MoPeT: Software for modified permutation test of significance of mean Ellenberg values

Introduction

This program enables to calculate modified test of significance for analysis of mean Ellenberg indicator values (EIVs) with other variables derived from species composition. The test of significance is based on modified permutation schema (see more deatails here).

Download

MoPeT_v1.2.zip (last update: 24.10.2013)

How to use it

MoPeT is based on R program - you need to have R installed on your computer, but you don't need any knowledge of R program. All data can be loaded from spreadsheet format. Optionally, you can run the program also from JUICE software (http://www.sci.muni.cz/botany/juice/) - in this case you need to have both JUICE and R installed on your computer. JUICE is a free program for editing vegetation data, developed by Lubomír Tichý from Masaryk University Brno, Czech Republic.

Running the program with data in spreadsheet format

Installation

- install R program (if not yet installed on your computer) from http://www.R-project.org (the most recent version of R program can be downloaded here)
- download the MoPeT.zip file from the download link above
- unzip the MoPeT.zip file into the R folder, where `R.exe` file is located - typically it's in `c:\Program Files\R\R-2.X.X\bin` folder[1)], where `R-2.XX.X` is installed version of R (e.g. R-2.14.1)

Launch the program

- launch the `MoPeT_vX.X.bat` executable file (vX.X in the name is the current version of the script you are using) - the interface of the MoPeT application should open.

Running the program from JUICE

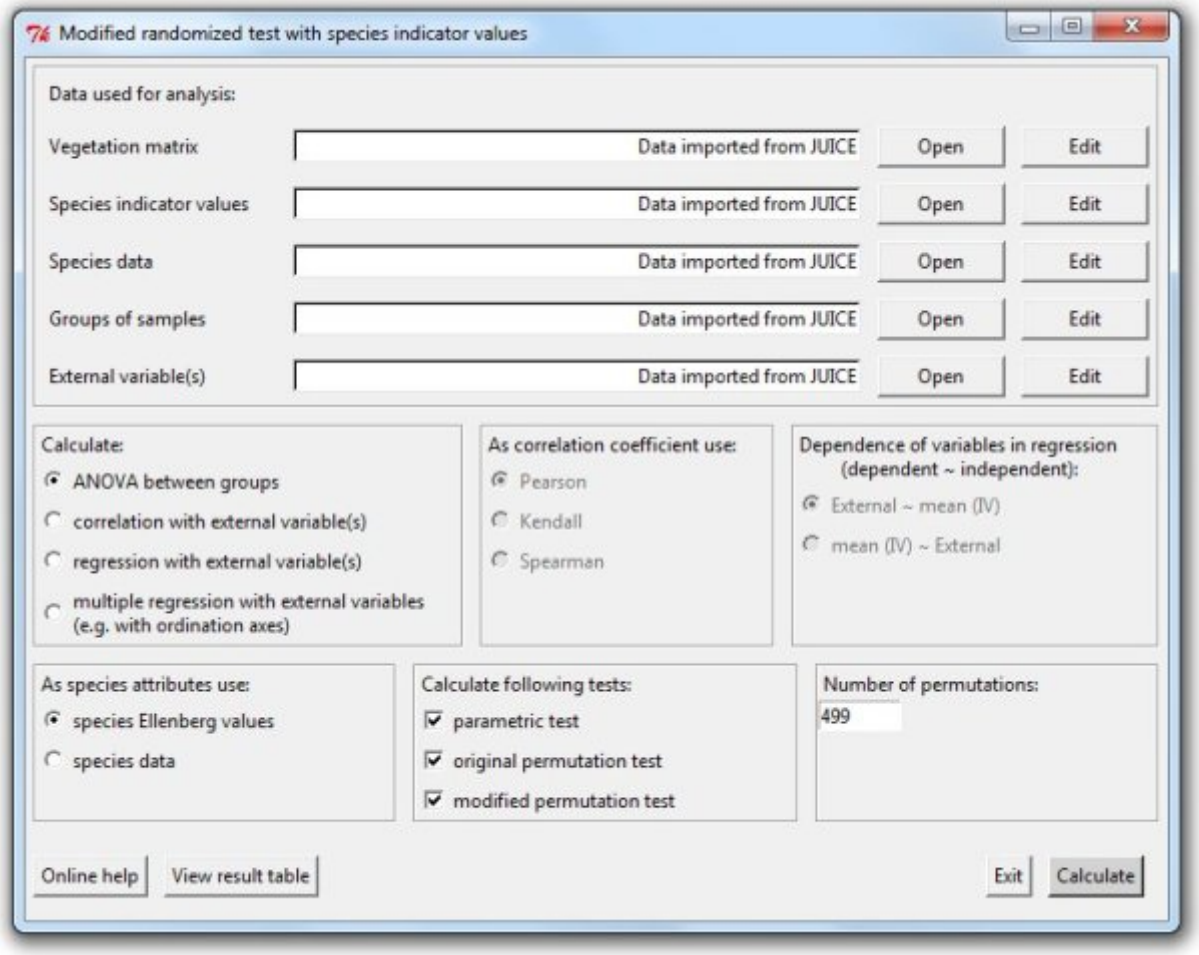
Installation

- install R program (if not yet installed on your computer) from http://www.R-project.org (the most recent version of R program can be downloaded here)
- install JUICE program (if not yet installed) from http://www.sci.muni.cz/botany/juice/?idm=3 - you need JUICE version 7.0.64 or higher (if you have JUICE already installed on your computer, check the version -

in JUICE menu *Help* > *Program Info*, and update if necessary)
- allow JUICE program to use R program - in JUICE menu go to *File* > *Options* > *External Program Paths*, and in the *R-PROJECT* row choose the path to Rgui.exe file (typically in `C:\Program Files\R\R-x.xx.x\bin` or `bin\x64` folder)
- download the MoPeT.zip file from the download link above and unzip it anywhere on your computer

Launch the program

- open JUICE, import the vegetation data (see JUICE manual for help), initiate Ellenberg indicator values (menu *Indicator values* > *Initiation* > *Initiation of indicator values* (and open file ELLENB.txt located in JUICE folder in Program Files)
- in JUICE menu go to *Analysis* > *JUICE-R function* (or **simply click on CTRL+W**) - JUICE-R interface will open
- choose the color of *Relevés Used in Analysis* (use All if you did not set any) and click the button *Append R script*
- find the MoPeT_vX.X.r file on your computer (**NOT** MoPeT_vX.X.bat file) and open it
- click the *Run* button to export data from JUICE and launch the R application.



Available functions

ANOVA between groups

Use this function if you want to test the differences between mean EIVs for groups of samples, assembled together e.g. by cluster analysis. One-way ANOVA is calculated, returning p-values based on modified

permutation test.

Data needed

- species composition matrix
- species indicator values
- groups of samples

Correlation with external variables

Use this function to calculate significance of correlation of mean EIVs with external variable, which is either derived from species composition (e.g. species richness of samples, scores of samples along ordination axes etc.) or is measured environmental variable. Modified permutation test is based on comparing real correlation coefficient with null distribution of correlation coefficients, generated by modified permutation of mean EIVs. Pearson, Kendall and Spearmann correlation coefficients are available.

Data needed

- species composition matrix
- species indicator values
- external variable

Regression with external variables

Calculates regression of mean EIVs with external variable, which is either derived from species composition (e.g. species richness of samples, scores of samples along ordination axes etc.) or is measured environmental variable. Regression can be calculated either with mean EIVs as dependent or independent variable, modifying the setting of `Dependence of variables in regression (dependent ~ independent)`. Modified permutation test is based on comparing real *F*-value of the regression with null distribution of *F*-values generated by modified permutation of mean EIVs.

Data needed

- species composition matrix
- species indicator values
- external variable

Multiple regression with external variables (e.g. with ordination axes)

This option is suitable, if you aim to calculate relationship of mean EIVs with ordination axes from unconstrained ordination analysis. It is based on an algorithm used in `vegan` package in R, written by Jari Oksanen, which calculates relationship between ordination axes and mean EIVs using multiple regression, where mean EIVs is dependent variable and sample scores on ordination axes (1, 2 or more) are independent variables. Advantage of this analysis, compared to calculating correlation of mean EIVs with each axis separately, is lower number of tests (in correlation, number of tests = number of mean EIVs × number of ordination axes, in multiple regression number of tests = number of mean EIVs). Significance is based on F-values as testing criteria, and is calculated by comparing multiple regression F-values of real mean EIVs with null distribution of F-values generated by modified permutations. Importance of individual mean EIVs can be judged by comparing their R^2. If you want to use absolute values of R^2 or R^2_{adj}, note that these are biased (overestimated) due to the similarity

issue between mean EIVs and ordination axes (if external variables are ordination axes); not-biased values are represented by variable `R2.adj.iv` (but the theory at the background is not yet elaborated, so the use of these values is not recommended).

Data needed

- species composition matrix
- species indicator values
- external variable(s) (e.g. sample scores on DCA)

Example data

Forest vegetation plots, located in even distances along transects following the steep valley slopes of Vltava river valley (close to Zlatá Koruna, Czech Republic) and collected during 2001-2003 (Zelený & Chytrý 2007). Each transect started at the valley bottom and end up at the upper part of the valley slope. Plots are of the size 10×15 m. In each plot, all species of tree, shrub and herb layer were recorded and their abundance was estimated using 9-degree ordinal Braun-Blanquette scale (these values were consequently transformed into percentage). At each plot, various topographical and soil factors were measured or estimated - soil pH is reported here. The dataset contains 27 transects with alltogether 97 samples.

Data in spreadsheet format

- vltava-spe.txt - **species composition matrix** (97 samples and 275 species)
- vltava-spec.eiv.txt - **species indicator values** (275 species and 6 Ellenberg indicator values [2])
- vltava-5-groups-cluster.txt - **groups of samples** (numerical clustering of samples into 5 groups, using Euclidean distance and Ward algorithm)
- vltava-scores-dca-1-2.txt) - **external variables** (scores of samples along first and second DCA axis - DCA calculated from square-rooted cover data, using detrending by segments)

All files are in plain text format, with cells separated by tabulators. First row usually contains variable names (or species names in case of vegetation matrix), first column contains plot numbers or species names. See preview for details.

Data in JUICE program format

- vltava-example.zip

The zip file needs to be unzipped first and than the file with *.wct extension can be opened in JUICE (use `File > Open` in JUICE menu, to click on the icon of the file is not enough). Species Ellenberg indicator values are already initiated, i.e. they are assigned to species (you may check this in JUICE menu Indicator values > Initiation). The table is separated into 5 groups, resulting from cluster analysis (Ward algorithm and Euclidean distances).

References

- Zelený D. & Chytrý M. (2007): Environmental control of vegetation pattern in deep river valleys of the Bohemian Massif. *Preslia*, 79: 205-222
- Zelený D. & Schaffers A.P. (2012): Too good to be true: pitfalls of using mean Ellenberg indicator values in vegetation analyses. *Journal of Vegetation Science*, 23: 419-431

Known bugs, which will be fixed in the next release

- clicking on “Online help” button redirect to an old, non functional website - needs to change the link into http://davidzeleny.net/wiki/doku.php/eiv:software

If something does not work

please write me (zeleny@sci.muni.cz), I will try to fix it. I really appreciate constructive feedback!

[1)]
or `c:\Program Files\R\R-2.X.X\bin\x64\` folder if you are using 64-bits version of Windows
[2)]
note that missing values must be as empty values, not zeroes!