Combining the data from individual mapped replicate plots into mean, weighted *O*(*r*), *K*(*r*), and *g*(*r*) functions

Background and formulas

For statistical analysis it is common to map several replicate plots of a larger point pattern under identical conditions. In this case the resulting test statistics of the individual replicate plots can be combined into average test statistics (Diggle 2003: page 123; Illian et al. 2008: page 263). This is of particular interest if the number of points in each replicate plot is relatively low. In this case the confidence limits of individual analyses would become wide, but combining the data of several replicate plots into average test statistics increases the sample size and thus narrows the simulation envelopes. Average test statistics are also an effective way of summarizing the results of several replicate plots.

When the patterns are strict replicates of an underlying process, the corresponding estimates $\hat{K}_i(r)$ of the *K*-functions from plots *i* are identically distributed and a reasonable overall estimate can be obtained by simply averaging the individual *K*-functions (Diggle 2003: equation 4.20 at page 52). Using the grid-based estimator of Programita, the resulting estimator of the *O*-ring statistic is given in the appendix of Riginos et al. (2005).

However, because the pair-correlation function *g*(*r*) and Ripley's *K*-function *K*(*r*) are defined as a ratio of expected number of points in rings or circles [= $\lambda g(r)$ and = $\lambda K(r)$] divided by the intensity λ, a better strategy may be to pool separately estimates of λ, $O(r) = g(r)$, and $\lambda K(r)$. The resulting average function are weighted averages of the individual estimates, where the weight is the number of points in plot *i* divided by the total number of points in all replicate plots (Diggle 2003: equation 8.11, page 123). Note that the resulting average pair-correlation and *K*-function are also appropriate estimators if the replicates would be differentially thinned versions of a common underlying process (Diggle 2003: page 123).

Using the grid-based estimators of Programita and following the notation in Wiegand and Moloney (2004) (equation 11), the numerical estimator of the bivariate *O*-ring statistic $O_{12}(r) = \lambda_2 g(r)$ is calculated as:

$$\hat{O}_{12}^{w}(r) = \frac{\frac{1}{n_1}\sum_{i=1}^{n_1}\mathbf{Points}_2[R_{1,i}^{w}(r)]}{\frac{1}{n_1}\sum_{i=1}^{n_1}\mathbf{Area}[R_{1,i}^{w}(r)]} \qquad \text{(A1)}$$

where n_1 is the number of points of pattern 1, $R_{1,i}^{w}(r)$ is the ring with radius *r* and width *w* centered in the *i*th point of pattern 1, $\mathbf{Points}_2$[X] counts the points of pattern 2 in a region X, and the operator **Area**[X] determines the area of the region X.

To integrate the data of *N* different replicates into a single weighted *O*-ring statistic, the formula for one replicate (eq. A1) is extended by calculating, for each spatial scale *r*, the average weighted number of points of pattern 2 taken over all *N* replicates and the average weighted area taken over all N replicates:

$$\hat{O}_{12}^{w}(r) = \frac{\left(\frac{n_1^1}{N}\left(\frac{1}{n_1^1}\sum_{i^1=1}^{n_1^1}\mathbf{Points}_2[R_{1,i^1}^{w}(r)]\right) + \ldots + \frac{n_1^N}{N}\left(\frac{1}{n_1^N}\sum_{i^N=1}^{n_1^N}\mathbf{Points}_2[R_{1,i^N}^{w}(r)]\right)\right)}{\left(\frac{n_1^1}{N}\left(\frac{1}{n_1^1}\sum_{i^1=1}^{n_1^1}\mathbf{Area}[R_{1,i^1}^{w}(r)]\right) + \ldots + \frac{n_1^N}{N}\left(\frac{1}{n_1^N}\sum_{i_1^N=1}^{n_1^N}\mathbf{Area}[R_{1,i^N}^{w}(r)]\right)\right)} \qquad \text{(A2)}$$

where i^j is the ith point of pattern 1 and replicate j, n_1^j is the number of points of pattern 1 and replicate j, and $N = \sum_j n_1^j$ is the total number of points of pattern 1 in all replicates. Equation A2 simplifies to:

$$\hat{O}^w_{12}(r) = \frac{\sum_{i^1=1}^{n_1^1} \mathbf{Points}_2[R^w_{1,i^1}(r)] + \ldots + \sum_{i^N=1}^{n_1^N} \mathbf{Points}_2[R^w_{1,i^N}(r)]}{\sum_{i^1=1}^{n_1^1} \mathbf{Area}[R^w_{1,i^1}(r)]) + \ldots + \sum_{i_1^N=1}^{n_1^N} \mathbf{Area}[R^w_{1,i^N}(r)]} \qquad \text{(A3)}$$

Following the notation in Wiegand and Moloney (2004) (equation 6), the numerical estimator of the bivariate K-function is given as:

$$\lambda_2 \hat{K}_{12}(r) = \pi r^2 \frac{\frac{1}{n_1}\sum_{i=1}^{n_1} \mathbf{Points}_2[C_{1,i}(r)]}{\frac{1}{n_1}\sum_{i=1}^{n_1} \mathbf{Area}[C_{1,i}(r)]} \qquad \text{(A4)}$$

where n_1 the total number of points of pattern 1, $C_{1,i}(r)$ is the circle with radius r centered on the ith point of pattern 1, the operator $\mathbf{Points}_2$[X] counts the points of pattern 2 in a region X, and the operator **Area**[X] determines the area of the region X. In analogue to equations A2 and A3, the combined estimator of $\lambda_2 \hat{K}_{12}(r)$ yields:

$$\lambda_2 \hat{K}_{12}(r) = \frac{\sum_{i^1=1}^{n_1^1} \mathbf{Points}_2[C_{1,i^1}(r)] + \ldots + \sum_{i^N=1}^{n_1^N} \mathbf{Points}_2[C_{1,i^N}(r)]}{\sum_{i^1=1}^{n_1^1} \mathbf{Area}[C_{1,i^1}(r)]) + \ldots + \sum_{i_1^N=1}^{n_1^N} \mathbf{Area}[C_{1,i^N}(r)]} \qquad \text{(A5)}$$

Following the strategy of Diggle (2003) to pool separately estimates of λ_2 and $\lambda_2 K_{12}(r)$ [and analogously estimates of λ_2 and $\lambda_2 O_{12}(r)$ for estimating the pair-correlation function] the overall intensity λ_2 is estimated as $\lambda_2 = N_2/A$ where

$$N_2 = \sum_{j=1}^{N} n_2^j \qquad \text{(A6)}$$

is the total number of points of pattern 2 in all N replicates j, and A the total area

$$A = \sum_{j=1}^{N} A^j \qquad \text{(A7)}$$

of all replicates j with area A^j. The estimator of the L-function is then calculated analogously to equation 10 in Wiegand and Moloney (2004), and the estimator of the pair-correlation function yields $\hat{g}^w_{12}(r) = \hat{O}^w_{12}(r)/\lambda_2$.

The univariate estimators of $O(r)$, $L(r)$, and $g(r)$ are calculated in a manner analogous to the bivariate functions by setting pattern 1 equal to pattern 2. In this case, however, the focal points of the circles are not counted. Thus, equation A6 is replaced by

$$N_2 = \sum_{j=1}^{N} (n_2^j - 1)\,. \tag{A8}$$

The accumulated distribution functions of the distances to the nearest neighbor of plots m [= $D^m{}_{12}(y)$] can be aggregated into a mean weighted distribution function by

$$\hat{D}_{12}(y) = \frac{\sum_{m=1}^{M} n_1^m D_{12}^m(y)}{\sum_{m=1}^{M} n_1^m} \tag{A9}$$

where $n^m{}_1$ is the number of points of pattern 1 and plot m (Illian et al. 2008: page 261).

Literature cited

Diggle, P.J. 2003. Statistical analysis of point patterns. Second edition. Arnold, London.

Illian, J., A. Penttinen, H. Stoyan, and D. Stoyan. 2008. Statistical Analysis and Modelling of Spatial Point Patterns. Wiley, Chichester.

Riginos, C., S. J. Milton, and T. Wiegand. 2005. Context-dependent negative and positive interactions between adult shrubs and seedlings in a semi-arid shrubland. *Journal of Vegetation Science* 16:331-340.

Wiegand T.,and K. A. Moloney 2004. Rings, circles and null-models for point pattern analysis in ecology. *Oikos* 104: 209-229.

Instruction for combine replicate feature in Programita

Before running an individual analysis enable the checkbox "Calculate replicates" (Figure 1), then run all analyses with replicate plots of the same treatment with the same settings (this is important!), i.e., do not change the maximal scale analyzed or the grid size.

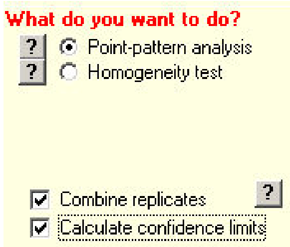
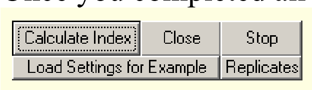


Figure 1

Programita creates specific results files containing the information necessary for combining the replicates. If your data file was named "name.dat", the corresponding results file will be called "WM_name.rep" if you select the method "Ring" (i.e., O-ring statistic and pair-correlation function) and "R_name.rep" if you select the method "Circle" (i.e., Ripley's *L*-function). In case of using the random labeling null model, more information is needed since e.g., g_{12} as well as g_{21} has to be calculated. In this case the corresponding results files are named "WM_name_1.rep" and "WM_name_2.rep".

Once you completed all analyses press the button "Replicates" below the "stop" button:



A window opens where you can select the files you like to combine (Figure 2). Highlight a file and press the button "select". Once you selected all files (Figure 2), press "Calculate joined statistic" and the result of the combined analysis appears (Figure 3).

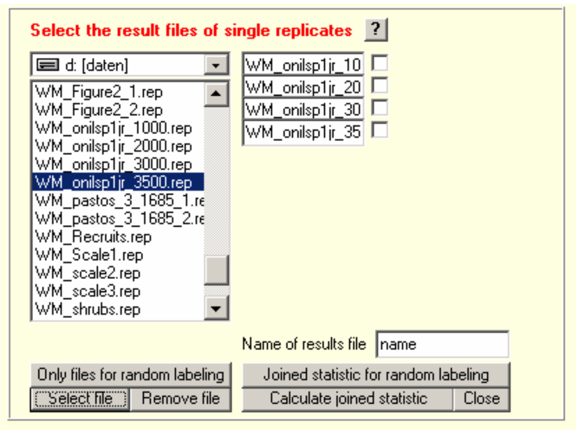


Figure 2. A list with result of replicate analyses.

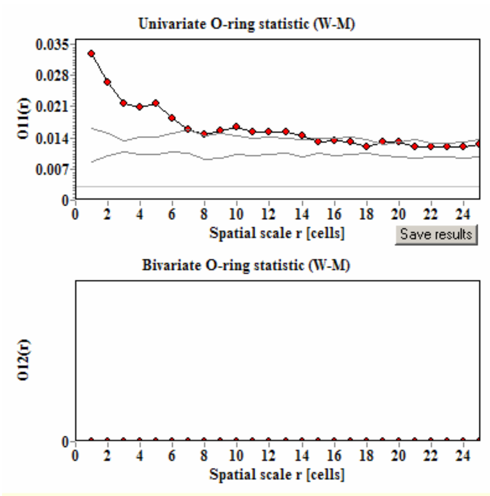
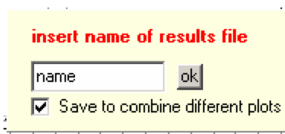


Figure 3. The results of the combined analyses.

You can save the results using the "Save results button". Insert a name. If "name" stands for the selected name, the results file "name.res" gives you the mean weighted $O(r)$ and $g(r)$ functions (if you selected the method "Ring") and the mean weighted $L(r)$ functions (if you selected the method "Circle").



Additionally, *Programita* saves the *.env results files
Uni_g_name.env, (or Uni_L_name.env,)
Uni_NN_name.env,
Bi_g_name.env (or Bi_L_name.env), and
Bi_NN_name.env
that contain the values of the test statistics for the data and all simulations of the null model.

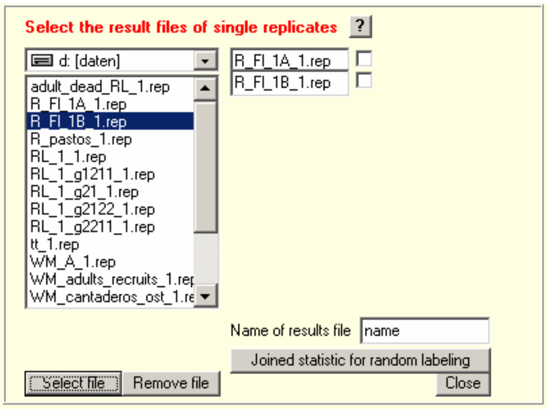
For method "Ring" the output in the *.res is:

`O11(r)`	`E11-`	`E11+:`		univariate *O*-ring statistic and confidence limits
`O12(r)`	`E12-`	`E12+:`		bivariate *O*-ring statistic and confidence limits
`g11min`	`g11max`	`g12min`	`g12max:`	minimal and maximal values of pair-correlation functions taken from *N* replicate plots
`g11(r)`	`E11-`	`E11+:`		univariate pair-correlation function
`g12(r)`	`E12-`	`E12+:`		bivariate pair-correlation function

For method "Circle" the output is:

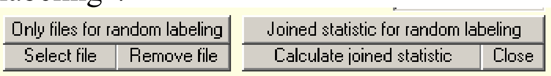
`L11(r)`	`E11-`	`E11+:`		univariate *L*-function and confidence limits
`L12(r)`	`E12-`	`E12+:`		bivariate *L*-function and confidence limits
`L11min`	`L11max`	`L12min`	`L12max:`	minimal and maximal values of **L**-functions taken from *N* replicate plots

Random labeling



If the null model used in your analyses was random labeling, *Programita* saved two results files per analyses. To simplify selection of results files and to tell Programita that you will combine replicates that used random labeling, you need to proceed as following.

To open the window for selection of replicate analyses data press as before the button "Replicates" below the "Stop" button. However, to indicate that you used random labeling press the "Only files for random labeling":



Now select the replicates as before, but press "Joined statistic for random labeling".

Remove files from the list

If you want to remove a replicate already selected (e.g., if one replicate appeared accidentally twice as in the example) enable the check bock on the right and press “Remove file”.