Bias due to autocorrelation where no one expects it: the relationship between community-weighted mean traits and environmental variables

D. Zelený

Department of Botany and Zoology
Faculty of Science
Masaryk University
Brno, Czech Republic
zeleny@sci.muni.cz

Keywords: autocorrelation of residuals, functional traits, inflated type I error rate, Monte Carlo permutation test.

Abstract:
Community-weighted mean of species trait values (CWM), one of functional diversity indices, became a standard tool to scale-up functional traits from individual to community or ecosystem level. Common practice is to test the relationship between CWM and environmental variables or ecosystem properties, e.g. using regression or correlation. In this study, I point out the bias inherited in these analyses, including inflated type I error rate and overestimated explained variation, leading to results more optimistic than would be justified by data. To explain the cause of this bias and to seek a potential remedy, I use a surprisingly helpful analogy with an analysis of spatially autocorrelated variables. I demonstrate this property on regression between set of CWM and set of environmental variables, using artificial and real community datasets. I propose using the Monte Carlo permutation test to correct an inflated type I error rate.

Results clearly show that CWM calculated from randomly generated species functional trait values has several times higher probability of being significantly related with environmental variables than would correspond to nominal significance level. Moreover, this inflation of type I error rate increases with increasing effect of environmental variables on species composition. Proposed Monte Carlo permutation test gives unbiased levels of type I error rate. The example with real community data shows that uncorrected and corrected results as well as interpretation can differ dramatically.

This is the first time to report that overly optimistic results may be expected when analysing the relationship between CWM and other variables, derived from or influencing a species composition. In future, this bias needs to be acknowledged and corrected, to avoid excessive reporting of significant relationships between community-aggregated traits and environmental variables or ecosystem properties.