Ecological Modelling 244 (2012) 20–27

Contents lists available at SciVerse ScienceDirect

Ecological Modelling

journal homepage: www.elsevier.com/locate/ecolmodel

A null model algorithm for presence–absence matrices based on proportional resampling

Werner Ulrich [a,*], Nicholas J. Gotelli [b,1]

[a] Nicolaus Copernicus University in Toruń, Department of Animal Ecology, Gagarina 9, 87-100 Toruń, Poland
[b] Department of Biology, University of Vermont, Burlington, VT 05405, USA

ARTICLE INFO

Article history:
Received 19 April 2012
Received in revised form 25 June 2012
Accepted 26 June 2012
Available online 31 July 2012

Keywords:
Ecological presence–absence matrix
Statistical inference
Benchmark testing
Meta-community
Sums of square reduction

ABSTRACT

Ecological presence–absence matrices capture information of species occurrences among a number of sites. Statistical inference of matrix structure often used a fixed–fixed (FF) null model in which matrix entries are randomized, but the row and column total of each random matrix match those of the original matrix. However, in a stochastically assembled meta-community, row and column totals of a random assemblage might be expected to vary among matrices. Here we introduce a 4-step proportional–proportional (PP) algorithm that creates null matrices in which the row and column vary randomly, but the average row and column totals in a set of PP matrices are unbiased and match those of the original matrix. We tested the performance of the PP algorithm with 5 sets of artificial matrices and one large set of 288 published empirical matrices. Compared to the FF algorithm, the PP algorithm has better power to detect segregated and nested matrices, but it is vulnerable to Type I errors if row and column sums have small variances. The PP algorithm identified only 9% of empirical matrices as significantly segregated, compared with 30% identified by the traditional FF algorithm. The choice between whether to use the PP or the FF algorithm is similar to the distinction between random and fixed effects in a mixed-model ANOVA. For robust analysis, it may be desirable to use both the PP and the FF algorithms with the same data matrix.

© 2012 Elsevier B.V. All rights reserved.

1. Introduction

A major focus in community ecology for the past 40 years has been the analysis of community assembly rules (Weiher and Keddy, 1999). Such "rules" ultimately refer to successional mechanisms by which an empty patch acquires a functioning community. Examples of such rules include habitat filtering, local colonization, priority and historical effects, orderly extinctions, and species interactions, including "negative" interactions such as competition and predation, and "positive" interactions such as mutualism and facilitation.

However, ecologists are rarely able to directly observe the temporal assembly of an entire assemblage, except in laboratory studies of small sets of species with short life spans. Operationally, community assembly has come to mean the analysis of repeated patterns of species associations in replicated, censused assemblages. Diamond (1975) popularized the approach with his analyses of the distribution of 141 land-bird species on islands of the Bismarck Archipelago. Diamond introduced "rules" such as checkerboard distributions (pairs of species that never occur together in the same site) and missing species combinations (particular sets of species that never occur among replicated assemblages) and attributed both patterns to the effects of interspecific competition. Other kinds of community assembly rules include matrix-wide patterns of nestedness (Ulrich et al., 2009), guild organization, and food-web structure (Bascompte et al., 2003). More recently, community assembly rules have been extended to similar phylogenetic patterns of over- and under-dispersion (Webb et al., 2002).

The data for such an analysis usually consist of only a binary presence–absence matrix, in which rows represent species or taxa, columns represent sites or samples, and the entries represent the presence (1) or absence (0) of a particular species in a particular site. Connor and Simberloff (1979) argued that ostensible patterns in such presence–absence matrices must be compared to those that would be expected in the absence of community assembly rules. They championed the use of explicit null model randomizations of observed presence–absence matrices to test for such patterns. Although null models have had a long and controversial history in ecology (Gotelli and Graves, 1996) they have been widely adapted to the analysis of many patterns in ecology and evolution (Gotelli and Ulrich, 2012).

Early null models based on parametric tests for species associations (Schluter, 1984) assumed equal probabilities of occurrences

* Corresponding author. Tel.: +48 56 611 4469.
E-mail addresses: ulrichw@umk.pl (W. Ulrich), ngotelli@uvm.edu (N.J. Gotelli).
[1] Tel.: +1 802 656 0450.

0304-3800/$ – see front matter © 2012 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.ecolmodel.2012.06.030

among matrix cells and controlled only for the grand total of occurrences (equiprobable–equiprobable algorithm EE). Connor and Simberloff (1979) introduced a null model algorithm in which randomized matrices have row and column totals identical to those of the original matrix. Although there have been some complications in how to create a truly random sample of such matrices (Zaman and Simberloff, 2002; Miklós and Podani, 2004), this fixed–fixed (FF) algorithm has proved very popular in null model analysis in the past 30 years.

The FF algorithm has two chief advantages: first, by preserving row and column sums, it retains differences among species in the number of sites they occupy (row sums) and it retains differences among sites in the number of species they harbor (column sums). These constraints reflect the intuition of field biologists, which is that widespread heterogeneity in species richness and in species occurrences may reflect factors that are not related to species interactions. Non-random patterns must be above and beyond those that are determined by the marginal totals of the matrix. A second advantage of the FF algorithm is that it appears to have good statistical properties. During the past 10 years, a variety of benchmark tests with artificial matrices have shown that the FF algorithm performs well when confronted with heterogeneous, but random, matrices (Gotelli, 2000; Ulrich and Gotelli, 2007a,b, in press). Because the FF randomization infrequently leads to an incorrect rejection of the null hypothesis (Type I statistical error), it is a good choice for analyses of assembly rules, which are rarely based on experimental data.

However, there are two aspects of the FF algorithm that are unsatisfying for a general null model. First, because row and column sums are strictly maintained, there is a type of zero-sum dependence, and the placement of species within such a matrix is not strictly random. Second, if communities were assembled through processes of random colonization and extinction, we would not expect row and column sums to be precisely maintained from one meta-community to the next, although we might expect the average row and column sums to match those of the observed matrix.

The simplest way to introduce some variation into row and column totals but still maintain differences among species and sites is to use a probabilistic placement algorithm to fill the null matrix (sim8 in Gotelli, 2000; Jonsson, 2001). Unfortunately, this algorithm is consistently biased, and results in row and column totals that are too even compared to the original matrix. Parametric analogs to this kind of algorithm can be derived from contingency table analysis (Diamond and Gilpin, 1982; Navarro-Alberto and Manly, 2009), but they have proven difficult to implement for ecological matrices.

In this paper, we develop a new proportional–proportional (PP) algorithm for creating null matrices that vary in their row and column totals. We show that the average row and column totals across a set of such matrices are unbiased, and match those of the original data matrix. We provide benchmark tests of the PP algorithm against a set of artificial random and non-random matrices to assess its propensity towards Type I and Type II statistical errors. We next test a large set of empirical matrices for patterns of nestedness and species segregation, and find interesting differences compared to the results of the traditional FF algorithm. In the discussion, we provide guidelines for users to decide when to use the PP and FF algorithms for null model analysis of community assembly rules.

2. Methods

2.1. Proportional resampling of presence–absence matrices

Our PP algorithm to create random matrices with varying row and column totals has 4 steps: (1) assignment of matrix row and column totals; (2) adjustment of marginal totals; (3) placement of matrix cell occurrences; (4) adjustment of matrix cell occurrences. The result is a random matrix for which the expected row and column totals match those of the original matrix.

2.1.1. Assignment of matrix row and column totals

In the first step, each row and column total is assigned from a binomial distribution centered around the observed total for each species and site (Fig. 1). The simplest approach would be to use a binomial distribution in which the probability of occurrence within a particular row = (total row occurrences)/(number of columns), and the number of trials in the binomial is the number of columns. For example, suppose there are ten columns in a matrix, and the number of occurrences observed in a particular row is 7. We would set the number of occurrences in a row by taking 10 draws from a binomial distribution with $p = 0.7$. This procedure would generate a minimum of 0 and a maximum of 10 occurrences, with an expectation of 7.0.

However, the inclusion of "empty" row or column sums is problematic. In a row with 7 occurrences, the binomial probability of obtaining a 0 is only 5.9×10^{-6}. However, if the row contained only a single occurrence, the probability of obtaining a 0 is 0.35. Although such missing species or empty islands might indeed arise by stochastic colonization processes, this kind of null model would contain an inherent bias: the number of filled rows and columns could be less than observed, but could never greater than observed. It is not appropriate to simply discard the trials in which the fill is zero because this will bias the row and column sums, which would then be consistently over-estimated.

Therefore, we constrained the binomial, by setting $p = 0.5$ and centering the distribution around the observed marginal total, with the restriction that the total number of occurrences cannot be greater than the maximum or less than 1. For example, with an observed row total of 7 out of 10, we take 6 draws from a symmetric binomial distribution with $p = 0.5$. The range of possible row occurrences is 0–6, with an expectation of 3. We then shift this distribution by adding 4 to all values, and the range becomes 4–10, with an expectation of 7. Note that if the number of occupied cells in a particular row or column is either 1 or the maximum, then those values are held constant in the simulation, because using any other distribution (and still excluding 0s) would lead to a bias. Thus, the mean μ of each row or column marginal distribution is given by the number of observed occupancies r_i and c_i within a given row and column of size r and c, respectively. Minimum (min_i) and maximum (max_i) values of each of the row binomial distributions are given by:

$$min_i = 1; \quad max_i = 2r_i - 1 \quad n_i \le \frac{r}{2} \tag{1}$$

$$min_i = 2r_i - r; \quad max_i = r \quad n_i \le \frac{r}{2} \tag{2}$$

The variance σ_i^2 is given by $(max_i - min_i + 1)/4$. The same equations hold for columns.

2.1.2. Adjustment of marginal totals

As in many other randomization algorithms (Diamond and Gilpin, 1982; Gotelli and Graves, 1996; Wright et al., 1998) we hold constant the observed total number of occurrences O in the matrix. In most cases, our algorithm generates small differences ΔO in observed and simulated row and column totals. The relative difference ΔO_{rel} between assigned and observed occurrences totals will asymptotically decrease by

$$\Delta O_{rel} \propto \sqrt{\frac{\sum_1^r \sigma_i^2 \sum_1^c \sigma_i^2}{r + c}} \tag{3}$$

To explore the magnitude of such differences in empirical matrices, we used the 288 presence–absence matrices of the Atmar and

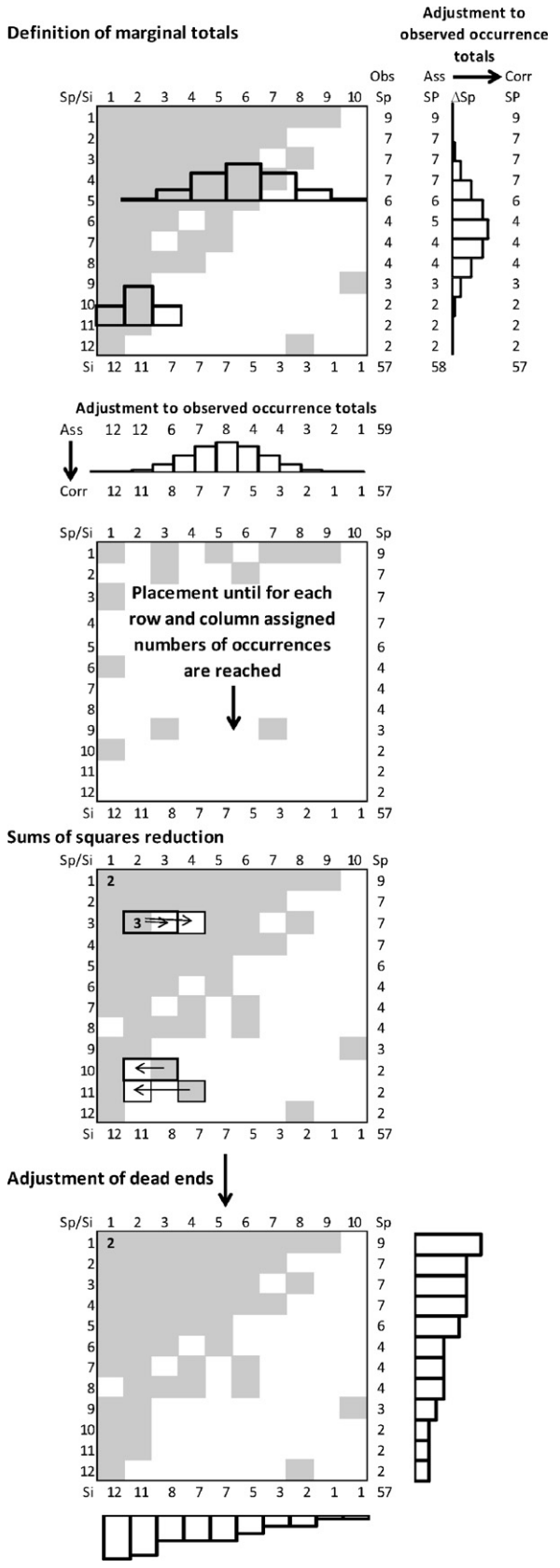

Fig. 1. The four step PP algorithm for resampling matrix entries with probabilities proportional to row and column marginal totals.

Patterson (1995) data set (Fig. 2). In 237 of the matrices (82.3%), the absolute difference between the observed and simulated *O* was less than 5 occurrences. The average difference was 2.5 occurrences (1.25%). Although the maximum observed difference was 7.25% in matrices with more than 500 occurrences, the relative difference was at most 3%.

Next, the few missing or exceeding occurrences in the row and column totals are again assigned from two binomial distributions with $p = 0.5$ and with the lower constraint of 1 and the upper constraint equaling the number of rows r or columns c, respectively (Fig. 1). The rationale for using a binomial distribution in this second step is that the variance in the number of occurrences is necessarily larger for species/sites of intermediate numbers of occurrences than for species/sites with either a high or a low number of occurrences. In summary the above two assignment steps produce row and column marginal distributions that are centered around the observed distributions and that have a total number of occurrence that equals that of the observed matrix.

2.1.3. Placement of matrix cell occurrences

In the third step, occurrences are placed step by step into the matrix. Following the classic proportional–proportional null model (sim8 in Gotelli, 2000), a row is selected randomly with probability $p_i = R_i/N$, where R_i is the row total and N is the total number of occurrences in the matrix. A column is selected randomly with probability $p_j = C_j/N$, where C_j is the column total. Thus, the probability of placing an occurrence in cell ij is $p_{ij} = (R_i)(C_j)/(N^2)$. The cells most likely to be chosen are the ones with the largest row and column totals, and the cells least likely to be chosen are the ones with the smallest row and column totals. In the classic sim8 model, only empty cells are filled this way, but in our model, we allow multiple entries to accumulate, which preserves the expected frequencies associated with the marginal totals established in Steps One and Two. These multiple entries are then reduced by the sum-of-squares reduction algorithm (SSR) of Miklós and Podani (2004) in which submatrices with entries $k, l > 1$ are reduced according to

$$\begin{pmatrix} l & i \\ j & k \end{pmatrix} \rightarrow \begin{pmatrix} l-1 & i+1 \\ j+1 & k-1 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} i & k \\ l & j \end{pmatrix} \rightarrow \begin{pmatrix} i+1 & k-1 \\ l-1 & j+1 \end{pmatrix} \tag{4}$$

with $i < l$, k and $j < l$, k until all $k, l \leq 1$.

The sum of squares reduction begins by randomly choosing a matrix cell for which the entry is >1. Next, a random row and random column is chosen until the resulting submatrix can be reduced according to the SSR algorithm. Repeated submatrices are chosen until the cell total has been reduced to 1. For example, in Fig. 1 the submatrix {{3,0},{0,1}} is first reduced to {{2,1},{1,1}} and then a new submatrix {{2,0},{0,1}} is reduced to the checkerboard {{1,0},{0,1}}. After reduction we performed an afterburn of 10*c*r (c columns, r rows) standard checkerboard swaps.

If, by chance, all assigned marginal totals from step two were to equal the original matrix totals, our method should produce a random distribution identical to that of the FF null model. This statement is equivalent to a proof that the SSR algorithm has the same sample space as the FF algorithm. Although there is no formal proof of this (Miklós and Podani, 2004), we used the Atmar–Patterson data set (Atmar and Patterson, 1995) to compare the results of FF and SSR when applied to the *C*-score (Stone and Roberts, 1990) as a metric of species associations and to *NODF* (nestedness from overlap and decreasing fill) (Almeida-Neto et al., 2008) as a metric of nestedness. Matrices generated by the FF and the SSR algorithms *with the same row and column totals* had virtually identical null distributions with highly correlated standardized effect

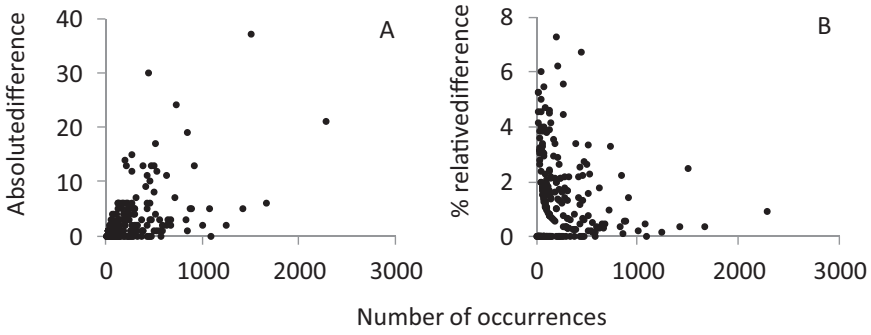

Fig. 2. The absolute (A) and percentage of relative differences (B) of assigned and observed numbers of occurrences in dependence of the total number of occurrences in 288 presences–absence matrices of the Atmar–Patterson (1995) compilation.

sizes ($r = 0.99$ for the correlation between SES_{FF} and SES_{SSR} for both the *C*-score and *NODF*). Moreover, classifications by the two methods were highly concordant (100 of 102 statistically significant *C*-score matrices, and 61 of 75 statistically significant *NODF* matrices were jointly identified by both the FF and SSR algorithms). This benchmark comparison shows that, *for a given set of matrix margin totals*, the FF algorithm and the SSR algorithm are very similar. Therefore, any differences between the behavior of the FF model and the PP model must be caused by the variability in row and column sums that is generated by the PP model, and not by differences in how the matrix cells are filled in each model.

2.1.4. Adjustment of matrix cell occurrences

Occasionally, the assigned row and column marginal totals in the PP algorithm can define an impossible matrix state, in which case the SSR algorithm will run into a dead end. For each multiple cell entry, we attempted $10 \times O$ trials with the SSR algorithm before terminating the search. In a fourth step, these irreducible multiple entries (if present) were placed into empty cells proportional to the predefined row/column totals (thus according to sim8). For example, in Fig. 1, the double entry in cell (1,1) cannot be eliminated with the SSR algorithm. Thus, one entry would be re-assigned to an empty cell with probability proportional to both marginal total distributions.

To estimate the frequency of irreducible entries and thus of the potential bias in empirical matrices, we used again the Atmar/Patterson data set. Of the 288×200 null model matrices generated 17,344 (=30.1%) could not be completely reduced with the SSR algorithm. In these irreducible matrices, there were, on average, 1.4 occurrences per matrix that needed to be re-assigned. These entries constituted an average of 0.6% of the total number of occurrences. In 39% of the empirical matrices, all null matrices were completely reducible by the SSR algorithm. Reducibility was uncorrelated with matrix size or matrix fill ($p(U) > 0.05$). For example, the Åland bird matrix of Fig. 3 (Haila et al., 1980) contains only two completely filled columns and irreducible entries were never encountered during the construction of 200 random matrices. The maximum number of irreducible occurrences was 20 (=4.6% of the total number of occurrences) and occurred in the Brazilian bird data matrix (Willis, 1979) in which 86 of its 216 rows are completely filled. Because rows that are completely filled are unchanged in the FF model and in our PP model, they can simply be eliminated before any analyses. Thus, the very small number of re-assignments does not introduce a bias and influence the null model performance.

Our assignment of row and column totals and the total number of occurrences gives unbiased random variates centered around the observed values (Fig. 3). None of the observed marginal totals in Fig. 3A and C was outside the 95% confidence limit of the null distribution. The average skewness of the row distributions (Fig. 3A) was −0.04 and that of the column distributions (Fig. 3C) was 0.004. Both values do not significantly differ from zero. In contrast, the classic proportional null model (Gotelli, 2000) yields consistently biased row and column totals (Fig. 3B and D), with over-estimates of occurrences for rare species and species-poor sites, and under-estimates of occurrences for common species and species-rich sites (Gotelli and Graves, 1996). 15% of the marginal totals in Fig. 3 fell outside the 95% confidence limits of the distribution generated by the classic proportional model.

In summary, our PP algorithm provides for the first time a set of null matrices with the following useful properties: (1) each null matrix has the same fill as the original matrix, the same matrix dimensions, and contains no empty rows or columns; (2) in contrast to the popular FF algorithm, the row and column sums do not match the empirical matrix exactly and vary randomly from one random matrix to the next; (3) the average row and column sums for a set of random matrices are unbiased and match the observed row and column sums of the empirical matrix.

2.2. Artificial and empirical matrices

We created five sets of artificial presence–absence matrices with specified amounts of randomness and structure using the software application *Matrix* (Ulrich and Gotelli, 2007a). Similar to previous approaches (Ulrich and Gotelli, 2007a,b), we constructed two types of random matrices (M_{equi} and M_{prop}) designed to span the range of empirical matrices that are typically generated by field ecologists who sample replicated assemblages at local and regional spatial scales. We generated 100 M_{equi} matrices with uniform, randomly drawn numbers of rows ($10 \leq m \leq 100$) and numbers of columns ($10 \leq n \leq 50$), and a uniform distribution of the percentage of matrix cells that were occupied (matrix fill; $0.1 \leq fill \leq 0.9$). Next, to generate 100 equally dimensioned M_{prop} matrices, we sampled row sums (=species occurrences) with placement probabilities from an exponential distribution, and column sums (=site richness) with placement probabilities from a uniform random distribution. Thus, the first simulation produced matrices with relatively uniform row and column sums (M_{equi}), whereas the second simulation produced matrices with relatively uniform column sums and highly heterogeneous row sums (M_{prop}). A third simulation was used to generate a set of matrices (M_{rand}) in which both the row and the column totals were sampled from exponential distributions.

These three sets of matrices (M_{equi}, M_{prop}, and M_{rand}) differed in the whether the row and column sums were sampled from uniform or exponential distributions, but the cell entries were all random. We also generated two sets of non-random matrices, one in which some of the species pairs were segregated (M_{segr}), and one in which some of the species pairs were nested (M_{nest}). In the M_{segr} matrices, we filled the matrix with 2×2 checkerboard submatrices $\{\{1,0\},\{0,1\}\}$ until a predefined matrix fill (drawn from a random uniform distribution between 0.1 and 0.9) was reached. In the M_{nest}

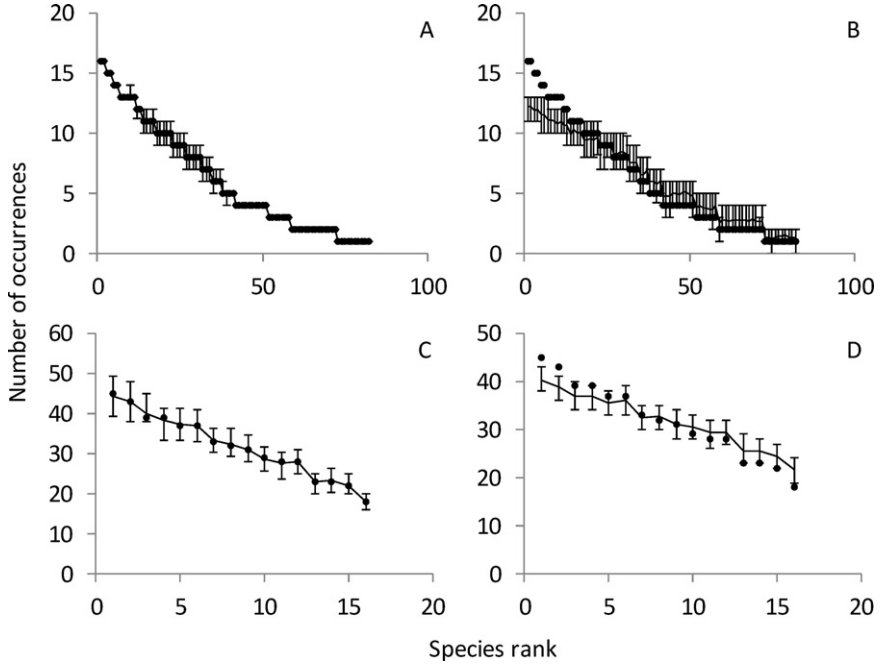

Fig. 3. Row (A and B) and column (C and D) totals (black dots) of the Åland island breeding bird matrix (Haila et al., 1980). The line gives the mean values assigned by the PP algorithm (A and C) and by a traditional placement algorithm (sim8 in Gotelli, 2000) with probability proportional to marginal totals (B and D). Error bars denote the upper and lower 95% confidence limits of the null distribution ($N = 200$).

matrices, we first filled the matrix with row and column placement probabilities drawn from two random exponential distributions. After sorting the matrix according to row and column totals, we randomly introduced a small number (5–10% of matrix fill) of unexpected absences in the upper left corner of the matrix and a few unexpected presences in the lower right corner. The M_{nest} matrices were therefore moderately to highly nested. Both matrix types represent a pattern of species segregation or nestedness in which species differ greatly in their occurrence frequency, and sites differ greatly in their suitability, but species within individual pairs tend to segregate or co-occur in classic checkerboard or nested patterns.

For empirical analyses, we compared the performance of the PP, FF, and EE null models when applied to the well-known set of 288 biogeographical presence–absence matrices compiled by Atmar and Patterson (1995).

2.3. Benchmark tests and empirical comparisons

We used two metrics of metacommunity structure proposed to account for pattern in presence–absence matrices. First, we estimated matrix wide species segregation (and aggregation) with a modified version of Stone and Robert's (1990) *C*-score (Ulrich and Gotelli, in press), which is a normalized count of the number of checkerboard submatrices ({{1,0},{0,1}} or {{0,1},{1,0}}).

$$C\text{-score} = \frac{4\sum_{i,j}\begin{pmatrix} 1 & \dots & 0 \\ \dots & \dots & \dots \\ 0 & \dots & 1 \end{pmatrix} \vee \begin{pmatrix} 0 & \dots & 1 \\ \dots & \dots & \dots \\ 1 & \dots & 0 \end{pmatrix}}{mn(m-1)(n-1)} \quad (5)$$

Because the original normalization of the *C*-score regards the number of species pairs only (Stone and Roberts, 1990) the *C*-score has been positively correlated with the number of sites (Ulrich and Gotelli, 2007b). Our modification accounts for the number of site combinations $[n(n-1)/2]$ and corrects this bias (Ulrich and Gotelli, in press). Second, we used the *NODF* index (Almeida-Neto et al., 2008) to estimate the degree of nestedness.

We used these two metrics with our PP null model to compare the performance with the FF null model implemented with the independent swap algorithm (Gotelli, 2000) and 10^*c^*r swaps. Null distributions of metric scores were based on 200 randomized matrices each. Because the null model distributions were in most cases approximately symmetrical, we converted each estimated probability value to a *Z*-transformed score ($Z = (x - \mu)/\sigma$, where x is the observed metric, μ is the average metric of the 1000 simulated matrices, and σ is the standard deviation of the 1000 simulated matrices). Assuming normality, *Z*-transformed scores should have approximate values of ± 2 at the two-sided 95% confidence limits. We compared the performance of the two metrics (*C*-score and *NODF*) in combination with the three null models (EE, PP, and FF).

3. Results

Both the FF and the PP models correctly identified the M_{prop} random matrices as being random, but the PP model failed when applied to the simple M_{equi} matrices (Table 1). The PP model incorrectly identified 76 (*C*-score) and 65 (*NODF*) of the M_{equi} matrices as non-random. The FF model correctly identified this same set of matrices as being random. The reason for the failure of the PP model to identify random matrices with equiprobable row and column totals appears to be a systematic bias in all randomizations that relax the marginal totals (see Section 4).

The PP algorithm identified more matrices as being non-random than FF when applied to the M_{segr} and M_{nest} matrices (Table 1). PP identified 73% of the M_{segr} matrices and 65% of the M_{nest} matrices as being not segregated while FF did so in only 37 and 21% of the cases, respectively. In line with previous work (Gotelli, 2000; Ulrich and Gotelli, 2007a,b) EE failed to identify segregated matrices but tend to classify them as being aggregated (Table 1). PP worked properly when applied to the M_{nest} matrices and correctly identified nestedness in 62% of the matrices. The behavior of the *C*-score was not clearly opposite to *NODF*. When used with FF the *NODF* had a much lower power than PP while the *C*-score identified 95% of the matrices as being segregated. This fact demonstrates again

Table 1
Numbers of presence–absence matrices below (<LCL) or above (>UCL) the two sided 95% confidence limits (CL) in five artificial matrix sets (*M*), and one empirical matrix set (AtPa: Atmar and Patterson, 1995). *N*: total number of matrices; fixed–fixed (FF), equiprobable–equiprobable (EE), proportional–proportional (PP) algorithms. *C*-score = Stone and Roberts (1990) *C*-score to measure species segregation. *NODF* =) index to measure nestedness.

	N	FF				EE				PP			
		C-score		*NODF*		*C*-score		*NODF*		*C*-score		*NODF*	
		<LCL	>UCL	<LCL	>UCL	<LCL	>UCL	<LCL	>UCL	<LCL	>UCL	<LCL	>UCL
M_{equi}	100	1	2	1	2	5	2	1	3	0	76	65	0
M_{rand}	100	4	7	17	0	98	0	0	88	3	1	12	1
M_{prop}	100	1	4	18	0	100	0	0	99	2	0	8	0
M_{segr}	100	1	37	21	2	60	4	6	55	0	73	65	0
M_{nest}	100	0	95	9	22	100	0	0	100	35	31	4	62
AtPa	288	2	108	70	5	224	12	5	233	24	26	29	14

that nestedness and species segregation are not mutually exclusive matrix patterns as has been claimed by recent work on metacommunity structure (Almeida-Neto et al., 2007; Presley et al., 2010, but see Gotelli and Ulrich, 2012).

The differences in performance can be understood by examining the null distributions from the three algorithms when applied to the same matrix. Fig. 4A and B shows the null distributions of the *C*-score and *NODF* obtained from PP, FF, and EE randomizations of one of the M_{rand} matrix with 50 rows and 15 columns. The three null model distributions differ in both their average and their standard deviation. In all of the 500 theoretical matrices (M_{equi}, M_{prop}, M_{rand}, M_{segr}, and M_{nest}) the FF distribution of the *C*-score and *NODF* had the smallest σ whereas the standard deviations of the PP distribution were greatest in 83% (*C*-score) and 64% (*NODF*) of the matrices. Thus, more than 60% of the matrices followed the inequality $\sigma_{FF} < \sigma_{EE} < \sigma_{PP}$ (Fig. 4). Further, in 74% of the matrices the EE expectation μ of the *C*-score was largest and in 73% of the matrices the expectation μ of *NODF* was the smallest. The PP expectations of both metrics were either lower than those of FF (*C*-score: 66%, *NODF*: 68%) or ranged intermediate between FF and EE (*C*-score: 34%, *NODF*: 32%). Hence for the *C*-score about 2/3 of the matrices followed the inequality $(\mu_{FF} \wedge \mu_{PP}) < \mu_{EE}$ (Fig. 4). Because the expectation and standard deviation of all 3 algorithms are different, they can give contrasting results when applied to the same matrix. For example, FF identified the Åland bird matrix of Fig. 3 (Haila et al., 1980) as being significantly segregated and anti-nested (Fig. 4C and D), whereas PP and EE identified this matrix as being aggregated and nested.

When applied to the empirical Atmar–Patterson data set, the PP algorithm identified a substantially lower number of significantly structured biogeographic matrices than the FF algorithm (Table 1). Only 9% of the matrices were identified as being significantly segregated whereas previous studies (Gotelli and McCabe, 2002; Ulrich et al., 2009) and our present reevaluation based on the FF null model pointed to more than 30% of the matrices as significantly segregated. For nested patterns, the PP algorithm identified 5% of the matrices as significant, compared to approximately 9% in other recent studies (Ulrich and Gotelli, 2007a; Ulrich et al., 2009). We note that 229 SESs of the *C*-score were positive with FF and 194 with PP pointing in both cases to a prevalence of "moderate" segregation in the empirical data set.

4. Discussion

The PP algorithm has the desirable property of introducing variation in row and column totals among a set of null matrices (Fig. 2), but avoids the biases that were present in previous algorithms (sim8 in Gotelli, 2000). Temporal variability in species richness of sites (column totals) and total numbers of occurrences of species (row totals) is of course a general feature of ecological communities spread among a number of sites. Our null model thus introduces more ecological realism into the randomization of matrices. This variation naturally leads to a wider range of possible values in the null distribution, but also changes the null expectation compared to the classic FF model (Fig. 4). The new algorithm has surprisingly good power for detecting non-random patterns (Table 1), although it identifies a substantially smaller number of the empirical Patterson–Atmar data matrices as non-random.

The choice between the two algorithms is not clear-cut, and depends, in part, on the kind of data being analyzed. The PP algorithm seems especially well-suited to small-scale survey data (such as invertebrate or plant occurrences in quadrats or traps), in which sampling regions (such as quadrat areas) are arbitrary, and occurrence records would be expected to vary substantially in subsequent surveys. On the other hand, the FF algorithm might be more appropriate for large-scale species occurrence records on islands (such as checklists of Darwin's finches on all of the Galapagos islands), for which it might seem more realistic to fix the number of species per site and the number of sites per species in the null matrices. In many ways, the choice between the FF and PP algorithm is analogous to the choice of a fixed-factor versus a random-factor in an analysis of variance model (Quinn and Keogh, 2002).

At least compared to the artificial structured matrices used in our analyses, the PP algorithm has better statistical power to detect effects than does the traditional FF algorithm. This feature does not depend on the predefined confidence limit (two sided 5% in this paper). Similar to the FF and EE null model distributions (Ulrich and Gotelli, 2007a,b) the respective PP distributions are approximately normally distributed and confidence limits of 1% or less do not qualitatively change the results reported in Table 1. However, the PP algorithm suffers from serious Type I errors (incorrect rejection of a true null hypothesis) when confronted with random matrices with equiprobable row and column totals (Table 1). This bias appears to be a general property of any null model that relaxes row and column totals (Appendix). In the FF algorithm, deviations from randomness reflect only the "internal" structure of the matrix, that is the pattern of co-occurrence conditioned on the marginal totals. In the PP algorithm, deviations from randomness reflect both the internal structure of the matrix, as well as the "external" structure – the row and column totals. Because negative species interactions will tend to reduce the variance in species richness among sites (Schluter, 1984), the PP algorithm can potentially detect non-random matrices that would be missed by the FF algorithm.

Because the choice between the PP and FF algorithm is not obvious, it might be prudent to analyze the same matrix with both methods. If both methods give qualitatively similar results, the pattern is robust to the details of the null model analysis. For the *C*-score 34% of the Atmar–Patterson matrices were classified as segregated by the FF or PP algorithm, 6% by both algorithms, and 60% of the matrices were classified as random by both algorithms. If the PP algorithm gives a significant result, but the FF algorithm

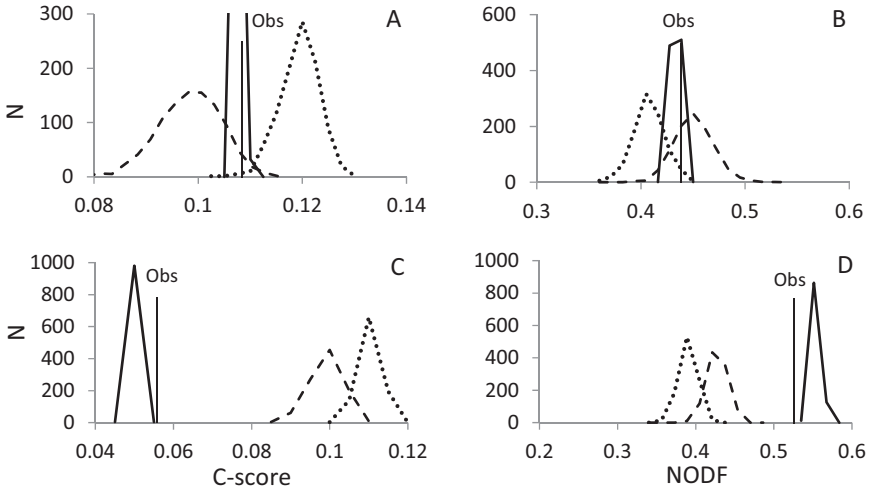

Fig. 4. Null model distributions (1000 null matrices each) of the of the FF (full line), PP (broken line), and EE (dotted line) applied to one of the M_{rand} matrices (50 species 15 sites) (A and B) and the empirical Åland island breeding bird matrix (Haila et al., 1980, C and D). The vertical lines mark the respective observed *NODF* and *C*-scores. Note that the peak of the FF distribution in A is truncated to increase readability.

does not, then the distribution of row and column sums should be carefully scrutinized. If the variance in row and column totals matches the expectation of a zero-truncated Poisson distribution then it might be prudent to use the FF results because the pattern may reflect random variation in row and column totals. We note that in the Atmar–Patterson data set 61% of the row and 69% of the column marginal distributions deviated from a truncated Poisson model at the 5% error level in having too high variances. The variance–mean relationship in these 288 meta-communities rather followed a power function ($\sigma^2 \propto \mu^z$) according to Taylor's power law (Taylor, 1961) with exponents of $z = 1.61$ for the column (site) distribution and $z = 2.24$ for the row (species) distribution. If the variance in row and column totals is greater than expected from a zero-truncated Poisson, then the significant result with PP cannot be attributed to random, equiprobable occurrences. Conversely, if the variance in row and column totals is substantially less than expected by chance, the result may indeed indicate significant species segregation, which will lead to this effect (see discussion in Appendix). For the Atmar–Patterson matrices, only one matrix classified as significant by PP, but non-significant by FF.

Finally, if the FF model is significant but the PP is not, it is instructive to examine the null distributions directly to see whether the result is due solely to the greater variation generated by the PP model, or whether it is due to shifts also in the expected value (as is the case for the Åland archipelago data in Fig. 4). For the Atmar–Patterson matrices, 28% were classified as significant by the *C*-score–FF algorithm combination, but were not significant when tested with the PP algorithm. These kind of differences reflect the added variability in co-occurrence metrics that is associated with variation in row and column totals.

There are two additional issues to consider. First, the PP algorithm mimics a random colonization process by allowing for variation in row and column totals, but it does not represent an explicit mechanistic meta-community model (Gotelli and Ulrich, 2012). If independent data are available, row and column probabilities can be conditioned on other variables. For example, Jenkins (2006) constructed null models in which occurrence probabilities were weighted by habitat area, and Gotelli et al. (2010) constructed null models in which occurrence probabilities were proportional to the biomass or population size of a species. These algorithms represent more mechanistic null models, although it is very likely in both cases that the simulated marginal distributions will not match those of the original matrix.

One final point is that null model analyses that rely on traditional summary metrics (such as the *C*-score or *NODF*) may not be reliable for classifying non-random patterns as segregated, nested, or aggregated. Recent analyses have demonstrated that a single matrix may simultaneously contain embedded structures that reflect different kinds of patterns (Ulrich and Gotelli, in press; Gotelli and Ulrich, 2012). It may be more informative to examine patterns of segregation, aggregation, or nestedness for the individual pairs of species in a presence–absence matrix (Gotelli and Ulrich, 2010).

Acknowledgments

WU was in part supported by grants from the Polish Science Ministry (NCN N 304 372839). NJG was supported by the U.S. National Science Foundation (DEB-136703) and the US Department of Energy (US-DoE; DE-FG02-08ER64510).

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.ecolmodel.2012.06.030.

References

Almeida-Neto, M., Guimarães Jr., P.R., Lewinsohn, T.M., 2007. On nestedness analyses: rethinking matrix temperature and anti-nestedness. Oikos 116, 716–722.
Almeida-Neto, M., Guimarães, P., Guimarães Jr., P.R., Loyola, R.D., Ulrich, W., 2008. A consistent metric for nestedness analysis in ecological systems: reconciling concept and measurement. Oikos 117, 1227–1239.
Atmar, W., Patterson, B.D., 1995. The Nestedness Temperature Calculator: A Visual Basic Program, Including 294 Presence–Absence Matrices. AICS Research, University Park, NM and Field Museum, Chicago, http://aics-research.com/nestedness/tempcalc.html.
Bascompte, J., Jordano, P., Melián, C.J., Olesen, J.M., 2003. The nested assembly of plant–animal mutualistic networks. Proceedings of the National Academy of Sciences of the United States of America 100, 9383–9387.
Connor, E.H., Simberloff, D., 1979. The assembly of species communities: chance or competition? Ecology 60, 1132–1140.
Diamond, J.M., 1975. Assembly of species communities. In: Cody, M.L., Diamond, J.M. (Eds.), Ecology and Evolution of Communities. Harvard University Press, Cambridge, pp. 342–444.
Diamond, J.M., Gilpin, M.E., 1982. Examination of the 'null' model of Connor and Simberloff for species co-occurrences on islands. Oecologia 52, 64–74.
Gotelli, N.J., Graves, G.R., 1996. Null Models in Ecology. Smithsonian Institution Press, Washington, DC, USA.

Gotelli, N.J., Graves, G.R., Rahbek, C., 2010. Macroecological signals of species interactions in the Danish avifauna. Proceedings of the National Academy of Sciences of the United States of America 107, 530–535.
Gotelli, N.J., 2000. Null model analysis of species co-occurrence patterns. Ecology 81, 2606–2621.
Gotelli, N.J., McCabe, D.J., 2002. Species co-occurrence: a meta-analysis of J.M. Diamond's assembly rules model. Ecology 83, 2091–2096.
Gotelli, N.J., Ulrich, W., 2010. The empirical Bayes distribution as a tool to identify non-random species associations. Oecologia 162, 463–477.
Gotelli, N.J., Ulrich, W., 2012. Statistical challenges in null model analysis. Oikos 121, 171–180.
Haila, Y., Jarvinen, O., Vaisanen, R.A., 1980. Habitat distributions and species associations of land bird populations on the Åland Islands, SW Finland. Annales Zoologici Fennici 17, 87–106.
Jenkins, D.G., 2006. In search of quorum effects in metacommunity structure: species co-occurrence analyses. Ecology 87, 1523–1531.
Jonsson, B.G., 2001. A null model for randomisation tests of nestedness in species assemblages. Oecologia 127, 309–313.
Miklós, I., Podani, J., 2004. Randomization of presence–absence matrices: comments and new algorithms. Ecology 85, 86–92.
Navarro-Alberto, J.A., Manly, B.F.J., 2009. Null model analyses of presence–absence matrices need a definition of independence. Population Ecology 51, 505–512.
Presley, S.J., Higgins, C.L., Willig, M.R., 2010. A comprehensive framework for the evaluation of metacommunity structure. Oikos 119, 908–917.
Quinn, G.P., Keogh, M.J., 2002. Experimental Design and Data Analysis for Biologists. Cambridge University Press, New York.
Schluter, D., 1984. A variance test for detecting species associations, with some example applications. Ecology 65, 998–1005.
Stone, L., Roberts, A., 1990. The checkerboard score and species distributions. Oecologia 85, 74–79.
Taylor, L.R., 1961. Aggregation, variance and the mean. Nature 332, 721–722.
Ulrich, W., Gotelli, N.J., 2007a. Null model analysis of species nestedness patterns. Ecology 88, 1824–1831.
Ulrich, W., Gotelli, N.J., 2007b. Disentangling community patterns of nestedness and species co-occurrence. Oikos 116, 2053–2061.
Ulrich, W., Almeida-Neto, M., Gotelli, N.J., 2009. A consumer's guide to nestedness analysis. Oikos 118, 3–17.
Ulrich, W., Gotelli, N.J. Pattern detection in null model analysis. Oikos, in press.
Willis, E.O., 1979. The composition of avian communities in reminiscent woodlots in Southern Brazil. Avulsos de Zoologia 33, 1–25.
Webb, C.O., Ackerly, D.D., McPeek, M.A., Donoghue, M.J., 2002. Phylogenies and community ecology. Annual Review of Ecology and Systematics 33, 475–505.
Weiher, E., Keddy, P.A. (Eds.), 1999. Ecological Assembly Rules: Perspectives, Advances, Retreats. Cambridge University Press, New York.
Wright, D.H., Patterson, B.D., Mikkelson, G.M., Cutler, A., Atmar, W., 1998. A comparative analysis of nested subset patterns of species composition. Oecologia 131, 1–20.
Zaman, A., Simberloff, D., 2002. Random binary matrices in biogeographical ecology – instituting a good neighbor policy. Environmental and Ecological Statistics 9, 405–421.