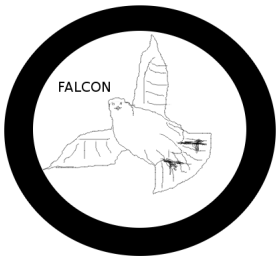
FALCON - Framework for Adaptive ensembLes for the Comparison Of Nestedness

Stephen J. Beckett*, Hywel T. P. Williams

October 10, 2013

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



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

1 Introduction

This document relates to FALCON a free software package available to download from https://github.com/sjbeckett/FALCON. FALCON stands for Framework for Adaptive ensembLes for the Comparison Of Nestedness. This software is designed to allow users interested in the nestedness properties of bipartite networks to compare the measures and statistical properties of some of the plethora of different nestedness measures and null models available in the literature using an adaptive sampling method. We aim to use this package to perform a comparison of these measures and null models using ecological datasets and question which measures do and do line up with each other. This document outlines and describes the methodology behind the functions used in the FALCON package. For details on how to run FALCON see the `Instructions_for_FALCON.pdf` file.

There exists a plethora of different methods to measure the nestedness, the 'amount of order/disorder', in a matrix as first envisaged by [3]. Their original concept was based on extending the ideas of Island Biogeography theory [8], such

that as well as expecting that smaller islands that are further from mainland have lower species diversity and abundance due to immigration/emmigration processes that these islands contain a species composition that is a subset of the species compositions of islands closer to the mainland. In a perfect setting this would mean that the mainland would contain all the species found on any of the islands, such that the perfectly nested matrix contains an upper triangular pattern. This site-species bipartite network concept has been extended by others to cover many species-species bipartite networks, which are undergoing mutualistic or antagonistic interactions [5, 7] and also species-time relationships and other types of networks and therefore the concept of nestedness should be regarded as a network structure in its own right, applicable to any type of bipartite network.

2 Procedure

The user inputs the matrix they wish to test for nestedness and several options that specify how they want to measure it. These are: which measurement method they wish to measure nestedness with, whether they are interested in binary or quantitative nestedness, which null model they wish to use in significance testing as well as which kind of solver they wish to use to perform the ensemble of null models. FALCON sorts the user input matrix into a nested configuration and removes any empty rows/columns before finding the nestedness of this matrix against the users choice. It determines whether a higher measure score is related to greater nestedness (or vice versa) in the chosen measure by comparing the measure taken from a highly nested network, a weighted upper triangular matrix(below left) using [13]'s weighted nestedness definition against a highly un-nested weighted near checkerboard configuration(below right) in which the occupation and element sums are equal. This information is used to calculate the statistical significance of nestedness as a p-value in an ensemble which is assembled for each of the null models that the user wishes to test. Output is returned to the user in the form of the nested configuration of the input matrix, the measure of the input matrix, the expected value of nestedness under the null model (as the mean measure of matrices created in the ensemble), the number of ensemble members, the statistical significance of the nestedness of the input matrix against the null model as a p-value as well as the standard deviation and sample z-scores of the measure in the ensemble.

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

3 Nested Configuration

For binary matrices the process of sorting matrices into a more nested configuration is relatively simple when (as we do here) we define nestedness in the same way as [1], such that the most nested configuration is that when rows and columns are assembled such that the next row/column is a subset of the preceeding one. Then the most nested configuration will be that such that the rows and columns are ordered in terms of their degree i.e. the number of connections between a node of one set and a node of the other set. Then rows and columns with a degree of zero are removed.

The ordering of rows and columns makes a difference to many, but not all of the measures (spectral radius) described below. We decide to sort input matrices (as well as the matrices produced by the null models) to this most nested ordering on the grounds that we are not investigating underlying features of the network, but rather the structure of nestedness itself, though understanding the processes involved in creating nested structures is important[14]. In many cases any underlying features may be unknown, or it may not be known how much they contribute to the nestedness structure. In practise, nestedness probably has more than one causal mechanism. However, some users may wish to test for the nestedness of a matrix in a set order rather than that of the nested configuration of their matrix. Enabling significance tests under this setup is work for furture versions of FALCON, however users can attain the raw measure score by running the required functions directly from the MEASURES folder.

Steps to address the ordering of a quantitative matrix may need to be considered. There may be several ways of determining the best way to sort for the most nested quantitative matrix when there are two or more rows or columns with the same degree. Here we decide that if two rows have the same degree that it is the row which has more element values greater than the other - i.e. looking at the pairwise comparisons of non-zero column elements shared between the two rows, which row has the largest number of higher values. The same concept is applied to columns that share the same degree.

4 Solvers

Fixed

The majority of significance testing is performed using what we term a 'fixed' solver. That is the number of null matrices used to make up the ensemble is determined by the user and is a fixed number. For the purposes of deciding whether a matrix is likely to be significantly nested or anti-nested at an arbitrary threshold (i.e. $p<0.05$ or $p>0.95$) this method is probably good enough providing that enough members are included in the ensemble to give it some statistical power. The larger the ensemble, the more power the test has and the closer the answer will be to the true p-value. In the literature among others [7] use 1,000 null models in their ensembles, whilst [9] use 10,000. However, there is no robust process in determining these numbers. It is up to the user to decide how well to sample from the underlying (and often unkown) distribution of the possible values of the null model.

Adaptive

We approach this unsatisfying quality of the fixed solver, by introducing an adaptive approach. We illustrate one method here, but other adaptive methods and variations on what we present are possible. In our approach we take two sampling groups from the underlying distribution simultaneously. After the ensembles for each group have reached some initial minimum number of members (we use 250 in each group) we start to compare a summary statistic calculated from each of the two samples in an effort to question how alike the two sampling groups are. Once the summary statistic (here we use what we call the normalised temperature - defined as the nestedness measure of the input divided by the mean average nestedness measurement in the ensemble) from each group are within some threshold of each other we say the distribution is at least reasonably well sampled and then can combine the two group ensembles together before running statistics on them. The comparison steps are performed for every 25 new members to each group; and if the condition is not fulfilled when the maximum number of ensemble members is reached (10,000 in total here) then the members of each group are combined at this stage before running the statistics. In effect the adaptive solver is playing off between the degree of precision of the statistical properties found according to the threshold difference between the two groups and a large sample size. If the threshold limit is achieved then reliably good statistics (never perfect!) can be given (dependent on the threshold choice) without having to run so many null models. In the case that threshold limit is not achieved then the adaptive solver is doing just as well as the fixed solver with an ensemble size of 10,000.

5 Meaures

5.1 Binary

NODF

The nestedness measure based on overlap and decreasing fill(NODF) was first described by [1] and has since become one of the most popular methods for describing the nestedness of a matrix. This is in part due to its ease of use, as it is deterministic and does not need to be statistically evaluated against an ensemble of null models and also to the transparency of its methods as its calculation is explicitly shown and is not subject to difficult calculations or algorithmic procedures. As we do not (and in general it isn't) know of any ecological gradient across our matrix, we use the form of NODF known as $NODF_{MAX}$ [11]. We first find the degree of each row and column by finding the number of elements in each row/column with a value greater than zero. To achieve $NODF_{MAX}$ the matrix must be sorted by the degree of the rows and the columns such that the rows with highest degree are at the top of the matrix and the rows with lowest degree are at the bottom; and repeat with the columns such that the rows(columns) with largest degree have the lowest row(column) index. Then $NODF_{COLUMNS}$ can be found by pairwise evaluating columns. This works by taking the first column(highest degree) and comparing it to the second column. If both columns have same degree there cannot be overlap, so a zero score is assigned, else the score is the percentage of elements in the second column which also appear in the first column i.e. how much of a subset the second column is of the first. This process is repeated to compare the first column against all the other columns; then the second column is compared against the remaining columns, of which it should be a superset; and so on. The sum of all these scores is the column score, N_{col}. $NODF_{COLUMNS}$ is found by dividing N_{col} by the number of comparisons, the number of columns (c) multiplied by the number of columns minus one(columns aren't assessed against themselves) all divided by two $\left(\frac{c(c-1)}{2}\right)$. $NODF_{ROWS}$ can be found in a similar manner. Putting these two measures together $NODF_{MAX}$ can be found as:

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

Values for each of the NODF indexes is measured between 0, low nestedness and 100, perfect nestedness.

τ-Temperature and Manhattan Distance

The τ-Temperature[4] measure is an example of a nestedness measure based on relative distances in a matrix of the present and absent elements. Unlike other distance based measures such as Atmar and Patterson's NTC, [3] and its better described successors, BINMATNEST[12] and AININHADO[6], the

τ-Temperature is clearly mathematically defined and does not rely on Genetic Algorithms to sort the data. It also has the property that it can be evaluated just by rows, or columns which makes it appealing and it apparently behaves well against nestedness found by NODF[4]. The τ-Temperature is found by shuffling the network matrix by the degree of rows and columns as described for finding NODF_{MAX} and then 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 that are filled. Once D is found, a null model is chosen (cf. section §6) and an ensemble of matrices with similar properties to the studied matrix network are created and measured in a similar way. Each of these measurements is denoted D_{rand}. By finding the mean average D_{rand}, denoted $< D_{rand} >$ from a large ensemble of null models τ-Temperature can be calculated as:

$$\tau = \frac{D}{< D_{rand} >} \tag{2}$$

That is the nestedness of the network for which D was calculated against the expected nestedness of similar networks. As Manhattan distance is lower in more nested networks (due to being able to shuffle rows and columns to make many of the elements appear in the upper left where row and column indices are low and fewer elements appear in the lower right where row and column indices are greater), then $\tau > 1$ implies D is greater than $< D_{rand} >$ and so the network is less nested than expected for a network with the properties defined in the null model. τ can be described as a test statistic of the τ-Temperature method. We note it is also possible to calculate a p-value for nestedness from the τ-Temperature by counting how many of the null models that make up the ensemble have a Manhattan distance less than that in our original network; i.e. are more nested than the network being studied, and divide this by the number of networks making up the ensemble. Then a low p-value(p$\rightarrow$ 0) corresponds to the matrix network being significantly nested.

5.2 Quantitative

WNODF

The weighted NODF measure, WNODF[2] is found using a similar algorithm for 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, by filled positions, for a row/column to be a weighted subset the preceding row/column must have greater values. As such, WNODF is essentially a stricter version of NODF and the highest score that could be achieved in WNODF is equal to that made under binary constraints in NODF. WNODF's assumptions consist in that a more nested matrix is one where the generalists have higher abundance than the specialists. Whilst this makes sense for species-site relationships in island biogeography theory, it makes less sense for antagonistic species-species relationships. Indeed it may well be expected that a tradeoff between generalism and specialism exists such that each specialist interaction should

have a greater interaction than that of more generalist compatriots. Hence, we also look at WNODF in reverse such that the gradient is evaluated such that contributions to the WNODF-Reverse score only count if specialists have stronger interactions than that of generalists. We also investigated a mixture of WNODF row with WNODF-Reverse column scores and vice-versa. For example it may be expected that specialist phage (on generalist hosts) have high adsorption rates, whilst specialist hosts (with generalist phage) have low adsorption rates(WNODF rows, WNODF-Reverse columns). Both these mixed scores were higher than WNODF by the end of the simulation, but not as strong as that given by WNODF-Reverse. Indeed if we denote WNODF-Equal to only score when subsets have the same interaction strengths, NODF can be rewritten in terms of WNODF scores as:

$$NODF = WNODF + WNODF^{REVERSE} + WNODF^{EQUAL} \tag{3}$$

5.3 Both

Spectral Radius

The spectral radius (SR) is defined as the absolute value of the maximum real eigenvalue for a matrix and provides a link between the concept of nestedness and linear algebra and dynamical systems theory which draws back on earlier ideas from [3] who thought about defining nestedness to be a highly ordered state and were thinking in terms of ideas about entropy. [13] show in their paper that highly nested matrix structures[1, 3] tend to have a greater SR than that of other matrix permutations in binary matrices and similarly that less nested structures have smaller SR. They measure the significance of the nestedness of a particular matrix by evaluating the p-value, counting up the number of ensemble matrices built from null models that have a greater SR than the original network. They also go on to show their approach is applicable to networks containing quantitative measures and that the SR approach measures quantitative nestedness in a similar way to WNODF, such that more nested matrices are those where the generalists have greater interaction strengths than specialists.

6 Null models

This list of null models is by no means exhaustive, many other possible null models have been designed and are still to be designed.

6.1 Binary

Equiprobable-Equiprobable (E-E)

The equiprobable rows - equiprobable columns (EE) null model is one which conserves the same size structure(same number of rows and columns) as the

input matrix and the same number of filled elements (the same number of ones), but these are placed at random within the matrix. It is the least strictly defined null model we consider here and is well known to suffer from type I errors [7] (a tendancy to produce false positives).

Fixed-Fixed (F-F)

The fixed rows - fixed columns (FF) null model conserves size structure, fill and degree distribution of the original matrix. The argument is that as such a degree-distribution is found in nature (the observed input matrix), natural processes may have led to such a pattern emerging. It is the most strict null model we consider here and is well known to suffer from type II errors [7] (a failure to detect positives). Null models instantiations are found by searching the input matrix for 2x2 submatrices and comparing them to the 2x2 identity matrix and its reflection in one dimension (see below). If one of these configurations is found the corresponding rows and columns are swapped changing the submatrix into the other configuration shown. This is known as the Miklós-Podani row swap algorithm[10]. It should be noted that as these permutations do not change the Manhattan distance of the matrix, as such the Manhattan distance and associated τ-Temperature measure are invariant under this null model.

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

Relatively Fixed-Relatively Fixed (RF-RF)

The relatively fixed rows - relatively fixed columns null model conserves size structure and fill as in the EE null model, but also conserves some of the core structure found in the observed input matrix. Thus it is somewhere between the EE and the FF null models. It is found by randomly removing elements from the input matrix which by doing so will not reduce the corresponding row or column degree to zero until a minimal core structure is attained. This ensures that the size structure is conserved. Then the removed elements are randomly reassigned into the remaining empty spaces.

6.2 Quantitative

Binary Shuffle

This null model was employed by [13] 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.

Conserve Row Totals

This null model conserves binary structure and the row sum totals, but the values of the elements on each row are changed.

Conserve Column Totals

This null model conserves binary structure and the column sum totals, but the values of the elements in each column are changed.

7 Output Statistics

p-value

The p-value is the probability of a matrix composed under the rules of the null model being more nested than the input matrix. Thus a low p-value p$\rightarrow$ 0 indicates that the chance of attaining a matrix (under the null model) that is more nested than the input is low, whilst a high p-value p$\rightarrow$ 1 indicates that the chance of attaining a more nested matrix is high. We can say that matrices that have p-values less than some threshold are statistically significantly nested at the threshold value, whilst matrices that get a p-value above some threshold are statistically significantly anti-nested at that threshold. It is calculated by counting the number of null model matrices that return a score more nested than that of the input matrix and dividing this by the number of null model matrix members in the ensemble. In the cases where the p-value is found to be p= 0 we conservatively assign it the value of 1 out of the total ensemble size.

Normalised Temperature

The normalised temperature is taken from ideas of the τ-Temperature and describes the relationship between nestedness measure found for the input matrix and the expected nestedness measure found from the null model ensemble. It is described as:

$$N = \frac{Measure}{< Measure >} \tag{4}$$

In measures where increasing nestedness is found with increasing measure score e.g. NODF, if the normalised temperature is found to be greater than 1 it indicates that the input matrix is more nested than expected, whilst if the normalised temperature is found to be less than 1 the input matrix is less nested than expected. Conversely in measures where increasing nestedness is found with decreasing measure score e.g. τ-Temperature, the opposite is true such that a normalised temperature (the τ-Temperature in this particular case) greater than 1 indicates the input matrix is less nested than expected, whilst a normalised temperature less than 1 indicates the input matrix is more nested than expected.

Adjusted Normalised Temperature

In order to compare this measure across different nestedness measurements we choose to create the adjusted normalised temperature, which is equivalent to the

normalised temperature if nestedness increases with increasing measure score, but otherwise is:

$$N' = \frac{1}{N} \tag{5}$$

Thus the adjusted normalised temperature score always describes the tendancy of nestedness of the input matrix measure in the same direction, such that N' scores greater than 1 indicate that the matrix is more nested than would be expected under the rules of the null model, whilst a score less than 1 indicates it is less nested than would be expected.

Mean

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

Standard Deviation

The standard deviation (σ) 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:

$$z = \frac{Measure - < Measure >}{\sigma} \tag{6}$$

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.

References

[1] 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.

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

[3] 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.

[4] G. Corso and N.F. Britton. Nestedness and τ-temperature in ecological networks. *Ecological Complexity*, 2012.

[5] 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.

[6] P.R. Guimarães and P. Guimarães. Improving the analyses of nestedness for large sets of matrices. *Environmental Modelling & Software*, 21(10):1512–1513, 2006.

[7] 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.

[8] R.H. MacArthur and E.O. Wilson. *The theory of island biogeography*, volume 1. Princeton University Press, 1967.

[9] C.F. McQuaid and N.F. Britton. Host-parasite nestedness: A result of co-evolving trait-values. *Ecological Complexity*, 13:53 – 59, 2013.

[10] I. Miklós and J. Podani. Randomization of presence-absence matrices: comments and new algorithms. *Ecology*, 85(1):86–92, 2004.

[11] J. Podani and D. Schmera. A comparative evaluation of pairwise nestedness measures. *Ecography*, 2012.

[12] 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.

[13] P.P.A. Staniczenko, J.C. Kopp, and S. Allesina. The ghost of nestedness in ecological networks. *Nature Communications*, 4:1391, 2013.

[14] W. Ulrich and M. Almeida-Neto. On the meanings of nestedness: back to the basics. *Ecography*, 2012.