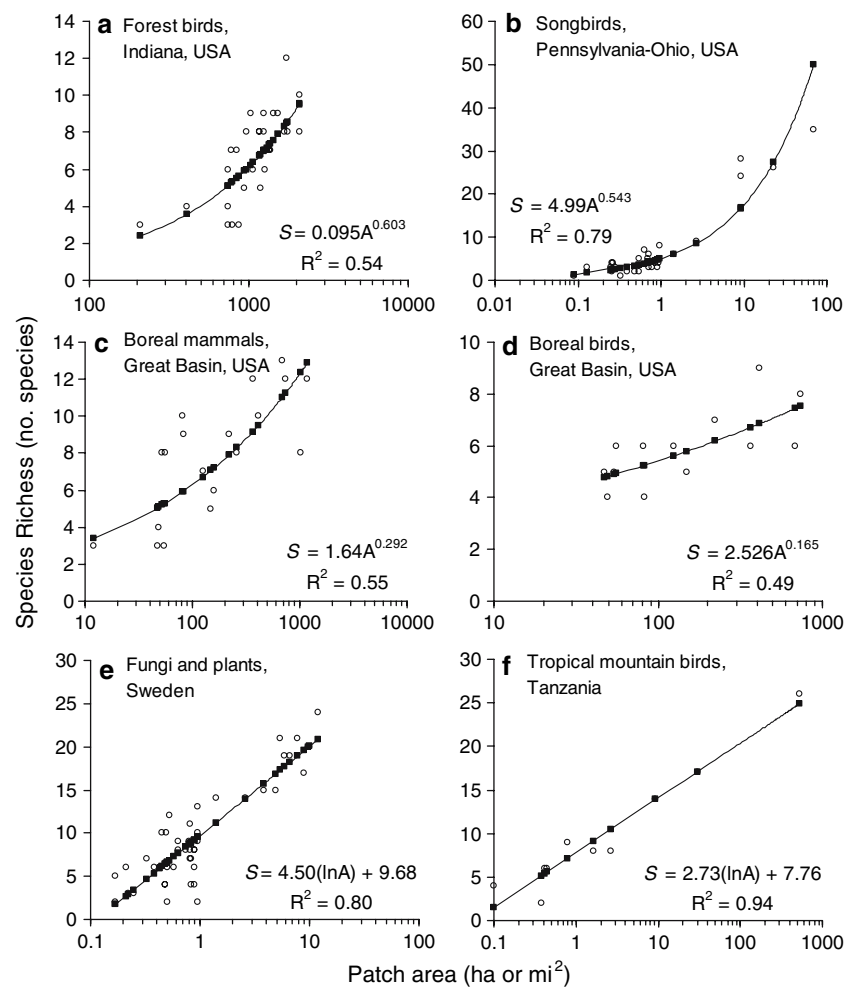
nested under null model *R0′* ($P < 0.01$; Table 2). The contrasting result for Great Basin birds can be explained by the lack of variation in richness across sites for this system. Richness only varied from four to nine species per island, conforming to a nearly uniform distribution, which is an assumption of model *R0′*. Expected species richness values for these four datasets, estimated by fitting area–richness (Fig. 4c–f), habitat diversity–richness, or area/isolation–richness models to the data, were used as row sum expectations for R_{ecol} null models. Tested against these null models, only Red List understory vascular plants and fungi were significantly nested ($P = 0.01$; Table 2). In summary, Great Basin boreal bird communities were no more nested than expected for a system with similar variation in occurrence frequencies, uniform variation in species richness, and random assembly otherwise. The degrees of nestedness observed in Great Basin boreal mammal communities and tropical mountain bird communities in Tanzania could be explained by relationships between species richness and environmental factors. Plant and fungal communities in Sweden appeared to be more strongly

structured than could be explained by simple species–area relationships. Importantly, these results relied on using observed incidence frequencies as the occurrence (column sum) expectations, which thereby built some unknown amount of community structure into the null model. Thus, we view these tests as conservative. No communities were significantly nested under null model R_{ff} (Table 2).

Discussion

Since the analysis of nestedness first became formalized (Patterson and Atmar 1986), dozens of published studies, using several different metrics, have examined whether species distributions across a set of sites are ''significantly nested.'' The collective results, as reported, suggest that nestedness is extremely common, as most studies have found systems to be significantly nested (Wright and Reeves 1992; Wright et al. 1998). This apparent trend was widely recognized in the early history of nested subset analysis, and led Simberloff and Martin (1991) to assert
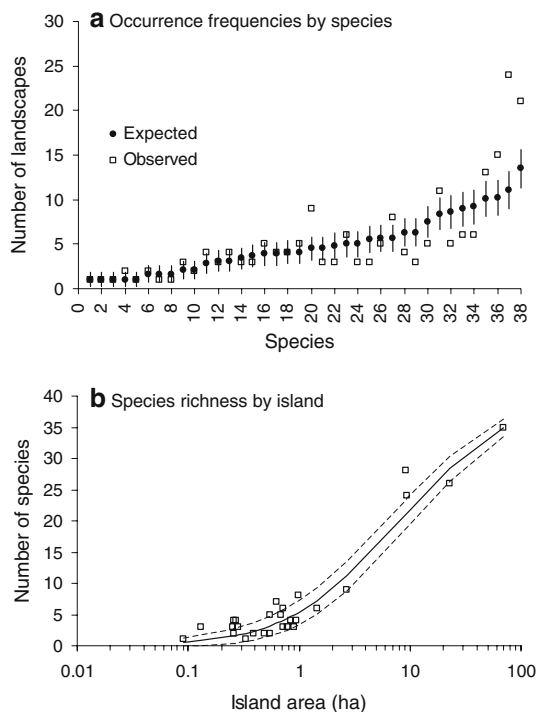
Fig. 5a–b Comparison of observed values (*open squares*) to those expected under a random placement model for: **a** number of study islands in Pymatuning Lake, Pennsylvania–Ohio, where each nesting bird species occurs, and **b** species richness on each island, as a function of island area. *Error bars* and dotted lines depict one SD. See Coleman et al. (1982) for species included in study

that the interesting and informative outcome of a nestedness analysis is only that in which species or systems do *not* conform to the pattern. Moreover, it led additional researchers (Skaggs and Boecklen 1996; Bird and Boecklen 1998) to doubt the utility of nestedness theory as the useful conservation tool that some suggested it might be (e.g., Brown 1986; Patterson 1987; Bolger et al. 1991). This is because even supposedly highly nested systems typically contain a large number of idiosyncracies (i.e., species occurrences or absences that deviate from the nested pattern), such that a collection of several less species-rich sites often harbors a greater total number of species than one or a few of the most species-rich sites (Quinn and Harrison 1988; Boecklen 1997; Fischer and Lindenmayer 2005). Thus, nested subset theory did not resolve the SLOSS (single large or several small) debate on reserve design that it revived. Moreover, Maron et al. (2004) found that a nestedness matrix was a poor predictor of future colonization or extinction. However, as we have demonstrated, widespread conclusions about communities being significantly nested, and the ecological and conservation debates that have centered on these findings, are critically dependent on the null models used in nestedness analyses and the interpretation of null model rejections.

Cutler (1994) noted that nestedness implies a strong species–area relationship but not vice versa. We agree with this perspective, and emphasize again that nested-subset theory, if useful as an ecological concept or conservation tool, should predict patterns which are more structured than those predicted by fundamental predictors of species richness. However, statistical analyses will not reflect this concept unless a proper set of null models is used. Under model *R0′*, for example, a strong species–area relationship *does* yield a statistically significant nestedness result. Therefore, we suspect that many systems historically found to be nested are in fact no more so than predicted by more basic relationships. This was the case for five of six datasets we analyzed in this study, and it helps explain much of the criticism that nestedness analyses have received with respect to their utility for ecology and conservation. That is, application of nested subsets theory will not usefully inform conservation decisions if so-called nested systems are not very strongly nested. Alternatively, useful conservation insight may come from nestedness tests such as those we have proposed, which are more rigorous and ecologically explicit. For example, in the single dataset we analyzed where nestedness was stronger than predicted by a species–area relationship (Red List plants and fungi in boreal forest patches of northern Sweden), 32 of 35 (91%) species were encompassed by the eight largest (≥5.4 ha) of 46 patches. This was also the only one of seven datasets analyzed by Berglund and Jonsson (2001) for which a set of large patches contained the same or more species than a set of small patches with the same total area, even though all were nested according their null model RANDNEST.

In reviews by Cutler (1994), Worthen (1996), and Wright et al. (1998), nestedness was said to be possible if species are ordered in their extinction risk as island size decreases, if habitat diversity itself is nested (smaller islands have a subset of habitats found on larger islands), if species are ordered in their ability to colonize more distant sites, or because of passive sampling, whereby nestedness occurs as an artifact of a stochastic process described by a species–abundance relationship (also see Higgins et al. 2006). The null model approach we have proposed helps alleviate the problem of passive sampling, because the stochastic processes linked to species–area or similar relationships are explicitly controlled for in the null model. Thus, the probability that passive sampling "causes" nestedness is only the α-level for the test. Some authors have claimed that implementing a null model such as *R0* or *R0′* controls for passive sampling (e.g., Fischer and Lindenmayer 2002, 2005; Azeria 2004; McAbendroth et al. 2005), but as we have described, these models only control for variation in richness *or* incidence, whereas passive sampling predicts variation in both.

In contrast with many datasets, which only contain information on species presence–absence for a set of sites, our null model optimally requires data on the number of individuals detected at each site so that null expectations for species incidence and richness can be estimated. Higgins et al. (2006) also suggested that null models should be based on random placement of individuals. Of course, this requirement is problematic for existing datasets that do not include detailed information about individuals, and is a more costly strategy in terms of time and resources required to collect the data. We have proposed a less demanding alternative to using abundance data, in which presence–absence data are used to calculate expected richness at a site as a function of some environmental predictor of interest. However, it may not be possible to generate a priori expectations of species occurrence frequencies under this framework, in which case observed frequencies must be used as null expectations. This undoubtedly builds some amount of unwanted structure into the null model, and so is probably an overly conservative approach for detecting nestedness. Nonetheless, it is still ecologically explicit, is preferable in our view to simpler unrealistic models, and is far less conservative than null model R_{ff} or any other null model that accurately constrains both row and column sums according to observed frequencies.

In our case study of forest songbirds in north-central Indiana, we found that bird communities in different landscapes were no more nested than expected, given species–area relationships generated by random placement of individuals or estimated by a power–function fit to habitat–area versus richness data. We are not surprised by this and similar results from the datasets we analyzed. Strongly nested patterns are unexpected where habitats are not homogeneous or nested across sites. For example, in the dataset for mountain birds in Tanzania, Newmark (1991) noted that the largest forest patch was too high in elevation, and therefore too cold, to support four species that occurred in smaller but lower elevation forest patches. Additional phenomena expected to contribute to non-nestedness include island endemism (in the case of true archipelagos), interspecific interactions such as competition that reduce the likelihood of species co-occurrence at a site, or high rates of colonization between isolated sites that partially counteract area-dependent patch-extinction risks.

One important caveat for our Indiana study is that nondetection error could have affected inference about nestedness of forest songbirds (Cam et al. 2000). Using methods of Bouliner et al. (1998) to estimate species richness, we estimated that an average of 1.7 species were present but not detected across the total area sampled in each landscape, and additional species could have been present in areas not sampled. Grayson and Livingston (1993) suggested that the inclusion of missed species would usually increase the estimate of nestedness. Indeed, if we replace ten 0s in the site–species matrix that we strongly believe reflect nondetection error to 1s (a 5% fill increase), the P value for nestedness temperature under the power-law species–area null model decreases from 0.17 to 0.08. Thus, we stress the importance of trying to collect as complete a record as possible of species occurrence for nestedness analyses.

We suggest the following general hierarchical approach for null hypothesis testing of nestedness. First, the researcher should make a decision as to whether nonuniform variations in richness and incidence are expected for the system in the absence of nestedness. This decision may be informed by testing variation against environmental gradients such as site area, isolation, habitat heterogeneity, etc. If such relationships do not exist, then a relatively unrestrictive null model such as $R0$ or $R0'$ may be appropriate. If a relationship between richness or incidence and environmental factors does exist, then a null model should somehow constrain both row and column totals in the randomizations, preferably in accordance with values that are calculated from a random placement model or from the fitted relationships. We suggest testing the observed data matrix against successive null models that incorporate increasing amounts ecological structure; this will yield a hierarchy of information (e.g., McAbendroth et al. 2005) and will help isolate factors contributing to the observed nestedness. More robust inference about community structuring could be obtained from analyzing the same site–species matrices using both nestedness and co-occurrence metrics (e.g., Feeley 2003). Systems with fewer co-occurring species than expected by chance are common (Gotelli and McCabe 2002); such systems should probably not be strongly nested, and vice versa. We believe our conclusions have important implications for the way in which ecologists view community structuring, the extent to which nestedness is perceived as an interesting ecological phenomenon, and the utility of nested subset theory as a conservation tool for systems in which nested structuring truly occurs.

References

Andrén H (1994) Can one use nested subset pattern to reject the random sample hypothesis? Examples from boreal forest communities. Oikos 70:489–491

Atmar W, Patterson BD (1993) The measure of order and disorder in the distribution of species in fragmented habitat. Oecologia 96:373–382

Atmar W, Patterson BD (1995) The nestedness temperature calculator: a visual basic program, including 294 presence–absence matrices. AICS Research, Inc., University Park, NM, and the Field Museum, Chicago, IL. Accessed February 21, 2007. http://www.aics-research.com/nestedness/tempcalc.html)

Azeria ET (2004) Terrestrial bird community patterns on the coralline islands of the Dahlak archipelago, Red Sea, Eritrea. Global Ecol Biogeogr 13:177–187

Bascompte J, Jordano PJ, Melian CJ, Olesen JM (2003) The nested assembly of plant–animal mutualistic networks. Proc Natl Acad Sci USA 100:9383–9387

Berglund H, Jonsson BG (2003) Nested plant and fungal communities; the importance of area and habitat quality in maximizing species capture in boreal old-growth forests. Biol Conserv 112:319–328

Bird BM, Boecklen WJ (1998) Nestedness and migratory status of avian assemblages in North America and Europe. Biodivers Conserv 7:1325–1331

Boecklen WJ (1997) Nestedness, biogeographic theory, and the design of nature reserves. Oecologia 112:123–142

Bolger DT, Alberts AC, Soulé ME (1991) Occurrence patterns of bird species in habitat fragments: sampling, extinction, and nested species subsets. Am Nat 137:155–166

Bouliner T, Nichols JD, Sauer JR, Hines JE, Pollock KH (1998) Estimating species richness: the importance of heterogeneity in species detectability. Ecology 79:1018–1028

Brown JH (1978) The theory of insular biogeography and the distribution of boreal birds and mammals. Great Basin Nat Memoirs 2:209–227

Brown JH (1986) Two decades of interaction between the MacArthur–Wilson model and the complexities of mammalian distributions. Biol J Linnean Soc 28:231–251

Brualdi RA, Sanderson JG (1999) Nested species subsets, gaps, and discrepancy. Oecologia 119:256–264

Cam E, Nichols JD, Hines JE, Sauer JR (2000) Inferences about nested subsets structure when not all species are detected. Oikos 91:428–434

Coleman BD, Mares MA, Willig MR, Hsieh Y-H (1982) Randomness, area, and species richness. Ecology 63:1121–1133

Cook RR, Angermeier PL, Finn DS, Poff NL, Krueger KL (2004) Geographic variation in patterns of nestedness among local stream fish assemblages in Virginia. Oecologia 140:639–649

Cook RR, Quinn JF (1998) An evaluation of randomization models for nested species subsets analysis. Oecologia 113:584–592

Connor EF, Simberloff D (1979) The assembly of species communities: chance or competition? Ecology 60:1132–1140

Cutler AH (1994) Nested biotas and biological conservation: metrics, mechanisms, and meaning of nestedness. Landsc Urban Plann 28:73–82

Diamond JM (1975) Assembly of species communities. In: Cody ML, Diamond JM (eds) Ecology and evolution of communities. Harvard University Press, Cambridge, MA, pp 342–444

Diamond JM, Gilpin ME (1982) Examination of the ''null'' model of Connor and Simberloff for species co-occurrences on islands. Oecologia 52:64–74

Feeley K (2003) Analysis of avian communities in Lake Guri, Venezuela, using multiple assembly rule models. Oecologia 137:104–113

Fernández-Juricic E (2000) Bird community composition patterns in urban parks of Madrid: the role of age, size and isolation. Ecol Res 15:373–383

Ficetola GF, De Bernardi F (2004) Amphibians in a human-dominated landscape: the community structure is related to habitat features and isolation. Biol Conserv 119:219–230

Fischer J, Lindenmayer DB (2002) Treating the nestedness temperature calculator as a ''black box'' can lead to false conclusions. Oikos 99:193–199

Fischer J, Lindenmayer DB (2005) Perfectly nested or significantly nested—an important difference for conservation management. Oikos 109:485–494

Fleishman E, RacNally R (2002) Topographic determinants of faunal nestedness in Great Basin butterfly assemblages: applications to conservation planning. Conserv Biol 16:422–429

Johnson NK (1975) Controls of number of bird species on montane islands in the Great Basin. Evolution 29:545–567

Martín HG (2006) On the origin and robustness of power-law species–area relationships in ecology. Proc Natl Acad Sci 103:10310–10315

Gotelli NJ (2000) Null model analysis of species co-occurrence patterns. Ecology 81:2606–2621

Gotelli NJ, Entsminger GL (2001) Swap and fill algorithms in null model analysis: rethinking the knight's tour. Oecologia 129:281–291

Gotelli NJ, Entsminger GL (2004) EcoSim: Null models software for ecology. Version 7. Acquired Intelligence Inc. and Kesey-Bear, Jericho, VT. Accessed February 21, 2007. http://www.garyentsminger.com/ecosim/index.htm

Gotelli NJ, Graves GR (1996) Null models in ecology. Smithsonian Institution Press, Washington, DC

Gotelli NJ, McCabe DJ (2002) Species co-occurrence: a meta-analysis of J. M. Diamond's assembly rules model. Ecology 83:2091–2096

Grayson DK, Livingston SD (1993) Missing mammals on Great Basin mountains: holocene extinctions and inadequate knowledge. Conserv Biol 7:527–532

Higgins CL, Willig MR, Strauss RE (2006) The role of stochastic processes in nested patterns of species distributions. Oikos 114:159–167

Jonsson BG (2001) A null model for randomization tests of nestedness in species assemblages. Oecologia 127:309–313

MacArthur RH, Wilson EO (1967) The theory of island biogeography. Princeton University Press, Princeton, NJ

Manly BFJ (1995) A note on the analysis of species co-occurrences. Ecology 76:1109–1115

Maron M, MacNally R, Watson DW, Lill A (2004) Can the biotic nestedness matrix be used predictively? Oikos 106:433–444

Martínez-Morales MA (2005) Nested species assemblages as a tool to detect sensitivity to forest fragmentation: the case of cloud forest birds. Oikos 108:634–642

McAbendroth L, Foggo A, Rundle SD, Bilton DT (2005) Unravelling nestedness and spatial pattern in pond assemblages. J Anim Ecol 74:41–49

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

Moore JE, Swihart RK (2005) Modeling patch occupancy by forest rodents: incorporating detectability and spatial autocorrelation with hierarchically structured data. J Wildl Manage 69:933–949

Newmark WD (1991) Tropical forest fragmentation and the local extinction of understory birds in the eastern Usambara Mountains, Tanzania. Conserv Biol 5:67–78

Patterson BD (1987) The principle of nested subsets and its implications for biological conservation. Conserv Biol 1:323–334

Patterson BD, Atmar W (1986) Nested subsets and the structure of insular mammalian faunas and archipelagos. Biol J Linnean Soc 28:65–82

Quinn JF, Harrison SP (1988) Effects of habitat fragmentation and isolation on species richness: evidence from biogeographic patterns. Oecologia 75:132–140

Rizkalla CE, Swihart RK (2006) Community structure and differential responses of aquatic turtles to agriculturally induced habitat fragmentation. Landsc Ecol 21:1361–1375

Rodríguez-Gironés MA, Santamaría L (2006) A new algorithm to calculate the nestedness temperature of presence–absence matrices. J Biogeogr 33:924–935

Rosenzweig ML (1995) Species diversity in space and time. Cambridge University, New York

Sanderson JG, Moulton MP, Selfridge RG (1998) Null matrices and the analysis of species co-occurrences. Oecologia 116:275–283

Simberloff D, Martin J-L (1991) Nestedness of insular avifaunas: simple summary statistics masking complex species patterns. Ornis Fennica 68:178–192

Skaggs RW, Boecklen WJ (1996) Extinctions of montane mammals reconsidered: putting a global-warming scenario on ice. Biodivers Conserv 5:759–778

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

Swihart RK, Moore JE (2004) Conserving biodiversity in agricultural landscapes. Purdue University Press, West Lafayette, IN

Wethered R, Lawes MJ (2005) Nestedness of bird assemblages in fragmented Afromontane forest: the effect of plantation forestry in the matrix. Biol Conserv 123:125–137

Worthen WB (1996) Community composition and nested-subset analyses: basic descriptors for community ecology. Oikos 76:417–426

Wright DH, Patterson BD, Mikkelson GM, Cutler A, Atmar W (1998) A comparative analysis of nested subset patterns of species composition. Oecologia 113:1–20

Wright DH, Reeves JH (1992) On the meaning and measurement of nestedness of species assemblages. Oecologia 92:416–428