

SOA

- spatial overlap assessor -

Manual

Software:

SOA (spatial overlap assessor) is written in PERL and requires the Tk and Tk::BrowseEntry modules which are available at <http://www.cpan.org/>. Executing PERL scripts requires installation of a PERL distribution which is preinstalled on Linux and Macintosh systems. Windows users must download and install a PERL distribution manually. Free PERL distributions are available at <http://www.perl.org/get.html>. A standalone executable file which runs on windows systems without any further requirements is available upon request. Anyone can use the source codes, documents or the executable file of SOA free of charge for non-commercial use.

Scope:

SOA is a software to evaluate spatial overlap of two groups (1/2) of individuals based on pairwise spatial distances. For each group, SOA calculates the average number of individuals assigned to group 1/2 within a defined number

of nearest neighbors. The resulting values indicate the degree of spatial overlap of the two different groups. In order to assess and eliminate the effect of two groups of individuals comprising different sample size, SOA allows to generate resampled datasets in which the number of individuals of the larger group is adjusted to the number of individuals of the smaller sized group by random selection of an appropriate number of individuals.

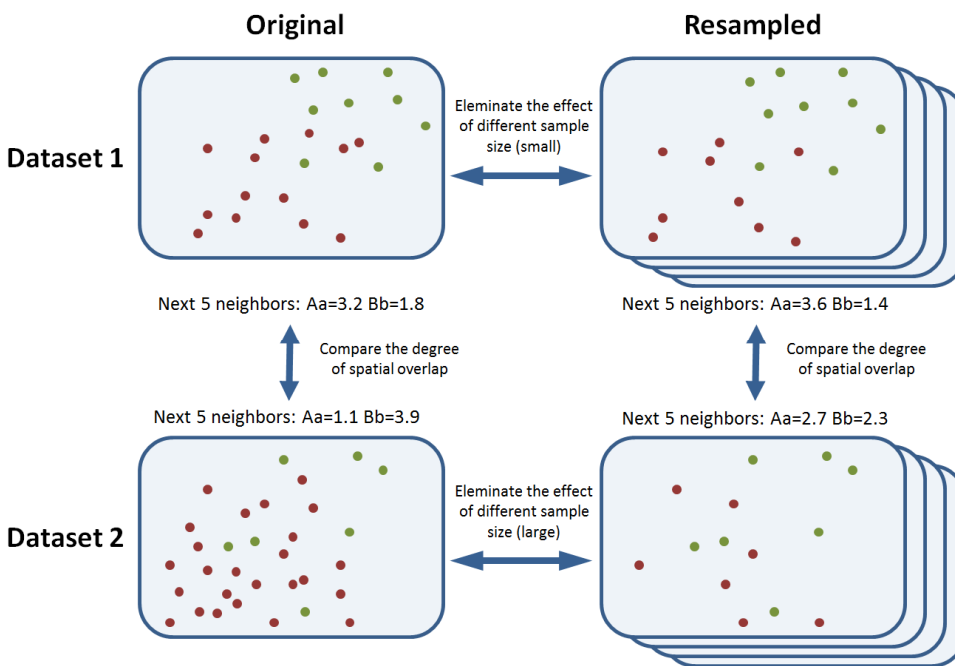


Fig. 1: Methodological overview. Spatial distribution of individuals of two different groups (group 1: green, group 2: red) at two different sites (dataset 1 and dataset 2).

Preparing your input file:

1. Build a table of pairwise spatial distances including individuals of the two groups (Aa and Bb). The table has to start with the individuals of the smaller sized group, followed by the individuals of the larger sized group. Distances must be four characters in length and must be followed by the two letter group name abbreviation. Values on the diagonal must be negative (e.g. -1.00)

	Aa 1	Aa 2	Aa 3	Bb 1	Bb 2	Bb 3	Bb 4	Bb 5
Aa 1	-1.00Aa	1.67Aa	0.37Aa	1.56Aa	1.77Aa	0.09Aa	1.01Aa	0.12Aa
Aa 2	1.67Aa	-1.00Aa	0.25Aa	1.82Aa	0.54Aa	1.67Aa	1.68Aa	0.39Aa
Aa 3	0.37Aa	0.25Aa	-1.00Aa	0.87Aa	1.67Aa	0.32Aa	1.09Aa	0.21Aa
Bb 1	1.56Bb	1.82Bb	0.87Bb	-1.00Bb	0.78Bb	1.76Bb	0.34Bb	1.71Bb
Bb 2	1.77Bb	0.54Bb	1.67Bb	0.78Bb	-1.00Bb	0.90Bb	1.51Bb	0.94Bb
Bb 3	0.09Bb	1.67Bb	0.32Bb	1.76Bb	0.90Bb	-1.00Bb	0.09Bb	1.82Bb
Bb 4	1.01Bb	1.68Bb	1.09Bb	0.34Bb	1.51Bb	0.09Bb	-1.00Bb	0.48Bb
Bb 5	0.12Bb	0.39Bb	0.21Bb	1.71Bb	0.94Bb	1.82Bb	0.48Bb	-1.00Bb

Tab. 1: Table of pairwise spatial distances. To prepare your input file, delete the red marked part of the table.

2. Delete everything but the green marked part from the table with the result that the top row of the table will include only individuals of the smaller sized group.
3. Save the table as tab delimited table in a text file without any blank lines.

Running the program:

After selecting your input file, you have to choose the number of nearest neighbors of each individual of the first group (which is the smaller sized group if sample size is different) to be analyzed.

If the sample size of the two groups is different, you can generate an arbitrary number of resampled datasets (see Scope) which will be analyzed and can be compared to the original dataset.

Note, that if you chose to avoid redundant resampled datasets, the maximum number of possible resampled datasets is limited. SOA will return an error message if you inadvertently chose to perform more resampling steps than possible.

The results can be found in one output file which will be saved in the current working directory.

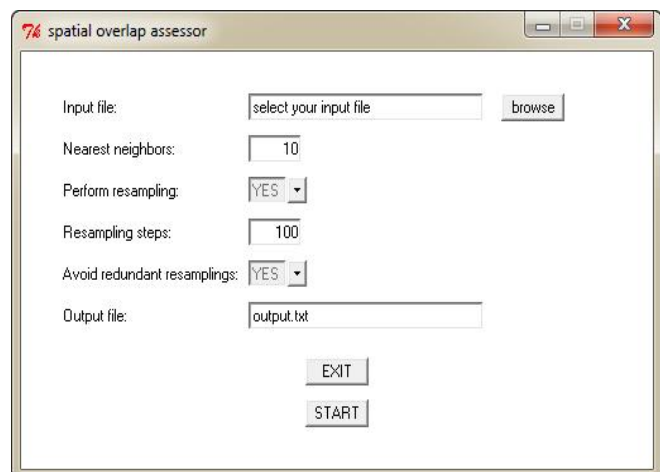


Fig. 2: SOA graphical user interface.

Output file:

If you do not perform any resampling steps, the output file will look like this:

```
Input file: C:/example.txt
Individuals in group Aa: 3
Individuals in group Bb: 51
Number of next neighbors to be analyzed: 5

Results individual 1:   Bb: 5   Aa: 0   Average distance: 0.154   Highest distance: 0.20
Results individual 2:   Bb: 4   Aa: 1   Average distance: 0.082   Highest distance: 0.11
Results individual 3:   Bb: 4   Aa: 1   Average distance: 0.102   Highest distance: 0.13
Average results original dataset: Bb: 4.33 Aa: 0.67 Total average distance: 0.11 Highest distance average: 0.15
```

For each individual of group Aa, the five next neighbors were analyzed and the average number of neighbor individuals per group is calculated for all individuals of group Aa. In this example, 4.33 of the five nearest neighbors of each individual from group Aa belong to group Bb. The average distance to any of the five nearest neighbors is 0.11 and the average distance to the fifth nearest neighbor is 0.15.

If you choose to perform resampling steps, the output file will look like this:

```
Input file: C:/example.txt
Individuals in group Aa: 3
Individuals in group Bb: 51
Number of next neighbors to be analyzed: 5
Resampling steps: 3
Avoid redundant resamplings: YES

Results individual 1:   Bb: 5   Aa: 0   Average distance: 0.154   Highest distance: 0.20
Results individual 2:   Bb: 4   Aa: 1   Average distance: 0.082   Highest distance: 0.11
Results individual 3:   Bb: 4   Aa: 1   Average distance: 0.102   Highest distance: 0.13
Average results original dataset: Bb: 4.33 Aa: 0.67 Total average distance: 0.11 Highest distance average: 0.15

resampling for individual 1
resampling 1:   Bb: 3   Aa: 2   Average distance: 0.402   Highest distance: 0.45
resampling 2:   Bb: 3   Aa: 2   Average distance: 0.446   Highest distance: 0.66
resampling 3:   Bb: 3   Aa: 2   Average distance: 0.49   Highest distance: 0.88
Average for resampling individual 1: Bb: 3Aa: 2 Average distance: 0.45 Highest distance (average): 0.66

resampling for individual 2
resampling 1:   Bb: 3   Aa: 2   Average distance: 0.284   Highest distance: 0.60
resampling 2:   Bb: 3   Aa: 2   Average distance: 0.468   Highest distance: 1.28
resampling 3:   Bb: 3   Aa: 2   Average distance: 0.444   Highest distance: 0.75
Average for resampling individual 2: Bb: 3Aa: 2 Average distance: 0.40 Highest distance (average): 0.88

resampling for individual 3
resampling 1:   Bb: 3   Aa: 2   Average distance: 0.286   Highest distance: 0.48
resampling 2:   Bb: 3   Aa: 2   Average distance: 0.186   Highest distance: 0.37
resampling 3:   Bb: 3   Aa: 2   Average distance: 0.316   Highest distance: 0.54
Average for resampling individual 3: Bb: 3Aa: 2 Average distance: 0.26 Highest distance (average): 0.46

Total average for resampling 3 individuals: Bb: 3 Aa: 2 Average distance: 0.37 Highest distance (average): 0.67
```

The first lines are similar to the first example but hold some additional information which is the number of resampling steps and whether redundant resampled datasets were rejected or not. The following blocks contain the results for each individual of group 1 and resampled dataset. The last line of each block contains the respective average values (Note: Resampling is performed separately for every individual, which means that resampling 1 for different individuals does not refer to the same resampled dataset). The last line of the output file contains the average values for all blocks with resampled results taken together.

Contact:

Please address any questions concerning this software to:

David Rosenkranz

rosenkrd@uni-mainz.de