**Design & analysis of ecological experiments: homework 5**

Here is the specification of your homework. Please submit your solution to Moodle before Saturday midnight (23:59) to have it commented and scored. Save this Word document using your surname only, without diacritical marks and append your answers / results at its end. Please send me also your Canoco 5 project file (which is the one with the *.c5p* extension). Ideally, its name would be also based on your surname. So, for example, I would sent myself two files - *Smilauer.docx* and *Smilauer.c5p*. Do not append homework number or date, different homeworks do not get mixed even if you name them all the same. Just to be sure, I must mention that you are expected to accomplish this task in an independent way, not by copying the solution from another student.

Data in the Excel file (attached to this homework task) describe the response of algal community (sheet *Algae*) to the pollution of water (quantified by the variables stored in the *Water* sheet: those variables that needed log-transformation were already log-transformed there). Algal populations were quantified using an estimation scale ranging from 0 to 5. **BEWARE:** Second table (*Water*) has a factor as its first variable (*BOD*) - make sure it is imported as a variable, rather than a second set of case labels (which is what the import wizard unfortunatelly suggests) ...

Your specific tasks/questions for this HW are as follows:

(1) Using stepwise selection method, choose **two** pollution characteristics best explaining algal community composition and give the names of selected variables. Select just two even when more of them are significant ...

(2) What size has the part of the total community variation, which is explained by these two variables (together)?

(3) Create ordination diagram summarizing the relation of individual species of algae to the selected two predictors and specify in its caption how much variation is explained by the first and the second axis, respectively.

(4) Find five algal species, which are best explained by the first ordination axis. Display their response curves in relation to the first constrained axis. These curves should use GLM with quadratic specification of predictor effect, optionally with the model complexity selected by stepwise selection. Include resulting graph into your HW document.

I look forward to receive your solution

===============================================================

YOUR SOLUTION:

1. From the stepwise selection method, the two pollution characteristics best explaining algal community composition are: **NH4** and **BOD.medium**.
2. The total community variation explained by these two variables is **26.17%.**
3. Response data are compositional and have a gradient 4.6 SD units long, so the linear method is not appropriate. A **CCA ordination diagram** is then realized.

A graph with a red line and black text

Description automatically generated with medium confidence

*Figure 1: CCA plot of the effect of NH4 and medium biological oxygen demand on algae community (only 20 species out of the 34 total species are displayed)*

1. CFit.1 enables to determine the percent of explained variation by axis 1 for each species. The five algal species which are best explained by the first ordination axis are: ***Navicryp, Hantamph, Nizspale, Phorfove, Oscilimo.***

A screenshot of a computer

Description automatically generated

*Figure 2: CFit.1 of the 5 species best explained by the first ordination axis*

The response curves in relation to the first constrained axis of these 5 species are displayed below.

*A graph of different colored lines

Description automatically generated*

*Figure 3: Response curves in relation to the first constrained axis of the species* *Navicryp, Hantamph, Nizspale, Phorfove, Oscilimo*