**Instructions for the RTEI program**

**(Randomization Tests for assessing Edge Influence)**

**Harper, Karen A. and Macdonald, S. Ellen**

The RTEI program is an updated version of the Critical Values Approach program (Harper and Macdonald 2002) that was used for Harper and Macdonald (2011). The main difference is that RTEI includes edge and reference values in the randomization whereas the Critical Values Approach only included the reference values. RTEI for a single edge type is described in the accompanying article that introduces this program (Harper and Macdonald 201x). We plan to expand this program to include comparisons among edge types later. Here we present the instructions on how to use the program for a single edge type. If you have any problems, questions, comments or suggestions about this program, please e-mail Karen.Harper@dal.ca or Ellen.Macdonald@ualberta.ca.

General description

The RTEI program (RTEI.xla) is an Excel macro written in Visual Basic. The RTEI program calculates means at different distances from the edge and in the reference forest, as well as the magnitude of edge influence (MEI) for different distances from the edge. MEI for each distance from the edge is tested for significance using randomization tests. Distance of edge influence (DEI) can then be determined based on these results (see Output). Any number of variables can be tested at one time, as long as the number of transects and the distances are the same.

Sampling design

RTEI was designed for a sampling design that consists of edge treatment data collected at sampling points at different distances from the edge along transects established perpendicular to edges and reference data collected within a reference forest. If the edge treatment data were collected at various distances from the edge but not along transects, they could be grouped into distance classes and assigned to transects (the layout of transects does not affect the analyses if the reference data are not blocked). The reference data can be blocked which means the sampling points are paired with the edge treatment transects. This is most commonly done by having the reference data collected along the same transects as for the edge treatment but at distances far enough from the edge to be considered the reference forest. Alternatively, there could be no blocking of the reference data.

Data entry

Data should be entered on two Excel worksheets labelled “Edge” and “Reference”. The program will only work on data entered into worksheets with these names.

*Edge data*. The edge data should be input in the worksheet “Edge” as follows: column A - names of the response variables starting at A4, column B - transect #s or names starting at B4, row 2 - distances from the edge starting at D2, cells D4+ should contain the data at different distances from the edge. There should be one value per transect for each distance from the edge. Subsamples at each sampling location can be averaged. If there are species composition data, we suggest performing an ordination and then entering the ordination scores along the first two axes as variables. Enter zeros where appropriate and do not leave blank cells. See below for Missing values.

*Reference data.* The reference data should be input in the worksheet “Reference”.

If there is no blocking, the data should be input as follows: column A – names of the response variables starting at A4, row 2 – sampling location numbers or names starting at C2, cells C4+ should contain the data at the different reference sampling locations. Unlike the edge data and the reference data with blocking, there should be one row of data per variable.

If there is blocking, the data should be input as follows: column A – names of the response variables starting at A4, column B – transect #s or names starting at B4, row 2 – sampling location numbers or names (which could be distances from the edge) starting at D2, cells D4+ should contain the data at the different reference sampling locations along each transect. With blocking, the first two columns should be identical for the edge and the reference data worksheets; however, the number of plots will most likely be different.

There should be one value at each sampling location; subsamples can be averaged. If there are species composition data, we suggest performing an ordination and then entering the ordination scores along the first two axes as variables. Enter zeros where appropriate and do not leave blank cells. See below for Missing values. The variance in the reference data set must not be zero (i.e., all values cannot be the same) for a given response variable.

*Missing values.* If you have missing values, enter ‘M’. The program will simply skip this value (e.g., the edge data from 6 transects at a given distance would be considered rather than 7 if one value was ‘M’ for 7 transects). However, missing values are not allowed for the following situations:

* in edge data for a sampling design with blocking,
* for all reference values along one transect with blocking or
* in edge or reference data for exact permutation (see below) with blocking.

In these situations we recommend using the average of the other edge data at a given distance or the average of the other reference sampling locations along the same transect. An alternative is to perform a separate analysis with only the transects that do not have missing data.

Program execution

Select RTEI from the menu and then ‘single edge type’. Input the relevant information. See Sampling design, above, for a description of edge transects, reference plots and blocking. The recommended minimum number of permutations for randomization tests is 1000 and 5000 at significance levels of 5% and 1%, respectively (Crowley 1992). See below for Exact permutation. It may take a while. You can figure out how long it will take by running the program for one variable and then multiplying the number of variables. If you want to run the program again, you must delete or rename the "results" worksheets first. The program will only work if the input worksheets have not been renamed and if there are no "results" worksheets.

*Exact permutation.* Exact permutation is performing a randomization test with all possible permutations. We recommend choosing this option when the number of possible permutations (due to smaller sample sizes) is less than the number of permutations desired (usually less than 5000). The number of permutations with no blocking is (Rt + T)!/(Rt!T!) where Rt = the total number of reference sampling points and T = the number of edge transects. The number of permutations with blocking is (1+R)T where R = the number of reference sampling points per transect and T = the number of transects. There must be 3-7 edge transects, and at least 3 reference sampling points per transect with blocking or at least 5 total reference sampling points with no blocking, otherwise the program will not work.

Output

The output is presented on the "Results" worksheet and includes the following for each variable: 1) mean values (across transects) for each distance from the edge and for the reference data set and 2) MEI = (e-r)/(e+r), where e = average at a given distance from the edge and r = average of the reference data set, for each distance from the edge. P-values for testing the significance of the averages and MEIs (the p-values are the same for averages and MEIs) are given in the row below these statistics. P-values that are statistically significant for negative edge influence are below the chosen significance level (in %) for a one-tailed test or below half this level for a two-tailed test. P-values that are statistically significant for positive edge influence are above 1 – the chosen significance level (in %) for a one-tailed test or below 1 – half this level for a two-tailed test. Statistically significant refers to averages that are significantly lower or higher near the edge as compared to the reference data set. Because RTEI usually involves testing more than one variable at more than one distance from the edge, researchers are advised to take multiple testing into account when interpreting significance. One option is the sequential Bonferroni test (Rice 1989).

Acknowledgements

Thank-you to Don Coady, anonymous advice from VisualBasic Excel experts on the internet and to students who tried our program.

References

Crowley, P.H. 1992. Resampling methods for computation-intensive data analysis in ecology and evolution. Annual Review of Ecology and Systematics. 23: 405-447.

Harper, K.A. and S.E. Macdonald. 2002. The critical values program for assessing edge influence. Bulletin of the Ecological Society of America. 83(1): 61-62.

Harper, K.A. and S.E. Macdonald. 2011. Quantifying distance of edge influence: A comparison of methods and a new randomization method. Ecosphere. 2: article 94.

Rice, W.R. 1989. Analyzing tables of statistical tests. Evolution. 43: 223-225.