Pairs – a FORTRAN program for studying pair wise species associations in ecological matrices

Version 1.0

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: 10.06.2008

1. Introduction

The study of species co-occurrences has a long tradition in ecology (Weiher and Keddy 1999). Particularly the long lasting and still ongoing discussion around community assembly rules (Diamond 1975, Diamond and Gilpin 1982, Gilpin and Diamond 1982, Connor and Simberloff 1979, 1983, 1984, Gotelli and McCabe, 2002, Ulrich 2004) has inspired the development of statistical tools to infer non random patterns in community assembly (Gotelli and Graves 1996, Gotelli 2000, 2001, Ulrich and Gotelli 2007a, b).

Community assembly is often studied in terms of nestedness (Patterson and Atmar 1986, Wright et al. 1998, Ulrich and Gotelli 2007a) and non-random patterns of species co-occurrences (Diamond 1975, Gotelli 2000, 2001, Ulrich and Gotelli 2007b) and the respective standard software is widely used: The *Nestedness Temperature Calculator* (Atmar and Patterson 1995), *EcoSim* (Gotelli and Entsminger 2002, and *Nestedness* (Ulrich 2006).

The present program *Pairs* extends these approaches and implements beside standard nestedness and co-occurrence metrics a new metric to study pairwise species associations: The software implements:

The discrepancy metric of Brualdi and Sanderson (1999),

the species combinations score (Pielou and Pielou 1968),

the C-score (Stone and Roberts 1990),

the checkerboard score (Gotelli 2000),

the Soerensen metric,

the togetherness score (Stone and Roberts 1992),

the species absences score (Stone and Roberts 1992),

the variance test (Schluter 1984),

a pairwise correlation test.

Pairs contains therefore similar metrics than the software *CoOccurrence* (Ulrich 2006) and is designed for the study of multiple matrices in null model analysis, the analysis of the statistical behaviour of certain metrics, and in studies of neutral models.

2. Metrics

Nestedness

A proper metric (Ulrich and Gotelli 2007a) to measure nestedness is the discrepancy *BR* (Brualdi and Sanderson 1999) that counts the number of discrepancies (absences or presence) that must be erased to produce a perfectly nested matrix. BR is standardized by dividing its values through the total number of occurrences in the matrix (the matrix fill) (Greve and Chown 2006).

Co-occurrence

The present program implements five matrix wide measures of co-occurrences:

1. The species combinations score (COMBO) screens the columns of the presence absence or abundance matrix for unique species combinations (Pilou and Pielou 1968). Hence, it counts the number of species that always co-occur.

2. The checkerboard score (Gotelli 2000) screens the matrix for checkerboards. These are 2*2 submatrices of the structure $\begin{matrix}1 & 0\\ 0 & 1\end{matrix}$ or $\begin{matrix}0 & 1\\ 1 & 0\end{matrix}$. The score is a simple count of the numer of such submatrices.

3. The C-score (Stone and Roberts 1990) is the average number of checkerboards for two species i and j. The score is calculated from

$$CS = \frac{2\sum_{i=1}^{S(S-1)/2}(n_i - N_{ij})(n_j - N_{ij})}{S(S-1)}$$

where S is the number of species and n_i and n_j are the row totals (numbers of occurrences) of species i and j and N_{ij} is the number of co-occurrences of both species.

4. the togetherness score (Stone and Roberts 1992) is based on presences and absences and calculated from

$$t = \frac{2}{S(S-1)}\sum_{1}^{S(S-1)/2}\frac{4p_{ij}a_{ij}}{sites^2}$$

Where p_{ij} and a_{ij} are the numbers of pairs presences and absences, respectively.

5. The absences score equals the C-score but counts joint absences instead of joint presences. For comparing matrices of different size and shape the metrics have to be standardized. This is done by dividing the effect size through the expected value.

6. The Soerensen score is calculated from

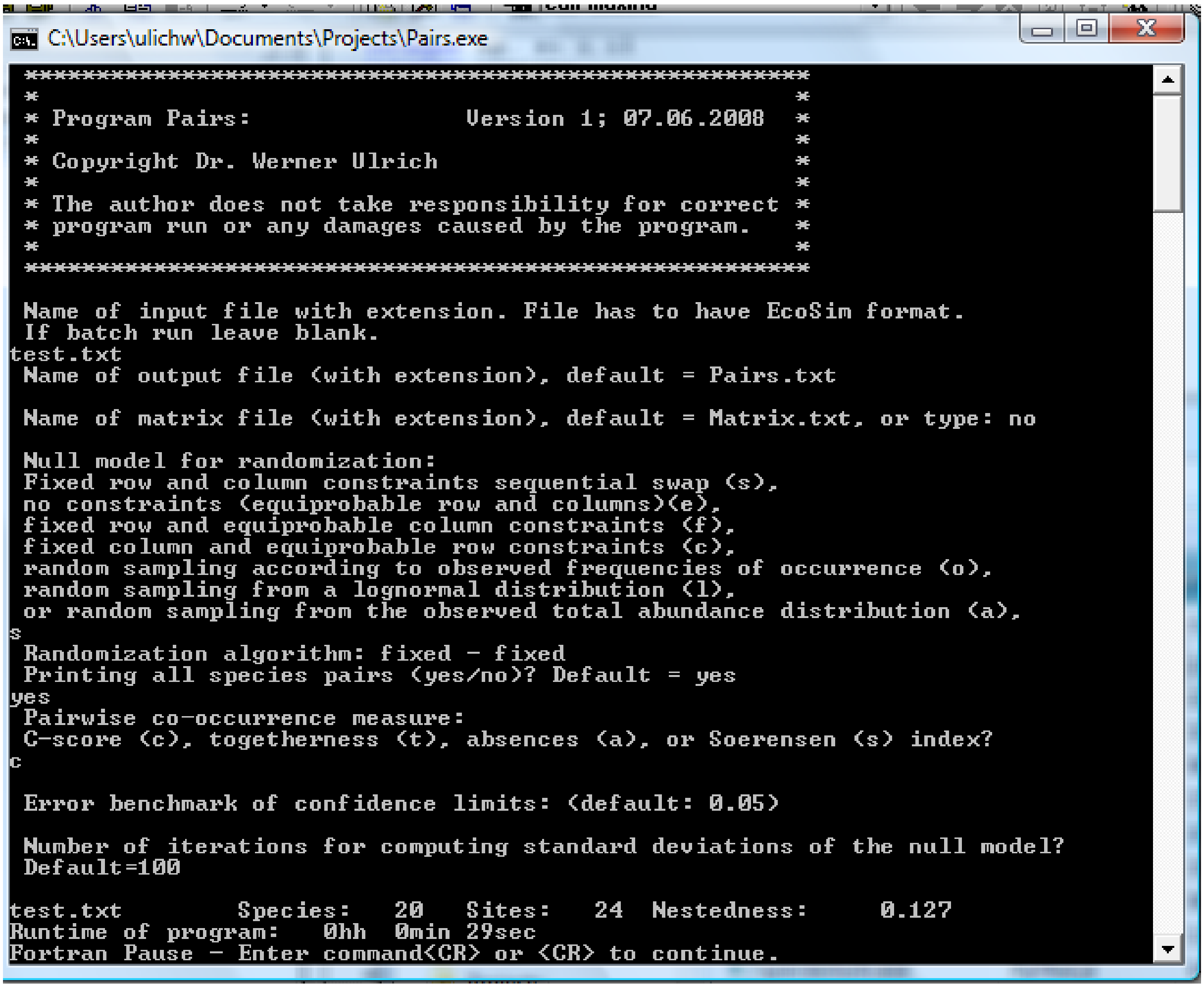
$$Soe = \frac{2}{S(S-1)}\sum_{1}^{S(S-1)/2}\frac{2n_{ij}}{n_i + n_j}$$

7. The variance test of Schluter (1984)

8. A matrix wide correlation coefficient calculated as the mean of the Spearman rank order correlations between all pair-wise site correlations. This option is only available for matrices that contain abundance data.

3. Species pairs

Pairs not only studies matrix wide patterns. It uses a Bayesian approach to detect non-random associations of pairs of species. The number of species pairs of a matrix is S(S-1/2. Hence even for medium sized matrices many ‘significantly non-random’ species pairs are expected at the 1% or 5% error level. For instance in a matrix of 50 species 61 significant pairs are expected just by chance at the 5% error benchmark. To reduce this high false detection error rate *Pairs* calculates first the expected empirical Bayes distribution of co-occurrence scores (C-score, togetherness score, Soerensen score and joint absences score) and compares this expectation with the observed distribution



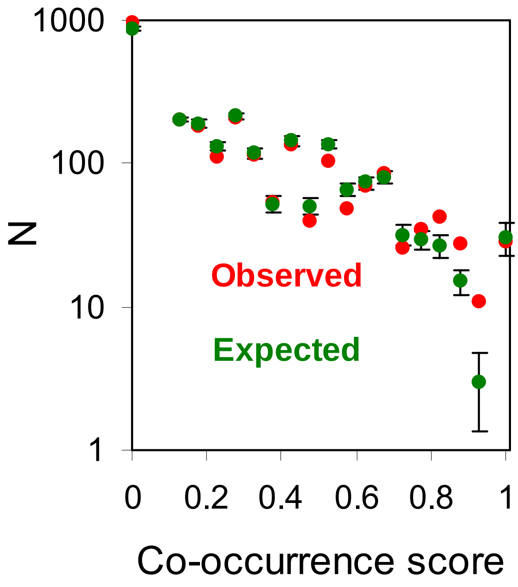
```
 ***********************************************************
 *                                                         *
 * Program Pairs:               Version 1; 07.06.2008      *
 *                                                         *
 * Copyright Dr. Werner Ulrich                             *
 *                                                         *
 * 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 = Pairs.txt

 Name of matrix file (with extension), default = Matrix.txt, or type: no

 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 from a lognormal distribution (l),
 or random sampling from the observed total abundance distribution (a),
s
 Randomization algorithm: fixed - fixed
 Printing all species pairs (yes/no)? Default = yes
yes
 Pairwise co-occurrence measure:
 C-score (c), togetherness (t), absences (a), or Soerensen (s) index?
c

 Error benchmark of confidence limits: (default: 0.05)

 Number of iterations for computing standard deviations of the null model?
 Default=100

test.txt        Species:   20   Sites:   24  Nestedness:     0.127
Runtime of program:   0hh  0min 29sec
Fortran Pause - Enter command<CR> or <CR> to continue.
```

of scores. This is done from the predefined number of random matrices. To compares observation and expectation scores of all metrics are standardized in the range of 0 to 1 and classified into 22 groups: 0, 0-0.04999,0.05-0.09999,….,0.9-0.94999, 0.95-0.99999,1.

The scores are calculated as above but instead dividing through S(S-1) they are divided through n_in_j for each pair .

The Fig. above shows such a comparison. In many cases the observed number of scores will be well within the confidence limits of the null expectation irrespective of whether the pair wise scores are later identified as being significant or not. *Pairs* chooses only those scores for further analysis where the number of observed instances is either larger than the mean expectation (Bayes M criterion) or larger than the upper 5% or 1% confidence limit (Bayes CL criterion). In the Figure the number of observed scores having values between 0.8 and 0.95 are well above the null expectation and pairs having such scores are first candidates to look for on-random associations.

For each pairs a Z-transformed value (Obs-Exp)/StDev. is calculated. For each of the above defined score classes *Pairs* calculates the two odds ratios = (Obs-Exp)/Obs. The first uses the mean of the Bayes distribution, the second its upper confidence limit. This value equals 1- false detection error rate. For all classes with positive odds ratios it chooses those pairs with scores above the respective pair wise confidence limit of the null model and prints it Z-scores. A further selection step (the Bayes M criterion) involves the odd



ratios. Only those pairs with the highest Z-scores are chosen. The benchmark is the number of pairs in the score class multiplied with the ML odds ratio. The most conservative third criterion (Byes CL) uses the CL odds ratio.

In the next analysis step the program calculates all S(S-1)/2 species pair scores and compares these with the null expectation. As a default null expectations are calculated from 100 randomized matrices for each pair.

An example: In the Figure above in the score class 0.5-0.55 106 pairs were observed but 136 pairs expected. The pairs of this class are not further considered as candidates for non-random associations. In the class from 0.85-0.9 28 pairs were observed but only 15 expected with an upper confidence limit of 22. Hence the Bayes M odds ratio is (28-15)/28 = 0.46 and the respective Bayes CL odd ratio = 0.21. That means only the 46% (ME) or 21% (CL) of species pairs with the highest pair wise significant Z-scores are considered as candidates for non-random association.

4. Data structure

Pairs 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 12 species distributed over 6 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 2002). The number of species is not limited, the maximum number of sites is about 150.

```
S  1  2  3  4  5  6
1  1.00  1.00  1.00  1.00  1.00  1.00
2  1.00  1.00  1.00  0.00  1.00  0.00
3  1.00  0.00  1.00  1.00  1.00  0.00
4  1.00  1.00  1.00  1.00  0.00  0.00
5  1.00  1.00  1.00  1.00  0.00  0.00
6  1.00  1.00  1.00  1.00  0.00  0.00
7  1.00  1.00  1.00  0.00  0.00  0.00
8  1.00  1.00  1.00  0.00  1.00  0.00
9  1.00  1.00  0.00  1.00  0.00  0.00
10  1.00  1.00  0.00  0.00  0.00  0.00
11  1.00  1.00  0.00  0.00  0.00  1.00
12  1.00  1.00  0.00  0.00  0.00  0.00
```

- The batch file format

Test1.txt
Test2.txt
Test3.txt
Test4.txt
Test5.txt

5. Program run

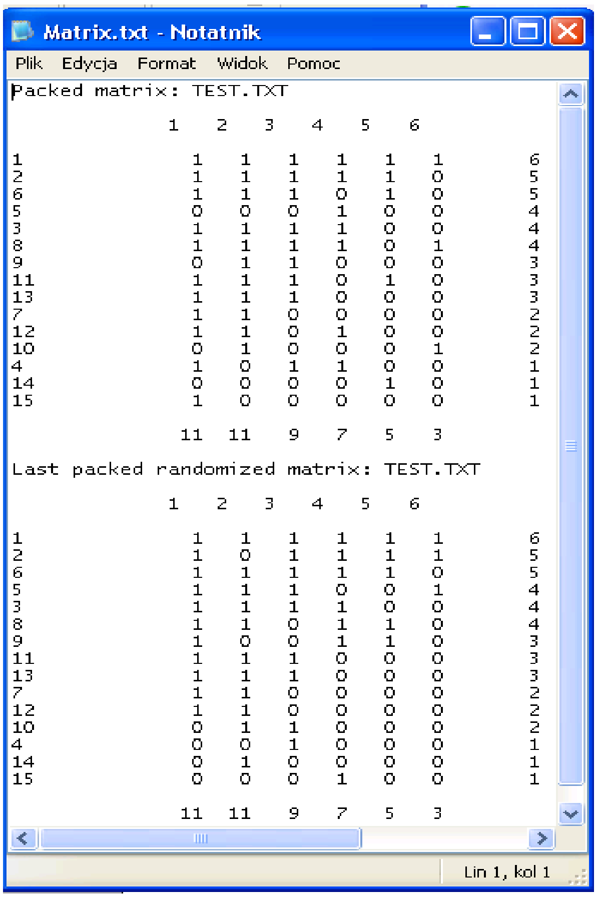
First, the program asks for the files names. The default output file names are *Pairs.txt*, *SignPairs.txt*, and *Matrix.txt*. You get the default values after returning *enter*. If you don’t give the name of the data file and return *enter* the program expects a batch run and a file name with the data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

Next, the program asks for the model for randomization. You have seven possibilities: A null model with 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) and Ulrich and Gotelli (2007a, b). The sequential swap model uses ten times the matrix size (10*rows*columns) 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 [$S=S_0Exp(-a(R-R_0)^2$] 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 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.

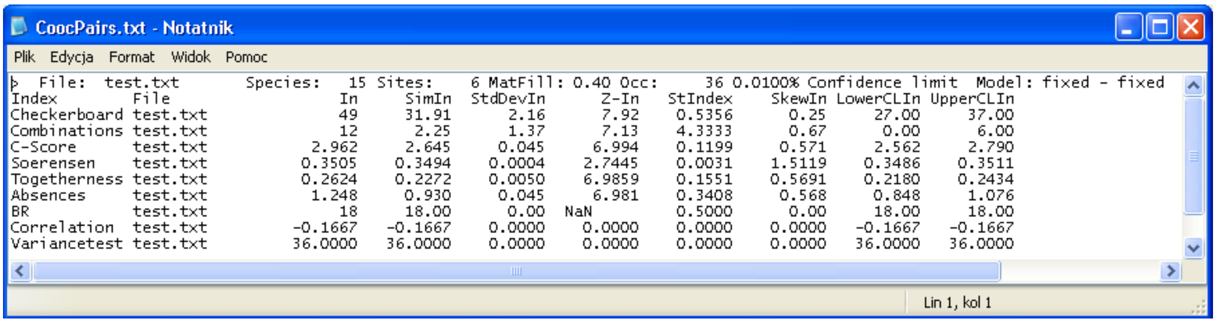
Next the program asks for the number of randomizations to compute the null model means and standard deviations, as well as upper and lower confidence limits. In most cases 100 such randomizations will be enough.



```
Packed matrix: TEST.TXT

                   1    2    3    4    5    6

1                    1    1    1    1    1    1          6
2                    1    1    1    1    1    0          5
6                    1    1    1    0    1    0          5
5                    0    0    0    1    0    0          4
3                    1    1    1    1    0    0          4
8                    1    1    1    1    0    1          4
9                    0    1    1    0    0    0          3
11                   1    1    1    0    1    0          3
13                   1    1    1    0    0    0          3
7                    1    1    0    0    0    0          2
12                   1    1    0    1    0    0          2
10                   0    1    0    0    0    1          2
4                    1    0    1    1    0    0          1
14                   0    0    0    0    1    0          1
15                   1    0    0    0    0    0          1

                    11   11    9    7    5    3

Last packed randomized matrix: TEST.TXT

                   1    2    3    4    5    6

1                    1    1    1    1    1    1          6
2                    1    0    1    1    1    1          5
6                    1    1    1    1    1    0          5
5                    1    1    1    0    0    1          4
3                    1    1    1    1    0    0          4
8                    1    1    0    1    1    0          4
9                    1    0    0    1    1    0          3
11                   1    1    1    0    0    0          3
13                   1    1    1    0    0    0          3
7                    1    1    0    0    0    0          2
12                   1    1    0    0    0    0          2
10                   0    1    1    0    0    0          2
4                    0    0    1    0    0    0          1
14                   0    1    0    0    0    0          1
15                   0    0    0    1    0    0          1

                    11   11    9    7    5    3
```

6. The output files

Pairs produces four output files. The first file (*CoocPairs.txt*) contains basic information about the matrix and the measurements. First it gives species and site numbers, matrix fill, the total number of occurrences, the confidence limit benchmark, and the null model algorithm. Then observed metric values, simulated values, the respective standard deviations, Z-scores, standardized values, skews of the null model distribution, and upper and lower confidence limits of



```
þ  File:  test.txt       Species:   15  Sites:    6 MatFill: 0.40 Occ:     36 0.0100% Confidence limit  Model: fixed - fixed
Index        File                 In      SimIn  StdDevIn       Z-In   StIndex    SkewIn LowerCLIn UpperCLIn
Checkerboard test.txt             49      31.91      2.16       7.92    0.5356      0.25     27.00     37.00
Combinations test.txt             12       2.25      1.37       7.13    4.3333      0.67      0.00      6.00
C-Score      test.txt          2.962      2.645     0.045      6.994    0.1199     0.571     2.562     2.790
Soerensen    test.txt         0.3505     0.3494    0.0004     2.7445    0.0031    1.5119    0.3486    0.3511
Togetherness test.txt         0.2624     0.2272    0.0050     6.9859    0.1551    0.5691    0.2180    0.2434
Absences     test.txt          1.248      0.930     0.045      6.981    0.3408     0.568     0.848     1.076
BR           test.txt             18      18.00      0.00  NaN        0.5000      0.00     18.00     18.00
Correlation  test.txt        -0.1667    -0.1667    0.0000     0.0000    0.0000    0.0000   -0.1667   -0.1667
Variancetest test.txt        36.0000    36.0000    0.0000     0.0000    0.0000    0.0000   36.0000   36.0000
```

this distribution are provided.

The second file *Matrix.txt* contains the packed original data matrix and the last randomized packed matrix. The examples above show both files.

The third output file *Pairs.txt* contains the observed and the empirical Bayes distribution, its mean, standard deviation, skewness and the lower and upper confidence limits. The two last columns contain the odds ratios of the respective score with regard to the Bayes M and the Bayes CL criterion. The odds ratio is the proportion of pairs above expectation = (Obs-Exp)/ Obs.

Next the program gives all species pairs. The last six columns contain Z-transformed scores [= (obs.-exp.)/ StDev] and associated probability levels. In the Z-score case it gives the Z-scores for those species with observed scores greater or smaller than the upper or lower confidence limit for that pair and the associated probability level. In the MeanScore and CLScore case it selects further according to the above defined Bayes M and Bayes CL criteria.

The last two columns contain false error rate corrected Z-scores and probability levels according to the method of Benjamini and Yekutieli (2001).This refinement modifies the test wise H_0 probability benchmark a from the ordered sequence (largest to smallest) of H_0 of the species pairs *r* probabilities p_k to

$$p_k^* = \alpha \frac{k}{r} \frac{1}{\sum_{i=1}^{r} \frac{1}{i}} \qquad (1)$$

where the *k* = 1 to *r* probability values p_k are ordered from largest to smallest, and $p*_k$ is the adjusted probability benchmark. The second last column contains the associated Z-score.

The last output file *SignPairs.txt* contains the significant species pairs. It contains also counts of species pairs and significant species (lower and upper 95% confidence limits) pairs for each score class and for the total matrix.

The file contains also expected numbers of significant values (simple CL criterion) obtained from 100 random matrices.

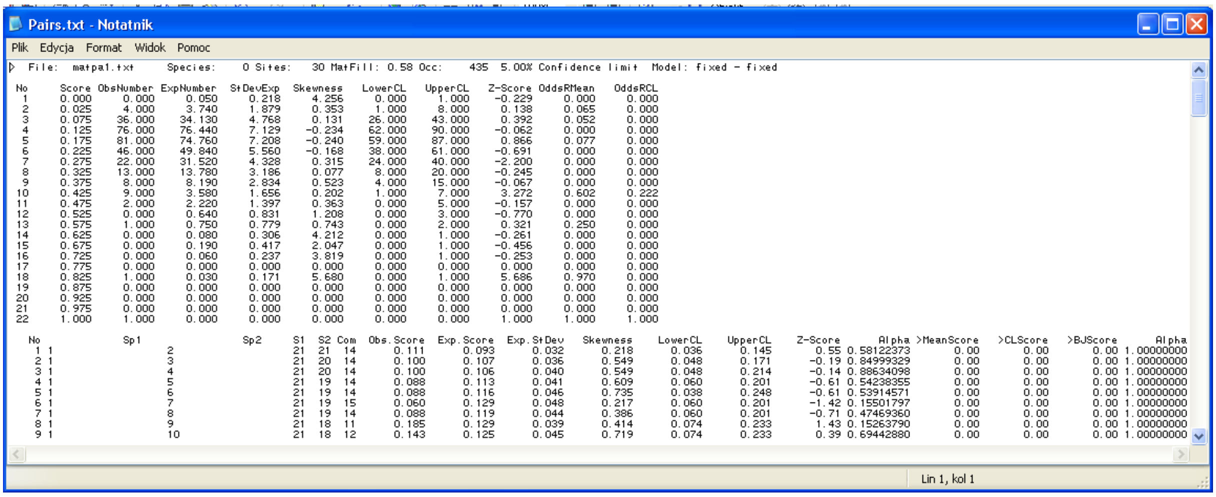
7. Citing *Pairs*

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

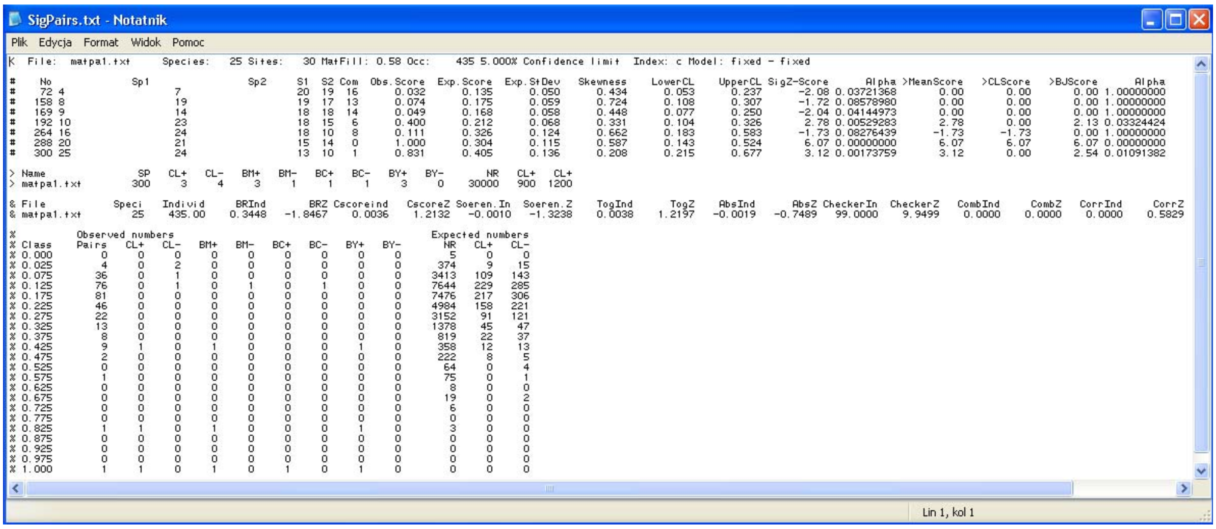
Ulrich W. 2008. Pairs – a FORTRAN program for studying pair-wise species associations in ecological matrices. www.uni.torun.pl/~ulrichw

8. System requirements

Pairs is written in FORTRAN 95 and runs under Windows 9.x, XP, and Vista. Computation abilities are



```
 File:   matpa1.txt      Species:     0 Sites:    30 MatFill: 0.58 Occ:     435   5.00% Confidence limit   Model: fixed - fixed

 No     Score  ObsNumber  ExpNumber   StDevExp   Skewness    LowerCL    UpperCL    Z-Score  OddsRMean     OddsRCL
  1     0.000      0.000      0.050      0.218      4.256      0.000      1.000     -0.229      0.000       0.000
  2     0.025      4.000      3.740      1.879      0.353      1.000      8.000      0.138      0.065       0.000
  3     0.075     36.000     34.130      4.768      0.131     26.000     43.000      0.392      0.052       0.000
  4     0.125     76.000     76.440      7.129     -0.234     62.000     90.000     -0.062      0.000       0.000
  5     0.175     81.000     74.760      7.208     -0.240     59.000     87.000      0.866      0.077       0.000
  6     0.225     46.000     49.840      5.560     -0.168     38.000     61.000     -0.691      0.000       0.000
  7     0.275     22.000     31.520      4.328      0.315     24.000     40.000     -2.200      0.000       0.000
  8     0.325     13.000     13.780      3.186      0.077      8.000     20.000     -0.245      0.000       0.000
  9     0.375      8.000      8.190      2.834      0.523      4.000     15.000     -0.067      0.000       0.000
 10     0.425      9.000      3.580      1.656      0.202      1.000      7.000      3.272      0.602       0.222
 11     0.475      2.000      2.220      1.397      0.363      0.000      5.000     -0.157      0.000       0.000
 12     0.525      0.000      0.640      0.831      1.208      0.000      3.000     -0.770      0.000       0.000
 13     0.575      1.000      0.750      0.779      0.743      0.000      2.000      0.321      0.250       0.000
 14     0.625      0.000      0.080      0.306      4.212      0.000      1.000     -0.261      0.000       0.000
 15     0.675      0.000      0.190      0.417      2.047      0.000      1.000     -0.456      0.000       0.000
 16     0.725      0.000      0.060      0.237      3.819      0.000      1.000     -0.253      0.000       0.000
 17     0.775      0.000      0.000      0.000      0.000      0.000      0.000      0.000      0.000       0.000
 18     0.825      1.000      0.030      0.171      5.680      0.000      1.000      5.686      0.970       0.000
 19     0.875      0.000      0.000      0.000      0.000      0.000      0.000      0.000      0.000       0.000
 20     0.925      0.000      0.000      0.000      0.000      0.000      0.000      0.000      0.000       0.000
 21     0.975      0.000      0.000      0.000      0.000      0.000      0.000      0.000      0.000       0.000
 22     1.000      1.000      0.000      0.000      0.000      0.000      0.000      1.000      1.000       1.000

   No            Sp1               Sp2     S1   S2  Com  Obs.Score  Exp.Score  Exp.StDev   Skewness    LowerCL    UpperCL    Z-Score      Alpha >MeanScore    >CLScore    >BJScore      Alpha
    1 1          2                         21   21   14      0.111      0.093      0.032      0.218      0.036      0.145       0.55 0.58122373       0.00        0.00        0.00 1.00000000
    2 1          3                         21   20   14      0.100      0.107      0.036      0.549      0.048      0.171      -0.19 0.84999329       0.00        0.00        0.00 1.00000000
    3 1          4                         21   20   14      0.100      0.106      0.040      0.549      0.048      0.214      -0.14 0.88634098       0.00        0.00        0.00 1.00000000
    4 1          5                         21   19   14      0.088      0.113      0.041      0.609      0.060      0.201      -0.61 0.54238355       0.00        0.00        0.00 1.00000000
    5 1          6                         21   19   14      0.088      0.116      0.046      0.735      0.038      0.248      -0.61 0.53914571       0.00        0.00        0.00 1.00000000
    6 1          7                         21   19   15      0.060      0.129      0.048      0.217      0.060      0.201      -1.42 0.15501797       0.00        0.00        0.00 1.00000000
    7 1          8                         21   19   14      0.088      0.119      0.044      0.386      0.060      0.201      -0.71 0.47469360       0.00        0.00        0.00 1.00000000
    8 1          9                         21   18   11      0.185      0.129      0.039      0.414      0.074      0.233       1.43 0.15263790       0.00        0.00        0.00 1.00000000
    9 1          10                        21   18   12      0.143      0.125      0.045      0.719      0.074      0.233       0.39 0.69442880       0.00        0.00        0.00 1.00000000
```



```
K  File:  matpa1.txt     Species:    25 Sites:    30 MatFill: 0.58 Occ:    435 5.000% Confidence limit  Index: c Model: fixed - fixed

#    No            Sp1              Sp2     S1  S2 Com  Obs.Score  Exp.Score  Exp.StDev   Skewness    LowerCL    UpperCL SigZ-Score      Alpha >MeanScore   >CLScore   >BJScore      Alpha
#    72 4            7                      20  19  16      0.032      0.135      0.050      0.434      0.053      0.237      -2.08 0.03721368       0.00       0.00       0.00 1.00000000
#   158 8           19                      19  17  13      0.074      0.175      0.059      0.724      0.108      0.307      -1.72 0.08578980       0.00       0.00       0.00 1.00000000
#   169 9           14                      18  18  14      0.049      0.168      0.058      0.448      0.077      0.250      -2.04 0.04144973       0.00       0.00       0.00 1.00000000
#   192 10          23                      18  15   6      0.400      0.212      0.068      0.331      0.104      0.326       2.78 0.00529283       2.78       0.00       2.13 0.03324424
#   264 16          24                      18  10   8      0.111      0.326      0.124      0.662      0.183      0.583      -1.73 0.08276439      -1.73      -1.73       0.00 1.00000000
#   288 20          21                      15  14   0      1.000      0.304      0.115      0.587      0.143      0.524       6.07 0.00000000       6.07       6.07       6.07 0.00000000
#   300 25          24                      13  10   1      0.831      0.405      0.136      0.208      0.215      0.677       3.12 0.00173759       3.12       0.00       2.54 0.01091382

> Name            SP    CL+    CL-    BM+    BM-    BC+    BC-    BY+    BY-       NR   CL+   CL+
> matpa1.txt     300      3      4      3      1      1      1      3      0    30000   900  1200

& File         Speci   Individ     BRInd       BRZ Cscoreind   CscoreZ Soeren.In  Soeren.Z     TogInd       TogZ     AbsInd       AbsZ CheckerIn   CheckerZ    CombInd      CombZ    CorrInd      CorrZ
& matpa1.txt      25    435.00    0.3448   -1.8467    0.0036    1.2132   -0.0010   -1.3238     0.0038     1.2197    -0.0019    -0.7489   99.0000     9.9499     0.0000     0.0000     0.0000     0.5829

%           Observed numbers                                                 Expected numbers
% Class     Pairs   CL+   CL-   BM+   BM-   BC+   BC-   BY+   BY-        NR   CL+   CL-
% 0.000         0     0     0     0     0     0     0     0     0         5     0     0
% 0.025         4     0     2     0     0     0     0     0     0       374     9    15
% 0.075        36     0     1     0     0     0     0     0     0      3413   109   143
% 0.125        76     0     1     0     1     0     1     0     0      7644   229   285
% 0.175        81     0     0     0     0     0     0     0     0      7476   217   306
% 0.225        46     0     0     0     0     0     0     0     0      4984   158   221
% 0.275        22     0     0     0     0     0     0     0     0      3152    91   121
% 0.325        13     0     0     0     0     0     0     0     0      1378    45    47
% 0.375         8     0     0     0     0     0     0     0     0       819    22    37
% 0.425         9     1     0     1     0     0     0     1     0       358    12    13
% 0.475         2     0     0     0     0     0     0     0     0       222     8     5
% 0.525         0     0     0     0     0     0     0     0     0        64     0     4
% 0.575         1     0     0     0     0     0     0     0     0        75     0     1
% 0.625         0     0     0     0     0     0     0     0     0         8     0     0
% 0.675         0     0     0     0     0     0     0     0     0        19     0     2
% 0.725         0     0     0     0     0     0     0     0     0         6     0     0
% 0.775         0     0     0     0     0     0     0     0     0         0     0     0
% 0.825         1     1     0     1     0     0     0     1     0         3     0     0
% 0.875         0     0     0     0     0     0     0     0     0         0     0     0
% 0.925         0     0     0     0     0     0     0     0     0         0     0     0
% 0.975         0     0     0     0     0     0     0     0     0         0     0     0
% 1.000         1     1     0     1     0     1     0     1     0         0     0     0
```

only limited by the computer’s memory.

9. Acknowledgements

The development was supported by grants from the Polish Science Committee (KBN, 3 P04F 034 22, KBN 2 P04F 039 29).

10. References

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.

Benjamini, Y., and D. Yekutieli. 2001. The control of false discovery rate in multiple testing under dependency. Annals of Statistics 29: 1165-1188.

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

Connor E. F., Simberloff D. 1979. The assembly of species communities: chance or competition. Ecology 60: 1132-1140.

Connor E. F., Simberloff D. 1983. Interspecific competition and species co-occurrence patterns on islands: null models and the evaluation of evidence. Oikos 41: 455-465.

Connor E. F., Simberloff D. 1984. Neutral models of species co-occurrence’ patterns. In: Strong D. R., Simberloff D., Abele L. G., Thistle A. B., (eds.), Ecological Communities: Conceptual Issues and the Evidence, Princeton Univ. Press, pp. 316-331.

Diamond J. M. 1975. Assembly of species communities. In: Cody, M. L., Diamond, J. M. (eds.), Ecology and Evolution of Communities, Harvard Press, Cambridge, pp. 342-444.

Diamond J. M., Gilpin M. E. 1982. Examination of the ‘null’ model of Connor and Simberloff for species co-occurrences on islands. Oecologia 52: 64-72.

Gilpin M. E., Diamond J. M. 1982. Factors contributing to non-randomness in species co-occurrences onislands. Oecologia 52: 75-84.

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. Global Ecology and Biogeography Letters 10: 337-343.

Gotelli N. J., McCabe D. 2002. Species co-occurrence: a meta-analysis of J. M. Diamond's assembly rules model. Ecology 83: 2091-2096.

Gotelli N.J., Entsminger G.L. 2002. EcoSim: Null models software for ecology. Version 7. - Acquired Intelligence Inc. & Kesey-Bear. Burlington, VT 05465. http://homepages.together.net/~gentsmin/ ecosim.htm.

Gotelli N. J., Graves G. R. 1996. Null Models in Ecology. Smithsonian Institution. Press, Washington D.C.

Greve M., Chown S. L. 2006. Endemicity biases nestedness metrics: a demonstration, explanation and solution. Ecography 29: 347–356

May R. M. 1975. Patterns of species abundance and diversity. In: Cody, M. L., Diamond J. M. (Eds) Ecology and evolution of communities. Harvard University Press, Cambridge, Massachusetts, USA, 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.

Pielou D. P., Pielou E. C. 1968. Association among species of infrequent occurrence: the insect and spider fauna of *Polyporus betulinus* (Bulliard) Fries. - Journal of Theoretical Biology 21: 202-216.

Schluter D. 1984 A variance test for detecting species associations, with some example applications. Ecology 65: 998-1005.

Stone L. and Roberts A. 1990. The checkerboard score and species distributions. - Oecologia 85: 74-79.

Stone L., Roberts A. 1992. Competitive exclusion or species aggregation? An aid in deciding. Oecologia 91: 419-424.

Ulrich W. 2004. Species co-occurrences and neutral models: reassessing J. M. Diamond's assembly rules. Oikos 107: 603-609.

Ulrich W. 2006. Nestedness—a Fortran program for program for calculating ecological matrix temperatures. Online available at http//www.uni.torun.pl/ ~ulrich.

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.

Weiher E., Keddy P. A. (eds.) 1999. Ecological Assembly Rules: Perspectives, Advances, Retreats. Cambridge Univ. Press, New York.

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