

MULTI-PATTERN 1.00. PROGRAM PACKAGE TO ANALYZE AND SIMULATE COMMUNITY-WIDE PATTERNS

Zs. Erdei and B. Tóthmérész

Erdei, Zs. and Tóthmérész, B. (1993): MULTI-PATTERN 1.00. Program package to analyze and simulate community-wide patterns. - Tiscia 27, 45-48.

Abstract: MULTI-PATTERN is a program package developed for IBM-compatible PC's to simulate multi-species point patterns and to calculate direct and indirect spatial statistics for real and simulated point patterns. The program is entirely menu controlled. The program package is supported by detailed graphical display of the results.

Keywords: multispecies point patterns, direct and indirect spatial series analysis, computerized sampling, simulation.

Zs. Erdei, B. Tóthmérész, Ecological Institute of Kossuth L. University, Debrecen, Pf. 14, H-4010, Hungary.

General description of the package

MULTI-PATTERN is written in Turbo Pascal. It may be run on IBM-compatible PC's with at least 640 Kbytes available memory. Because of the extensive input and output manipulations and of the large amount of simulated data, a hard disk is inevitable. At least 2 Mbytes free space is recommended to store the simulated data sets. The graphics routines can run on any graphics card, but EGA or VGA are preferred. To carry out an extensive simulation study a mathematical coprocessor is highly recommended. Each procedure is planned to minimize computing time and maximize the size of data set.

The package has two separate goals; one of these is to analyze the pattern of multispecies communities in direct and indirect spatial series analysis and the second one is to simulate multispecies point patterns which may be useful to test other techniques which are frequently used in ecology.

Graphical display of the field and/or simulated point patterns (Fig. 1) and the result of spatial series analysis is included in the package (Fig. 2).

Spatial statistics included in the package

There is a long tradition of the analysis of one- and two-species patterns in botany and ecology (Greigh-Smith, 1952, 1983; Kershaw, 1957, 1961). It is incredibly time-consuming, however, to study

multispecies spatial patterns quantitatively; Erdei et al. (1993) published results about computing times. Before the widespread availability of fast computers, the calculations even for species poor communities would have been a formidable task.

During a spatial series analysis the spatial pattern is analyzed using a series of different sized plots. For each plot size n plots are placed randomly to the studied community. The importance of such spatial analysis of the vegetation, when a series of increasing plot size is used, was recognized and emphasized very early by Juhász-Nagy (1967, 1976). According to the method developed by him the number of plots having different species list are counted first and after that the statistics describing

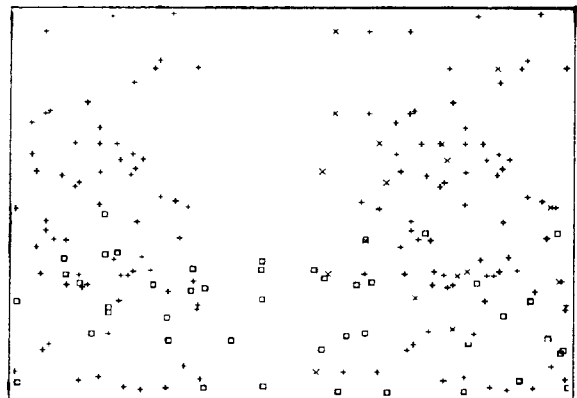


Fig. 1. A community containing four species. The distribution of the individuals are strongly aggregated for each species.

the spatial structure of the vegetation are calculated from this derived data set. Therefore, it might be mentioned as an indirect spatial series analysis. In the case of direct spatial series analysis for each plot a characteristics, like number of species or diversity, is calculated and the spatial pattern is described this way (Tóthmérész, 1993a).

Two statistics are included in the package for direct spatial series analysis; the average and variance of the number of species (Tóthmérész, 1993a). There are five indirect spatial statistics in the program: number of species combinations, florula diversity, florula evenness, preferential information heterogeneity or local distinctiveness and associatum (Juhász-Nagy and Podani, 1983; Podani, 1984a, 1984b).

Neutral models and complete spatial randomness

It is natural to use as a neutral or null model the hypothesis of Complete Spatial Randomness (CSR). CSR is frequently used in ecology and elsewhere to the formulation of hypotheses concerning the pattern and its genesis (Diggle, 1983; Ripley, 1981). The CSR hypothesis can be stated as follows: Given the occurrence of m individuals in a finite region A , the location of those individuals are a random sample of size m from the uniform distribution on A . There are alternative characterizations of CSR related to Poisson distribution.

Multispecies point patterns can be constructed by extending univariate models. The simplest such models are those compound of independent components. In many respects, the role of independence is analogous to the role of CSR in the univariate case (Cressie, 1991).

Juhász-Nagy stressed very early the importance of null models in ecology (Juhász-Nagy, 1970). Nowadays, the significance of these models is generally accepted (Caswell, 1976; Harvey et al., 1983). There are two possibilities to generate neutral models for a spatial series analysis. It is possible to calculate the "theoretical" value of a spatial statistics for a random community. The theoretical null model is calculated for "infinitely large" community where the ratio of the species is the same as for the studied community. Another possibility is to simulate a random community where the number of species and the number of individuals is the same as for the studied community, and then to calculate the statistics for that pattern; that produce a realized null model (see Fig. 2). Because of the finite number of individuals the realized and the

theoretical null model may produce strikingly different result. In the case of rare and subdominant species Tóthmérész and Erdei (1992) demonstrated that the association, which should be zero for a random community, is much higher than zero for realized random models.

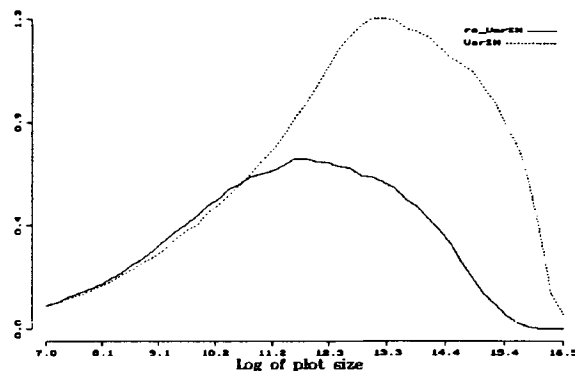


Fig. 2. Variance of the average number of species (re_varSN) in a direct spatial series analysis for the community presented in Fig. 1 and the realized null model of it ($varSN$).

Both theoretical and realized random models can be calculated by the package. Calculation of the theoretical values is included in the package as an option. The realized random model also can be simulated by the package and then, the realized values of the spatial statistics can be calculated.

Format of data files

The program works with the two-dimensional euclidean coordinates of the individuals. The default extension of the data files is "*.COO". The first line of the data file is a label which helps to identify the data-file. The second line contains the x-size and y-size of the studied area. The third line contains the number of individuals. In the following lines are the coordinates of the individuals; one individuals in one line. The first figure in the line is an integer value to identify the species. The second figure is the x-coordinate and the third figure is the y-coordinate of the individual. This data format is compatible with the format used by the SYN-TAX III (Podani, 1988); the only difference is that the first three lines of the files, which provide additional information from the data, are missing in the case of SYN-TAX and the file directly starts with the species identification and the coordinates of the data.

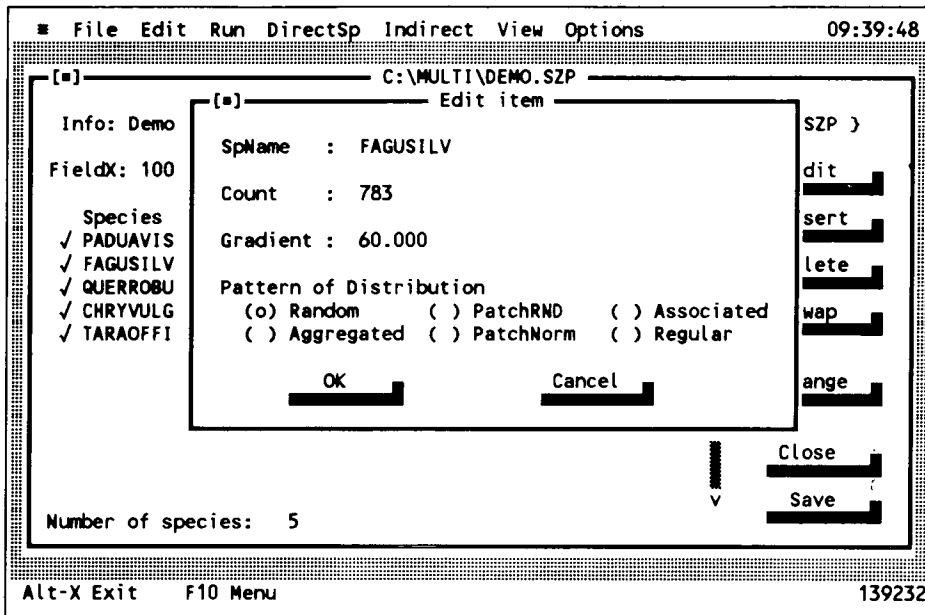


Fig. 3. The menu-controlled built-in editor of the program package to create a file containing the parameters of simulation.

Parameter files of simulation and sampling

Parameters of a simulation are proposed to store in a file having a ".SIM" extension. This is the default extension. The first line of the file contains information about the data included in the file. The second line contains two figures: the length and width of the area where the individuals are located. The third line contains the number of species. Then each consecutive line contains the number of individuals of a species and some other parameters describing the distributional pattern of the individuals. The following patterns can be generated: (i) uniformly random; (ii) uniformly random along a gradient; (iii) aggregated pattern; (iv.a) patchy distribution with 2-dimensional normal distribution within the patches; (iv.b) patchy distribution with uniformly random distribution within the patches; (v) association between the species; (vi) regular distribution. The file can be generated by any general-purpose editor or it is generated automatically by the program using the menu options (see Fig. 3). It is stored as an ASCII file.

It is also possible to create communities with pre-defined abundance-dominance structure; the following abundance-dominance distributions are included: (i) MacArthur's broken stick model; (ii) Zipf model; (iii) Zipf-Mandelbrot model. The

abundance-dominance structure can be modified or entirely created by the built-in editor of the package.

The point pattern resulted in by the simulation can be displayed by symbols of different form and colour using the "View" option of the program.

The program package is able to simulate sampling procedure. It can be used by the data set generated by the package or any other data set (e.g. by a real point map of a forest or any other communities). Total random and regular sampling may be carried out. Sampling units with the shape of square, rectangle, circle and ellipse can be used.

".PLP" is the default extension of a parameter file which describes the sampling technique of the spatial series analysis. This is also generated automatically by the program. The shape and size of the plots are included in the file.

There are two options for sampling. In the case of binary sampling just the presence/absence of species is recorded while in the case of quantitative sampling the number of individuals are also recorded. The result of sampling is saved to disk in a file with a user-defined file-name in the form of a two-way table. Thus it can be used by other programs or program packages. Data format of this file is identical with the data-format used by the NuCoSA program package (Tóthmérész, 1993c).

Availability.

Executable program and demo files for IBM-compatible PC's are available from the authors on request sending a formatted disk (3.5" or 5.25"). It is also available on E-mail.

Acknowledgements

We are indebted to Tamás Czárán and Sándor Bartha for helpful proposals. The research was supported by the Hungarian Research Fund (OTKA) F 6082 to the first author and T 5066 to the second author.

Literature

- Caswell, H. (1976): Community structure: A neutral model analysis. - *Ecol. Monogr.* 46, 327-354.
- Cressie, N. (1991): *Statistics for Spatial Data*. John Wiley and Sons, New York.
- Diggle, P.J. (1983): *Statistical Analysis of Spatial Point Patterns*. - Academic Press, New York.
- Erdei, Zs., Tóthmérész, B. and Erdei A. (1993): Linear algorithm to calculate indirect spatial statistics for completely random multi-species communities. - Preprint of the Ecological Institute of Kossuth L. University, Debrecen
- Greig-Smith, P. (1983): *Quantitative Plant Ecology*. 3rd ed. - Blackwell Scientific Publications, Oxford.
- Greig-Smith, P. (1952): The use of random and contiguous quadrats in the study of structure in plant communities. - *Annals Botany (London)* N.S. 16, 293-316.
- Harvey, P.H., Colwell, R.K., Silvertown and May, R.M. (1983): Null models in ecology. - *Ann. Rev. Ecol. Syst.* 14, 189-211.
- Juhász-Nagy, P. (1967): On some "characteristic areas" of plant community stands. *Proceedings of the Colloquium on Information Theory*. - Bolyai Mathematical Society (ed. A. Rényi). Akadémiai Kiadó.
- Juhász-Nagy, P. (1976): Spatial dependence of plant populations. Part 1. - *Acta Botanica Hungarica* 22, 61-78.
- Juhász-Nagy, P. (1970): Egy operatív ökológia hiánya és szükséglete. - *MTA Biol. Oszt. Közl.* 12, 441-464.
- Juhász-Nagy, P. and Podani, J. (1983): Information theory methods for the study of spatial processes and succession. - *Vegetatio* 51, 129-140.
- Kershaw, K.A. (1957): The use of cover and frequency in the detection of pattern in plant communities. - *Ecology* 38, 291-299.
- Kershaw, K.A. (1961): Association and co-variance analysis of plant communities. - *J. Ecol.* 49, 643-654.
- Podani, J. (1984a): Analysis of mapped and simulated vegetation pattern by means of computerized sampling techniques. - *Acta Bot. Hung.* 30, 403-425.
- Podani, J. (1984b): Spatial processes in the analysis of vegetation: theory and review. *Acta Botanica Hungarica* 30, 75-118.
- Podani, J. (1988): SYN-TAX III. - *Abstracta Botanica* 12 (supplement), 1-183.
- Ripley, B.D. (1981): *Spatial Statistics*. - Wiley, New York.
- Tóthmérész, B. (1993a): Direct Spatial Series Analysis of Multispecies Point Patterns. - Preprint of the Ecological Institute of Kossuth L. University, Debrecen, Hungary.
- Tóthmérész, B. (1993b): Spatial structure of multispecies community: A review of the state of the art and further perspectives. - Preprint of the Ecological Institute of Kossuth L. University, Debrecen, Hungary.
- Tóthmérész, B. (1993c): NuCoSA 1.00: Number cruncher for community studies and other ecological applications. - *Abstracta Botanica* 17, 155-158.
- Tóthmérész, B. and Erdei, Zs. (1992): The effect of species dominance on information theory characteristics of plant communities. - *Abstracta Botanica* 16, 43-47.