NEW FEATURES OF MULTI-PATTERN 1.10: ROBUST NONLINEAR SMOOTHING

B. Tóthmérész and Zs. Erdei

Abstract. New features of the MULTI-PATTERN program package, to analyse the spatial structure of multispecies point patterns, are discussed. Robust nonlinear smoothers are included in the package to improve the estimation of the maximum area of a community and the curve of spatial statistics. Eight robust nonlinear smoothers are mentioned and their performances are discussed. For spatial series analyses the smoothings related to the locally weighted robust regression are performed exceptionally well.

Keywords: robust nonlinear smoothing, maximum area, spatial series analyses, multispecies spatial point patterns.

B. Tóthmérész, Zs. Erdei, Ecological Institute, Kossuth L. University, Debrecen, POB 71, H-4010
Fax: (36) 52 431-148, E-mail: tothmerb@tigris.klte.hu and zserdei@tigris.klte.hu

Introduction

It was Juhász-Nagy who proposed the indirect spatial series analysis to study the spatial pattern of communities (Juhász-Nagy, 1967, 1976). He developed a systematic, quite formal theory. From technical point of view, spatial series analyses are based on the graphing of a spatial statistics against the plot size (Bartha, 1992; Juhász-Nagy and Podani, 1993; Podani et al., 1993). This is the so-called spatial series curve. Juhász-Nagy preferred such spatial statistics which were based on the presence-absence pattern of communities (Juhász-Nagy, 1984, 1993). During the interpretation of results the characteristics of the spatial series curve is vital; especially the global maximum and the local maxima are important. The main difficulty with the indirect spatial series analysis is that it requires huge sample size. Even with large sample sizes the oscillation of spatial series curve may be significant (Tóthmérész and Erdei, 1993; Tóthmérész, 1994b).

In this paper robust smoothings are used to identify the maximum area of spatial statistics for multispecies communities. Robust nonlinear smoothers are ideal tools to smooth heavily oscillating sequences and let us see regularities with the elimination of "random noise." Tóthmérész (1994a) proposed to use robust nonlinear smoothers to improve the estimation of spatial series curves and especially to estimate the maximum area. For small and medium sample sizes these techniques are reducing heavily the variance of the estimation of maximum area (Tóthmérész, 1994c); this was the basic motivation to include these methods into the MULTI-PATTERN program package (Erdei and Tóthmérész, 1993).

Methods

There are three groups of nonlinear smoothers in the new release of MULTI-PATTERN (Fig. 1). The first one includes composed running median smoothers. 3RSSH and 4253H are included. These are standard running median smoothers of the statistical literature (Velleman and Hoaglin, 1981; Tukey, 1977). Robust locally weighted regression (LOWE, Cleveland, 1978, 1979, 1981) and a special version of it (RANK-LOWE) proposed by Tóthmérész (1994c) comprise the second group. In the case of RANK-LOWE smoothing the steps of the spatial series analysis are considered to be equally scaled. We can also say that just the rank-order of the plot sizes is used and the actual size of the plot is not used during the computations. Finally, there is a third group of smoothers included in the package. These are based on the composition of a running median

smoothers and the robust locally weighted regression. These were also proposed and assessed by Tóthmérész (1994c).
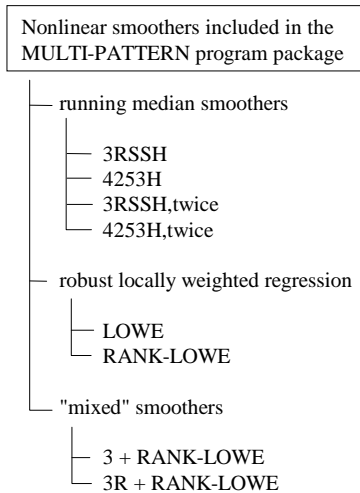
```
┌─────────────────────────────────────┐
│ Nonlinear smoothers included in the  │
│ MULTI-PATTERN program package        │
└─────────────────────────────────────┘

├── running median smoothers

          ├── 3RSSH
          ├── 4253H
          ├── 3RSSH,twice
          └── 4253H,twice

├── robust locally weighted regression

          ├── LOWE
          └── RANK-LOWE

└── "mixed" smoothers

          ├── 3 + RANK-LOWE
          └── 3R + RANK-LOWE
```

Figure 1. Three diagram of the nonlinear smoothers included in the package.

How to use the new options

The new "Smoothing RESult files" option is included in the "Edit" menu. This is the last item in the pull down menu (Fig. 2). With a click on the left button of the mouse to the option, the screen of the nonlinear smoothers appears (Fig. 3). You can smooth spatial series curves. This means that you have to have a "*.RES" file which contains the result of a direct or indirect spatial series analysis if you would like to use the nonlinear smoothers. The first step is to choose a file which contains the result of a spatial series analysis. With a click on the "Input file" button you receive a list of "*.RES" files of the directory and thus, you can choose a file to be smoothed or you can change the directory to find a file. The result of the smoothing is written into another file. Thus, you have to define the file name clicking on the "Output file" button and typing the file name. Then you have to choose one of the nonlinear smoothers by clicking to the name of the smoothers. Your choice is indicated by a "($\bullet$)" sign. In the case which is demonstrated in Fig. 3 we have chosen the RANK-LOWE smoother. Finally, when everything is all right, you have to click on the "Smooth" button or on the "Cancel" button if you decided not to smooth et al. The program shows the progression of the calculation and informs you when it is ready. Then you can use the "View" option of the program to draw the smoothed spatial series curves.

For 3RSSH and 4253H smoothers there is a possibility for twice-ing using the "Twice" button. This procedure is based on the smoothing and adding of the rough to the smoothed series (Tukey,
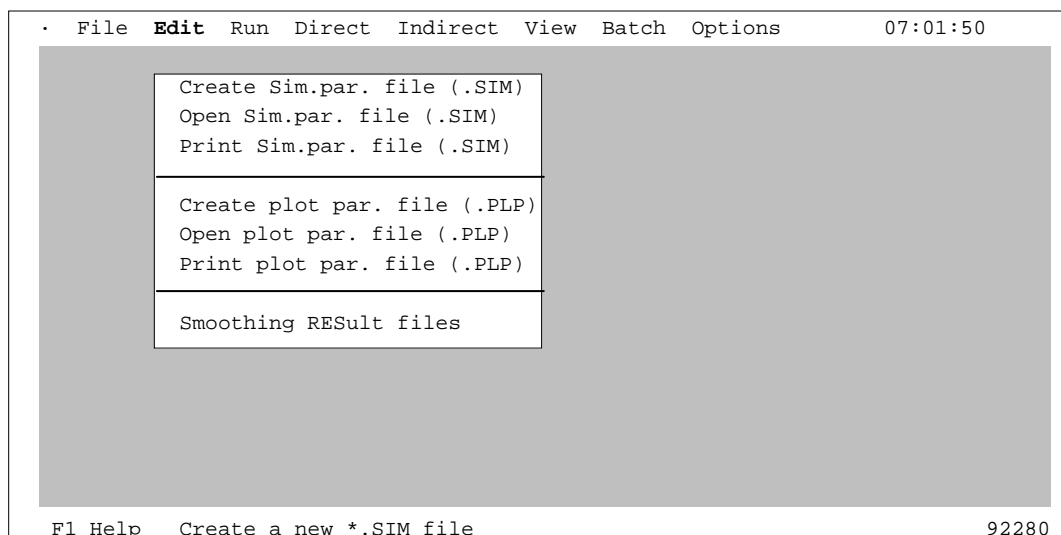
```
·  File  Edit  Run  Direct  Indirect  View  Batch  Options        07:01:50
  ┌─────────────────────────────────────┐
  │  Create Sim.par. file (.SIM)         │
  │  Open Sim.par. file (.SIM)           │
  │  Print Sim.par. file (.SIM)          │
  │                                      │
  │  Create plot par. file (.PLP)        │
  │  Open plot par. file (.PLP)          │
  │  Print plot par. file (.PLP)         │
  │                                      │
  │  Smoothing RESult files              │
  └─────────────────────────────────────┘



   F1 Help    Create a new *.SIM file                              92280
```

Figure 2. The "Smoothing RESult files" option is included in the "Edit" menu.

```
  ·   File   Edit   Run   Direct   Indirect   View   Batch   Options          07:01:58

  ┌[■]──────────────────────── Smooth a RES file ─────────────────────────┐
  │                                                                        │
  │       Input file ■          C:\MULTI\DEMO.RES                          │
  │                                                                        │
  │       Output file ■         C:\MULTI\SMOOTH.RES                        │
  │                                                                        │
  │       Select smoothing method                                         │
  │       ( )  3RSSH                        [ ] Twice                      │
  │       ( )  4253H                                                       │
  │       (•)  RANK-LOWE                                                   │
  │       ( )  LOWE                                                        │
  │       ( )  3 + RANK-LOWE           LOWE parameter                      │
  │       ( )  3R + RANK-LOWE              0.25                            │
  │               Smooth ■                  Cancel ■                       │
  │                                                                        │
  └────────────────────────────────────────────────────────────────────────┘



  Alt-X Exit      F10 Menu     F1 Help                                   90848
```
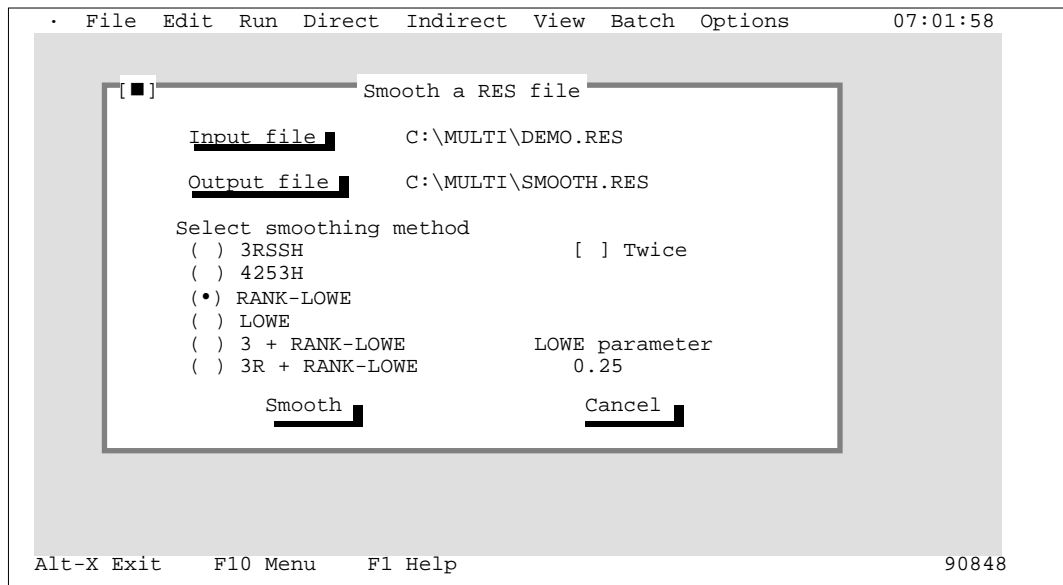
Figure 3. The screen of the nonlinear smoothers in the package.

1977). This may prevent the elimination of some of the trends or fluctuations that are really present in the data sequence.

Discussion

There is a rather elementary solution to "eliminate" the oscillation of the spatial series curve. Simply to choose a large step size for the spatial series analysis as it is demonstrated by the Fig. 4. In this case the curve looks smooth enough for moderate plot sizes. For small step sizes the oscillation is too heavy even for large sample sizes. Large step size, however, prevents the correct identification of the maximum area. Decreasing the step size, we have to cope with the oscillation. The problem of oscillation was also reported by Tóthmérész and Erdei (1992).

Figs. 5. and 6. show the spatial series curve of a 4-species completely spatially random community using the diversity of species combinations as a spatial statistics and the smoothed versions of the curve. The number of sample plots were 100 for each step of the spatial series analysis. The 4253H is not a drastic smoother; it does not eliminate the oscillation of the curve just the high peaks and low valleys (Fig. 5). The 3R+RANK-LOWE produces a very smooth curve which eliminates almost all the oscillations (Fig. 6).

Robust nonlinear smoothers are very powerful in extracting regularities from data sequences. Therefore these are ideal tools to improve the

estimation of the maximum area and especially to improve the quality of indirect spatial series curves. Our experience with these smoothers shows that the 3RSSH and 4253H smoothers are moderately effective in smoothing. They do not eliminate the
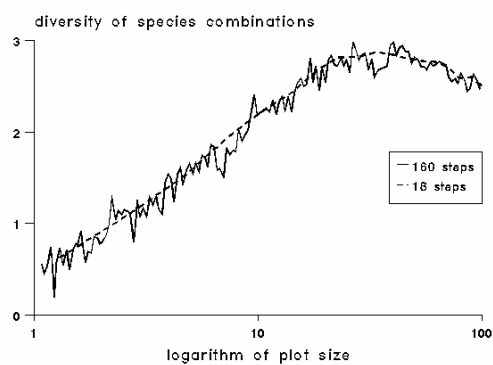

Figure 4. Spatial series analysis of a 4-species community using the diversity of species combinations as a spatial statistics. One of the spatial series curve is produced by large steps while the other is calculated by small steps between the plot sizes. The sample size was 1000 plots for each plot size.

oscillations as heavily as the LOWE techniques. Therefore, these can be used to eliminate the heavy peaks and valleys and still there are some mild oscillations in the series. These techniques proved not very effective in reducing the variance of the estimation of the maximum area (Tóthmérész,

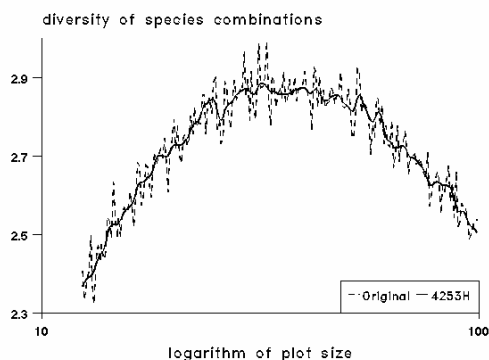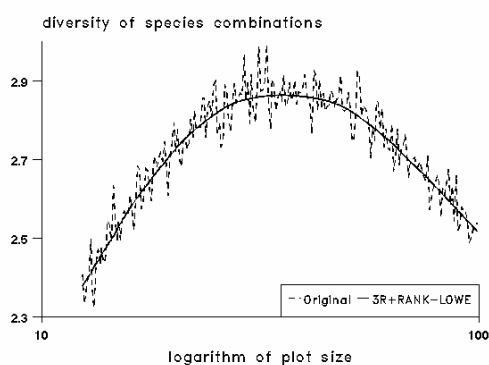1994b). There was almost no effect of twice-ing in the studied cases.

Figure 5. Spatial series analysis of a 4-species community using the diversity of species combinations as a spatial statistics. The sample size was 100 plots for each plot size. The spatial series curve is smoothed by the 4253H smoother.

Figure 6. Spatial series analysis of a 4-species community using the diversity of species combinations as a spatial statistics. The sample size was 100 plots for each plot size. The spatial series curve is smoothed by the 3R+RANK-LOWE smoother.

A dramatic improvement was produced by the LOWE and RANK-LOWE techniques. The smoothed curves are without oscillation and the variance of the estimation of the maximum area are heavily reduced even for relatively small number of sample plots. In the case of LOWE the computing time is much longer and this is usually less effective in reducing the variance of the estimation than the RANK-LOWE. Therefore, the favourite choise is probably the RANK-LOWE smoother. The 3+RANK-LOWE and 3R+RANK-LOWE techniques eliminate the extremely high peaks and very low valleys by 3 and 3R smoothing and than they perform a RANK-LOWE smoothing. These smoothers are designed especially robust against "peculiar" values and also to produce very smooth spatial series curves without oscillations. These techniques are designed to incorporate the beneficial features of both the running median smoothers and the RANK-LOWE.

References

Bartha, S. (1992): Preliminary scaling for multi-species coalitions in primary succession. - Abstracta Botanica 16, 31-41.
Cleweland, W.S. (1978): The Elements of Graphing Data. - Wadsworth Publ., Monterey, California.
Cleveland, W.S. (1979): Robust locally weighted regression and smoothing scatterplots. - J. Am. Stat. Ass. 74, 829-836.
Cleveland, W.S. (1981): LOWESS: A program for smoothing scatterplots by robust locally weighted regression. - The American Statistician 35, 54.
Erdei, Zs. and Tóthmérész, B. (1993): MULTI-PATTERN 1.00. A program package to analyze and simulate community-wide patterns. - Tiscia 27, 45-48.
Juhász-Nagy, P. (1967): On associations among plant populations I. Multiple and partial associations: a new approach. - Acta Biol. Debrec. 5, 43-56.
Juhász-Nagy, P. (1976): Spatial dependence of plant populations. Part 1. Equivalence analysis (An outline for a new model). - Acta Bot. Hung. 22, 61-78.
Juhász-Nagy, P. (1984): Spatial dependence of plant populations. Part 2. A family of new models. - Acta Bot. Hung. 30, 363-402.
Juhász-Nagy, P. (1993): Notes on compositional diversity. - Hydrobiologia 249, 173-182.
Juhász-Nagy, P. and Podani, J. (1983): Information theory methods for the study of spatial processes and succession. - Vegetatio 51, 129-140.
McNeil, D. R. (1977): Interactive Data Analysis. - Wiley, New York.
Podani, J., Bartha, S. and Czárán, T. (1993): Pattern, area and diversity: the importance of spatial scale in species assemblages. - Abstracta Botanica 17, 37-51.
Tóthmérész, B. (1994a): Diversity ordering and spatial series analyses. - Manuscript of DSC dissertation, pp. 196, Debrecen.
Tóthmérész, B. (1994b): Statistical analysis of spatial pattern in plant communities. - Coenoses 9, 33-41.
Tóthmérész, B. (1994c): Robust nonlinear smoothers in spatial series analysis. - Abstracta Botanica (in print).
Tóthmérész, B. and Erdei, Zs. (1992): The effect of species dominance on information theory characteristics of plant communities. - Abstracta Botanica 16, 43-47.
Tukey, J. W. (1977): Exploratory Data Analysis. - Addison-Wesley, Reading, Mass.
Velleman, P. F. and Hoaglin, D. C. (1981): Applications, Basics, and Computing of Exploratory Analysis. - Duxbury Press, Boston.