EPFL

ENV-521	**Multivariate statistics with R in environment**

Buttler Alexandre, Walker Thomas

Cursus	Sem.	Type
Environmental Sciences and Engineering	MA1, MA3	Opt.
Mineur STAS Russie	H	Opt.

Language of teaching	English
Credits	4
Session	Winter
Semester	Fall
Exam	Oral
Workload	120h
Weeks	14
Hours	**3 weekly**
Courses	2 weekly
Exercises	1 weekly
Number of positions	

Summary

Introduction to multivariate data analysis and modelling. The course helps for a critical choice of methods and their integration in a research planning. It prepares for complexe data analysis in various fields of environemental sciences. Use of dedicated R libraries

Content

- Biological and environmental descriptors, multidimensional data, coding and transformation
- Resemblance and dependence measures, association matrices
- Analysis of discontinuities: unsupervised clustering techniques
- Analysis of discontinuities: supervised clustering, regression and classification trees
- Gradient analysis: ordination techniques in reduced space (PCA, CA, PCoA, NMDS)
- Direct gradient analysis: symmetric coupling of tables (COIA, MFA)
- Direct gradient analysis: constrained ordination (RDA, CCA, pRDA, pCCA, db-RDA)
- Statistical tests for multivariable responses

Keywords

Multivariable analysis, statistics for complexe data sets

Learning Prerequisites

Recommended courses

Probabilities and statistics
Experimental Design and Data Analysis with R" (EDDAR - ENG 467)

Learning Outcomes

By the end of the course, the student must be able to:

- Select appropriately methods for data analysis knowing the basic principles of calculation et the field of their application
- Construct a plan for data analysis

- Interpret properly the results given by the different methods
- Apply the methods with exercices and a personal project
- Work out / Determine means for combining data from two or more independant data sets describing the same objects and test the relationship

Teaching methods

Lecture and exercises on computer, personel project for applying methods.

Expected student activities

Participating at the lecture and reading the hand-out
Applying the various methods with the exercices and provided data set
Personal project with report and defense

Assessment methods

40 % spot written control (2h) during the semester
10 % continuous control (exercises) during the semester
50 % oral exam (30 min) during exam session

Supervision

Office hours   Yes
Assistants   Yes

Resources

Bibliography
BIBLIOGRAPHY
Legendre, P., & Legendre, L. (2012) Numerical Ecology. *3e ed., Elsevier* ***
Jongman, R.H.G, Ter Braak, C.J.F. & Van Tongeren, O.F.R. (1987) Data analysis in community and landscape ecology. *PUDOC, Wageningen*
Borcard, D., Gillet, F. & Legendre, P. (2011) Numerical Ecology with R. Springer Verlag.*

*** for theory and fundamental concepts
* to work with R (codes)

Ressources en bibliothèque

- Numerical Ecology / Legendre
- Data analysis in community and landscape ecology / Jongman
- Numerical Ecology with R /Borcard

Notes/Handbook
Available on Moodle.epfl.ch

Websites

- http://www.r-project.org/
- http://cran.r-project.org/

Moodle Link

- http://moodle.epfl.ch/course/view.php?id=1361

Prerequisite for

Master project