

The influence of single elements on nested community structure

Lucas N. Joppa* and Rich Williams

Microsoft Research, 7 JJ Thomson Avenue, Cambridge CB3 0FB, UK

Summary

1. Asking whether or not an ecological community displays the structural property of *nestedness* is becoming increasingly common in ecological research. We note, however, a currently limited understanding of how individual matrix components (e.g. species as rows and sites as columns) and elements (e.g. a species' occurrence at a site) contribute to overall community structure.
2. Here, we investigate the influence of single elements in a community matrix on overall community nestedness. We derive a unique extension to the popular nestedness algorithm NODF that is necessary for this analysis.
3. For 210 empirical communities, we show that elements, on average, determine a little less than 50% of a community's nestedness value, with components determining the remainder. Looking at just matrix elements we find, on average, that 10% of elements determine 25% of a community's nested structure. Further, we find statistical conclusions change for more than 20% of the empirical communities with the addition or removal of a single element.
4. Generally, the search for ecological pattern occurs by calculating a single metric designed to quantify the structure of an entire community. By showing the sensitivity of a metric to changes in a single element, our results cast doubt on the usefulness of this approach. That said, the extension to NODF that we present should allow researchers to better understand how individual elements in a community influence patterns of nestedness.

Key-words: community ecology, ecological networks, nestedness, NODF

Introduction

Studying complex ecological communities usually requires simplification. The idea of *nestedness* (Darlington 1957), and its measurement, exemplifies this issue. The earliest attempts to quantify nestedness in nature dealt primarily with metacommunities in a biogeographic sense (Schoener & Schoener 1983; Simberloff & Levin 1985; Patterson & Atmar 1986). Here, a basic question was whether the sets of species present at sites with low diversity were proper subsets of the species at more highly diverse sites (Atmar & Patterson 1993a,b). More recently, the search for nestedness has moved out of the biogeographic realm and into networks of interacting species (Bascompte *et al.* 2003; Joppa *et al.* 2010). In these networks, the questions are often of the form: Are the resource species of specialist consumers proper subsets of the resource species of more generalist consumers?

The answers to these questions are used to infer an ecological or evolutionary process that causes patterns. In biogeographic data, mechanisms proposed to explain highly nested communities have attempted to discern colonization or extinction events along gradients of size or isolation. For networks of interacting species, nestedness may arise from ecological or morphological specialization. Ulrich *et al.*'s recent review of the nestedness concept is an excellent resource for understanding potential structuring mechanisms (Ulrich, Almeida-Neto, & Gotelli 2009).

While multiple metrics have been used to measure a community's nestedness, all report a single value as a simplification of a community's overall structure. This nestedness value can be very useful for purposes of analysis and comparison between communities, yet it is extremely important to understand where the information that leads to a particular value of nestedness is coming from. Until now, that understanding has been missing.

Here, we ignore hypothesized mechanisms of community construction and instead seek to understand how a particular level of nestedness arises within a given community matrix.

*Correspondence author. E-mail: lujoppa@microsoft.com
Correspondence site: <http://www.respond2articles.com/MEE/>

We argue that detailed ecological insight can be obtained by calculating a community-wide structural metric (i.e. nestedness) and then searching for those matrix components (columns and rows) and elements that contribute maximally to that metric. Species traits, temporal dynamics of the system, and other issues of ecological interest can all be tested in this way. Understanding the way that individual community elements contribute to overall structure can also assist in data set quality control. Ensuring that the most important elements are well understood (e.g. are not erroneous datapoints) is crucial for analytic confidence.

The problem then is to analyse the distribution of nestedness information within a community using a nestedness metric that is invariant to community size (Ulrich, Almeida-Neto, & Gotelli 2009). We chose to analyse the recent nestedness metric nested overlap and decreasing fill (NODF) (Almeida-Neto *et al.* 2008) because it is well behaved in this way. NODF is a metric that overcomes known problems with previous methods (Almeida-Neto *et al.* 2008), but does not provide element-specific information on nestedness. We provide an extension to the original NODF algorithm that gives this element-specific information for a given matrix. We then use this extended algorithm to analyse a set of empirical data (ED) comprised of 210 ecological metacommunities, testing the sensitivity of their nestedness values to minor changes in the distribution of presences and absences in the community matrix. To the best of our knowledge, both of these steps are unique contributions to the current understanding of pattern in ecological communities (although see Dormann *et al.* (2009) for another approach to robustness checks. In this study of pollinators and plants, the authors assessed changes in network properties by removing interactions that were only observed once).

Methods

Nestedness is a measure of the structure of a binary presence–absence matrix. For a metacommunity, one could put species as rows, and the sites at which they occur as columns. A ‘1’ in the matrix then records the occurrence of a given species at a given site, while a ‘0’ records an absence. This approach works for other types of communities as well. For example, one could analyse a plant–pollinator community with pollinator species as columns, plant species as rows and pollination interactions as 1s in the binary matrix. Throughout the manuscript, we refer to an ecological community as a ‘community’ and more simply as the ‘matrix’. Ecological occurrences (i.e. presence–absence or interaction–no interaction) are referred to as matrix elements. A matrix element recorded as a ‘1’ represents a link, while a ‘0’ is a missing link. We identify specific matrix elements with the notation M_{ij} , where M identifies the matrix, i the row the element occurs within, and j the column the element occurs within. Species and/or sites are referred to collectively as the community ‘components’.

Regardless of community type, ordering a matrix by decreasing row and column sums has the effect of packing the 1s into the upper left corner of the matrix. Consequently, nestedness has been commonly considered a measure of the concentration of 1s in the upper left corner. There are various methods for specifically quantifying that concentration (Atmar & Patterson 1993a,b; Rodríguez-Gironés & Santamaría 2006; Galeano, Pastor, & Iriando 2009; Levartoski 2010), all of which distill the total nestedness information within a

community into a single number. That value is reported as a community’s ‘nestedness’, and it is against a distribution of such values from a set of randomized communities that statistical significance is drawn (Gotelli 2001). Missing, however, is any analysis of the distribution of nestedness information across a community and of how sensitive a nestedness result might be to slight changes in community structure.

Previous nestedness metrics such as the original ‘Matrix Temperature’ algorithm (Atmar & Patterson 1993a,b) and its modification called ‘BINMATNEST’ (Rodríguez-Gironés & Santamaría 2006) would seemingly be ideal for this approach. Both of these metrics attempt a maximal packing procedure, creating a line above which a ‘1’ (i.e. occurrence) is to be expected, and below which ‘0’ (i.e. absence) is to be expected. Any ‘0’ above the line and ‘1’ below the line has a weighted distance to the line. These distances can be loosely thought of as an amount of ‘disorder’ in the community, as compared with a perfectly ordered one. These distances are then combined in different ways to calculate overall nestedness. What is important here is that each cell in the matrix or each potential occurrence within the community contains a measured amount of information that is passed on to the overall nestedness value in a known fashion. It is easy to display and analyse the distribution of how each potential occurrence contributes to a community’s nestedness.

However, temperature-based metrics are sensitive to the way one sorts the matrix when row or column sums are tied, making it difficult to draw appropriate conclusions from traditional null models [see also Ulrich (2006)]. Also, the original ‘Matrix Temperature’ algorithm incorrectly gave an increase in nestedness after a matrix was seeded with additional occurrences of rare, non-nested species (Greve & Chown 2006).

It was only partly for these reasons the more recent metric NODF was created (Almeida-Neto *et al.* 2008). The primary reason for creating NODF was the belief that other metrics are not fully consistent with the nestedness concept in that they return positive values for cases in which there is no nestedness. In any event, NODF also avoids the above-mentioned issues. Because of this, NODF has seen widespread use in nestedness studies (Bezerra, Machado, & Mello 2009; Graham *et al.* 2009). However, unlike previous metrics, NODF calculates deductions from perfect nestedness for each column and row, but not for each element (i.e. matrix cell). Here, we first present the original NODF algorithm and then provide an algorithmic extension that allows one to calculate element-specific deductions. Using this new information, we then characterize the way these deductions are distributed across 210 ED sets on ecological communities and test how sensitive a community is to the addition and removal of single elements. Source code for the R software platform for the algorithm and visualization is available as Data S1 and S2.

ORIGINAL NODF IMPLEMENTATION

In the original NODF algorithm, the matrix can be sorted by any criterion (e.g. area, isolation or disturbance for sites, and dispersal capacity or tolerance to environmental harshness for species). However, for simplicity of presentation, we assume that the matrix is sorted in descending order by both row and column marginal totals. A matrix’s NODF value is a function of the paired overlap between all combinations of the N_r rows and all combinations of the N_c columns. The number of possible combinations of row pairs or column pairs (NPr or NPc) is:

$$NPx = Nx(Nx - 1)/2$$

where x is either r or c .

The paired overlap ($PO_{x_{ij}}$) between any two ($i, j, i < j$) rows or columns is the fraction of 1s in row or column j that are in identical positions in column or row i . Note that the sorting described earlier means that it is always the case that $nx_i \geq nx_j$. Defining nx_i as the number of 1s in row or column i , NODF then considers the paired overlap in row or column pairs, defining

$$Npx_{ij} = PO_{x_{ij}} \text{ if } nx_i > nx_j$$

$$Npx_{ij} = 0 \text{ if } nx_i \leq nx_j.$$

It is by setting Npx_{ij} to zero when $nx_i \leq nx_j$ that NODF becomes insensitive to the ways that rows or columns with identical sums are arranged. NODF is defined as the average of the Npx_{ij} values:

$$NODF = \left(\sum Npr_{ij} + \sum Npc_{ij} \right) / NP \text{ where } NP = NPr + Npc.$$

Note that NODF scales from 0 to 1, where 1 indicates perfect nestedness. There are two causes of a matrix having an NODF value < 1 – either there are row or column pairs with identical sums and hence Npx_{ij} values of 0 or there are row or column pairs in which there is imperfect overlap of non-zero elements. See the original description of NODF (Almeida-Neto *et al.* 2008) for further details.

EXTENDING THE ALGORITHM

We now extend the NODF algorithm to understand how the arrangement of individual matrix elements contributes to decreasing a matrix's nestedness score from perfect nestedness. Figure 1 shows an example matrix (M) with six rows and five columns and is designed to resemble the example shown in the original description of the NODF algorithm (Almeida-Neto *et al.* 2008). The matrix in Fig. 1 and several aspects of the figure in

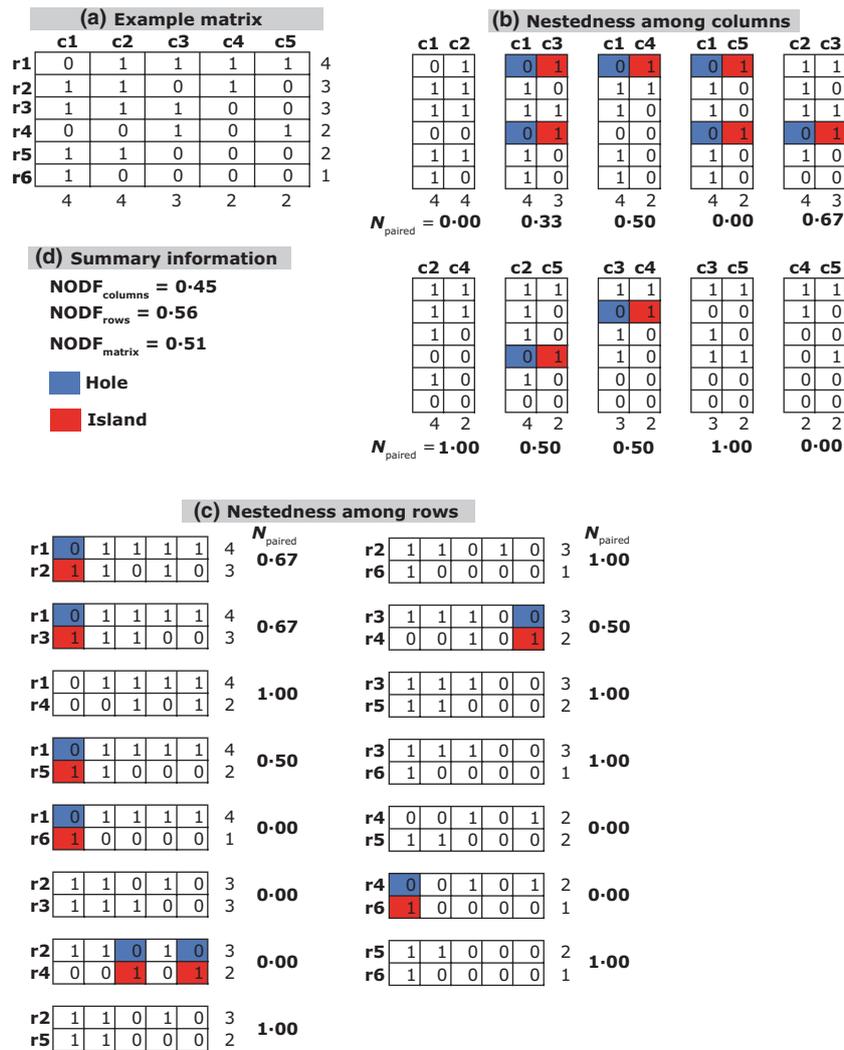


Fig. 1. Description of the original nested overlap and decreasing fill (NODF) algorithm and modelled after Fig. 2 in Almeida-Neto *et al.* 2008; (a) An example matrix with five rows and six columns. For each of the possible row and column combinations, we have shown the 'paired overlap' value from the original NODF algorithm. (b) Paired overlap values for comparisons between columns (see Methods). (c) Paired overlap values for comparisons between rows. A matrix element can decrease NODF in one of two related ways. The first is by being what we term an 'Island'. An Island is a cell with a 1 in row or column j where there is a 0 in row or column i . Islands are coloured red. The second way a cell can decrease NODF is by being what we term a 'Hole', which is a cell with a 0 in row or column i where there is a 1 in row or column j . Holes are shown in blue. (d) Summary information derived from original NODF algorithm.

general have been modified from the original version to better explain how the NODF algorithm can be extended to find the contribution of individual matrix elements.

In the example matrix in Fig. 1, $NPr = 15$ and $NPc = 10$. For each of the NPr and NPc combinations, we have shown the corresponding Npr_{ij} (Fig. 1c) or Npc_{ij} (Fig. 1b) value and have highlighted the cells that deduct from perfect nestedness. A cell can decrease NODF in one of two related ways. The first is by being what we term an 'Island'. An Island is a cell with a 1 in row or column j where there is a 0 in row or column i . Islands are shown in red in Fig. 1b,c. The second way a cell can decrease NODF is by being what we term a 'Hole', which is a cell with a 0 in row or column i where there is a 1 in row or column j . Holes are shown in blue in Fig. 1b,c. Islands and Holes directly complement each other in matrix space, but their proportional deductions from a matrix's NODF score are much less intuitive. Finally, recall that $Npx_{ij} = 0$ when row or column sums are identical, explaining why $M_{1,2}$ (where the subscripts refer to row and column, respectively) is not an Island and $M_{1,1}$ is not a Hole when comparing columns $c1$ and $c2$.

While determining the distribution of Holes and Islands throughout a matrix is interesting in itself, it does not provide the necessary information to determine which elements in the matrix maximally deduct from a perfect NODF score. Determining this requires two additional steps. In the first, we distribute the amount Npx_{ij} deviates from its maximum value of 1 proportionally across all Islands in row or column j by assigning each of the Islands in row or column j a score

$$ISr_{jk} = (1 - Npr_{ij})/NIr_j$$

$$ISc_{kj} = (1 - Npc_{ij})/NIC_j$$

where k is the index in row or column j of the Island and NIx_j is the total number of Islands in row or column j .

As an example, we compare columns 1 (with a sum of 4) and 3 (sum of 3) in Fig. 1. In column 3, there is only one 1 ($M_{3,3}$ in column 3) that is in the same position as a 1 in column 1. Thus, $Npc_{1,3} = 0.33$. There are two Islands in this pairwise comparison, in positions $M_{1,3}$ and $M_{4,3}$, making $NIC_3 = 2$. We distribute the deduction from perfect nestedness evenly across those two Islands:

$$ISc_{13} = (1 - 0.33)/2 \text{ and } ISc_{43} = (1 - 0.33)/2.$$

Each matrix position M_{ij} will potentially have more than one Island score as there can be more than one Island at each matrix position (across all Npx row or column pairs).

The Island found at matrix position $M_{4,3}$ is a good example of this process. Element $M_{4,3}$ is an Island three times ($[c1,c3]$, $[c2,c3]$, $[r2,r4]$) and so has three IS values. The value $Npc_{1,3} = 0.33$, and there are two Islands in column j , making $IS_2 = (1 - 0.33)/2 = 0.33$. The value $Npc_{2,3} = 0.67$, there is only one Island in column j , so $IS_1 = 0.33$. $Npr_{2,4} = 0.00$. Having two Islands in row j leads to $IS_3 = (1 - 0)/2 = 0.50$.

The same process can be carried out for the complementary Holes (H) in row or column i , giving

$$HSr_{ik} = (1 - Npr_{ij})/NHR_i$$

$$HSc_{ki} = (1 - Npc_{ij})/NHC_i.$$

Determining the degree to which each matrix element detracts from maximal nestedness requires two additional steps. First, one needs to sum the individual Hole ($\sum HS_{ij}$) and Island ($\sum IS_{ij}$) scores at each

matrix location. Using element $M_{4,3}$ as an example, $\sum IS_{4,3} = (0.33 + 0.33 + 0.50) = 1.17$. These ($\sum IS_{ij}$) and ($\sum HS_{ij}$) values for all elements in M are shown in Fig. 2c,g, respectively.

Finally, one normalizes with the total number of row and column pairs to produce proportional Island and Hole deductions from nestedness, $ID_{ij} = \sum IS_{ij}/NP$ and $HD_{ij} = \sum HS_{ij}/NP$. In the example above, $ID_{4,3} = \sum IS_{4,3}/NP = 0.047$. These final ID and HD values for the example matrix in Fig. 1 are shown in Fig. 2d,h, respectively. We show these are true proportional deductions from NODF through the simple equalities

$$\sum ID_{ij} = \sum HD_{ij}$$

and

$$(\sum ID_{ij} + NP_0)/NP = 1 - \text{NODF}$$

where NP_0 = the number of times $NP = 0$ because of the original NODF condition

$$Npx_{ij} = 0 \text{ if } nx_i \leq nx_j.$$

Thus, $\sum ID_{ij}$ or $\sum HD_{ij}$ report the nestedness information contained within Islands and Holes. This may not be all of the nestedness information for the community. The remainder would be contained in the NP_0 row and column pairs with equal sums. One can calculate the potential for NP_0 's influence as a fraction of all (NP) row or column pairs where $\sum X_i \leq \sum X_j$, where X is either a row or column. The fraction of information (FI) not contained within Islands and Holes is given by

$$FI = NP_0/NP$$

Note that this FI term would disappear if one removed the equality condition in the original NODF algorithm. This is simple to do, where instead of assigning a maximum deduction to instances where $\sum X_i = \sum X_j$ one just ignores that pairwise comparison. In that case, the $\sum ID_{ij}$ value would contain all of the nestedness information in a matrix.

Importantly, the matrix element with, for example, the most Islands is not necessarily the element that contributes the most to decreasing the NODF score. In a hypothetical matrix where $NP = 25$, element x might be an Island twice, where both $IS_{x1} = 1$ and $IS_{x2} = 1$, making $ID_x = 2/25 = 0.080$. Element y might then be an Island four times, with $IS_{y1} = 0.20$, $IS_{y2} = 0.50$, $IS_{y3} = 0.50$ and $IS_{y4} = 0.33$, making $ID_y = 1.53/25 = 0.061$. Although element y is an Island four times, its overall NODF deduction is less than element x .

Using the Island and Hole deduction matrices (Fig. 2d,h), one can visualize the degree to which each matrix element deducts from a perfect NODF score. In Fig. 2e, we plot the ID_{ij} and HD_{ij} values of the example matrix together, although they operate in tandem. This coloured matrix is scaled such that the darkest blue represents the Hole with the maximum NODF deduction, and the darkest red indicates the Island with the largest NODF deduction. Visualizing binary matrices in this way allows one to understand more intuitively how the distribution of 1s and 0s combine to form an overall community metric.

THE IMPORTANCE OF SINGLE MATRIX ELEMENTS

The above extension to the original NODF algorithm serves more purpose than to simply visualize the contributions of each matrix element to the total NODF value. We use the ID and HD values to analyse the distributions of element-specific NODF deductions across

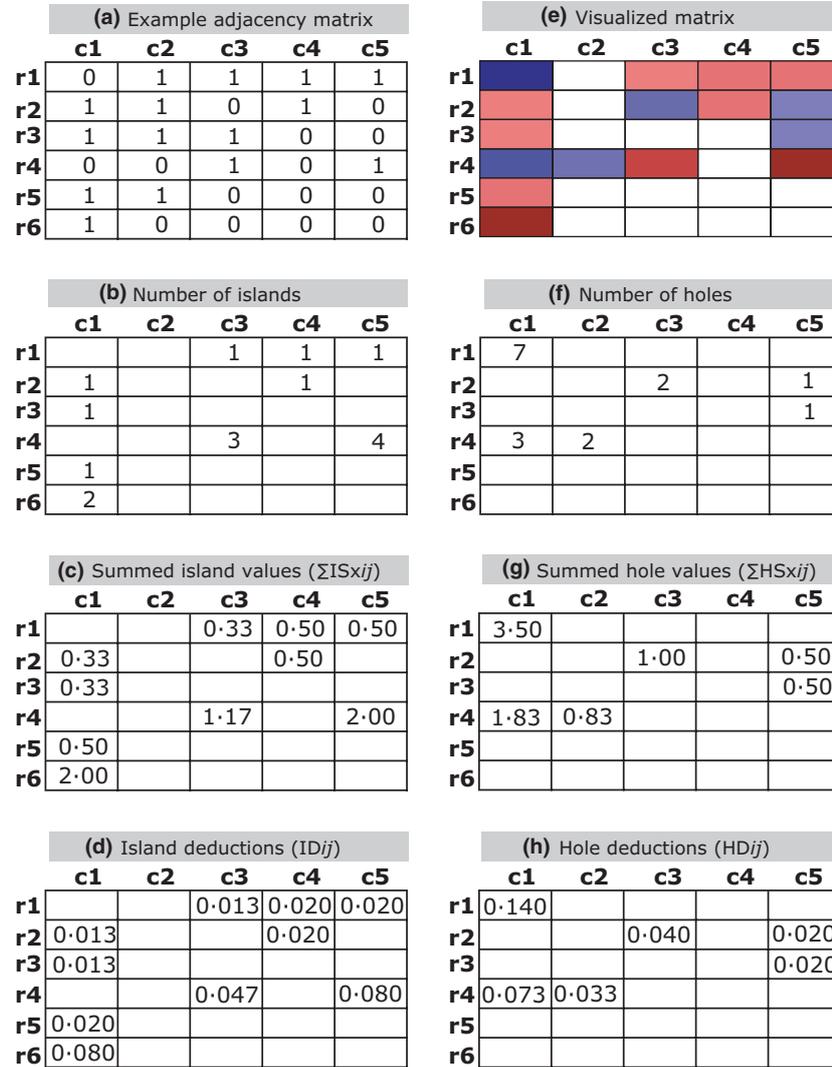


Fig. 2. Description of the extension to the original nested overlap and decreasing fill (NODF) algorithm. (a) Identical community matrix as that in Fig. 1. (b, f) The total number of Islands (b) and Holes (f) identified in Fig. 1 for each matrix element. An element can be a Hole, Island or neither, but never both an Island and Hole. (c, g) the sum of all non-normalized Island (c) and Hole (g) scores for each matrix element. See definition of $\sum IS_{ij}$ and $\sum HS_{ij}$ in Methods for further details. (d, h) The final, normalized deductions from perfect nestedness each Island (d) or Hole (h) contributes. See definition of $\sum ID_{ij}$ and $\sum HD_{ij}$ in Methods for further details. (e) The same matrix as in a, but here coloured according to the degree to which each element deducts from nestedness. Colours scale from light to dark, with reds representing Islands (values found in d) and blues representing Holes (values found in h).

210 ED sets in a manner similar to previous ‘idiosyncrasy’ methods (Almeida-Neto & Ulrich 2010). The data sets analysed contain information on ecological communities, with the majority of them describing species occurrence patterns with species as rows and sites as columns. These data sets come with the original temperature calculator (Atmar 1993a) and occur frequently in the nestedness literature (Ulrich & Gotelli 2007). The data sets we used ranged in size from 9 to 1000 elements (rows \times columns). We analysed each ED in two different ways.

In the first, we ‘flip’ each matrix element, with a one becoming a zero or vice versa and recalculate the overall NODF score. For each NODF calculation, the matrix was only one element away from its original state. If a row or column sum of the original matrix was 1 and as a result of flipping an element became 0, then that row or column was dropped from the matrix during that NODF calculation.

In the second analysis, we used the element-specific *ID* and *HD* scores of a matrix to create two additional matrices. In the first, which

we term ED+, we found the matrix element with the highest *HD* score and added a 1 to the binary matrix at that position. This removed the Hole with the largest *HD* value from the community but did not alter the matrix dimensions. For the second matrix, which we term ED-, we found the matrix element with the highest *ID* value and removed a 1 from the binary matrix at that position. If there were multiple elements with identical *HD* or *ID* values, we chose one of them randomly. Again, rows or columns with sums of zero were removed and the ED+ and ED- matrices were re-sorted by decreasing row and column sums.

SIGNIFICANTLY UNUSUAL?

The purpose of these two analyses is to assess how robust community-wide conclusions derived from a single metric are to the presence or absence of a single link. To draw any conclusions about an ecological community requires asking the question ‘compared to

what?' (Gotelli 2001). In this study, we chose to address that question by creating, for each ED, ED+ and ED-, a set of 100 randomized matrices with identical fill and fixed row and column sums. This highly constrained fixed-fixed null model was implemented using Miklos & Podani's (2004) trial-swap procedure, although we implemented only 1000 trial-swaps. This is adequate for our purposes of demonstration, but for exhaustive analysis we recommend following Miklos & Podani (2004) in implementing significantly more trial-swaps.

Determining where the NODF score of a data set falls within the distribution of randomized matrices allows one to determine the statistical significance of the observed matrix structure. We do this by asking whether a matrix is 'unusual' with respect to its NODF score (Joppa *et al.* 2010). An unusual matrix is one whose observed NODF score is greater than 95% of the NODF values obtained from randomized matrices. Of the 210 empirical matrices, 12 could not be randomized using the trial-sweep procedure because of a lack of swappable pairs and so were not included in this section of the analysis.

Results

INFORMATION DISTRIBUTION

On average, the elements in a matrix contribute 45% of a matrix's deductions (or 'information') from perfect nestedness, although across all matrices the value ranged from 5% to >80%. These element deductions correspond to the 'nested overlap' portion of the original NODF algorithm. The remaining deductions come from the 'decreasing fill' NODF requirement, which on average accounts for greater than 50% of all deductions from perfect nestedness. On average 10% of all links contain *c.* 25% of a matrix's deductions from perfect nestedness. Deductions caused by missing links (zeros in the matrix) are similarly unevenly distributed.

The fraction of matrix elements that contain either Islands or Holes spans the range of possibilities. Most matrices contain either a Hole or Island in 30–60% of their elements, although four have NODF deductions coming from more than 80% of elements.

By definition, the total deductions from Islands are equal to those from the Holes. The numbers of Islands and Holes in the community, and the magnitude of their respective deductions, however, need not be so complementary. Across the 210 communities we examine, significantly more matrix elements are Holes than Islands (paired *t*-test, $t = -8.12$, $P < 0.0001$), although it is also true that the majority (>75%) of communities have more absences (paired *t*-test, $t = -8.33$, $P < 0.0001$). By correcting for this bias by dividing the number of links in the community by the number of links that are Islands and comparing with the equivalent ratio for Holes, we show that the expected contribution of Holes and Islands is proportional (paired *t*-test, $t = 1.24$, $P = 0.22$).

Finally, in Fig. 3 we show that the element-specific information contained using two different metrics is not necessarily the same. Here, we compare a popular modification of the temperature nestedness metric (BINMATNEST, Rodríguez-Girónés

& Santamaría 2006) to NODF. Although the original implementation of BINMATNEST would attempt to find the matrix arrangement that maximizes nestedness, we ensured that both algorithms were calculating nestedness on the same matrix arrangement. Across the 210 data sets analysed here, we find that NODF distributes element-specific nestedness information more widely than does BINMATNEST. For the vast majority of communities analysed here (>80%), the set of elements identified by NODF as decreasing nestedness include 80% or more of the nestedness reducing elements identified by BINMATNEST. The reverse is not true. Instead, only 3% of the communities studied have BINMATNEST nestedness deduction sets that contain 80% or more of the elements identified by NODF. This finding is clear in Fig. 3, where the set of elements reducing nestedness under BINMATNEST are, except for one element, a perfect subset of those identified by NODF.

IMPORTANCE OF INFORMATION

In this section, we analyse the importance of individual elements in determining a community's overall nestedness score. We do this in two different ways. In the first, we simply 'flip' each matrix element (see Methods) and recalculate the community's NODF score. Figure 4a shows the minimum, maximum and mean per cent change in overall NODF score through the modification of only a single empirical matrix element. One can see that on average most matrix elements do not greatly influence the overall NODF score. However, *c.* 30% of

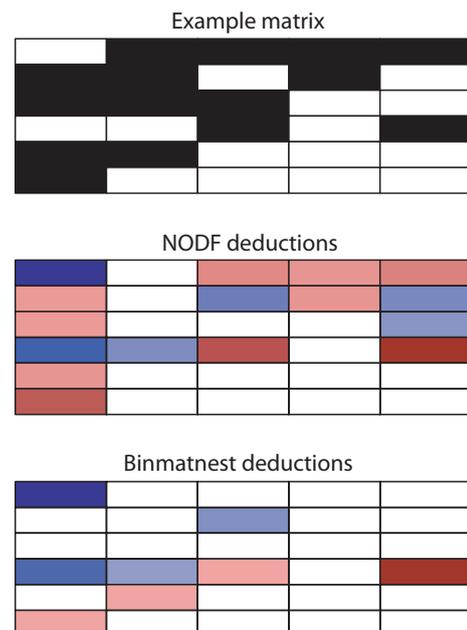


Fig. 3. Example of the distributions of deductions from perfect nestedness for different nestedness metrics. Top: Actual Community. Black represents species presence and white indicates absence. Middle: Distribution of deductions from nestedness under nested overlap and decreasing fill for each matrix element. Islands are coloured red and Holes blue. Colour intensity scales with the degree of deduction each element contains. Bottom: Same as middle, but this time the matrix colours represent the distribution of deductions from nestedness under the BINMATNEST version of the temperature algorithm.

empirical communities had at least one matrix element that changed the overall NODF score by more than 10% of the original. As expected, there is a strong decreasing trend in these results with increasing community size (Kendall's rank correlation test, $z = -8.7, P < 0.0001$) and with how close to perfect nestedness the community was to start with ($z = -4.1, P < 0.0001$), although matrix size has a larger effect than original nestedness. NODF itself is not correlated with matrix size in our data set ($z = 0.44, P = 0.66$).

How do these results compare to the influence single elements have on perfectly nested matrices? In Fig. 4b, we repeat the analysis above, but in this case with perfectly nested, artificially constructed, matrices of sizes that fall within the range of our ED set. Because the original matrix is perfectly nested, the nestedness score of the altered matrix can only be less than the original and all per cent changes will fall below zero. While the influence single elements can exert within ED sets ranges more widely than those within perfectly nested ones, Fig. 4b shows that even within

perfectly nested matrices, the change of a single element can alter the nestedness score by up to 20% in small matrices and 5% even in quite large data sets.

Figure 5 shows two examples of communities where the removal of a single species presence caused a change in NODF of 20% or more. Along with the large change in overall NODF value, there is a considerable change in the element-by-element pattern of NODF deductions.

These results, however, do not account for the fact that occasionally there are species presences (matrix elements) that when removed cause the loss of a species or site from the data set. In this case, the information about the importance of an element is clouded by the loss of that species or site. Out of 210 matrices, flipping a 1 to a 0 results in the loss of both a row *and* a column in six different matrices (only once in each of them). It caused the loss of a row in 74 matrices and the loss of a column in 186 of them. On average, 17% of all links in a matrix, when removed, caused the loss of a row or column. We tested the importance of this by repeating the 'flip'

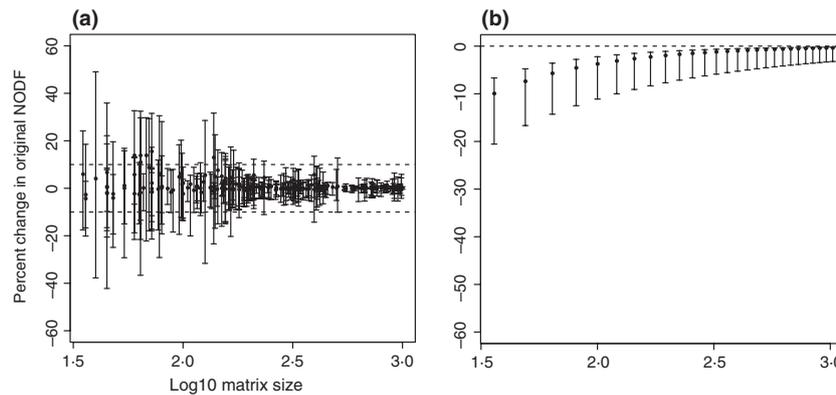


Fig. 4. The minimum, maximum and mean per cent of the original nested overlap and decreasing fill (NODF) score that occurs with the flip of a single matrix element across all possible flips. (a) Results for all 210 empirical data (ED) sets analysed in the study. Vertical bars mark the positive and negative maximum deviations from the original NODF score, while black circles show the mean deviation for all elements in the matrix. Dashed horizontal lines mark $\pm 10\%$. (b) Same as for (a), but this time with a set of artificially created, perfectly nested matrices that span the range of matrix sizes found in the ED.

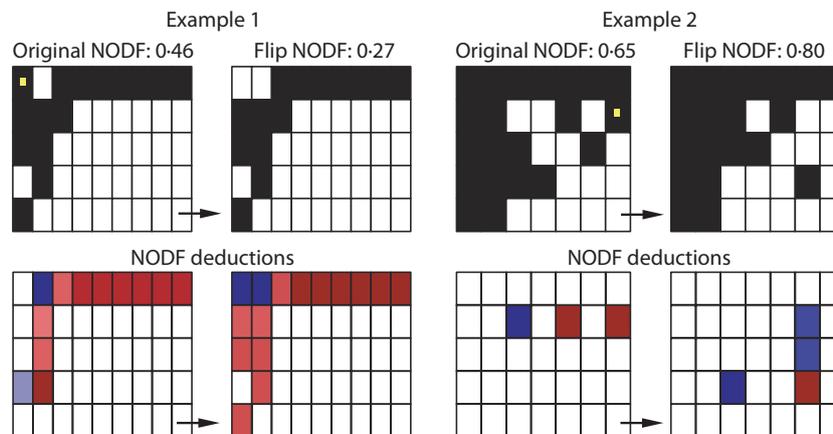


Fig. 5. Two examples of real communities, showing how flipping a single element (marked by yellow dot) can change the way nested overlap and decreasing fill deductions are distributed across the matrix. Small arrows indicate the change from original community structure to alternative structure after the yellow-marked element was flipped. Examples are the Britchah (reptiles and amphibians of the British Channel Islands) and Lmlb rept (terrestrial vertebrates on Islands in eastern Lake Michigan) data sets obtained from the original temperature calculator.

analysis above but only considering links that when removed do not alter the dimensions of the matrix. Doing this does not change the conclusions above. This implies that those links in the matrix that can change overall nestedness the most are not the species that occur at only one site or sites with only one species.

Additionally, of those elements that changed their overall community score the most; the majority (65%) were neither Islands nor Holes when computing the original NODF score. Of those that did come from an Island or Hole, there was no discernable difference (exact binomial test, $P = 0.19$) between the number of communities where the maximum change came from a Hole ($n = 30$) or from an Island ($n = 42$).

The results above highlight the fact that it is impossible to know, *a priori*, which matrix elements, after being changed, might alter the results of a test of statistical significance. This is because of the complex arrangement of the generally quite large null space. Even so, it is of general interest to know the extent to which changing a single element might change the results of a test of statistical significance. To test this, for each data set we chose to flip the Hole and Island that contained the largest deductions from perfect nestedness (see Methods). Across the 198 data sets analysed and using a highly constrained null model, 27 (14%) had a change in statistical significance with the removal of the single Island link and 24 (12%) showed the same result by changing the single Hole to a link (Table 1). Overall, 44 (22%) of the data sets we examined went from statistically unusual to no different than the random expectation or vice versa with the change of a single matrix element.

Discussion

When we analyse real ecological networks using our NODF extension, we find that for most communities, a large amount of nestedness information (*c.* 25%) comes from a surprisingly small percentage of links or missing links (10%). Community metrics are often thought of as describing, appropriately, the structure of an entire community. Here, we show that may not be the case. Instead, the way information about nestedness is contained within elements changes from one community to the next.

There is also the issue of how robust nestedness metrics are. We show that many communities contain at least one element that, when changed, can alter the entire community structure metric by 10% or more. Certain examples show upwards of a 40% change in the measured structure. Moreover, even perfectly nested matrices show this tendency, although it is not as pronounced. Our rather arbitrary approach to choosing an Island or Hole to change alters the statistical conclusions about the level of nestedness in a non-trivial fraction (22%) of all communities. This result, however, is likely to be highly dependent on the choice of a null model used for significance testing. The fixed-fixed model used here is a highly specific and constrained one. More liberal null models are likely to be more robust to the flipping of single matrix elements.

Table 1. Changes in statistical significance from the original community, relative to a community with a single Island removed (ED-) or a single Hole filled (ED+)

NODF		
Change in significance	ED- 'Unusual'	ED+ 'Unusual'
Insig – MoreNest	5	3
Insig – LessNest	2	5
LessNest – MoreNest	0	0
MoreNest – LessNest	0	0
MoreNest – Insig	4	4
LessNest – Insig	16	12
Total (%)	27 (14%)	24 (12%)
Total (ED+, ED-)	44 (22%)	

ED, empirical data; NODF, nested overlap and decreasing fill.

These findings highlight the need for researchers to carefully assess the certainty of their observations. Given the rather large variance seen in Fig. 4, it is reasonable to speculate that most small communities contain at least one matrix element that, if flipped, would result in a different statistical conclusion.

We stress that our findings are not a criticism of community metrics. Instead, we argue that the greatest insights into the structure of ecological communities will come from continuing past the quantification of community structure in a single metric. Our derivation of element-specific nestedness deductions for NODF should be useful in helping researchers gain insight into ecological communities. Understanding how nestedness information is distributed across a community can now be a fundamental step in any analysis of community structure, and this information can be linked to environmental variables, evolutionary processes or ecological interactions in ways similar to how those issues are addressed in traditional nestedness studies. For example, in ongoing work we are examining the ecological properties of matrix elements to see whether there are consistencies across different types of systems (e.g. mutualistic vs. antagonistic interactions where evolutionary processes are presumed to play different roles). In this way, using existing metrics in novel ways can enhance our understanding of ecological structure.

Acknowledgements

We thank Mário Almeida-Neto, Paulo Guimarães, Matthew Smith and Greg McInerny for thoughtful comments on previous versions of this manuscript and Stuart Pimm for motivating this work.

References

- Almeida-Neto, M. & Ulrich, W. (2010) A straightforward computational approach for measuring nestedness using quantitative matrices. *Environmental Modeling and Software*, **26**, 173–178.
- Almeida-Neto, M., Guimaraes, P., Guimaraes 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. (1993a) The measure of order and disorder in the distribution of species in fragmented habitat. *Oecologia*, **96**, 373–382.
- Atmar, W. & Patterson, B.D. (1993b) Nestedness temperature calculator program. Available from <http://www.aics-research.com/nestedness/tempcalc.html>.
- Bascompte, J., Jordano, P., Melian, 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.
- Bezerra, E., Machado, I. & Mello, M. (2009) Pollination networks of oil-flowers: a tiny world within the smallest of all worlds. *Journal of Animal Ecology*, **78**, 1096–1101.
- Darlington, P. (1957) *Zoogeography: The Geographical Distribution of Animals*. Wiley, New York, New York.
- Dormann, C., Frund, J., Bluthgen, N. & Gruber, B. (2009) Indices, graphs and null models: analyzing bipartite ecological networks. *The Open Ecology Journal*, **2**, 7–24.
- Galeano, J., Pastor, J. & Iriando, J. (2009) Weighted-interaction nestedness estimator (WINE): a new estimator to calculate over frequency matrices. *Environmental Modelling & Software*, **24**, 1342–1346.
- Gotelli, N.J. (2001) Research frontiers in null model analysis. *Global Ecology & Biogeography*, **10**, 337–343.
- Graham, S., Hassan, H.K., Burkett-Cadena, N.D., Guyer, C. & Unnasch, T.R. (2009) Nestedness of ectoparasite-vertebrate host networks. *PLoS ONE*, **4**(11), e7873.
- Greve, M. & Chown, S. (2006) Endemicity biases nestedness metrics: a demonstration, explanation and solution. *Ecography*, **29**, 347–356.
- Joppa, L., Montoya, J.M., Sole, R., Sanderson, J. & Pimm, S.L. (2010) On nestedness in ecological networks. *Evolutionary Ecology Research*, **12**, 35–46.
- Levartoski, I.A. (2010) *Nesting in Bipartite Networks*. PhD Thesis, Federal University of Rio Grande do Norte, Natal.
- Miklos, I. & Podani, J. (2004) Randomization of presence-absence matrices: comments and new algorithms. *Ecology*, **85**, 86–92.
- Patterson, B.D. & Atmar, W. (1986) Nested subsets and the structure of insular mammalian fauna and archipelagos. *Biological Journal of the Linnean Society*, **28**, 65–82.
- R Development Core Team (2010) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL <http://www.R-project.org/>.
- Rodríguez-Gironés, M. & Santamaría, L. (2006) A new algorithm to calculate the nestedness temperature of presence-absence matrices. *Journal of Biogeography*, **33**, 924–935.
- Schoener, T. & Schoener, A. (1983) Distribution of vertebrates on some very small islands. I. Occurrence sequences of individual species. *Journal of Animal Ecology*, **52**, 209–235.
- Simberloff, D. & Levin, B. (1985) Predictable sequences of species loss with decreasing island area: land birds in two archipelagos. *New Zealand Journal of Ecology*, **8**, 11–20.
- Ulrich, W. (2006) Nestedness – a FORTRAN program for calculating ecological matrix temperatures. Available from <http://www.home.umk.pl/~ulrichw/index.php?Research:Software:Nestedness>.
- 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. (2007) Null model analysis of species nestedness patterns. *Ecology*, **88**, 1824–1831.

Received 30 August 2010; accepted 5 February 2011
Handling Editor: Robert P Freckleton

Supporting Information

Additional Supporting Information may be found in the online version of this article.

Data S1. Annotated code for the algorithmic extension to NODF provided in this paper. Code is intended to be run in R, a freely available statistical software package.

Data S2. Annotated code for working through the example matrix detailed in Figures 1 and 2, and replicating Figure 2 (without legends). Code is intended to be run in R, a freely available statistical software package, and additionally requires the code provided in Data S1.

As a service to our authors and readers, this journal provides supporting information supplied by the authors. Such materials may be re-organized for online delivery, but are not copy-edited or typeset. Technical support issues arising from supporting information (other than missing files) should be addressed to the authors.