Nestedness – a FORTRAN program
for measuring order and disorder in ecological
communities
Version 2.0

Werner Ulrich

Nicolaus Copernicus University in Toruń

Department of Animal Ecology

Gagarina 9, 87-100 Toruń; Poland

e-mail: ulrichw @ uni.torun.pl

Latest update: 12.06.2008

1. Introduction

A nested structure is a basic feature of many communities spread over a series of islands or mainland sites (Atmar and Patterson 1993). Nestedness measures whether the faunal composition of different study sites form perfect subsamples from a larger common species pool (a nested pattern) or whether local peculiarities occur. These peculiarities might affect several ecological processes like extinction probabilities, species turnover, competition, or reintroduction success in bioconservation (Patterson and Atmar 2000).

Different mechanisms can cause a nested pattern of species occurrences. Most common seems to be a passive sampling (Andreń 1994, Cutler 1994). Sites of different species numbers form subsamples of the species pool. This situation should predominate for communities of highly dispersive species and for sites within a homogeneous area. Next, habitat isolation might create nested subsets through dispersal limitation and selected immigration or extinction (Patterson and Atmar 1986, Lomolino 1996, McAbendroth et al. 2005). Further, nested habitat types, disturbance regimes, and patterns of hierarchical niche overlap are supposed to create nested distributions (Kolasa 1996, Wright et al. 1998, Honnay et al. 1999, Atmar and Patterson 2000).

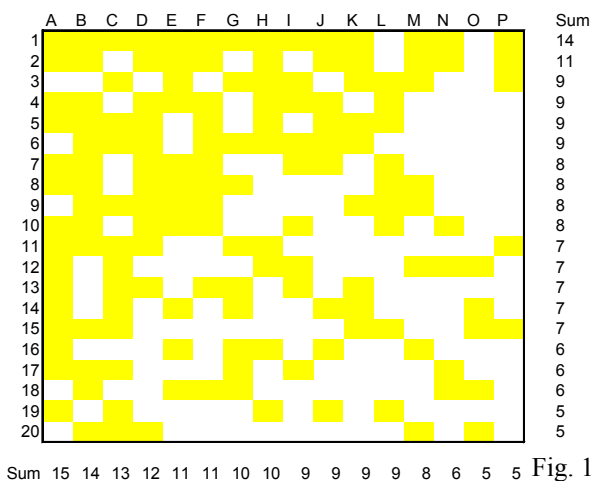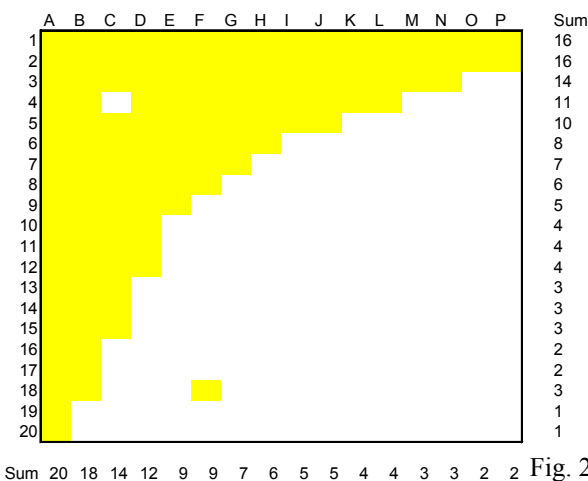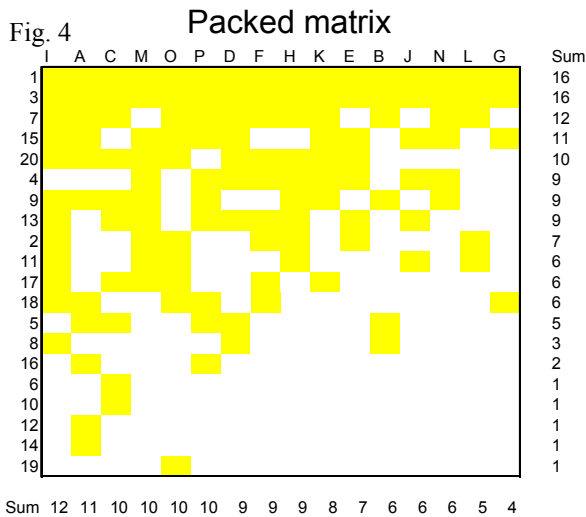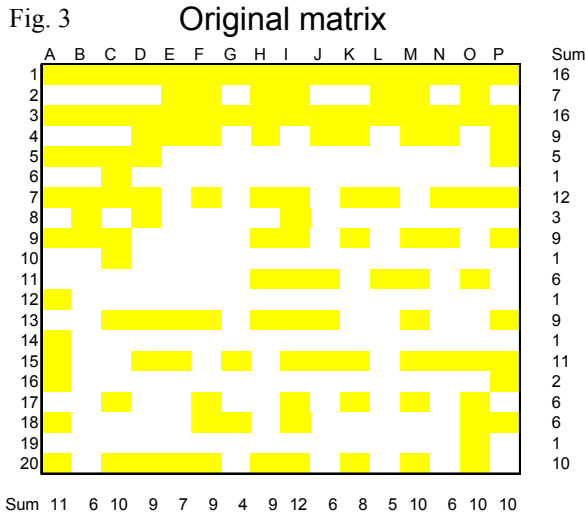The basic pattern is shown in Figs. 1 and 2.

Fig. 1

Fig. 2

Fig. 3   Original matrix

Fig. 4   Packed matrix

occurs (N0). N1 is the compliment of N0 and counts presences of a species at sites species poorer than the richest site it occurs (Cutler 1991). NC counts the number of pairwise species co-occurrences summed over all combinations of sites (Wright and Reeves 1992). UA, UP, and UT count unexpected absences (UA) from more species rich sites and presences (UP) from more depauperate sites (Cutler 1991). UT (the sum of UA and UP) is the total number of deviations from perfect nestedness (Wright et al. 1998). Brualdi and Sanderson (1999) proposed to use the number of discrepancies (absences or presence) that must be filled and erased to produce a perfectly nested matrix (BR). BR is therefore similar to the concept of UT.

The above measures treat all cells of the matrix equally. Atmar and Patterson (1993, 1995) introduced a concept that considers also the placement of a cell in a give matrix. They used an information approach to define the degree of ordering in a presence absence matrix and defined a matrix temperature T in the range between 0° (perfect order) and 100° (perfect disorder) in a three step process. The first step involves the rearrangment (the packing) of the matrix in such a way that the packed matrix has a minimum temperature. This will be in most cases (but by far not in all cases) a matrix sorted according to row (species) and column (sites) totals (Fig. 3, 4). T is now computed in the following way. A point [X,Y] is chosen that divides the matrix diagonal according to matrix fill (= number

Two presence absence matrices of 20 species (1 to 20) and 16 sites (A to P) have been sorted according to species (rows) and site (columns) totals so that species occurrences (the yellow squares) are concentrated in the upper left corner. In Fig. 1 this sorting is far from being perfect. In Fig. 2 a sharp boundary between occupied and unoccupied sites appears. There is only one unexpected occurrence and one unexpected absence. This latter matrix appears to be highly ordered. The basic idea behind the measuring of nestedness is therefore to assess its state of ordering.

Several measures have been proposed to measure the degree of nestedness (Ulrich and Gotelli 2007a, b). Patterson and Atmar (1986) proposed to count how often a species is absent from a site more species richness than the most impoverished site it

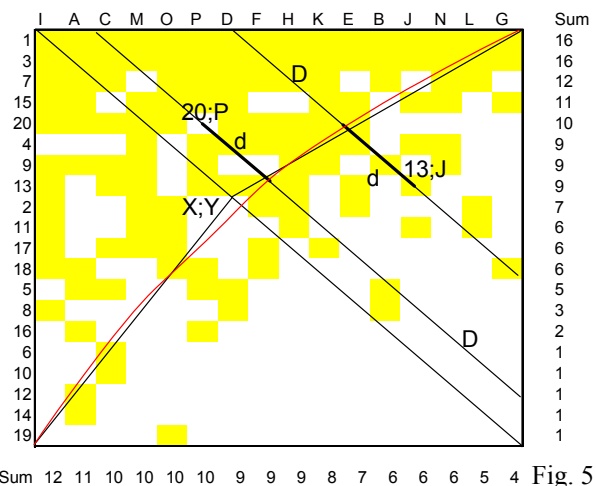
Fig. 5

of occupied cells divided through the total number of cells) . Hence X = fill *Sp and Y = fill * Si, with Sp and Si being the numbers of species and sites, respectively. Lines are drawn from the matrix edges to this point (Fig. 5). Below these are called the border lines. Atmar and Patterson then constructed an isocline (red in Fig. 5) that divides these lines and forms areas above and below them in such a way that the areas above equal the areas below. The isocline divides therefore the whole matrix in two parts with the upper part equaling the matrix fill (fill*Sp*Si) and the lower part the unoccupied area ([1-fill]*Sp*Si).

The distance d_{ij} (Fig. 5) of an empty cell [i;j] (unexpected absence) above the isocline or an occupied cell [i;j](unexpected presence) below the isocline is now a measure of disorder. Because these distances rely on matrix size and matrix form, the relative measure $(d_{ij} / D_{ij})^2$ is used, where D_{ij} is the total length of the diagonal through the focal cell [i;j]. T is now defined as

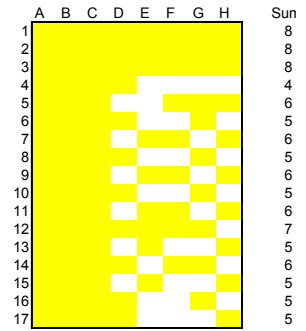
Fig. 6

note: 2 duplicate 'all-black' row(s) and 2 column(s) were removed from the top and the left of the matrix, thus all column totals will be 2 less than the original data, and all row totals will similarly be 2 less than the original data.

$$T = \frac{100}{0.04145} \frac{1}{SpSi} \left(\sum_{i=1}^{Sp} \sum_{j=1}^{Si} \left(\frac{d_{ij}}{D_{ij}} \right)^2 \right)$$

It is also possible to define so-called idiosyncratic temperatures for each species (T_{Sp}) and each site (T_{Si}) that measure the respective numbers of unexpected occurrences and absences.

$$T_{Sp}(i) = \frac{100}{0.04145} \frac{1}{Si} \left(\sum_{j=1}^{Si} \left(\frac{d_{ij}}{D_{ij}} \right)^2 \right)$$

and

$$T_{Si}(j) = \frac{100}{0.04145} \frac{1}{Sp} \left(\sum_{i=1}^{Sp} \left(\frac{d_{ij}}{D_{ij}} \right)^2 \right)$$

The value of $U_{max} = 0.04145$ is approximately the maximum possible disorder of any given matrix.

To assess, whether T is significant and deviates from a temperature that would occur in a given matrix just by chance temperatures have to be compared with those obtained from an appropriate null model. Atmar and Patterson used the Random00 model where occurrences are randomly assigned to matrix cells irrespective of the observed total numbers per site and species (equiprobable row and column constraints). The *nestedness temperature calculator* (NTC) designed by Atmar and Patterson (1995) to perform all these calculations is now a standard tool in the analysis of presence absence matrices.

Unfortunately, the concept used by Atmar and Patterson and the calculator have some well known disadvantages (Fischer and Lindenmayer 2002, 2005, Rodriguéz-Gironés and Santamaría 2006).

First the use of the Random00 model makes the calculator susceptible to detect nestedness as an artefact of passive sampling (Fischer and Lindenmayer 2002). Hence, NTC does not take into account that some species are more ubiquitous than others.

Second, complete rows or columns are eliminated by the calculator because the isocline has to be constructed from the edges. This leads to the following undesired situation. The matrix of Fig. 6 contains 17
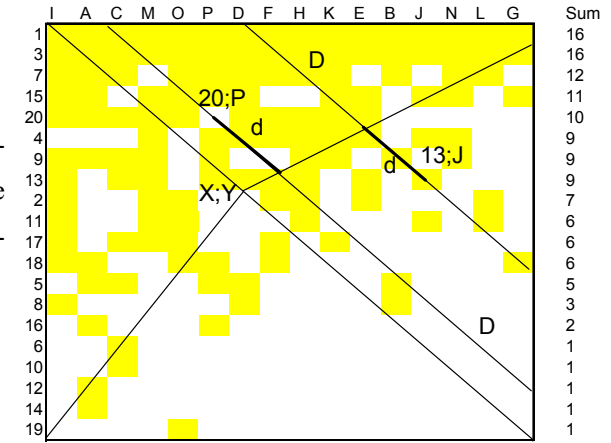
Fig. 7

species in 8 sites. Three species occur at all sites and three sites are occupied by all species. To construct the isocline that divides the occupied from the unoccupied area, two sites and two species have to be eliminated by the calculator. From the rest the program infers no significant nestedness although the original matrix appears to be highly ordered.

Third, the Random00 model tends to homogenize occurrences and points too often to a nested pattern. Further, due to the well known portefolio effect its variance is negatively correlated with matrix size. Hence, the model overestimates standardized effects sizes ($Z = [x-\mu]/\sigma$) of larger matrices.

Fourth, NTC behaves poor in cases of abundant checkerboard occurrences (Rodriguéz-Gironés and Santamaría 2006). Checkerboards form two sites were one species occurs at one site and the other not and vice versa.

Lastly the calculator does not allow for a batch run, necessary to test different null models. This feature makes it very time consuming to study multiple models. The present program tries to overcome these shortcomings by adding several features to the calculation of matrix temperatures.

Recently, Rodriguéz-Gironés and Santamaría (2006) published a new nestedness calculator (BINMATNEST) that overcomes some of these and other shortcomings of NTC. They implemented a new equation for the formerly weakly defined isocline and
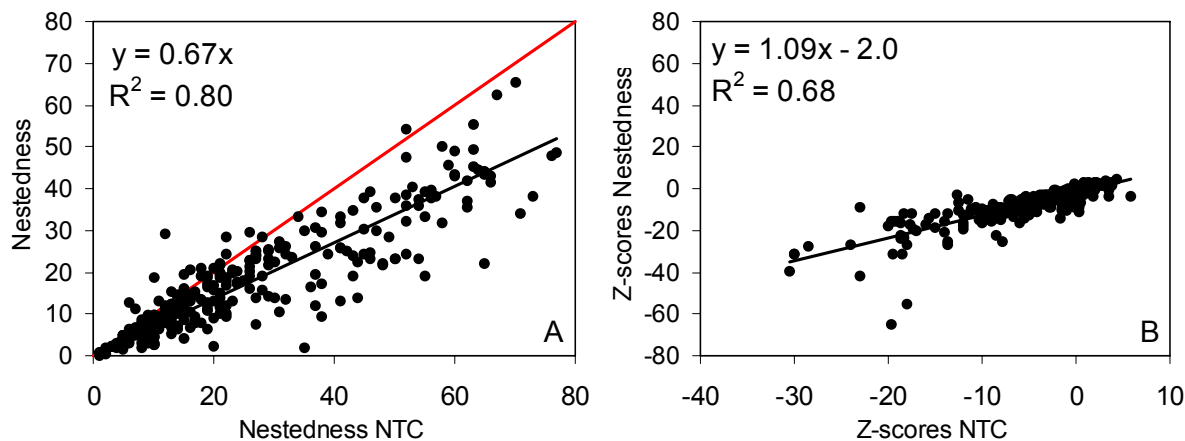
Fig.8: A: Comparison of matrix temperatures computed by NTC and by *Nestedness*. B: The same comparison using the respective Z-scores. In both cases Random00 was used. The results of this and of the following Figs are based on ten randomizations of the Random00 null model (due to runtime). Data sets are the 294 presence absence matrices provided together with the Atmar Patterson calculator.

implemented additional null models.

Recently, Almeida-Neto et al. (2008) introduced a new metric *NODF* (**n**estedness based on **o**verlap and **d**ecreasing **f**ill) that enables to differentiate between a part of overall nestedness introduced by species differences (*NODF*rows) and site differences (*NODF*sites). NODF is calculated as follows:

$$NODF = \frac{\sum NP_{\text{paired}}}{\left[\dfrac{n(n-1)}{2}\right] + \left[\dfrac{m(m-1)}{2}\right]}$$

where n is the number of columns, n the number of sites and NPpaired is the paired nestdedness degree in the matrix ordered by row and column totals. The lat-

ter is a standardized count how many joint pairs of species exist at those combinations of sites and how many joint pairs of sites exist for those combinations of species where the marginal total of site/species 2 is less than the marginal total of site/species 1 (for details cf. Almeida-Neto et al. 2008).

3. Computing the matrix temperature

Nestedness uses the same principle for defining the matrix temperature as Atmar and Patterson (1993). T, T_{Sp} and T_{Si} are computed with the above equations. In contrast to NTC and to BINMATNEST, however, *Nestedness* does not use the isocline that divides the

Fig. 9: A: Differences in the Z-scores of nestedness and NTC (Δ-Z-score) in dependence on matrix size (A) and matrix fill (B).

filled from the empty matrix area. The distances d_{ij} are based on the border lines from the point [X,Y] to the matrix edges (Fig. 7). In contrast to NTC *Nestednes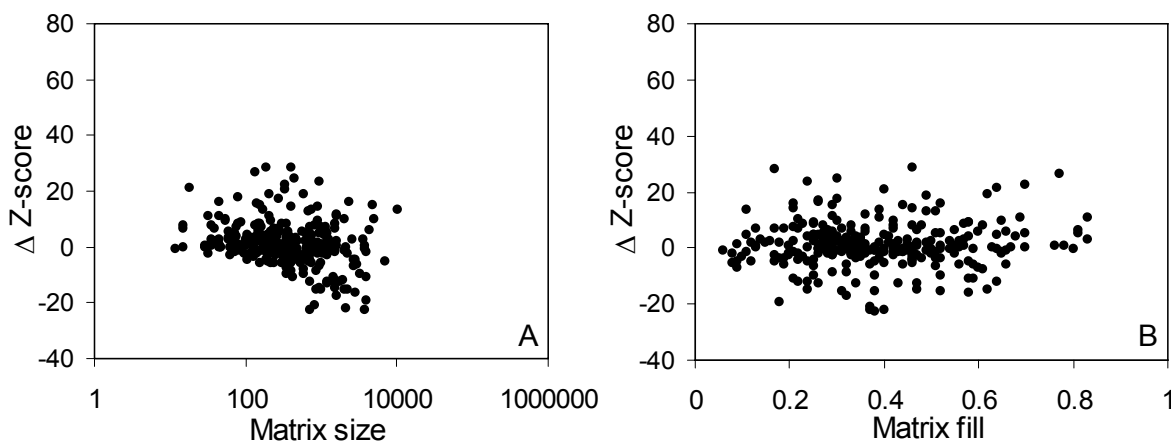s* does not eliminate full row and columns from the matrix in order to construct these lines. The program uses the last full row/column (if any exists) as end point of the border lines as shown in Fig. 7. Three species occupy all sites. The border line is drawn from the third full row.

A further difference to NTC and to BINMAT-NEST is that the user can choose whether cells adjacent to the border lines are considered for the calculation of T or not. The reason for this is that in matrices with very low or very high fill or in matrices with striking different numbers of rows and columns, the small values of d_{ij} might sum to significant temperatures although the matrix appears to be highly ordered. Further, if (as current theory assumes) the degree of disorder is correlated with extinction or introduction probability, cells near the borderline should be places with high species turnover. Hence presence or absence in these cells should be more affected by chance processes associated with point estimates. The effect on temperature might be particularly severe at the edges of the matrix were the values of D_{ij} are small.

Nevertheless the temperatures computed by *Nestedness* are similar to those of NTC (Fig. 8). The NTC temperatures are nearly always lower than the *Nestedness* temperatures (Fig. 8A). However, when comparing the respective Z-scores (using Random00) both program give very similar results (Fig. 8B).

Table 1 shows that both programs identified order in a similar way. They differed in about 11% of cases while NTC being slightly less conservative (199 significances than *Nestedness* (186 significances).

Fig. 10: Standard deviations (StdDev) of mean temperature of the Random00 null model of NTC (red) and *Nestedness* (black) in dependence on matrix size.

Tab. 1: A comparison of Z-scores of *Nestedness* and NTC. Given are numbers of significant temperatures ($p < 0.05$) indicating order or disorder of 291 of the Atmar Patterson data sets

| | | Nestedness | | |
	Order	Disorder	Insignificant	Sum
NTC Order	165	5	16	186
Disorder	1	4	17	22
Insignificant	34	4	45	83
Sum	200	13	78	291

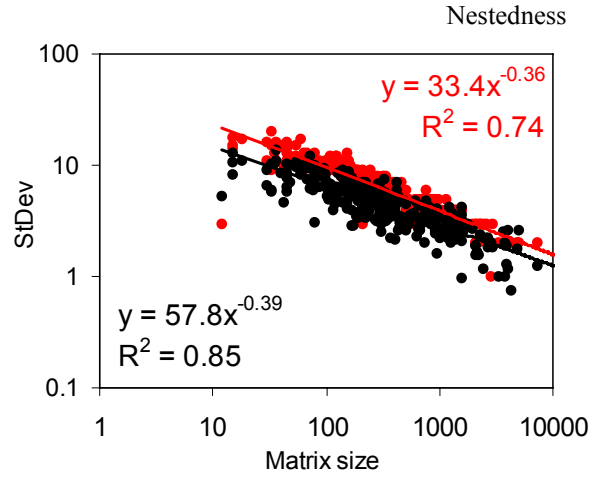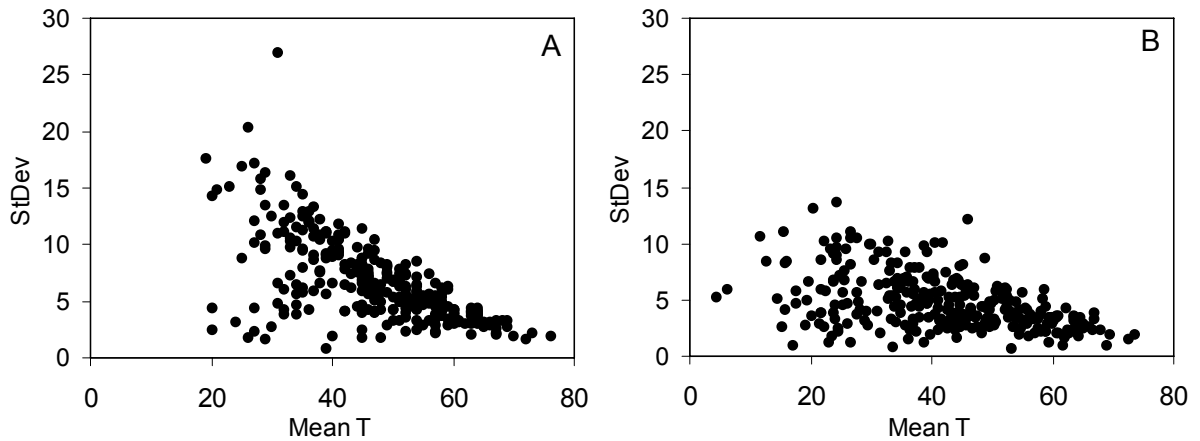

Fig. 11: Standard deviations (StdDev) of mean temperature of the Random00 null model of NTC (A) and *Nestedness* (B) in dependence on matrix temperature T.
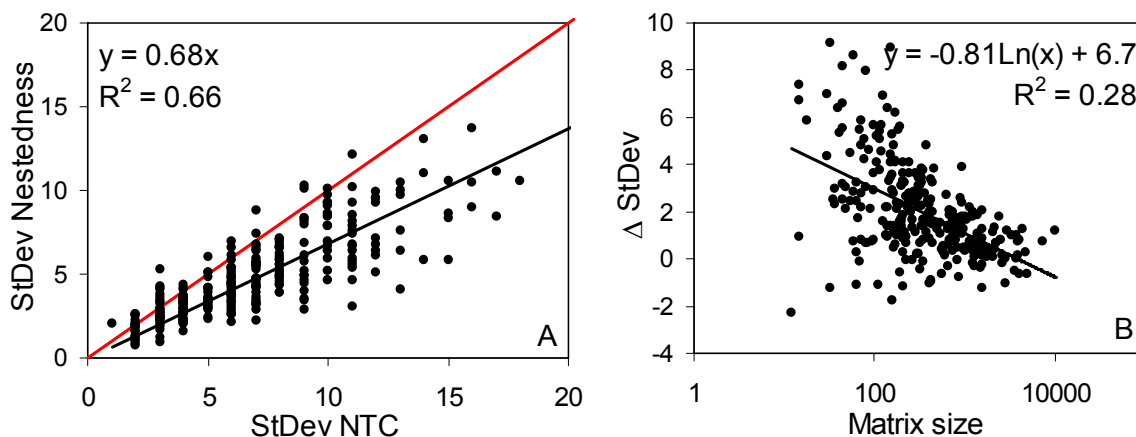
Fig. 12: A: Comparison of the standard deviations (StdDev) of mean temperature of the Random00 null model of NTC and *Nestedness*. B: The Difference Δ StdDev (= StdDev$_{NTC}$ - StDev$_{Nestedness}$) in dependence on the matrix size. The red line gives again the isometric relationship.

```
S 1 2 3 4 5 6 7 8 9 10 11 12
a 0.00 0.00 1.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 1.00 0.00
b 1.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 1.00 1.00 1.00
c 0.00 0.00 1.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00
d 1.00 0.00 1.00 0.00 0.00 0.00 0.00 1.00 1.00 0.00 0.00 0.00
e 1.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 1.00 0.00
f 0.00 1.00 0.00 0.00 0.00 0.00 0.00 1.00 1.00 0.00 1.00 1.00
g 1.00 1.00 1.00 0.00 1.00 0.00 1.00 1.00 0.00 1.00 1.00 0.00
h 0.00 0.00 1.00 1.00 1.00 1.00 0.00 1.00 0.00 0.00 0.00 0.00
i 1.00 1.00 0.00 1.00 0.00 0.00 0.00 1.00 1.00 0.00 0.00 0.00
j 0.00 0.00 0.00 0.00 0.00 1.00 1.00 1.00 0.00 1.00 1.00 0.00
k 0.00 0.00 1.00 1.00 1.00 1.00 1.00 1.00 0.00 0.00 0.00 0.00
l 1.00 1.00 0.00 1.00 1.00 1.00 0.00 0.00 1.00 0.00 1.00 1.00
m 0.00 1.00 1.00 1.00 1.00 1.00 0.00 1.00 1.00 0.00 0.00 0.00
n 0.00 1.00 1.00 0.00 0.00 1.00 1.00 1.00 0.00 1.00 0.00 0.00
o 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 0.00 0.00 1.00 0.00
p 1.00 1.00 1.00 1.00 0.00 1.00 1.00 1.00 0.00 0.00 1.00 0.00
q 0.00 0.00 1.00 1.00 0.00 0.00 1.00 1.00 0.00 0.00 1.00 0.00
r 1.00 1.00 1.00 0.00 1.00 0.00 1.00 0.00 0.00 0.00 0.00 0.00
s 0.00 1.00 0.00 1.00 1.00 1.00 1.00 1.00 1.00 0.00 0.00 0.00
t 0.00 1.00 0.00 1.00 0.00 1.00 1.00 1.00 1.00 0.00 1.00 0.00
```

NTC identified 21 matrices as being disordered while *Nestedness* only 13. In total *Nestedness* and the Atmar Patterson calculator disagreed in 77 out of 291 cases (26%).

Of course, the means of Random00 are for both calculators correlated with matrix size ($R^2 > 0.3$). The feature makes the model less appropriate for the study of nestedness. However, the difference in T and in the detection probability of order and disorder between *Nestedness* and NTC are only weakly correlated with matrix size (Fig. 9A) and matrix fill (Fig. 9B). It seems that the standard deviation of the simulated mean temperatures of the Random00 null model affects the results of NTC more than of *Nestedness*. Surprisingly, the standard deviations computed by NTC

are in the mean about 75% higher than those of *Nestedness* (Fig. 10, 11). This difference is most pronounced at lower temperature (Figs. 11, 12). A multiple regression with mean T and log transformed matrix size did also point to mean T as influencing the standard deviation of the NTC means ($p < 0.01$). The dependence vanished in the case of *Nestedness* ($p > 0.5$). Hence, NTC is biased towards a rejection of a nested pattern at smaller matrix temperatures of the Random00 model.

3. Data structure

Nestedness needs one main plain text data file of the following structure. The columns of the matrix are sites, the rows species. Hence the matrix above contains 20 species distributed over 12 sites. The data file has to be a simple ASCII file with data delimitated by one or more spaces. Accepted are either abundance or presences absence data of the integer (In) or real format (fn.k) The first row contains site names, the first column species names. The file has therefore the same format that is needed for *EcoSim* (Gotelli and Entsminger 2006). The number of species is not limited, the maximum number of sites is about 150.

Alternatively, you may run *Nestedness* in a

```
* Comment
Test1.txt
Test2.txt
Test3.txt
Test4.txt
Test5.txt
```

```
C:\Documents and Settings\Właściciel\Moje dokumenty\Models\Nestedness.exe    _ □ ✕

*******************************************************
*                                                     *
* Program Nestedness: Version 2; 11.06.2008           *
*                                                     *
* Copyright Dr. Werner Ulrich                         *
*                                                     *
* The author does not take responsibility for correct *
* program run or any damages caused by the program.   *
*                                                     *
*******************************************************

Name of input file with extension. File has to have EcoSim format.
If batch run leave blank.
test.txt
Name of output file (with extension), default = Output.txt

Name of matrix file (with extension), default = Matrix.txt

Name of cell temperatures file (with extension), default = Unexp.txt

Null model for randomization:
Fixed row and column constraints sequential swap (s),
no constraints (equiprobable row and columns)(e),
fixed row and equiprobable column constraints (f),
fixed column and equiprobable row constraints (c),
random sampling according to observed frequencies of occurrence (o),
random sampling according to observed species abundances (a),
or random sampling from a lognormal distribution (l)
a
Number of iterations for computing standard deviations of the null model
Default=100

Minimum distance to borderline. Cells nearer to the borderline will be
exlcuded from the computation of the matrix temperature (default = 0.5)

How to pack the matrix prior to analysis?
According to richness (s; default), according to matrix temperature (t),
or leaving unsorted (n)
n
test.txt         Species:   15   Sites:    8 Temperature:    92.08
Runtime of program:   0hh  0min  8sec
Fortran Pause - Enter command<CR> or <CR> to continue.
```

batch. In this case the input files must have a structure as shown above and the driver file contains only the names of the data files. The first line has to be a comment line.

4. Program run

First, the program asks for the files names. The default output file names are *Output.txt, Unexp.txt*, and *Matrix.txt*. You get the default values after returning *enter*. If you don't give the name of the data file and return *enter* the program expects a batch run and a file name with the data files.

Next, the program asks for the model for randomization. You have seven possibilities: fixed row and column constraints (input: s) using the sequential swap algorithm (Gotelli 2000, 2001), no constraints (equiprobable row and columns, input: e), or fixed row (input: f) or fixed column (input: c) constraints only.

For details of the null models used see Gotelli (2000, 2001). The sequential swap model uses 25000 single swaps to generate a randomized matrix.

The fifth null model (o) assigns species with a probability according to the number of site occurrences. This model is therefore identical to the Random 1 model of Patterson and Atmar (1986) and Wright et al. (1998).

The sixth null model is a sampling model, where the sites are filled with species using a random sampling of individuals from a common species pool that is structured according to a lognormal species abundance distribution. In this case the program asks for the shape generating parameter *a* of the lognormal model. This has the typical form $[S=S_0 Exp(-a(R-R_0)^2]$ and is computed using a normally distributed random number on a log scale. Preston's canonical lognormal has the parameter value a = 0.2 (May 1975). In the

Unexp.txt - Notatnik

Plik   Edycja   Format   Widok   Pomoc

```
test1.txt          Species:   20    Sites:     8
 RowF   ColF          X        Y       a1       a2      RowI      ColI
    0      0         9.88     3.95    -2.56     2.44    20.00      8.00

Packed matrix with u-values of unexpectedness:  Temperature:    26.54

               1       3       4       2       5       6       8       7

a         0.0000  0.0000  0.0000  0.0000  0.0000  0.0000  0.0000  0.0000       0.0000
q         0.0000  0.0000  0.1563  0.0000  0.0000  0.0901  0.0433  0.0000       0.2897
c         0.0000  0.0000  0.0000  0.0000  0.0737  0.0433  0.0000  0.0299       0.1469
j         0.0000  0.0000  0.0000  0.0658  0.0433  0.0182  0.0000  0.0591       0.1863
e         0.0000  0.0000  0.0000  0.0000  0.0000  0.0056  0.0000  0.0000       0.0056
g         0.0000  0.0000  0.0000  0.0000  0.0116  0.0000  0.0000  0.0000       0.0116
o         0.0000  0.0000  0.0314  0.0163  0.0000  0.0000  0.0197  0.0000       0.0674
i         0.0000  0.0000  0.0000  0.0000  0.0000  0.0027  0.0000  0.0000       0.0027
d         0.0000  0.0000  0.0000  0.0040  0.0000  0.0068  0.0000  0.0000       0.0108
m         0.0000  0.0000  0.0145  0.0015  0.0000  0.0000  0.0000  0.1521       0.1682
t         0.0000  0.0444  0.0094  0.0000  0.0027  0.0177  0.0000  0.0000       0.0742
l         0.0000  0.0352  0.0049  0.0000  0.0000  0.0238  0.0678  0.0000       0.1316
b         0.0000  0.0257  0.0000  0.0000  0.0113  0.0000  0.0759  0.0000       0.1129
p         0.0000  0.0000  0.0000  0.0064  0.0000  0.0000  0.0000  0.0000       0.0064
k         0.0000  0.0000  0.0000  0.0145  0.0000  0.0000  0.0000  0.0000       0.0145
f         0.0000  0.0000  0.0077  0.0000  0.0000  0.0000  0.0970  0.0000       0.1047
h         0.0000  0.0000  0.0217  0.0000  0.0000  0.0000  0.0000  0.0000       0.0217
n         0.0000  0.0113  0.0000  0.0807  0.0000  0.0000  0.0000  0.0000       0.0920
r         0.0000  0.0000  0.0978  0.0000  0.0000  0.0000  0.0000  0.0000       0.0978
s         0.0000  0.0000  0.0000  0.0000  0.2150  0.0000  0.0000  0.0000       0.2150

          0.0000  0.1166  0.3438  0.1729  0.2290  0.0354  0.1644  0.1548
```

Lin 1, kol 1

Output.txt - Notatnik

Plik   Edycja   Format   Widok   Pomoc

```
Null model: a   Packing type: n   Iterations:  100 Minimum distance:   0.500

x   File            NO      SimNO   StdDevNO    Z-NO     Z-Ind     SkewNO    L95%CNO    U95%CNO
NO test.txt          6       0.60      2.03     2.65     10.00      4.79       0.00       6.00
x   File            N1      SimN1   StdDevN1    Z-N1     N1-Ind    SkewN1     L95%CN1    U95%CN1
N1 test.txt         16       0.74      2.03     7.51     28.26      3.87       0.00       7.00
x   File            NC      SimNC   StdDevNC    Z-NC     NC-Ind    SkewNC     L95%CNC    U95%CNC
NC test.txt         84      63.36     12.22     1.69     38.22      0.88      49.00      92.00
x   File            UA      SimUA   StdDevUA    Z-UA     UA-Ind    SkewUA     L95%CUA    U95%CUA
UA test.txt       0.00       0.02      0.14    -0.14     -0.04      7.07       0.00       0.00
x   File            UP      SimUP   StdDevUP    Z-UP     UP-Ind    SkewUP     L95%CUP    U95%CUP
UP test.txt       0.00       0.06      0.28    -0.22     -0.11      5.18       0.00       1.00
x   File            UT      SimUT   StdDevUT    Z-UT     UT-Ind    SkewUT     L95%CUT    U95%CUT
UT test.txt       0.00       0.08      0.37    -0.22     -0.15      4.80       0.00       2.00
x   File            BR      SimBR   StdDevBR    Z-BR     BR-Ind    SkewBR     L95%CBR    U95%CBR
BR test.txt         23       3.29      4.58     4.30     36.50      2.34       0.00      18.00
x   File          NODF   SimNODF  StDevNODF   Z-NODF   NODF-Ind  SkewNODF L95%CNODF U95%CNODF
NF test.txt      27.11       5.81      5.49     3.88      0.23      0.81       0.00      18.97
x   File        NODFcol SimNODFc StDeNODFc   Z-NODFc NODF-Indc SkewNODFc L95%NODFc U95%NODFc
F1 test.txt      19.54      17.88     15.80     0.11      0.02      0.44       0.00      50.00
x   File        NODFrow SimNODFr StDeNODFr   Z-NODFr NODF-Indr SkewNODFr L95%NODFr U95%NODFr
F2 test.txt      29.13       2.59      4.54     5.85      0.27      2.87       0.00      17.59

x   File         Species    Sites Mat Fill  MatTemp  SimTemp    StdDev  Z-score Skewness L95%Conf U95%Conf
MT test.txt          15        8    0.45    92.08    79.13      6.69     1.94     0.54    70.57    91.06

Species    TotalOcc    Temp UnAbsOcc    Site      TotalOcc    Temp UnAbsOcc
1              5     100.00      3      1             7     100.00      7
2              3      74.58      4      2             6      31.57      6
3              5      55.07      4      3             8      61.25      5
4              6      45.31      3      4             6      28.36      4
5              6      54.39      3      5             7      79.61      6
6              4     100.00      4      6             5      85.84      6
7              2      49.31      2      7             6      75.44      4
8              3     100.00      5      8             9      97.44      6
9              2       6.60      1
10             3      98.23      4
11             2      53.84      3
12             3     100.00      4
13             2      52.96      3
14             4     100.00      4
15             4     100.00      3
```

Lin 1, kol 1

case of the lognormal null model column (site) species numbers are fixed to the observed values (fixed column constraint).

The seventh null model resamples rows according to the observed species abundance distribution calculated from row totals of abundance. This last null model, of course, needs abundance data as input.

Then the program asks for the number of randomizations to compute the null model means and standard deviations, as well as upper and lower 95% confidence limits. In most cases 100 such randomizations will be enough. Then you have to give the minimum distance of a cell to the border line to be included in the computation of T. The default is 0.5. That means cells adjacent to the line will be excluded.

```
Matrix.txt - Notatnik
Plik  Edycja  Format  Widok  Pomoc
Null model: s     Packing type: t     Iterations:     3 Minimum distance:   0.500
Packed matrix: test1.txt
            1   3   4   2   5   6   8   7
a           1   1   1   1   1   1   1   1       8
q           1   1   0   1   1   0   0   1       5
c           1   1   1   1   0   0   0   1       5
j           1   1   1   0   0   0   1   1       5
e           1   1   1   1   1   0   0   0       5
g           1   1   1   1   0   1   0   0       5
o           1   1   0   0   1   1   1   0       5
i           1   1   1   1   0   1   0   0       5
d           1   1   1   0   1   1   0   0       5
m           1   1   0   0   1   0   0   1       4
t           1   0   0   0   1   1   0   0       3
l           1   0   0   0   0   1   1   0       3
b           1   0   0   0   1   0   1   0       3
p           1   0   0   1   0   0   1   0       3
k           1   1   0   1   0   0   0   0       3
f           1   0   1   0   0   0   1   0       3
h           1   1   1   0   0   0   0   0       3
n           0   1   0   1   0   0   0   0       2
r           1   0   1   0   0   0   0   0       2
s           1   0   0   0   1   0   0   0       2

           19  14  10   9   9   7   6   5
Last packed randomized matrix: test1.txt       No. of Swaps:    5000
            1   3   4   2   5   7   6   8
a           1   1   1   1   1   1   1   1       8
m           1   0   0   0   1   0   0   1       5
d           1   1   1   1   0   0   1   0       5
i           1   0   1   1   1   1   0   0       5
g           1   1   1   1   1   0   0   0       5
s           1   0   0   1   0   0   0   1       5
k           1   1   1   1   0   1   0   0       5
p           1   0   1   1   0   0   0   0       5
f           1   1   1   1   0   1   0   0       5
b           1   0   1   1   0   0   1   1       4
e           1   1   1   0   0   0   0   0       3
j           1   0   0   0   0   1   0   1       3
t           1   0   1   0   1   0   1   0       3
q           0   1   1   0   0   0   0   0       3
o           1   0   1   0   1   1   0   1       3
c           1   1   1   0   0   0   0   0       3
n           1   1   1   0   1   0   1   0       3
r           1   0   0   1   0   0   0   0       2
l           1   0   0   0   1   0   0   0       2
h           1   0   0   0   1   1   0   0       2

           19  14  10   9   9   6   7   5

Lin 1, kol 1
```

The next step involves the sorting algorithm. Theoretically, the matrix has to be packed as to minimize unexpected occurrences or absences. Hence the matrix temperature has to become minimal. Nestedness solves this problem in a two step procedure. It first sorts according to row and column totals and then sorts rows and columns again according to temperature. This procedure gives in the vast majority of cases a good approximation to the desired state. However, while not sorting rows and columns simultaneously, it
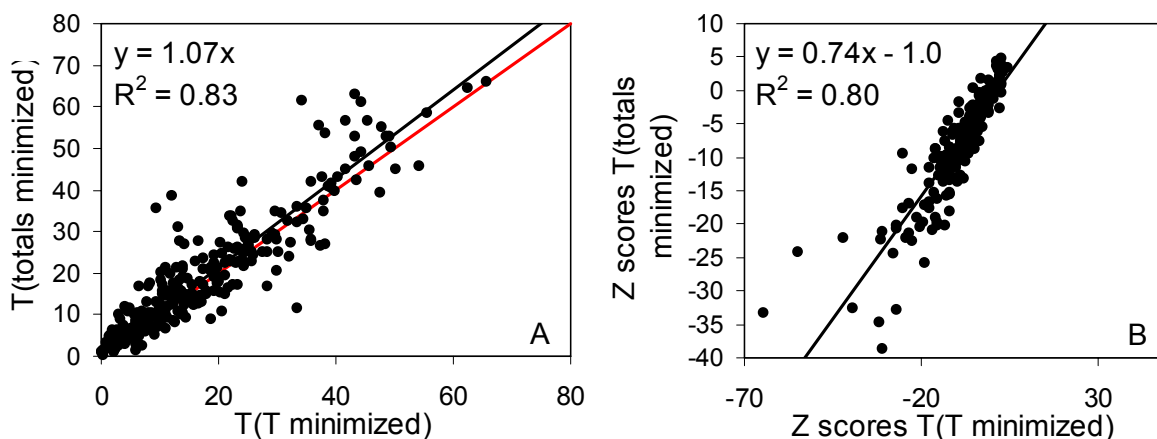

Fig. 13: A: The relation between matrix temperatures after packing according to matrix temperature and according to row and column totals. Both values are very similar although some outliners occur. B: The same plot using the respective Z-scores.
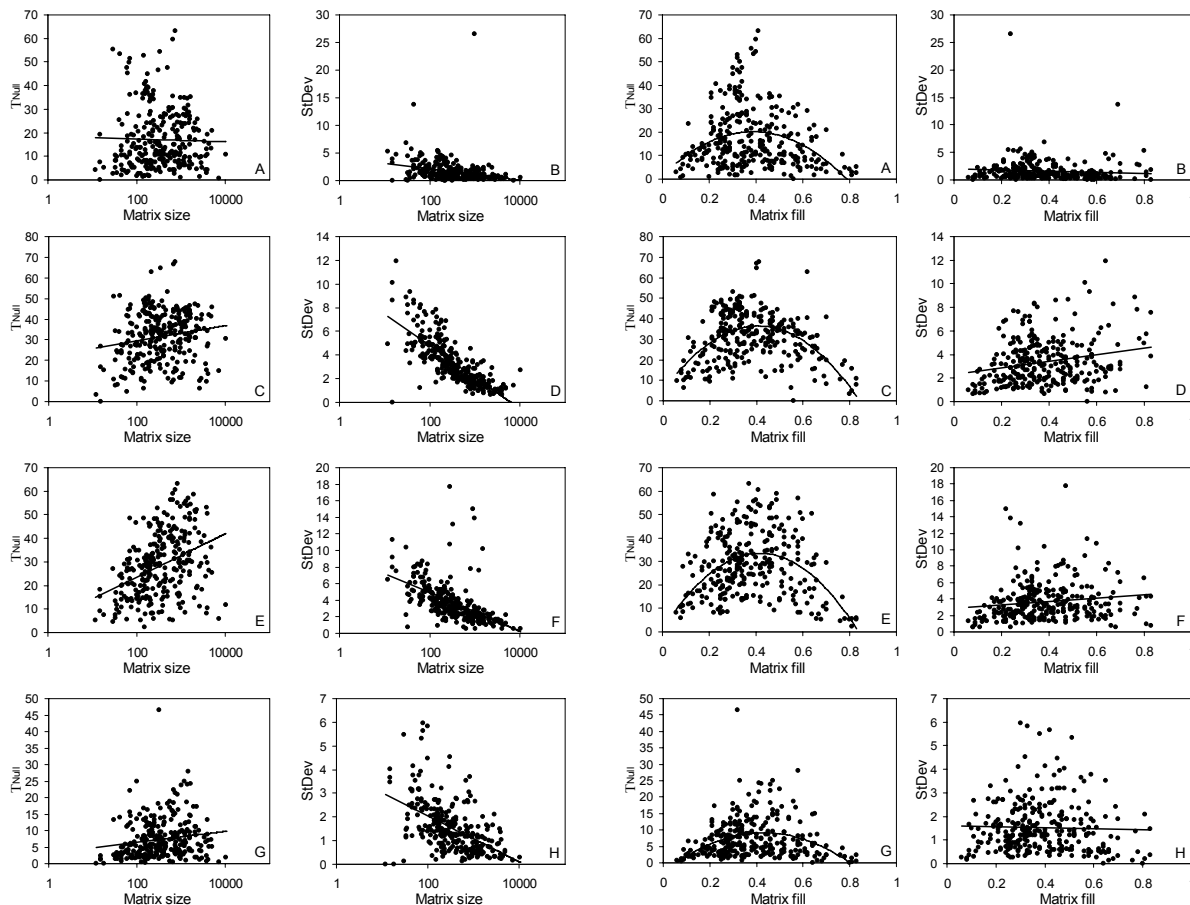
Fig. 14: Dependence of mean simulated matrix temperature and its standard deviation on matrix size for four null models implemented in *Nestedness*. A, B: sequential swap; C, D: fixed row equiprobable column; E, F: equiprobable row, fixed column; G, H: sampling from a lognormal (a = 0.2).

Fig. 14: Dependence of mean simulated matrix temperature and its standard deviation on matrix fill for four null models implemented in *Nestedness*. A, B: sequential swap; C, D: fixed row equiprobable column; E, F: equiprobable row, fixed column; G, H: sampling from a lognormal (a = 0.2).

is not optimal (Rodriguéz-Gironés and Santamaría 2006). Nevertheless, for practical purposes the algorithms appears to be sufficient.

Unfortunately, sorting time increases exponentially with matrix size (a feature of NTC too). Therefore, for larger matrices it might be desirable to approximate the optimal stage by a simple sorting according to column and row totals. Fig. 13 shows that the matrix temperatures after sorting according to row and column totals are very similar to those obtained from a sorting according to temperature. As expected they are slightly larger than those obtained from temperature minimalization. However, in about 5% of cases (particularly at very low temperatures) the algorithm runs into troubles and the simple sorting accord-

ing to row and column totals gives better results.

You can also leave the matrix unsorted. This feature is particularly helpful in studies of gradients that might influence the degree of nestedness.

5. The output files

Nestedness produces three output files. The first file (*Unexp.txt*) gives the packed matrix together with the d_{ij} values for each cell as well as row and column totals (Σd_i and Σd_i). From this file you can exactly infer the degree of unexpectedness of each cell and its contribution to total unexpectedness. The file gives further the total numbers of filled rows (RowF) and columns (ColF), the point [X,Y] where the border lines intersect the matrix diagonal, the slopes of the

two borderlines and the theoretical intercepts of these lines. If there is not more than one full column or row the borderlines origin at the matrix edges (RowN = species; ColN = sites) (cf. Fig. 7, there ColN will be > sites, because the borderline intersects outside the matrix).

The next file (*Output.txt*) contains basic information about the matrix and the measurements. For N0, N1, Nc, UA, UP, UT, NODF, and BR observed counts, simulatede counts, standard deviations, Z-scores, skews of the null model distribution, and lower and upper 95% confidence limits of this distribution are provided. Note that for NC the fixed - fixed null model cannot be applied while giving always the same number of counts. Note further that for all measures except NC negative Z-scores point to nestedness. For NC a positive score indicates nestedness.

Next the program gives the results for the matrix temperature (MT). Numbers of species and sites, and the matrix fill are basic properties. Given are also the observed matrix temperature and the simulated temperatures according to the predefined null model, the standard deviation and the respective Z-score. Again, the program prints the skewness of this distribution and the respective lower and upper 95% confidence limits.

Next the program gives total occurrences (TotalOcc) and unexpected occurrences/absences (UnAbsOcc) and idiosyncratic temperatures for each species and each site. Note that the total number of unexpected occurrences/absences might be higher than the total number of occurrences. In theory, the total number of unexpected occurrences should equal the total number of unexpected absences. However, occurrences or absences close to the border lines are omitted. Note also that the quotient of unexpected occurrences/absences to the total number of occurrences is only weakly correlated to the idiosyncratic temperature.

The third file *Matrix.txt* contains the packed original data matrix and the last randomized packed matrix. The examples above show these matrices. Note that the matrices are sorted according to matrix temperature. Hence reversals in species and site richness occur.

6. Null model properties

Fig. 14 shows that the null models used (Random00 is omitted because it was discussed earlier) depend more or less on matrix size and fill. Hence there might be biases in detecting order or disorder with respect to size and fill. Least affected seems to be the fixed row - fixed column model (sequential swap). However, standard deviations of this model are mostly smaller than those of the other models increasing thus the probability of pointing to order or disorder. Fixed row - equiprobable column and fixed column - equiprobable row constraints behave quite similar, which is not surprising due to the theoretical equivalence of matrix rows and columns. The sampling model appears to be robust against matrix fill but not against matrix size. Theoretically this model should behave similar to the fixed row - fixed column model.

There is no consensus about the question which null model is best suited to account for passive sampling. Because the lognormal model is a passive sampling model it should theoretically be preferred. However, we seldom know the underlying abundance distribution from which the individuals are drawn. It is sometimes argued (McAbendroth et al. 2004) that the fixed row equiprobable column constraints account best for passive sampling. However, there is still no comparative study and the presence results (Fig. 13, 14) do not corroborate this conclusion. In my view the lognormal model and the sequential swap should be preferred.

7. Citing *Nestedness*

Nestedness is freeware but nevertheless if you use *Nestedness* in scientific work you should cite *Nestedness* as follows:

Ulrich W. 2006 - *Nestedness* – a FORTRAN program for calculating ecological matrix temperatures. www.uni.torun.pl/~ulrichw

8. System requirements

Nestedness is written in FORTRAN 95 and runs under Windows 9.x, XP, and Vista. Computation abilities are only limited by the computer's memory.

9. Acknowledgements

10. References

Almeida-Neto M., Guimarães P., Guimarães Jr P. R., Loyola R. D., Ulrich W. 2008. A consistent metric for nestedness analysis in ecological systems: reconciling concept and quantification. Oikos: in press.

Andren H. 1994 - Effects of habitat fragmentation on birds and mammals in landscapes with different proportions of suitable habitat: a review - Oikos 71: 355-366.

Atmar W., Patterson B. D. 1993 - The measure of order and disorder in the distribution of species in fragmented habitat - Oecologia 96: 373-382.

Atmar W., Patterson B. D. 1995 - The nestedness temperature calculator: a visual basic program, including 294 presence absence matrices. AICS Research Incorporate and The Field Museum.

Brualdi R. A., Sanderson J. G. 1999 - Nested species subsets, gaps, and discrepancy - Oecologia 119: 256-264.

Cutler A. 1991 - Nested faunas and extinction in fragmented habitats - Conserv. Biol. 5: 496-505.

Fischer J., Lindenmayer D. B. 2002 - Treating the nestedness temperature calculator as a "black box" can lead to false conclusions. Oikos 99: 193-199.

Fischer J., Lindenmayer D. B. 2005 - Perfectly nested or significantly nested—an important difference for conservation management - Oikos 109: 485-494.

Gotelli N. J. 2000 - Null model analysis of species co-occurrence patterns - Ecology 81: 2606-2621.

Gotelli N. J. 2001 - Research frontiers in null model analysis - Glob. Ecol. Biogeogr. 10: 337-343.

Gotelli, N.J., Entsminger, G.L. 2006. EcoSim: Null models software for ecology. Version 7. - Acquired Intelligence Inc. & Kesey-Bear. Burlington, VT 05465.

Honnay O., Hermy M., Coppin P. 1999 - Nested plant communities in deciduous forest fragments: species relaxation or nested habitats? - Oikos 84: 119-129.

Kolasa J. 1996 - Nestedness and discontinuities in species range size distributions - Tr. Ecol. Evol. 11: 433.

Lomolino M. V. 1996 - Investigating causality of nestedness of insular communities: selective immigrations or extinctions? - J. Biogeogr. 23: 699-703.

May R. M. 1975 - Patterns of species abundance and diversity - In: Cody M. L., Diamond J. M. (eds.) *Ecology and evolution of communities,* Cambridge, Belknap, pp. 81-120.

McAbendroth L., Foggo A., Rundle D., Bilton D. T. 2005 - Unravelling nestedness and spatial pattern in pond assemblages - J. Anim. Ecol. 74: 41-49.

Patterson B. D., Atmar W. 1986 - Nested subsets and the structure of insular mammalian faunas and archipelagos - In: Heaney L. R., Patterson B. D. (eds) *Island biogeography of mammals*, Academic Press, London, pp. 65-82.

Patterson B. D., Atmar W. 2000 - Analyzing species composition in fragments - Bonn. Zool. Monogr. 46.

Rodriguéz-Gironés M. A., Santamaría L 2006 - A new algorithm to calculate the nestedness temperature of presence-absence matrices - J. Biogeogr. 33: 924-935.

Wright D. H., Patterson B. D., Mikkelson G. M., Cutler A., Atmar W. 1998 - A comparative analysis of nested subset patterns of species composition. Oecologia 113: 1-20.

Wright D. H., Reeves J. H. 1992 - On the meaning and measurement of nestedness of species assemblages - Oecologia 92: 416-428.

Ulrich W., Gotelli N. J. 2007a. Null model analysis of species nestedness patterns. Ecology: 88: 1824-1831.

Ulrich W., Gotelli N. J. 2007b. Disentangling community patterns of nestedness and species co-occurrence. Oikos 116: 2053-2061.