# FALCON: Efficient statistical analysis of nestedness in bipartite networks

Stephen J. Beckett*, Chris A. Boulton, Hywel T. P. Williams

April 15, 2014

College of Life and Environmental Sciences, University of Exeter, Exeter, UK

*author for correspondence: S.J.Beckett@exeter.ac.uk

**Abstract**

1. Nestedness is a statistical measure used to interpret bipartite interaction data in several ecological and evolutionary contexts, e.g. biogeography (species-site relationships) and species interactions (plant-pollinator and host-parasite networks). Multiple methods have been used to evaluate nestedness, which differ in how the metrics for nestedness are determined. Furthermore, several different null models have been used to calculate statistical significance of nestedness scores. The profusion of measures and null models, many of which give conflicting results, is problematic for comparison of nestedness across different studies.

2. We developed the FALCON software package to allow easy and efficient comparison of nestedness scores and statistical significances for a given input network, using a selection of the more popular measures and null models from the current literature. FALCON currently includes six measures and five null models for nestedness in binary networks, and two measures and four null models for nestedness in weighted networks. The FALCON software is designed to be efficient and easy to use. FALCON code is offered in three languages (R, MATLAB, Octave) and is designed to be modular and extensible, enabling users to easily expand its functionality by adding further measures and null models.

3. FALCON provides a robust methodology for comparing the strength and significance of nestedness in a given bipartite network using multiple measures and null models. It includes an "adaptive ensemble" method to increase computational efficiency and reduce under-sampling of the null distribution when calculating statistical significance. It can work with binary or weighted input networks.

4. FALCON is a response to the proliferation of different nestedness measures and associated null models in the literature. It allows easy and efficient calculation of nestedness scores and statistical significances using different methods, enabling comparison of results from different studies

and thereby supporting theoretical study of the causes and implications of nestedness in different biological contexts. We hope FALCON will be a useful tool for the research community and promote greater synthesis and theoretical insight about nestedness in bipartite networks.

**Key words:** null models, binary networks, ecological networks, species interactions, host-parasite, mutualism

**Running title:** Analysis of nestedness in bipartite networks

# 1 Introduction

Nestedness is a statistical property of systems where two kinds of entity interact. Originally used as a metric for species-site distributions [Atmar and Patterson, 1993, Brualdi and Sanderson, 1999], nestedness has recently gathered much attention as a metric for bipartite species interaction networks, e.g. plant-pollinator mutualisms [Bascompte et al., 2003, James et al., 2012] and host-virus interactions [Flores et al., 2011, 2013, Beckett and Williams, 2013]. Various discussions have considered the sources of nestedness in such systems and its implications for ecological dynamics [Thébault and Fontaine, 2010, Saavedra et al., 2011, Bustos et al., 2012, James et al., 2012, Staniczenko et al., 2013, McQuaid and Britton, 2013, Lever et al., 2014]. However, it is unclear how to systematically compare results for different datasets. Multiple methods for measuring nestedness have been used in different studies, along with multiple approaches to calculating statistical significance of the measured values. This provides a large number of ways in which nestedness could be evaluated [Gotelli, 2000, Ulrich and Gotelli, 2007, Csermely et al., 2013]. Before theoretical investigation of the mechanisms of nestedness can be properly investigated, robust measures and statistical tests for nestedness are required to allow comparison of results from different studies.

Here we present FALCON – a free software package that allows the user to easily compute several measures of nestedness and associated statistical significances based on a selection of null models. FALCON stands for "Framework for Adaptive ensembLes for the Comparison Of Nestedness". FALCON operates on any form of bipartite interaction data represented as a matrix of associations and is set up to be deliberately 'blind' to the source and interpretation of input data. FALCON is based on the assumption that nestedness is a general statistical property of matrices and therefore its measurement should be independent of context or interpretation. FALCON calculates nestedness as a context-free statistical property of a matrix, by returning the nestedness score for the most-nested configuration of the input matrix. Since calculating statistical significance of nestedness scores can be computationally demanding, involving generation of a large ensemble of matrices from a null distribution, FALCON uses a novel "adaptive ensemble" method to improve efficiency.

The FALCON package is available for three commonly used numerical analysis platforms: MATLAB, Octave and R. MATLAB (http://www.mathworks.co.uk/products/matlab/)

is a commercial software platform, while Octave (https://www.gnu.org/software/octave/) and R (http://www.r-project.org/) are both freely available open source platforms. FALCON can be freely downloaded on Github (http://github.com/sjbeckett/FALCON) or figshare[Beckett et al., 2014] and all code is open and accessible. A guide to downloading, installing and running FALCON accompanies the code. This document describes the assumptions on which FALCON is based, how it calculates nestedness and statistical significance, and gives details of the adaptive ensemble method used to improve computational efficiency.

## 2 What is nestedness?

Nestedness is a statistical property of bipartite interaction data presented in matrix form. In a perfectly nested matrix, the entries in each successive row are a strict subset of those in the previous row, while the entries in each successive column are a strict subset of those in the previous column (Figure 1). Interpretation of nestedness depends on context.

The concept of nestedness was first described in studies of how species distributions varied between sites [Hultén, 1937, Darlington, 1957, Daubenmire, 1975], and later defined quantitatively as measuring the 'amount of order/disorder' in matrices representing presence/absence of species in island communities [Atmar and Patterson, 1993]. Used in this way, nestedness is calculated from a matrix of presence-absence data where rows are species and columns are sampling sites along some environmental or spatial gradient. A perfectly "nested" matrix (see figure 1) would be achieved when the set of species present at each site along the gradient is a subset of the species present at the previous site. Since then the concept of nestedness has been extended in various directions; see Ulrich et al. [2009] for an historical overview of the nestedness concept. Nestedness has continued to be applied to spatial patterning (e.g. [Fraser et al., 2014]) and has been linked with $\beta$-diversity [Baselga, 2012], but has also been applied to study mutualistic or antagonistic species-species interactions [Fortuna et al., 2010, Joppa et al., 2010], species-time relationships for a single site [Chávez et al., 2013], and several other types of bipartite networks [Saavedra et al., 2011, Tinker et al., 2012, Bustos et al., 2012, Dalsgaard et al., 2013, Piepenbrink and Gaur, 2013, Melo et al., 2014]. For pairwise interactions (e.g. plant-pollinator or host-parasite systems), nestedness has been interpreted as placing species along a gradient of generalism-specialism in the number of partners they interact with; in this context, perfect nestedness is achieved when the most specialist species of one class interacts with the most generalist species of the other class.

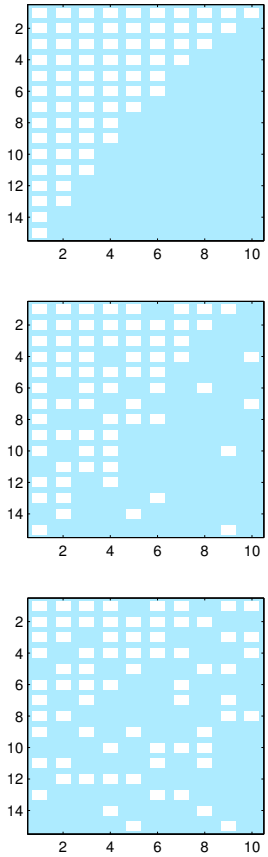

Figure 1: Perfectly nested, weakly nested, and randomly connected matrices. White squares indicate connections between two kinds of entity arranged in rows and columns.

Nestedness is calculated from a biadjacency matrix representing pairwise interactions between two kinds of entity. The order of rows and columns for a biadjacency matrix is arbitrary; rows and columns can be permuted without affecting the underlying topology of the interaction network. Any non-arbitrary ordering of rows and columns in the matrix representation necessitates the use of additional contextual information to specify which order rows and columns

should take. While some datasets may suggest a "natural" ordering to rows and columns in the matrix representation of data (e.g. when one of the dimensions represents an environmental/spatial/temporal gradient), for many applications of nestedness there is no natural ordering (e.g. species interactions).

As stated above, we consider that nestedness should be a context-free metric, so that it can be applied to data without requiring any supplemental information on row/column ordering. This assumption implies that the ordering of rows and columns should not affect the measurement of nestedness. While some nestedness measures are insensitive to row/column ordering, several of the most commonly used measures are highly sensitive to ordering, introducing indeterminacy to the quantification of nestedness when rows/columns are ordered arbitrarily. To avoid this indeterminacy and return a single robust nestedness score for a given input matrix, FALCON can sort the rows and columns such that nestedness (however calculated) is maximised. Since re-ordering rows/columns in a matrix representation does not alter the structural information (node adjacency) of the underlying data, this re-ordering is a reasonable approach and makes measurement of nestedness more consistent.

# 3 Measures of nestedness in FALCON

Nestedness is most commonly calculated for binary data representing presence/absence of an interaction between two entities, but nestedness can also be calculated for weighted data that also indicate the strength of the interaction. The methods used to calculate nestedness vary depending on whether binary or weighted interaction data is provided. The nestedness measures available in FALCON are shown and briefly described below and in table 1; further details are given in Appendix A.

The nestedness measures considered here are not trivial variations upon each other, but differ significantly in their derivations. However, some similarities can be drawn. Spectral radius (SR)[Staniczenko et al., 2013] and the measure of Johnson, Domínguez-García, & Muñoz[Johnson et al., 2013] (JDM) are invariant to the ordering of rows and columns in the network and are calculated using the adjacency matrix of the network. On the other hand, discrepancy (BR) [Brualdi and Sanderson, 1999], Manhattan distance (MD) [Corso and Britton, 2012] and nestedness based on overlap and decreasing fill (NODF) [Almeida-Neto et al., 2008] are all sensitive to row/column ordering and maximised when rows/columns are ranked by degree. The nestedness temperature calculator (NTC)[Atmar and Patterson, 1993, Oksanen et al., 2013] involves sorting of rows and columns against the 'isocline of perfect order' (see figure 5) such that it maximises connections above the isocline and minimises connections below the isocline. BR is similarly calculated relative to an idealised 'maximally packed' matrix. NODF is found through pairwise comparisons of overlap between subsequent rows and columns, whilst MD is found by assigning a weight to each connection as a sum of it's row and column indexes. The measures also differ in how nestedness is scored; the degree of nestedness in a network increases with

| Name | Shorthand | Binary/Weighted | Brief description | Reference |
|---|---|---|---|---|
| Nestedness based on overlap and decreasing fill | NODF | Binary | Pairwise row and column comparisons | Almeida-Neto et al. 2008 |
| Manhattan distance | MD | Binary | Sum of row and column indexes of connections | Corso and Britton, 2012 |
| Nestedness temperature calculator | NTC | Binary | Difference from an 'isocline of perfect order' | Atmar and Patterson, 1993 Oksanen et al. 2013 |
| Jonhson, Domínguez-García & Muñoz | JDM | Binary | Measure of dissassortivity using configuration model | Johnson et al. 2013 |
| Discrepancy | BR | Binary | Difference from a 'maximally packed' matrix | Brualdi and Sanderson, 1999 |
| Weighted NODF | WNODF | Weighted | Weighted version of NODF | Almeida-Neto and Ulrich, 2011 |
| Spectral radius | SR | Both | Maximum real eigenvalue of adjacency matrix | Staniczenko et al. 2013 |

Table 1: Nestedness measures available in FALCON.

increasing measure score for JDM, NODF and SR, but with decreasing measure score for BR, MD and NTC.

# 4 Comparison of nestedness scores

Nestedness is strongly sensitive to the size (number of rows and columns) and fill (number of non-zero entries) of the input matrix [Rodríguez-Gironés and Santamaría, 2006]. This is problematic in practical terms, since we often wish to compare nestedness of matrices that differ in these basic properties; in fact, cases where we compare empirically derived matrices with identical size and fill are an exception. Thus comparison of absolute values of nestedness metrics are not informative and may be misleading. To compare nestedness of matrices with differing size and fill, observed nestedness should always be interpreted in the context of a null distribution of matrices with similar properties. Measuring observed nestedness relative to expected nestedness derived from a null distribution of similar matrices allows determination of both effect size (e.g. as a $z$-score, which is commonly used to compare different nestedness schemes [Ulrich et al., 2009, Ulrich and Almeida-Neto, 2012])) and statistical significance (e.g. as a $p$-value giving the expected frequency of the observed score in the null distribution). This approach necessitates choice of a suitable null model and generation of a distribution of random matrices drawn from it.

In the present context, a null model is a method for creating a distribution of matrices that conserve some properties of the input matrix while varying other properties at random [Gotelli, 2001]. We continue the "context-free" approach in our treatment of null models; to allow comparison of nestedness across different scenarios, a good null model should not make assumptions about the mechanisms by which data were generated, but treat the matrix as an independent data structure. However, to be comparable to the input matrix, null matrices must conserve some key matrix properties (such as size and fill) on which nestedness depends. The null models available in FALCON are given in table 2; further detail is given in Appendix B. FALCON includes some of the more popular null models from the literature, alongside some additional null models that we feel can be useful. Null models vary in whether the original data is binary or quantitative, and in which properties of the original input matrix are preserved.

# 5 How FALCON works

## 5.1 Inputs and outputs

FALCON requires several inputs:

- an input network in the form of a bipartite matrix
- whether binary or quantitative nestedness should be investigated (quantitative matrices can be analysed using binary measures)

| Name | Description | Binary/weighted | Conserved features | Reference |
|---|---|---|---|---|
| SS | Shuffles positions randomly | Binary | Shape, Fill | Staniczenko et al. 2013 |
| FF | Permutations of structure with same node degrees | Binary | Shape, Fill, Degree | Miklós and Podani, 2004<br>Gotelli and Ulrich, 2011 |
| CC | Some structure is preserved, rest is shuffled | Binary | Shape, Fill | |
| DD | Determined probabilistically by node degree | Binary | Probabilistic Degree | Bascompte et al. 2003 |
| EE | Determined probabilistically by fill | Binary | Probabilistic Fill | |
| Binary Shuffle | Order of weighted links is swapped | Weighted | Binary positions, weights | Staniczenko et al. 2013 |
| CRT | Random weights where row totals conserved | Weighted | Binary positions, row totals | |
| CCT | Random weights where column totals conserved | Weighted | Binary positions, column totals | |
| RCTA | Average of conserve row totals and column totals | Weghted | Binary positions | |

Table 2: Null models in FALCON.

- whether to sort rows and columns to maximise nestedness score
- which nestedness measures should be used
- which null models nestedness should be tested under
- whether the ensemble of null models should be created with a fixed number or adaptively chosen
- whether or not to plot the distributions of nestedness scores

Output is returned to the user in the form of:

- the most nested configuration of the input matrix
- the nestedness measure(s) of the input matrix
- the expected value of nestedness under the null model(s) (as the mean measure of matrices created in the ensemble)
- the number of ensemble members used to calculate significance in each null model
- the statistical significance of the nestedness of the input matrix against each null model as a p-value
- the standard deviation and sample $z$-scores of the measure in the ensemble as well as other properties.

## 5.2 What FALCON does

FALCON follows the process shown in figure 2. First it sorts the user input matrix into a maximally nested configuration and removes any empty rows/columns before finding the nestedness of this matrix using the users chosen measures. Then FALCON goes through each of the user specified null models one by one, creating an ensemble of null matrices according to the rules of each null model. Each null matrix is then sorted and measured by each of the chosen nestedness measures. Thus, for each null model, nestedness measures are calculated for each of the null matrices in a single null ensemble, enabling direct comparison of results. The size of the null ensemble is determined by the input choice of using either the fixed or adaptive ensemble size (see section 5.5). Statistics are computed from the measures found in the null ensemble (and the direction in which that nestedness measure is calculated), before the next null model ensemble is instantiated. Once all null models have been computed the results are returned to the user.

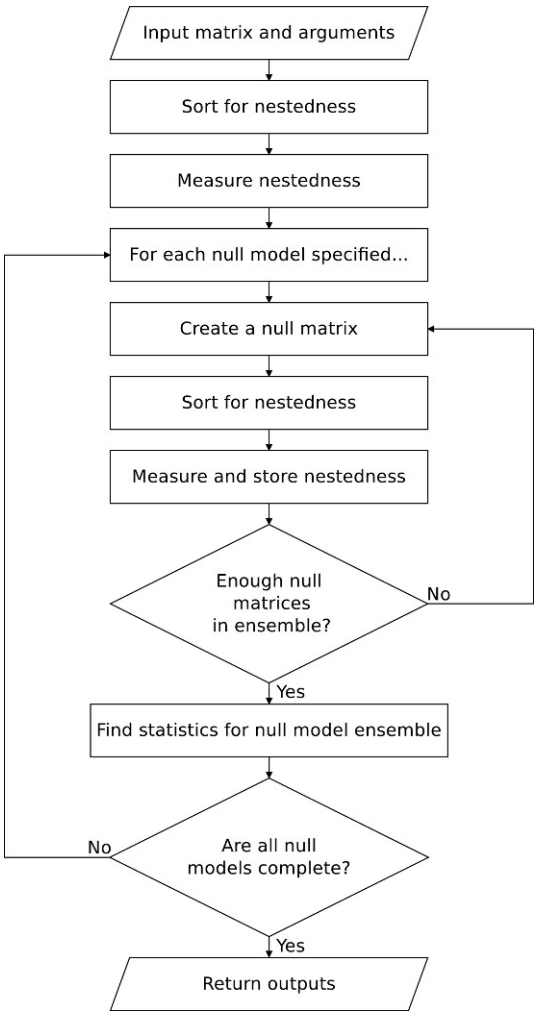


Figure 2: FALCON algorithmic procedure

### 5.3 Direction of increasing nestedness

For different nestedness measures, increasing scores can represent either increasing or decreasing nestedness as discussed in section §3. FALCON initially determines whether a higher measure score is related to greater nestedness (or vice versa) in the chosen measure by comparing the scores returned for a highly nested network (see figure 3A) and a highly non-nested network (a weighted checkerboard configuration; figure 3B), for which the fill (number of non-zero elements) and element sums are equal. The direction of increasing nestedness for a given measure is used during calculation of statistical significance. This method of determining direction each time the algorithm runs is included to

allow easy extensibility; if a new measure is added, FALCON will automatically determine which direction indicates increasing nestedness.

(A)

$$\begin{pmatrix} 10 & 9 & 8 & 7 & 6 & 5 & 4 & 3 & 2 & 1 \\ 9 & 8 & 7 & 6 & 5 & 4 & 3 & 2 & 1 & 0 \\ 8 & 7 & 6 & 5 & 4 & 3 & 2 & 1 & 0 & 0 \\ 7 & 6 & 5 & 4 & 3 & 2 & 1 & 0 & 0 & 0 \\ 6 & 5 & 4 & 3 & 2 & 1 & 0 & 0 & 0 & 0 \\ 5 & 4 & 3 & 2 & 1 & 0 & 0 & 0 & 0 & 0 \\ 4 & 3 & 2 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 3 & 2 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 2 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

(B)

$$\begin{pmatrix} 4 & 4 & 4 & 0 & 4 & 0 & 4 & 0 & 0 & 0 \\ 0 & 4 & 0 & 4 & 0 & 4 & 0 & 4 & 0 & 4 \\ 4 & 0 & 4 & 4 & 4 & 0 & 4 & 0 & 4 & 0 \\ 0 & 4 & 0 & 4 & 0 & 4 & 0 & 4 & 0 & 4 \\ 4 & 0 & 4 & 0 & 4 & 4 & 4 & 0 & 4 & 0 \\ 0 & 4 & 0 & 4 & 0 & 4 & 0 & 4 & 0 & 4 \\ 4 & 0 & 4 & 0 & 4 & 0 & 4 & 4 & 4 & 0 \\ 0 & 4 & 0 & 4 & 0 & 4 & 0 & 4 & 0 & 4 \\ 4 & 0 & 4 & 0 & 4 & 0 & 4 & 0 & 4 & 4 \\ 0 & 4 & 0 & 4 & 0 & 4 & 0 & 4 & 0 & 4 \end{pmatrix}$$

Figure 3: (A) Weighted nested matrix. (B) Weighted non-nested matrix (checkerboard). Both (A) and (B) have 55 non-zero elements that sum to 220.

### 5.4 Initial sort

For efficiency, FALCON is set up to initially sort the input matrix by row and column degrees for calculation of BR, MD and NODF, retains this sorted configuration for calculation of JDM and SR, and subsequently re-sorts for NTC in order to find the maximal nestedness of a binary matrix. For quantitative data, FALCON uses the same methods as for binary interactions, but also utilises weight data to break symmetry when two rows (columns) have the same degree; in this case, the row (column) which has greater values for most overlapping elements is ranked highest. Where two or more rows (columns) share the same degree and most overlapping elements, the rows (columns) are ranked according to the total sum of row (column) elements. This sorting does not affect the underlying topology or the relationships in the data. FALCON also allows the user to decide if any sorting is performed, enabling the "context free" assumption to be relaxed (e.g. for investigation of gradient-based nestedness [Ulrich and Almeida-Neto, 2012]).

### 5.5 Size of null ensemble

FALCON uses a bootstrap method to calculate the statistical significance of a given nestedness score, since the true null distributions of the test statistics are not known. The ensemble size used for this calculation can either be fixed or calculated adaptively by FALCON to improve computational efficiency and reduce undersampling effects. Note that the strongest significance that can be assigned is $p < \frac{1}{N}$ where $N$ is the ensemble size.

### Fixed

The number of null matrices used to make up the ensemble is fixed by the user. This method is effective providing that the ensemble is large enough to have statistical power; the larger the ensemble, the more power the test has and the closer the answer will be to the $p$-value for the (unknown) true null distribution. However, it is not obvious how large the ensemble needs to be; in the literature, amongst others, Joppa et al., 2010 use 1,000 null models in their ensembles, whilst McQuaid and Britton, 2013 use 10,000, and Flores et al., 2013 use 100,000. A large number of different null matrix configurations are possible for a given input matrix and we may wish to avoid undersampling [Poisot and Gravel, 2014]; however, at the same time very large ensembles can make the calculation of significance computationally intractable.

### Adaptive

FALCON includes a mechanism for adaptive determination of ensemble size. This is intended to ensure robust statistics are achieved, avoiding concerns about undersampling or oversampling [Poisot and Gravel, 2014], while minimising computational load. The adaptive method works by creating two ensembles in parallel using the same null model. Starting with a minimum ensemble size of 500 in each group, the ensembles are expanded until they show similar statistical properties. This condition is met when the null hypothesis (both ensembles come from the same distribution) of a Mann-Whitney U-test cannot be rejected at 10% significance. When this occurs, it suggests each group represents a good sample of the underlying distribution, and the two groups are combined to form a single null ensemble used to calculate final statistics. The expansion of the size of the ensemble has an upper limit of 100,000 members in case the null hypothesis is always rejected. The adaptive ensemble methods balances statistical precision with computational efficiency; we conservatively use 1,000 as a minimum final ensemble size such that a $p$-value as low as 0.001 can be assigned.

## 5.6 Output Statistics

### $p$-value

The $p$-value is the probability that a matrix drawn from the null distribution will be more nested than the input matrix. Low values ($p \to 0$) indicate that the input matrix is highly nested relative to the null distribution; commonly a threshold of $p \leq 0.05$ or $p \leq 0.01$ is used to denote a statistically significant level of nestedness.Here $p$ is calculated by counting the frequency of matrices in the null ensemble that are more nested than the input matrix; for cases where no member of the null ensemble is more nested than the input matrix we conservatively assign $p < \frac{1}{N}$ where $N$ is the ensemble size.

#### Normalised Temperature

The normalised temperature is inspired by the $\tau$-Temperature [Corso and Britton, 2012]. It describes the relationship between the nestedness measure found for the input matrix and the expected nestedness measure derived from the null model ensemble. It is described as:

$$T = \frac{Measure}{< Measure >} \tag{1}$$

where $< Measure >$ denotes the expected value. In simple terms, the normalised temperature indicates whether the input matrix is more or less nested than the expectation for a null distribution of similar matrices. Where the measure gives increasing scores with increasing nestedness, $T > 1$ indicates greater-than-expected nestedness. Where the measure gives decreasing scores with increasing nestedness, $T < 1$ indicates greater-than-expected nestedness.

#### Mean

The mean average of the set of nestedness measure found for each of the ensemble members is returned.

#### Standard Deviation

The standard deviation ($\sigma$) of the set of nestedness measures found for each of the ensemble members is returned.

#### Sample z-score

The $z$-score, or standard score, is calculated as the difference between the nestedness measure and its expected value divided by the standard deviation of the sample:

$$z = \frac{Measure - < Measure >}{\sigma} \tag{2}$$

It is a measure of the number of standard deviations the nestedness measure of the input matrix is above the expected value. Hence, the way it should be interpreted, as with the normalised temperature, depends on whether nestedness increases with increasing measure score.

### 5.7 Performance and complexity

A user's choice of which null models and nestedness measures to use for their analysis will depend on the questions they wish to ask of their data. However, these choices also need to consider the computational resources available for the task. In figure 4 we show the time complexity to compute some of the nestedness measures and null models that FALCON can perform. The graphs below show the average time to measure a single matrix or generate a single null matrix and the standard deviation from this. 10 trials were used starting

with different 10×10 matrices. In each trial each matrix was measured 1,000 times and 1,000 null matrices were generated from the initial trial matrix. The time taken to conduct each trial was then divided by 1,000 and the average and standard deviation was calculated.

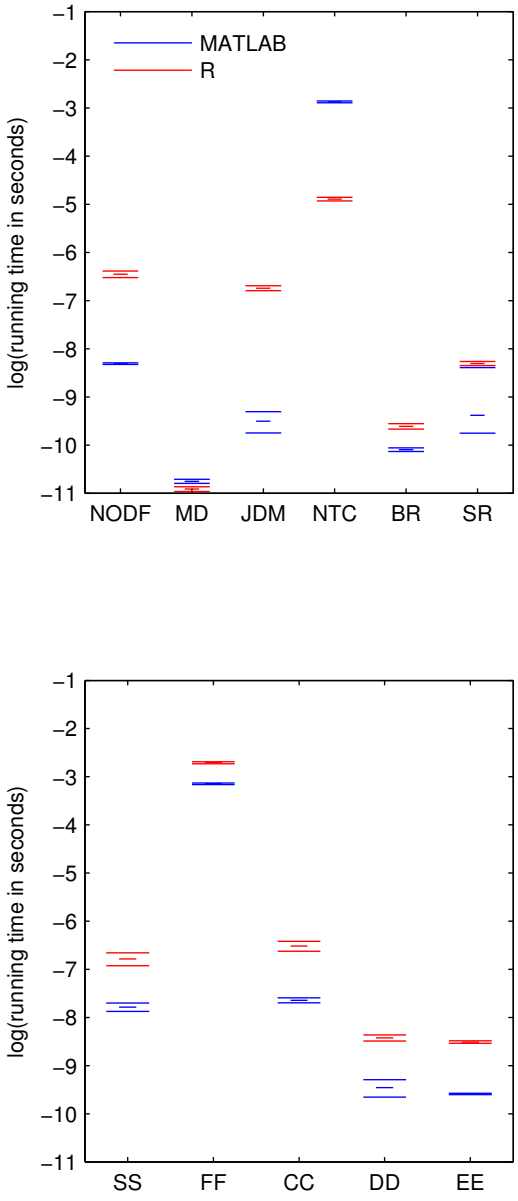


Figure 4: (A) Time taken to perform each of the binary nestedness measures on the same 10×10 matrix averaged over 1,000 measurements. SR - spectral radius, MD - Manhattan distance, NTC - nestedness temperature calculator, BR - discrepancy. (B) Time taken to generate a 10×10 null matrix using each of the binary null models averaged over 1,000 times.

# 6 Summary

In this paper we have presented FALCON, a software tool for reliable and efficient calculation of nestedness (and associated effect size and statistical significance) based on a selection of popular nestedness measures and null models used in the literature. FALCON treats nestedness purely as a statistical property of a bipartite matrix and removes any form of interpretation or contextual information from the analysis. This enables FALCON to be used to compare nestedness across a wide variety of application areas, noting that the concept of nestedness has already spread from its origin in island biogeography to include species-species interactions and other scenarios, and is likely to find further application in other domains. The contribution of FALCON is to enable easy cross-comparison of observed nestedness using different nestedness measures and null models. We hope that this functionality will allow greater methodological uniformity and comparability of studies of nestedness. Uniformity of measurement and comparability of empirical results is an important preliminary step that must be achieved to enable understanding of the mechanistic basis and ecological (and otherwise) implications of nestedness. We hope that FALCON will be of use to other researchers and help illuminate this intriguing property of bipartite networks in many natural systems.


# Acknowledgements

We thank Virginia Domínguez-García with implementation of JDM nestedness. S.J.B., C.A.B and H.T.P.W. acknowledge funding from the University of Exeter.


# 7 Appendix A: Detailed description of nestedness measures in FALCON

## 7.1 Binary

### NODF

The nestedness measure based on overlap and decreasing fill (NODF) was first described by [Almeida-Neto et al., 2008] and has since become one of the most popular methods for describing the nestedness of a matrix. NODF can be found as:

$$NODF = \frac{N_{col} + N_{row}}{\frac{(c(c-1)}{2} + \frac{(r(r-1))}{2}} = \frac{2(N_{col} + N_{row})}{c(c-1) + r(r-1)} \tag{3}$$

Here $N_{col}$ and $N_{row}$ are scores found by pairwise comparison of rows and columns, $c$ is the number of columns, and $r$ is the number of rows. $N_{col}$ is found as the sum of scores from pairwise comparisons of each column against all columns to its right. If both columns have the same degree then the score

is zero. If they have different degrees, the score is the percentage of elements in the second column which also appear in the first column. $N_{rows}$ is found similarly for pairwise comparisons of each row against all rows below it. The sum of $N_{col}$ and $N_{row}$ is then normalised by the total number of pairwise comparisons. Values for NODF are between 0 (zero nestedness) and 100 (perfect nestedness). If the input matrix is first sorted to maximise nestedness by rank ordering rows and columns by degree, the form of NODF known as $NODF_{MAX}$ is found [Podani and Schmera, 2012].

### $\tau$-Temperature and Manhattan Distance

The $\tau$-Temperature [Corso and Britton, 2012] is a nestedness measure based on relative distances between matrix elements. Unlike other distance-based measures (such as NTC [Atmar and Patterson, 1993] and its better described successors BINMATNEST [Rodríguez-Gironés and Santamaría, 2006] and AININHADO [Guimarães and Guimarães, 2006]), the $\tau$-Temperature does not use genetic algorithms to sort the data. The $\tau$-Temperature is found by measuring the Manhattan distance $D$ of the network matrix. This is the sum of the row and column indexes of all of the matrix elements $A_{ij}$ that are filled:

$$D = \sum_{A_{ij}>0} (i + j) \tag{4}$$

Manhattan distance is lower in more highly nested networks, since rows and columns can be shuffled so that many of the elements appear in upper-left positions where row and column indices are low. Once $D$ is found, a null model is chosen (cf. section §4) and an ensemble of null matrices are created. By finding the mean average Manhattan distance from the ensemble, denoted $< D_{rand} >$, $\tau$-Temperature can be calculated as:

$$\tau = \frac{D}{< D_{rand} >} \tag{5}$$

Values $\tau > 1$ imply that $D$ is greater than $< D_{rand} >$ and the network is less nested than expected for a network with the properties defined in the null model. $\tau$ is better described as a test statistic of the Manhattan distance, than as a measurement of nestedness itself.

### JDM Nestedness

The nestedness measure described in [Johnson et al., 2013], here termed JDM after author initials, treats nestedness as a measure of dissassortativity between the nodes, i.e., negative correlation between row and column degrees for non-zero elements of the input matrix. Their measure calculates the overlap (as the sum of the elements in the squared adjacency matrix which shows the minimum number of length two paths needed to connect any two nodes) of the input matrix and normalises it by the expected nestedness of the configuration model ( a random graph with the same empirical degree distribution as the input

network) and thus discounts the effect of degree heterogeneity. This nestedness score is unbounded, but when close to 1 it indicates that the matrix represents an uncorrelated random network. Unadjusted nestedness $\tilde{\eta}$ is calculated using the adjacency matrix $a$ formed from the input bipartite matrix with $r$ rows and $c$ columns, where $D$ is the node degrees in the adjacency matrix:

$$\tilde{\eta} = \frac{1}{(r+c)^2} \sum_{i}^{(r+c)} \sum_{j}^{(r+c)} \left( \frac{a_{ij}^2}{D_i D_j} \right) \tag{6}$$

Nestedness of the configuration model $\eta_{conf}$ can be calculated as:

$$\eta_{conf} = \frac{r \left( \frac{\sum_i^r k_i^2}{c} \right) + c \left( \frac{\sum_j^c d_j^2}{r} \right)}{\left( \frac{\sum_i^r k_i}{c} \right) \left( \frac{\sum_j^c d_j}{r} \right) (r+c)^2} \tag{7}$$

which can also be written as:

$$\eta_{conf} = \frac{r < k^2 >_c + c < d^2 >_r}{< k >_c < d >_r (r+c)^2} \tag{8}$$

where $k$ are the row degrees and $d$ are column degrees in the bipartite matrix. This leads to the normalised measure of nestedness for bipartite networks defined by [Johnson et al., 2013] as:

$$\eta_{bip} = \frac{\tilde{\eta}}{\eta_{conf}} \tag{9}$$

**Nestedness Temperature**

The original nestedness temperature calculator (NTC) Atmar and Patterson, 1993 was vaguely described and therefore difficult to re-implement, leading to several subsequent variations utilising similar underlying principles [Rodríguez-Gironés and Santamaría, 2006, Guimarães and Guimarães, 2006, Ulrich and Gotelli, 2007, Oksanen et al., 2013]. Here we have recoded the `nestedtemp` function from the R package *vegan* [Oksanen et al., 2013]. The nestedness temperature for an input matrix is based on the 'isocline of perfect order', a curve drawn from the lower-left corner of the matrix to the upper-right, with curvature defined by matrix fill (see figure 5). Row and column orderings are then permuted using a genetic algorithm to maximise the number of connections above the isocline and minimise connections below the isocline. The number of connections which violate these rules, termed 'surprises', are then counted and normalised to give a score between 0 (highly ordered) and 100 (highly disordered).

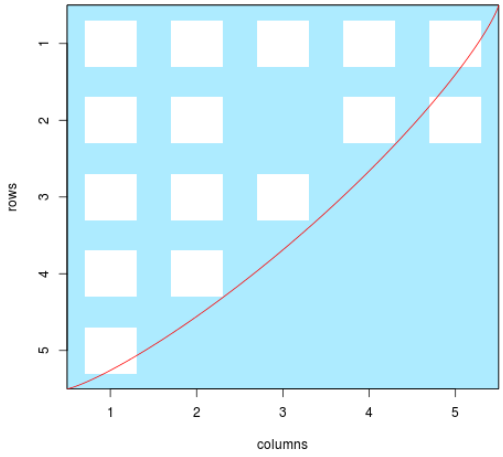


Figure 5: The NTC measure is based on an isocline of perfect order. Here there are two 'surprises'.

#### Discrepancy

Discrepancy [Brualdi and Sanderson, 1999], here denoted BR, quantifies nestedness as the difference between the input matrix and a perfectly nested matrix of the same dimensions and fill. A duplicate matrix $P$ with the same row degrees as the input matrix, but where the 1s in each row are pushed as far to the left as possible (ignoring the effect this has on underlying network topology). The discrepancy is then found by subtracting the input matrix from this perfectly nested matrix and counting the number of 1s that remain – the number of differences between $P$ and the input matrix. A different discrepancy score can be found by treating columns instead of rows, forming an alternative perfectly nested comparitor matrix $P'$ by pushing the 1s in each column to the top.(instead of the row) degrees of the input matrix to form $P'$ a different discrepancy score can be found. Here we modify the original method of [Brualdi and Sanderson, 1999], which looks at discrepancy only in respect to $P$, and instead define discrepancy as the minimum of the individual discrepancy scores found from $P$ and $P'$, to remove any bias towards row or column nodes.

### 7.2 Quantitative

#### WNODF

The weighted NODF measure, WNODF [Almeida-Neto and Ulrich, 2011], uses a similar algorithm to NODF, but is designed for use on quantitative rather than binary networks. In addition to asking which pairs of rows/columns are subsets of one another, WNODF utilises weight information by also requiring

that the preceding row/column has greater values in the overlapping elements. In effect, WNODF is a stricter version of NODF; the maximum WNODF score that can be achieved for a quantitative matrix is equal to the NODF score for the binary matrix.

### 7.3 Both

#### Spectral Radius

The spectral radius (SR) is defined as the absolute value of the maximum real eigenvalue from the adjacency matrix of a given input bipartite matrix. SR was proposed as a nestedness measure by Staniczenko et al. [2013] and can be applied to both binary and quantitative matrices.

## 8 Appendix B: Null models available in FALCON

### 8.1 Binary

#### Swappable-Swappable (SS)

The "swappable rows, swappable columns" (SS) null model conserves matrix dimensions (numbers of rows and columns) and fill. It is similar to 'test one' in Staniczenko et al. [2013], which works by shuffling elements at random within the matrix; however, it differs in that degenerate matrices (those containing rows/columns with no connections) are not permitted.

#### Fixed-Fixed (FF)

The "fixed rows, fixed columns" (FF) null model conserves dimensions, fill and degree distribution of the original matrix. It is the most strict null model we consider here and is known to suffer from Type II errors (i.e. a failure to detect nestedness) [Joppa et al., 2010]. Null matrix instantiations are found by searching the input matrix for 2×2 submatrices and comparing them to the 2×2 identity matrix and its mirror image:

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \qquad \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

If one of these configurations is found, the corresponding rows and columns are swapped, changing the submatrix into the other configuration shown. This has been called the Miklós-Podani trial swap algorithm [Miklós and Podani, 2004] as it counts the number of trials rather than the number of actual swaps (which could lead to a biased sample if the solution is periodic). We use an initial 30,000 (or row×column if this is larger) trial swaps as a spinup for this algorithm in order to 'escape' the initial matrix, then take a sample null matrix every additional 5,000 trial swaps [Gotelli and Ulrich, 2011]. It should be noted that as the Manhattan distance is invariant to these permutations.

### Cored-Cored (CC)

The "cored rows, cored columns" (CC) null model conserves dimensions and fill as in the SS null model, but also conserves some of the core structure found in the observed input matrix. Thus it is somewhere between the SS and the FF null models. It is found by trial-swapping a total of row×column elements from the input matrix and changing any 1's to 0's when this does not reduce the corresponding row or column degree to zero. This ensures that the size structure is conserved and preferentially preserves specialist interactions within the network. The removed elements are then randomly reassigned to the remaining empty spaces to preserve matrix fill.

### Degreeprobable-Degreeprobable (DD)

The "degreeprobable rows, degreeprobable columns" (DD) null model first described by Bascompte et al., 2003 has subsequently been a popular choice for application to species-species nested comparisons. Matrix elements are probabilistically determined depending on the degree distribution of the rows and columns of the initial matrix as:

$$p_{ij} = \frac{1}{2}\left(\frac{d_j}{r} + \frac{k_i}{c}\right) \tag{10}$$

where $p_{ij}$ is the probability of assigning a 1 to the ith row and jth column of the null matrix, $d_j$ is the column degree of the jth column, $k_i$ is the row degree of the ith row and $r$ and $c$ are the respective number of rows and columns. Due to the stochastic nature of this null model its output matrices will vary in size and fill.

### Equiprobable-Equiprobable (EE)

The "equiprobable rows, equiprobable columns" (EE) null model is probabilistic and assumes that the probability of a connection occurring between two nodes is related to the number of total connections in the input matrix. Hence for an input matrix with fill $M$, $r$ rows and $c$ columns, the probability of a connection being present between two nodes is $p_{ij} = \frac{M}{r \times c}$. Due to the stochastic nature of this null model its output matrices will vary in size and fill. It is the least strictly defined null model we consider here and is known to suffer from Type I errors (a tendency to falsely detect nestedness) [Joppa et al., 2010].

## 8.2 Quantitative

### Binary Shuffle

This null model was employed by Staniczenko et al., 2013 and conserves the entire binary structure of the input matrix and the values of the elements in the matrix, but shuffles the order of these values randomly across the binary structure.

### Conserve Row Totals (CRT)

This null model conserves binary structure and the row sum totals, but the values of the elements on each row are changed such that each connection in the row is assigned a random proportion of the row sum total.

### Conserve Column Totals (CCT)

This null model conserves binary structure and the column sum totals, but the values of the elements in each column are changed such that each connection in the column is assigned a random proportion of the column sum total.

### Row Column Totals Average (RCTA)

Both of the two above null models conserve information related to either the rows or the columns, giving this property precendent over that of the other entity. We also introduce the Row Column Totals Average (RCTA) null model which uses the average of a single null model made from each of the CCT and CRT null models. As information from both rows and columns is utilised in the creation of this null model it may better fit with the context free ethos of nestedness we persue than either of CRT or CCT alone.

## References


M. Almeida-Neto and W. Ulrich. A straightforward computational approach for measuring nestedness using quantitative matrices. *Environmental Modelling & Software*, 26(2):173–178, 2011. doi: 10.1016/j.envsoft.2010.08.003.

M. Almeida-Neto, P. Guimaraes, P. R. Guimarães, R. D. Loyola, and W. Ulrich. A consistent metric for nestedness analysis in ecological systems: reconciling concept and measurement. *Oikos*, 117(8):1227–1239, 2008. doi: 10.1111/j.0030-1299.2008.16644.x.

W. Atmar and B. D. Patterson. The measure of order and disorder in the distribution of species in fragmented habitat. *Oecologia*, 96(3):373–382, 1993. doi: 10.1007/BF00317508.

J. Bascompte, P. Jordano, C. J. Melián, and J. M. Olesen. The nested assembly of plant–animal mutualistic networks. *Proceedings of the National Academy of Sciences*, 100(16):9383–9387, 2003. doi: 10.1073/pnas.1633576100.

Andrés Baselga. The relationship between species replacement, dissimilarity derived from nestedness, and nestedness. *Global Ecology and Biogeography*, 21(12):1223–1232, 2012. doi: 10.1111/j.1466-8238.2011.00756.x.

S. J. Beckett and H. T. P. Williams. Coevolutionary diversification creates nested-modular structure in phage–bacteria interaction networks. *Interface Focus*, 3(6):20130033, 2013. doi: 0.1098/rsfs.2013.0033.

S.J. Beckett, C.A. Boulton, and H.T.P Williams. *FALCON: nestedness statistics for bipartite networks*, 2014. URL `http://dx.doi.org/10.6084/m9.figshare.999117`.

R. A. Brualdi and J. G. Sanderson. Nested species subsets, gaps, and discrepancy. *Oecologia*, 119(2):256–264, 1999. doi: 10.1007/s004420050784.

S. Bustos, C. Gomez, R. Hausmann, and C. A. Hidalgo. The dynamics of nestedness predicts the evolution of industrial ecosystems. *PloS one*, 7(11): e49393, 2012. doi: 10.1371/journal.pone.0049393.

Vasthi Alonso Chávez, C Patrick Doncaster, John A Dearing, Rong Wang, Jing-Lun Huang, and James G Dyke. Detecting regime shifts in artificial ecosystems. In *Advances in Artificial Life, ECAL*, volume 12, pages 625–632, 2013. doi: 10.7551/978-0-262-31709-2-ch089.

G. Corso and N. F. Britton. Nestedness and $\tau$-temperature in ecological networks. *Ecological Complexity*, 11:137–143, 2012. doi: 10.1016/j.ecocom.2012.05.003.

Peter Csermely, András London, Ling-Yun Wu, and Brian Uzzi. Structure and dynamics of core/periphery networks. *Journal of Complex Networks*, 1(2): 93–123, 2013. doi: 10.1093/comnet/cnt016.

B. Dalsgaard, K. Trøjelsgaard, A. M. M González, D. Nogués-Bravo, J. Ollerton, T. Petanidou, B. Sandel, M. Schleuning, Z. Wang, C. Rahbek, et al. Historical climate-change influences modularity and nestedness of pollination networks. *Ecography*, 36(12):1331–1340, 2013. doi: 10.1111/j.1600-0587.2013.00201.x.

Philip Jackson Darlington. *Zoogeography: the geographical distribution of animals*. Wiley, New York, 1957.

R Daubenmire. Floristic plant geography of eastern washington and northern idaho. *Journal of Biogeography*, pages 1–18, 1975.

C. O. Flores, J. R. Meyer, S. Valverde, L. Farr, and J. S. Weitz. Statistical structure of host–phage interactions. *Proceedings of the National Academy of Sciences*, 108(28):E288–E297, 2011. doi: 10.1073/pnas.1101595108.

C. O. Flores, S. Valverde, and J. S. Weitz. Multi-scale structure and geographic drivers of cross-infection within marine bacteria and phages. *ISME Journal*, 7:520–532, 2013. ISSN 1751-7370. doi: 10.1038/ismej.2012.135.

M. A. Fortuna, D. B. Stouffer, J. M. Olesen, P. Jordano, D. Mouillot, B. R. Krasnov, R. Poulin, and J. Bascompte. Nestedness versus modularity in ecological networks: two sides of the same coin? *Journal of Animal Ecology*, 79(4):811–817, 2010. doi: 10.1111/j.1365-2656.2010.01688.x.

Ceridwen I Fraser, Aleks Terauds, John Smellie, Peter Convey, and Steven L Chown. Geothermal activity helps life survive glacial cycles. *Proceedings of the National Academy of Sciences*, page 201321437, 2014. doi: 10.1073/pnas.1321437111.

N. J. Gotelli. Research frontiers in null model analysis. *Global ecology and biogeography*, 10(4):337–343, 2001. doi: 10.1046/j.1466-822X.2001.00249.x.

N. J. Gotelli and W. Ulrich. Over-reporting bias in null model analysis: a response to fayle and manica (2010). *Ecological Modelling*, 222(7):1337–1339, 2011. doi: 10.1016/j.ecolmodel.2010.11.008.

Nicholas J Gotelli. Null model analysis of species co-occurrence patterns. *Ecology*, 81(9):2606–2621, 2000. doi: 10.1890/0012-9658(2000)081[2606:NMAOSC]2.0.CO;2.

P. R. Guimarães and P. Guimarães. Improving the analyses of nestedness for large sets of matrices. *Environmental Modelling & Software*, 21(10):1512–1513, 2006. doi: 10.1016/j.envsoft.2006.04.002.

Eric Hultén. *Outline of the history of arctic and boreal biota during the Quaternary period: their evolution during and after the glacial period as indicated by the equiformal progressive areas of present plant species*, volume 1. Bokförlags Aktiebolaget Thule, 1937.

A. James, J. W. Pitchford, and M. J. Plank. Disentangling nestedness from models of ecological complexity. *Nature*, 487(7406):227–230, 2012. doi: 10.1038/nature11214.

S. Johnson, V. Domínguez-García, and M. A. Muñoz. Factors determining nestedness in complex networks. *PloS one*, 8(9):e74025, 2013. doi: 10.1371/journal.pone.0074025.

L. N. Joppa, J. M. Montoya, R. Solé, J. Sanderson, and S. L. Pimm. On nestedness in ecological networks. *Evolutionary Ecology Research*, 12:35–46, 2010. ISSN 1522-0613.

J Jelle Lever, Egbert H Nes, Marten Scheffer, and Jordi Bascompte. The sudden collapse of pollinator communities. *Ecology letters*, 17(3):350–359, 2014. doi: 10.1111/ele.12236.

C. F. McQuaid and N. F. Britton. Host-parasite nestedness: A result of co-evolving trait-values. *Ecological Complexity*, 13:53 – 59, 2013. ISSN 1476-945X. doi: http://dx.doi.org/10.1016/j.ecocom.2013.01.001. URL `http://www.sciencedirect.com/science/article/pii/S1476945X13000044`.

Adriano S Melo, Marcus V Cianciaruso, and Mário Almeida-Neto. treenodf: nestedness to phylogenetic, functional and other tree-based diversity metrics. *Methods in Ecology and Evolution*, 2014. doi: 10.1111/2041-210X.12185.

I. Miklós and J. Podani. Randomization of presence-absence matrices: comments and new algorithms. *Ecology*, 85(1):86–92, 2004. doi: 10.1890/03-0101.

J. Oksanen, F. Guillaume Blanchet, R. Kindt, P. Legendre, P. R. Minchin, R. B. O'Hara, G. L. Simpson, P. Solymos, M. H. H. Stevens, and H. Wagner. *vegan: Community Ecology Package*, 2013. URL `http://CRAN.R-project.org/package=vegan`. R package version 2.0-10.

Anke Piepenbrink and Ajai S Gaur. Methodological advances in the analysis of bipartite networks an illustration using board interlocks in indian firms. *Organizational Research Methods*, 16(3):474–496, 2013. doi: 10.1177/1094428113478838.

J. Podani and D. Schmera. A comparative evaluation of pairwise nestedness measures. *Ecography*, 35(10):889–900, 2012. doi: 10.1111/j.1600-0587.2011.07319.x.

T. Poisot and D. Gravel. When is an ecological network complex? connectance drives degree distribution and emerging network properties. *PeerJ*, 2:e251, 2014. doi: 10.7717/peerj.251.

M. A. Rodríguez-Gironés and L. Santamaría. A new algorithm to calculate the nestedness temperature of presence–absence matrices. *Journal of Biogeography*, 33(5):924–935, 2006. doi: 10.1111/j.1365-2699.2006.01444.x.

S. Saavedra, D. B. Stouffer, B. Uzzi, and J. Bascompte. Strong contributors to network persistence are the most vulnerable to extinction. *Nature*, 478(7368):233–235, 2011. doi: 10.1038/nature10433.

P. P. A. Staniczenko, J. C. Kopp, and S. Allesina. The ghost of nestedness in ecological networks. *Nature Communications*, 4:1391, 2013. doi: 10.1038/ncomms2422.

E. Thébault and C. Fontaine. Stability of ecological communities and the architecture of mutualistic and trophic networks. *Science*, 329(5993):853–856, 2010. doi: 10.1126/science.1188321.

M.T. Tinker, P. R. Guimarães, Mark Novak, Flavia Maria Darcie Marquitti, James L Bodkin, Michelle Staedler, Gena Bentall, and James A Estes. Structure and mechanism of diet specialisation: testing models of individual variation in resource use with sea otters. *Ecology letters*, 15(5):475–483, 2012. doi: 10.1111/j.1461-0248.2012.01760.x.

W. Ulrich and M. Almeida-Neto. On the meanings of nestedness: back to the basics. *Ecography*, 35(10):865–871, 2012. doi: 10.1111/j.1600-0587.2012.07671.x.

W. Ulrich and N. J. Gotelli. Null model analysis of species nestedness patterns. *Ecology*, 88(7):1824–1831, 2007. doi: 10.1890/06-1208.1.

W. Ulrich, M. Almeida-Neto, and N. J. Gotelli. A consumer's guide to nestedness analysis. *Oikos*, 118(1):3–17, 2009. doi: 10.1111/j.1600-0706.2008.17053.x.