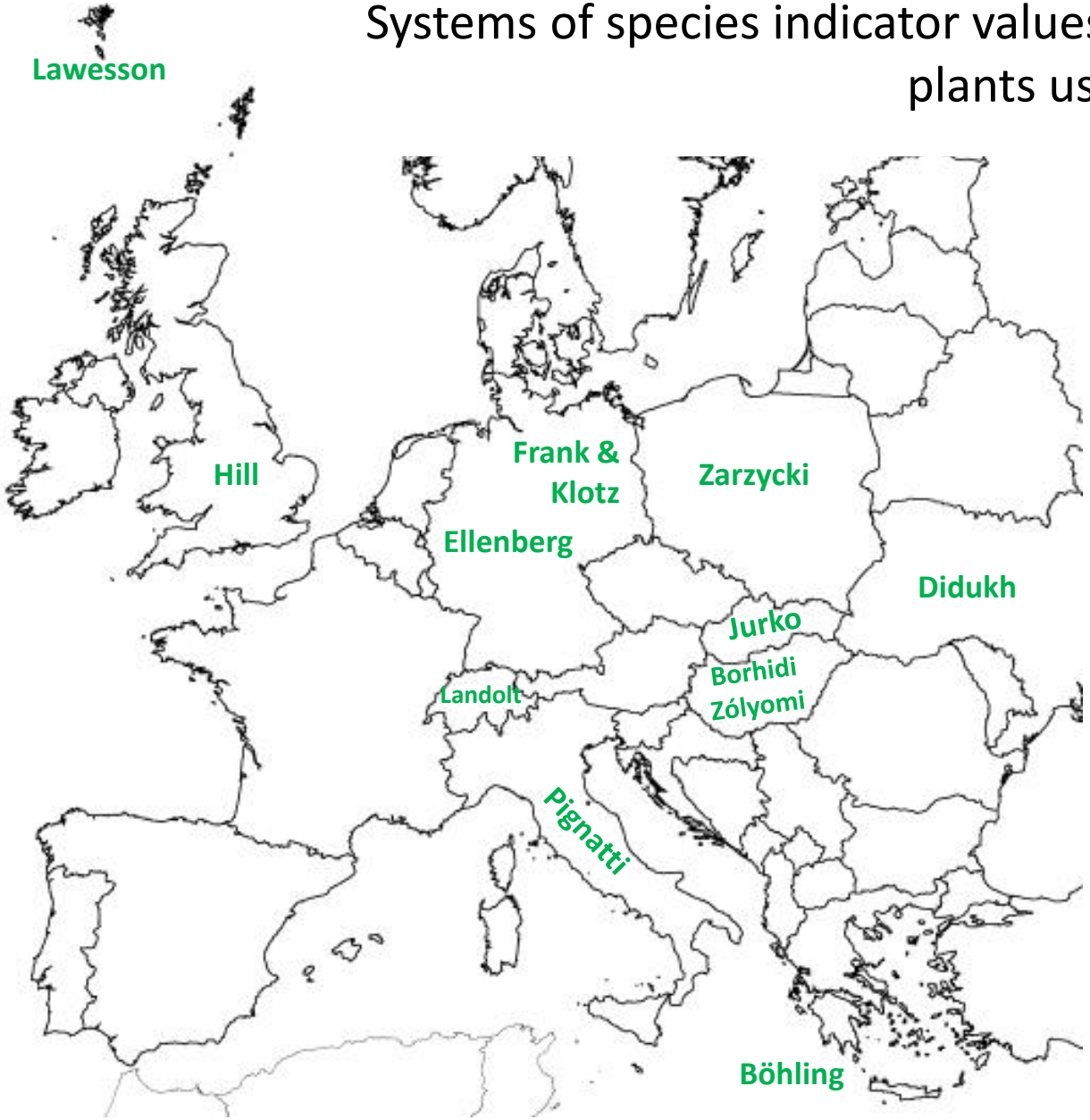
Use of **mean** Ellenberg indicator values revisited (again)

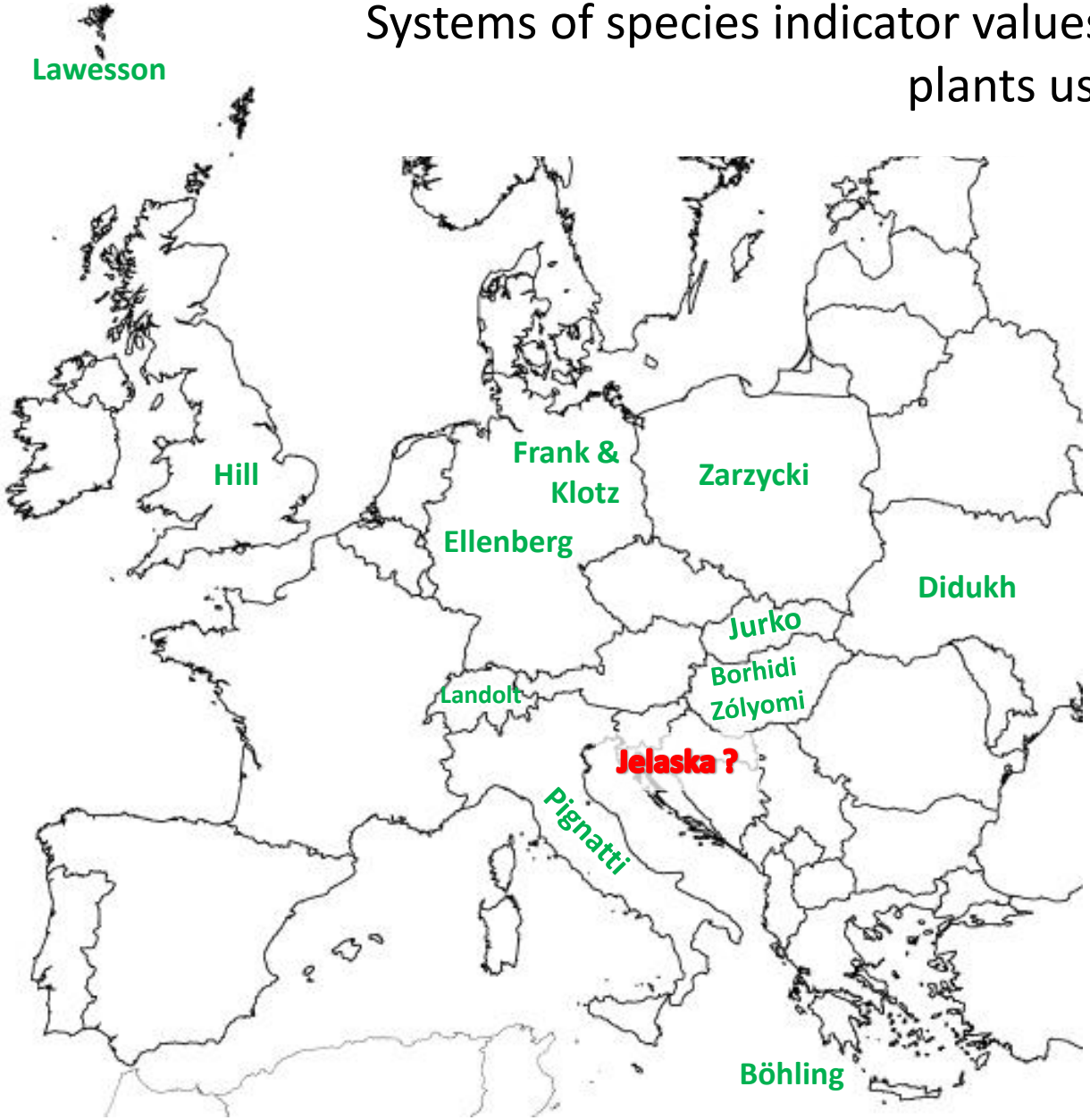
David Zelený

VEGETATION SCIENCE GROUP
MASARYK UNIVERSITY BRNO

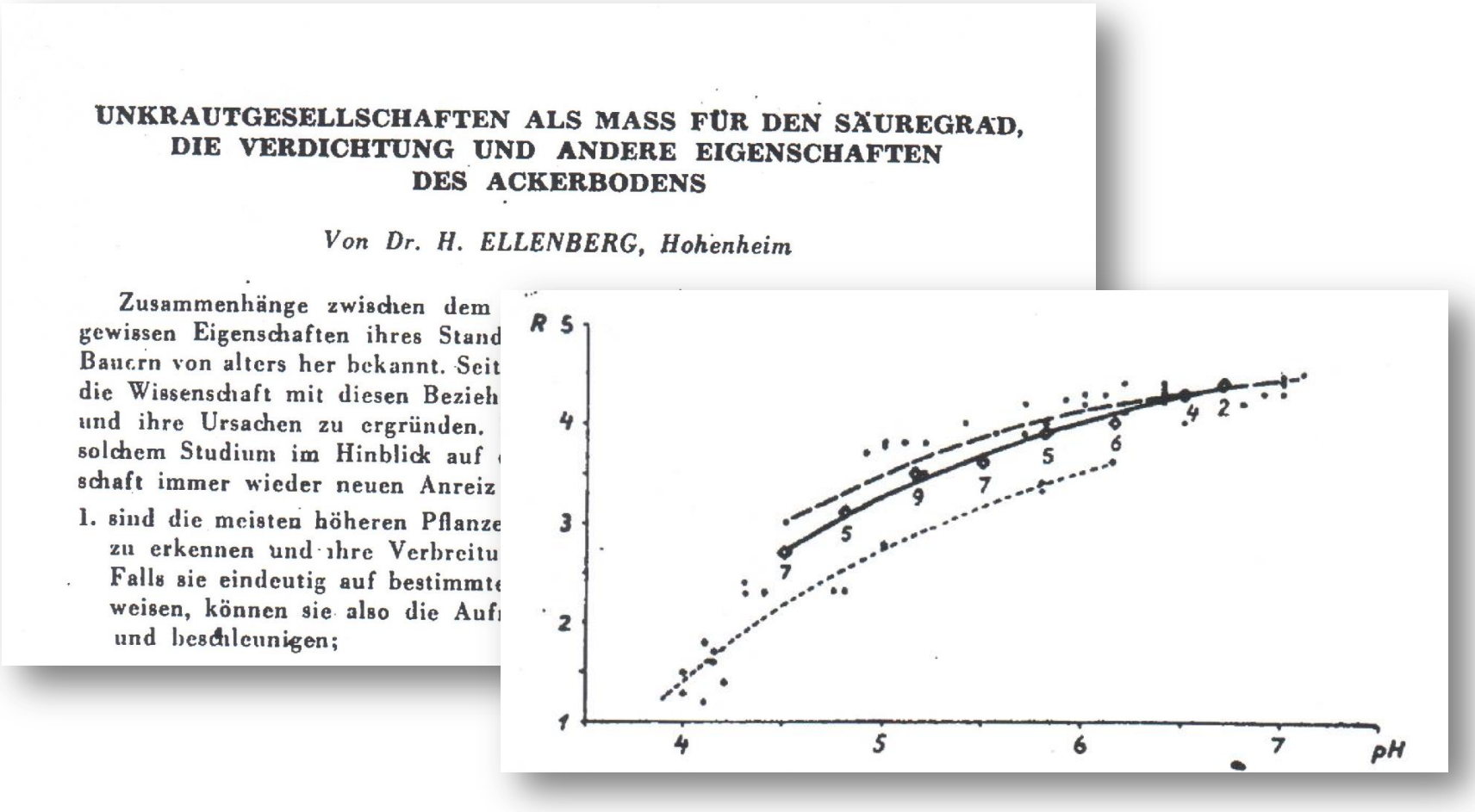
Systems of species indicator values for vascular plants used in Europe



Systems of species indicator values for vascular plants used in Europe



Relationship between mean indicator values and measured environmental variables



UNKRAUTGESELLSCHAFTEN ALS MASS FÜR DEN SÄUREGRAD, DIE VERDICHTUNG UND ANDERE EIGENSCHAFTEN DES ACKERBODENS

Von Dr. H. ELLENBERG, Hohenheim

Zusammenhänge zwischen dem
gewissen Eigenschaften ihres Stand
Bauern von alters her bekannt. Seit
die Wissenschaft mit diesen Bezieh
und ihre Ursachen zu ergründen.
solchem Studium im Hinblick auf
schaft immer wieder neuen Anreiz
1. sind die meisten höheren Pflanze
zu erkennen und ihre Verbreitu
Falls sie eindeutig auf bestimmte
weisen, können sie also die Auf
und beschleunigen;

Heinz Ellenberg (1948) *Berichte über Landtechnik*

Relationship between mean indicator values and measured environmental variables tested

SUCCESSION IN A SOUTH SWEDISH DECIDUOUS WOOD: A NUMERICAL APPROACH*

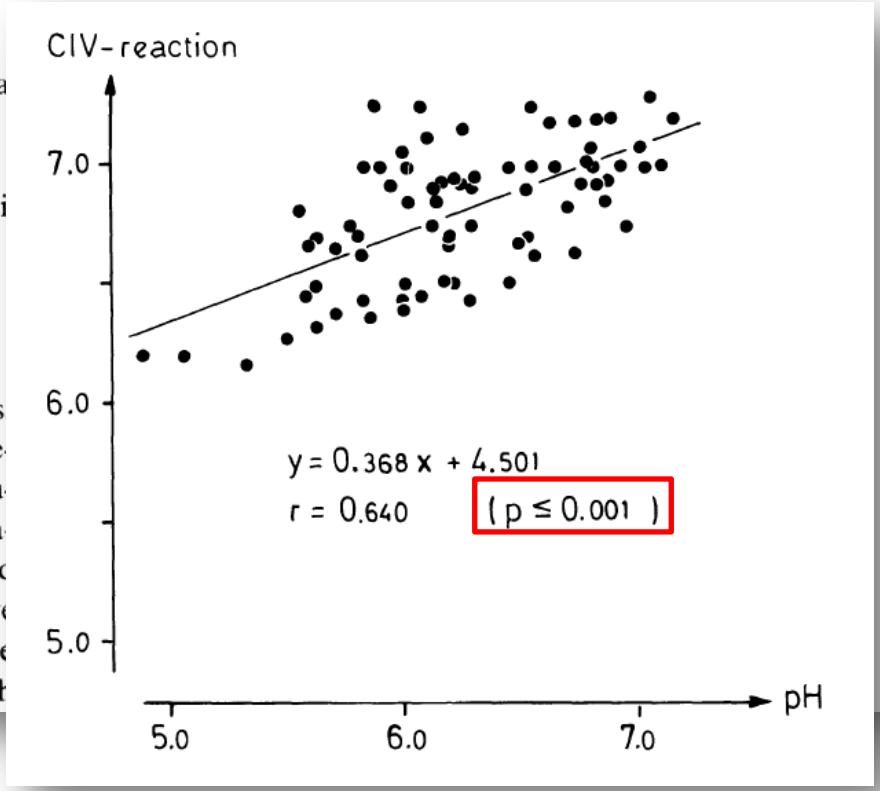
Stefan PERSSON**

Institute of Plant Ecology, University of Lund, ö Vallga

Keywords:
Classification, Ordination, Semipermanent plots, Successi

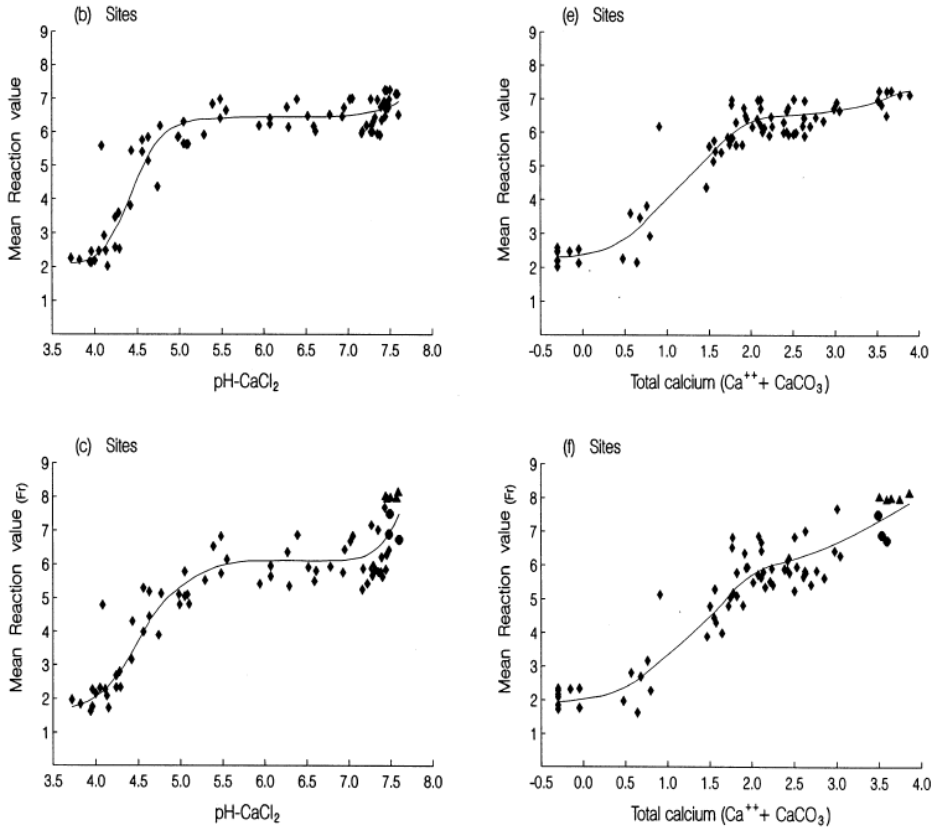
Introduction

Studies of successions in vegetation must, in many cases involve observation periods of considerable length. Espe- cially when successions are dependent on or strongly influ- enced by developments in tree- and shrub layers, observa- tions must be extended over decades. Very rarely one anc the same investigator is in a position to be able to observe his permanent plots for such a long time. An alternative to this dilemma is the possibility of reestablishing, with



Stefan Persson (1980) *Vegetatio*

Relationship between mean indicator values and measured environmental variables tested



Species included		Vascular plants			
Calculation type		Normal		Frequency wt.	
Transformation	*n*	Pres	Abun	Pres	Abun
Moisture					
Average groundwater	74	0.87	0.87	0.88	0.88
Lowest moisture	74	0.84	0.84	0.87	0.86
Highest pF	74	−0.77	−0.76	−0.77	−0.76
Nitrogen					
Soil C:N ratio	74	−0.54	−0.54	−0.52	−0.50
Soil total N	74	0.61	0.61	0.62	0.61
N-mineralization	74	0.39	0.40	0.41	0.42
Soil mineral N	74	0.50	0.51	0.53	0.54
Soil NO_3	74	0.68	0.69	0.69	0.69
Nitrification degree	74	0.66	0.67	0.65	0.65
Soil available PO_4	74	0.51	0.51	0.53	0.52
Soil available K	74	0.59	0.60	0.60	0.59
Tissue N-concentration	72	0.71	0.72	0.69	0.70
N-accumulation	57	0.83	0.85	0.83	0.85
Biomass production	57	0.82	0.85	0.84	0.86
Soil reaction					
pH-$CaCl_2$	74	0.79	0.79	0.78	0.78
Base saturation	62	0.94	0.94	0.87	0.86
Ca^{2+} saturation	62	0.94	0.94	0.88	0.87
Ca^{2+} amount	73	0.92	0.92	0.90	0.90
Total calcium	73	0.90	0.90	0.86	0.86
Specie					
pH-Ca					14
Base s					10
Ca^{2+} s					12
Ca^{2+} a					17
Total calcium	73	0.92	0.92	0.91	0.92
pH > 5.25					
pH-$CaCl_2$	45	0.36	0.37	0.32	0.36
Base saturation	37	0.38	0.39	0.31	0.37
Ca^{2+} saturation	37	0.43	0.45	0.37	0.43
Ca^{2+} amount	44	ns	ns	ns	ns
Total calcium	44	0.60	0.65	0.60.	0.68

[a] Bryophytes constituted between 0 and 80% of the total number of species wit only 14-15%.

correlation coefficients (all significant at $p < 0.001$)

Schaffers & Sýkora (2000) *Journal of Vegetation Science*

Calculation of mean Ellenberg indicator values

species composition matrix

	sp1	sp2	sp3	sp4	sp5	sp6	sp7
plot1	1	1	0	0	0	0	0
plot2	1	0	1	1	0	1	0
plot3	0	0	1	1	1	0	1
plot4	0	0	0	1	0	1	1

Calculation of mean Ellenberg indicator values

species composition matrix

	sp1	sp2	sp3	sp4	sp5	sp6	sp7
plot1	1	1	0	0	0	0	0
plot2	1	0	1	1	0	1	0
plot3	0	0	1	1	1	0	1
plot4	0	0	0	1	0	1	1

✖

species EIVs for soil reaction

R	1	1	3	4	4	6	7

Calculation of mean Ellenberg indicator values

species composition matrix

	sp1	sp2	sp3	sp4	sp5	sp6	sp7
plot1	1	1	0	0	0	0	0
plot2	1	0	1	1	0	1	0
plot3	0	0	1	1	1	0	1
plot4	0	0	0	1	0	1	1

×

species EIVs for soil reaction

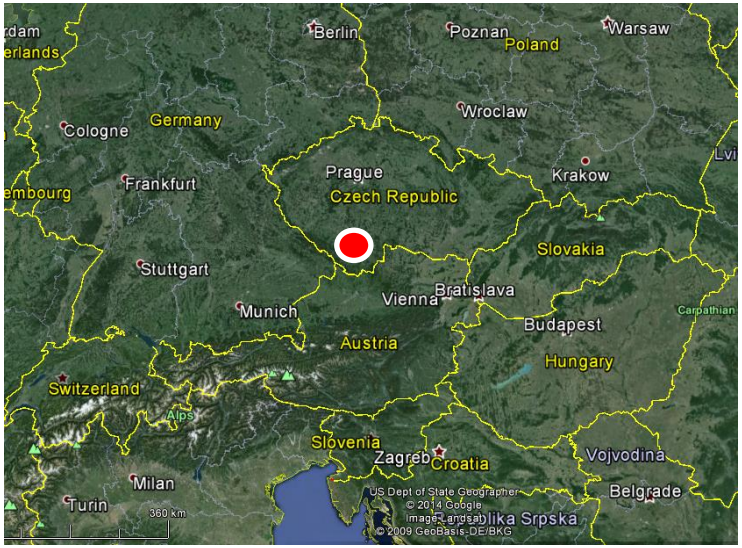
R	1	1	3	4	4	6	7

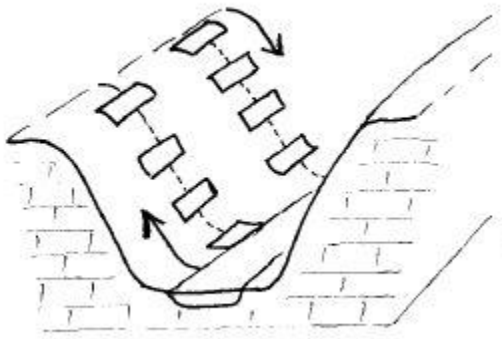
mean EIVs for soil reaction

mR
1
3.5
4.5
5.7

Relationship between mean EIV and measured pH

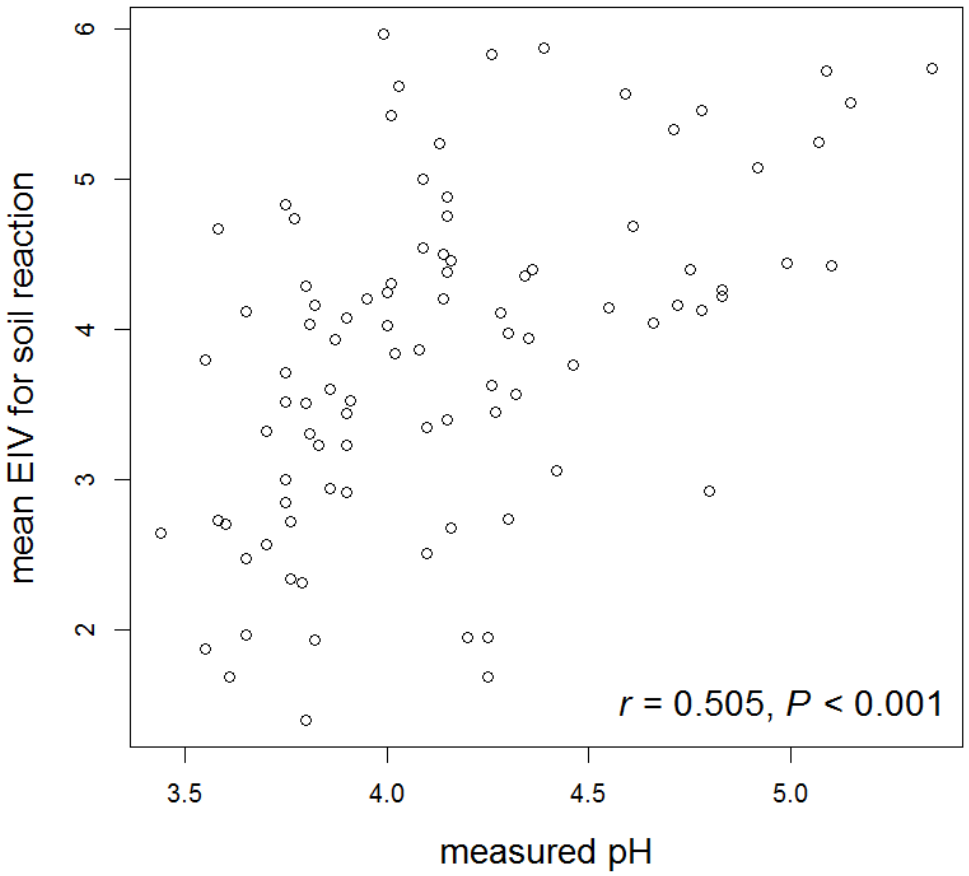
Case study: Vltava valley forest vegetation





Relationship between mean EIV and measured pH

Case study: Vltava valley forest vegetation



Calculation of mean Ellenberg indicator values

species composition matrix

	sp1	sp2	sp3	sp4	sp5	sp6	sp7
plot1	1	1	0	0	0	0	0
plot2	1	0	1	1	0	1	0
plot3	0	0	1	1	1	0	1
plot4	0	0	0	1	0	1	1

×

species EIVs for soil reaction

R	1	1	3	4	4	6	7

mean EIVs for soil reaction

mR
1
3.5
4.5
5.7

Randomization of Ellenberg indicator values

species composition matrix

	sp1	sp2	sp3	sp4	sp5	sp6	sp7
plot1	1	1	0	0	0	0	0
plot2	1	0	1	1	0	1	0
plot3	0	0	1	1	1	0	1
plot4	0	0	0	1	0	1	1

mean EIVs for soil reaction

mR
1
3.5
4.5
5.7

✖

species EIVs for soil reaction

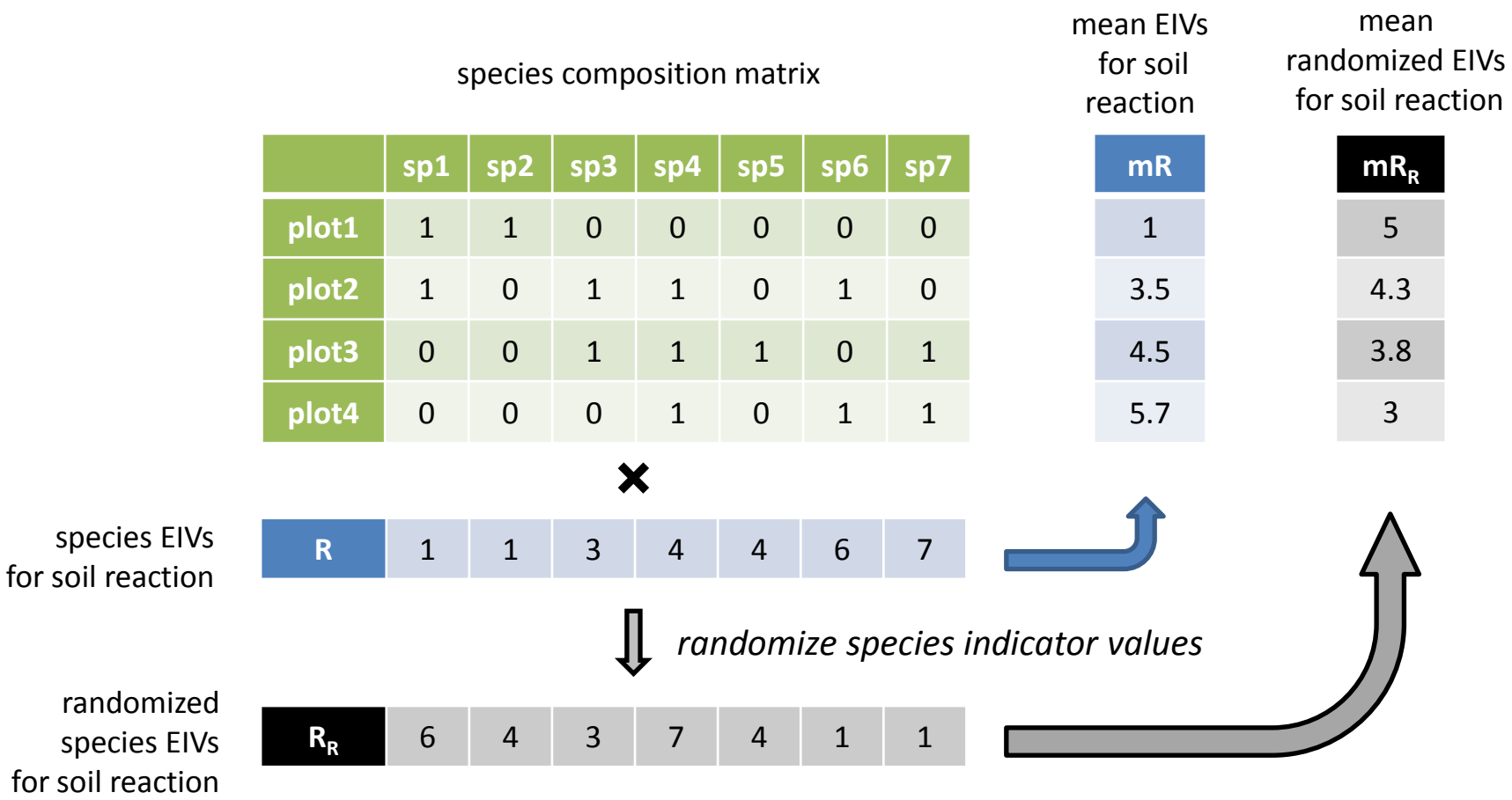
R	1	1	3	4	4	6	7

randomize species indicator values

randomized species EIVs for soil reaction

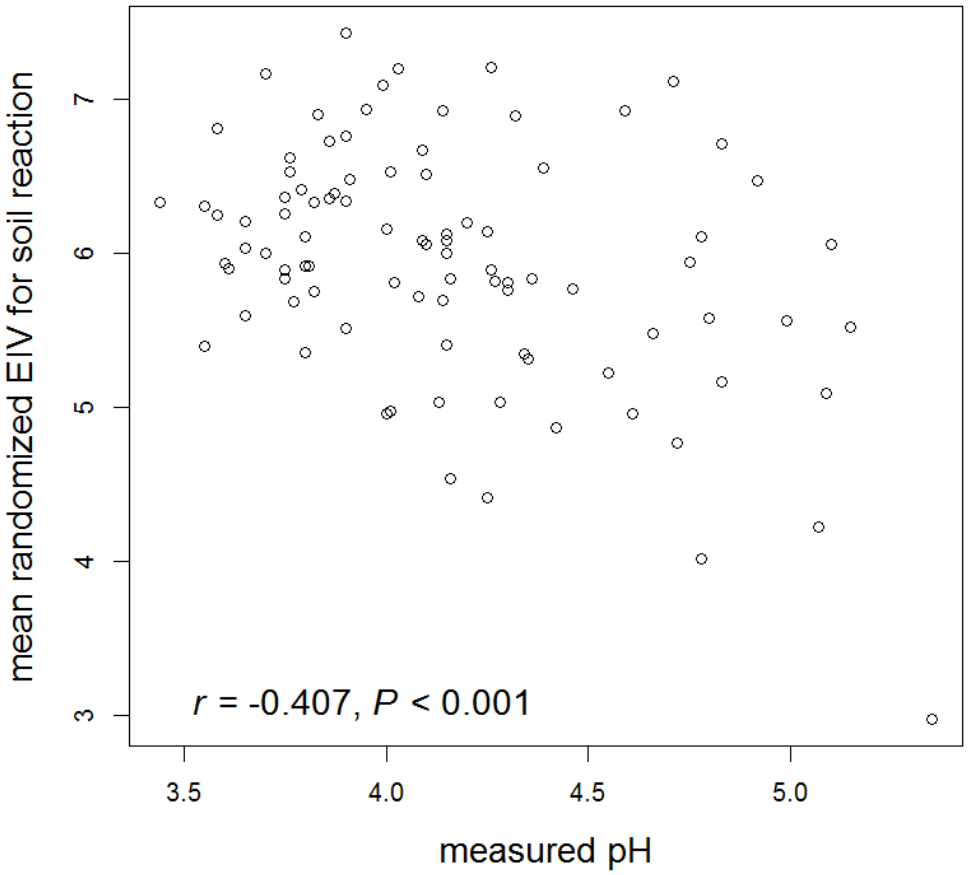
R_R	6	4	3	7	4	1	1

Randomization of Ellenberg indicator values



Relationship between mean randomized EIV and measured pH

Case study: Vltava valley vegetation



Relationship between mean randomized EIV and measured pH

Case study: Vltava valley vegetation

1000 randomly generated mean Ellenberg indicator values correlated with measured soil pH and tested the significance ($p < 0.05$)

- expect around 50 (5%) significant correlations

Relationship between mean randomized EIV and measured pH

Case study: Vltava valley vegetation

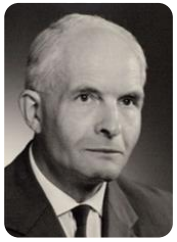
1000 randomly generated mean Ellenberg indicator values correlated with measured soil pH and tested the significance ($p < 0.05$)

~~- expect around 50 (5%) significant correlations~~

- in reality **327 (~33%) correlations** are significant

Use of mean Ellenberg indicator values in vegetation analysis

Compositional data



Heinz Ellenberg

Species indicator values

environmental variable

correlation, regression, ANOVA

mean Ellenberg indicator values

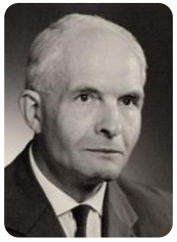
Use of mean Ellenberg indicator values in vegetation analysis

Compositional data

environmental variable

correlation, regression, ANOVA

Heinz Ellenberg



Species indicator values

mean Ellenberg indicator values

Reasons for inflated Type I error rate

- wrong null hypothesis is being tested
  - H_0: there is no relationship between environmental variable and mean EIVs (i.e. species composition *and* species indicator values)
  - if there is relationship between env. var. and species composition (which is likely), than this null hypothesis is easy to be rejected and to test it is not informative

- better to test an alternative null hypothesis
  - H_0: the relationship between env. var. and mean EIVs is not stronger than would be relationship between env. var. and mean value of randomized indicator values (without ecological meaning)

Modified permutation test

Journal of Vegetation Science **23** (2012) 419–431

Too good to be true: pitfalls of using mean Ellenberg indicator values in vegetation analyses

David Zelený & André P. Schaffers

Keywords
Bio-indication; Circularity of reasoning; Compositional similarity; Environmental calibration; Null models; Ordination analysis; Randomization; Species richness; Weighted averaging

Abbreviations
ANOVA = analysis of variance; CA = correspondence analysis; CCA = canonical correspondence analysis; DCA = detrended correspondence analysis; EIVs = Ellenberg indicator values; TWINSPAN = two-way indicator species analysis

Received 2 September 2010
Accepted 14 October 2011
Co-ordinating Editor: Michael Palmer

Zelený, D. (corresponding author, zeleny@sci.muni.cz): Department of Botany and Zoology, Masaryk University, Kotlářská 2, CZ-611 37, Brno, Czech Republic
Schaffers, A.P. (Andre.Schaffers@wur.nl):

Abstract

Question: Mean Ellenberg indicator values (EIVs) inherit information about compositional similarity, because during their calculation species abundances (or presence–absences) are used as weights. Can this *similarity issue* actually be demonstrated, does it bias results of vegetation analyses correlating mean EIVs with other aspects of species composition and how often are biased studies published?

Methods: In order to separate information on compositional similarity possibly present in mean EIVs, a new variable was introduced, calculated as a weighted average of randomized species EIVs. The performance of these *mean randomized EIVs* was compared with that of the *mean real EIVs* on the one hand and *random values* (randomized mean EIVs) on the other. To demonstrate the *similarity issue*, differences between samples were correlated with dissimilarity matrices based on various indices. Next, the three mean EIV variables were tested in canonical correspondence analysis (CCA), detrended correspondence analysis (DCA), analysis of variance (ANOVA) between vegetation clusters, and in regression on species richness. Subsequently, a modified permutation test of significance was proposed, taking the *similarity issue* into account. In addition, an inventory was made of studies published in the *Journal of Vegetation Science* and *Applied Vegetation Science* between 2000 and 2010 likely reporting biased results due to the *simi-*

Summary: variables and analyses, for which modified permutation test is advisable

Origin of the variable	**Type of the variable**	**Example of analysis**	**Example of the variables and references**
measured or observed variables related to species composition	quantitative	correlation, regression	measured environmental variables (e.g. soil depth, soil moisture, pH, successional age) ecosystem properties (e.g. standing biomass)
	qualitative	ANOVA type	experimental treatment (e.g. grazing, fertilizing, mowing)
numerically derived from species composition	quantitative	correlation, regression	species richness or indices of alpha diversity scores of samples on unconstrained ordination axes other mean Ellenberg indicator values CWM for functional traits
	qualitative	ANOVA type	sample clusters derived by cluster analysis groups of samples assembled manually according to similarity in observed species composition

Regressions between **measured environmental variables** and **mean EIVs** (table of p-values)

	Fluvisol	ASPSSW	SURFSL	pH.H	COVER32	SOILDPT	Lithic	Elevation	Skeletic	slope	XERSSW	Cambisol	SURFIS
LIGHT	0.176	**< 0.001**	0.175	0.479	**< 0.05**	**< 0.001**	**< 0.01**	0.067	**< 0.05**	0.090	**< 0.01**	0.197	**< 0.05**
TEMP	**< 0.001**	0.500	**< 0.001**	**< 0.001**	**< 0.001**	**< 0.05**	**< 0.01**	**< 0.001**	**< 0.01**	0.337	0.513	0.152	0.189
CONT	**< 0.001**	0.571	0.064	**< 0.05**	0.836	0.760	0.213	**< 0.05**	0.943	0.798	0.729	**< 0.01**	0.721
MOIST	**< 0.001**	**< 0.01**	**< 0.001**	**< 0.001**	**< 0.001**	**< 0.001**	**< 0.001**	**< 0.001**	**< 0.001**	**< 0.01**	**< 0.001**	0.489	**< 0.01**
REACT	**< 0.01**	0.060	**< 0.05**	**< 0.001**	**< 0.001**	**< 0.001**	**< 0.001**	**< 0.01**	0.264	**< 0.05**	0.271	0.407	0.796
NUTR	**< 0.01**	**< 0.001**	**< 0.001**	**< 0.001**	**< 0.001**	**< 0.001**	**< 0.001**	0.171	**< 0.001**	**< 0.01**	**< 0.001**	0.543	**< 0.05**
# signif	5	3	4	5	5	5	5	4	4	3	3	1	3

78 regressions, **50** of them are significant (64%)

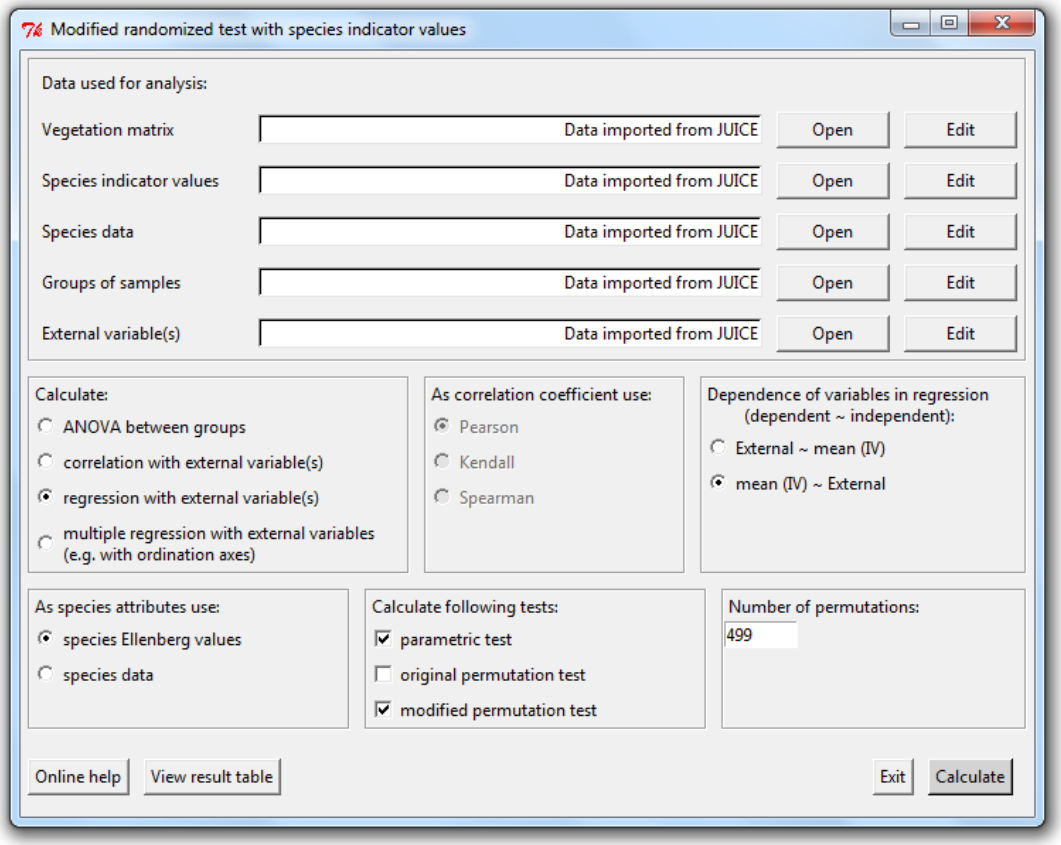
Regressions between **measured environmental variables** and **mean EIVs** (table of p-values after correction)

	Fluvisol	ASPSSW	SURFSL	pH.H	COVER32	SOILDPT	Lithic	Elevation	Skeletic	slope	XERSSW	Cambisol	SURFIS
LIGHT	0.537	**< 0.01**	0.477	0.744	0.199	0.079	0.177	0.321	0.172	0.217	**< 0.05**	0.361	0.067
TEMP	0.069	0.760	**< 0.01**	**< 0.01**	**< 0.05**	0.220	**< 0.05**	**< 0.01**	**< 0.05**	0.513	0.586	0.305	0.348
CONT	**< 0.05**	0.804	0.355	0.336	0.911	0.891	0.539	0.208	0.967	0.864	0.797	**< 0.05**	0.798
MOIST	**< 0.01**	0.141	**< 0.001**	**< 0.05**	**< 0.001**	**< 0.01**	**< 0.01**	**< 0.05**	**< 0.01**	**< 0.01**	**< 0.01**	0.629	**< 0.05**
REACT	0.209	0.410	0.189	**< 0.01**	**< 0.05**	**< 0.05**	0.063	0.094	0.544	0.056	0.384	0.600	0.833
NUTR	0.132	**< 0.01**	**< 0.01**	**< 0.05**	**< 0.01**	**< 0.001**	**< 0.05**	0.443	**< 0.05**	**< 0.05**	**< 0.001**	0.683	0.077
# signif	2	2	3	4	4	3	3	2	3	2	3	1	1

78 regressions, **35** of them are significant (45 %)

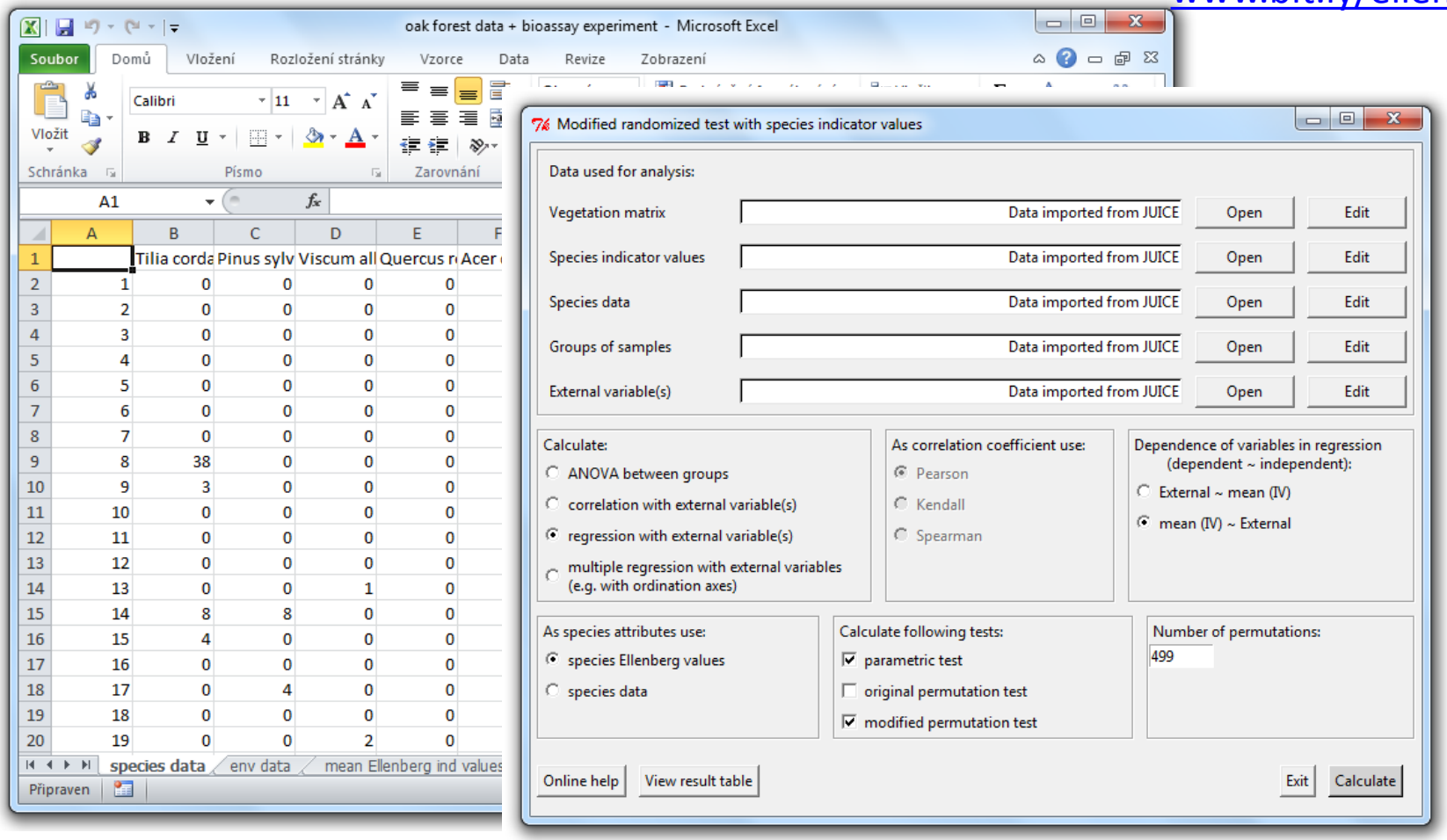
MoPeT: program (R-based) for calculation of modified permutation test

www.bit.ly/ellenberg



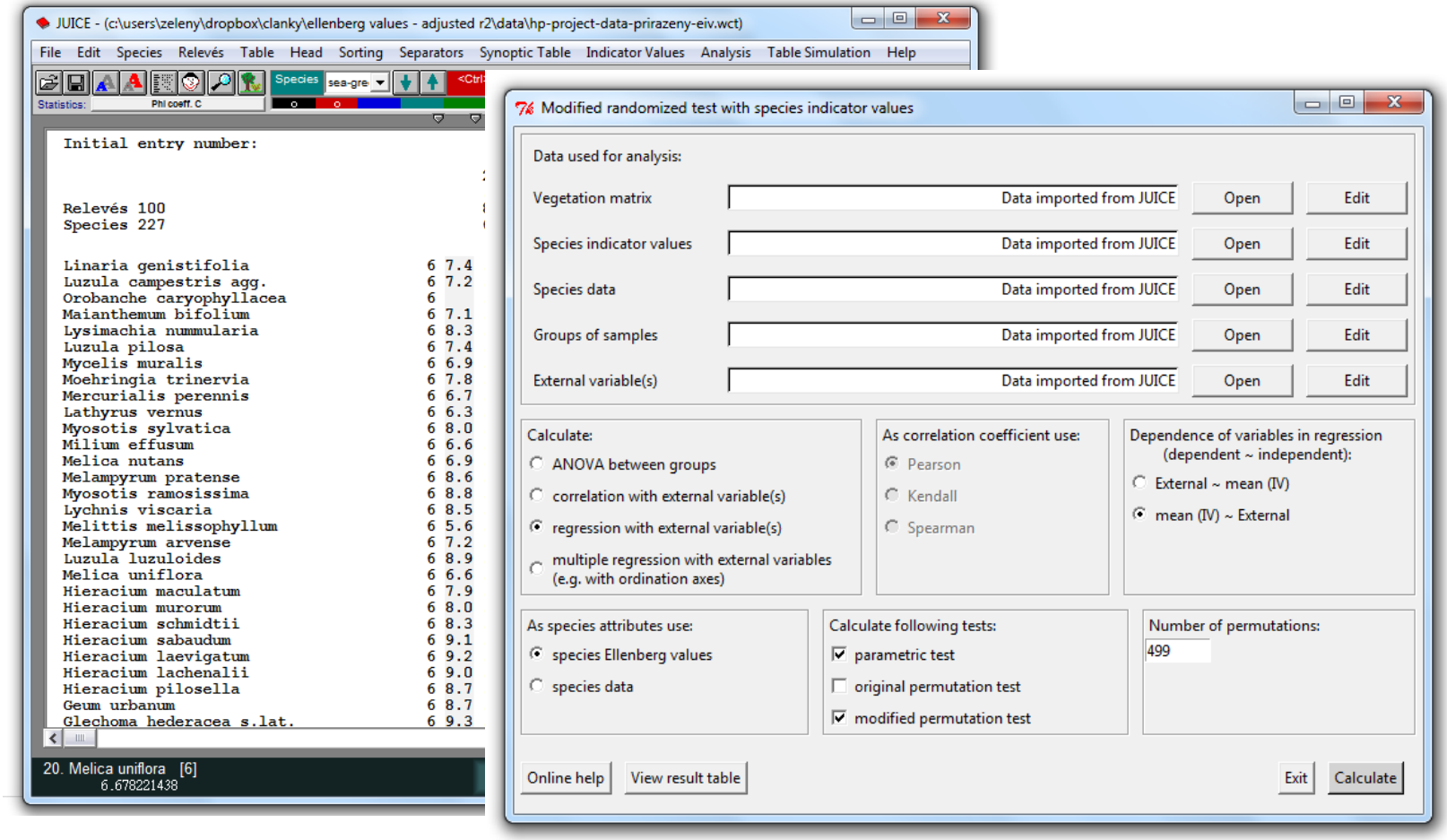
MoPeT: program (R-based) for calculation of modified permutation test

www.bit.ly/ellenberg



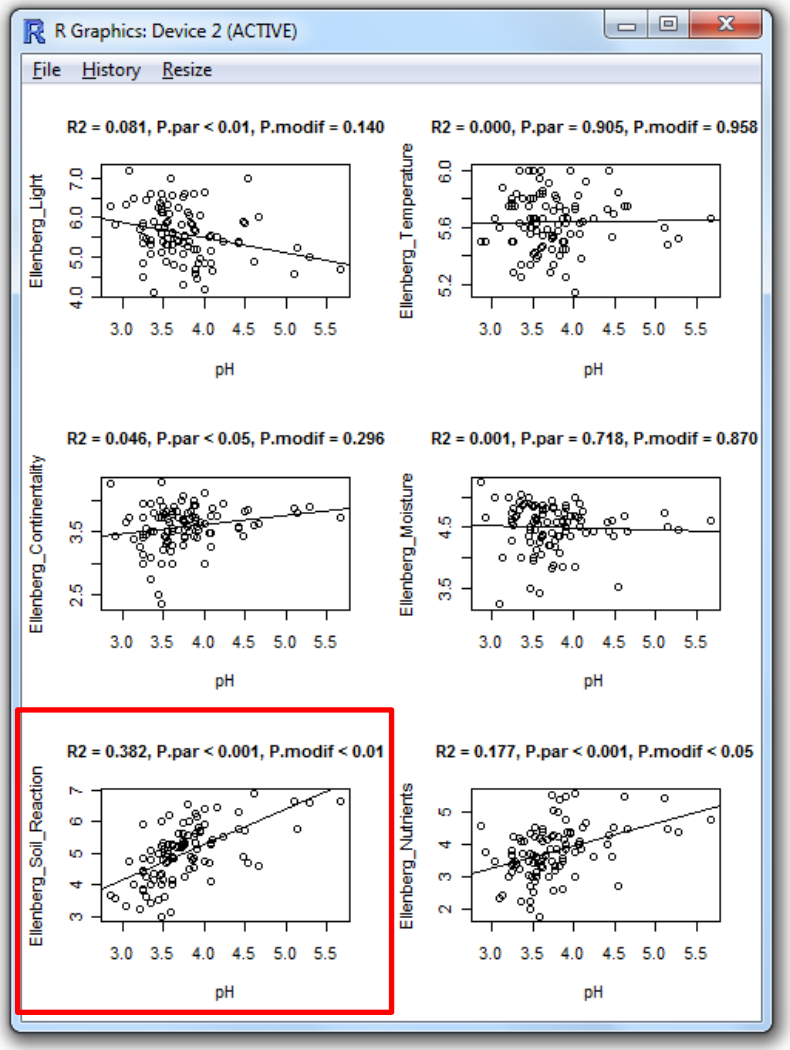
MoPeT: program (R-based) for calculation of modified permutation test

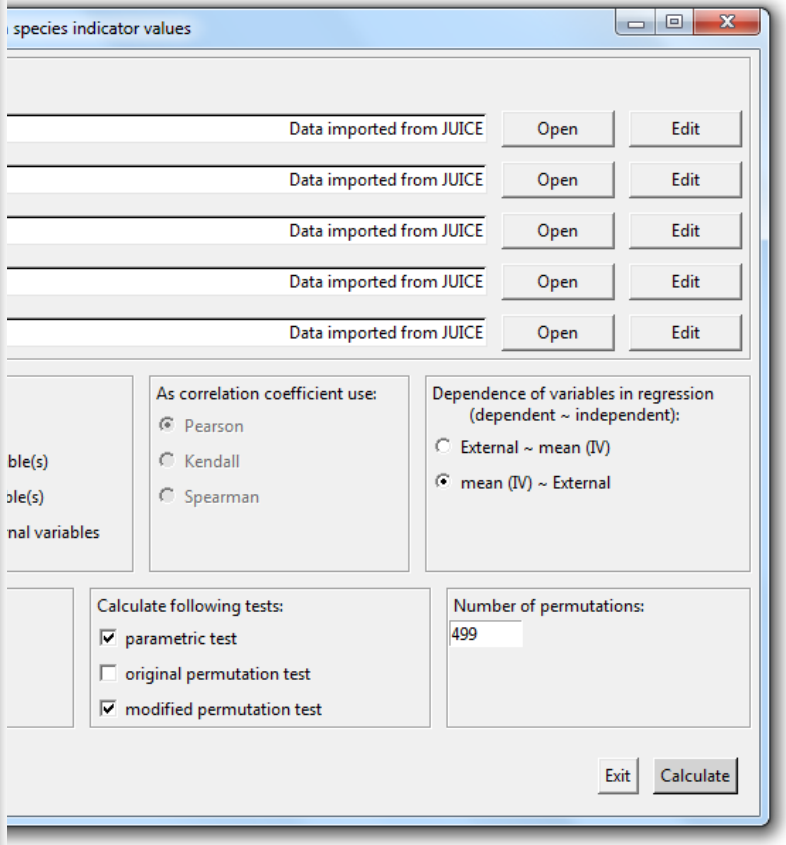
www.bit.ly/ellenberg



MoPeT: program (R-based) for calculation of modified permutation test

www.bit.ly/ellenberg





Conclusions

- If you test the significance of relationship between mean Ellenberg indicator values and other variables and there is a danger that this variable is linked to species composition (either because it influences it, like environmental variables, or because it's derived from it, like ordination scores), you may avoid inflated Type I error rate by using modified permutation test

- The same problem applies not only to mean Ellenberg values, but to all weighted means of species attributes (i.e. also to community weighted means of species functional traits, to diatom indices etc.)

Conclusions

- If you test the significance of relationship between mean Ellenberg indicator values and other variables and there is a danger that this variable is linked to species composition (either because it influences it, like environmental variables, or because it's derived from it, like ordination scores), you may avoid inflated Type I error rate by using modified permutation test

- The same problem applies not only to mean Ellenberg values, but to all weighted means of species attributes (i.e. also to community weighted means of species functional traits, to diatom indices etc.)

Thank you for your attention!