ENV-513

Multivariate statistics in R

Peter Hannes Markus

Cursus	Sem.	Type
Civil & Environmental Engineering		Opt.
Environmental Sciences and Engineering	MA1, MA3	Opt.

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

Summary

Environmental datasets often contain numerous parameters. Multivariate statistics allow us to simultaneously explore, understand and model such datasets. This course provides conceptual introduction and guidelines for applying multivariate statistical tools using the R platform.

Content

1. Biological and environmental data, multidimensional data, and the R platform
2. Resemblance, similarity and dependence measures
3. Unsupervised and supervised clustering techniques
4. Ordination techniques (PCA, CA, PCoA, NMDS)
5. Constrained ordination (RDA, CCA, db-RDA)
6. Statistical tests for multivariable responses (anosim, betadisper)

Keywords

Multivariable analysis, statistics for ecological data sets, ordination, clustering

Learning Outcomes

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

- Explore multivariate datasets
- Select appropriately the methods for multivariate data analysis
- Explain the basic principles of various tools
- Interpret obtained results
- Apply methods in exercices and in a personal project

Transversal skills

- Communicate effectively with professionals from other disciplines.

Teaching methods

Lectures and computer exercises. Personal projects.

Expected student activities

- Active participation in lectures and excercises.
- Application of methods to example and a personal dataset
- Presentation of results (oral and written)

Assessment methods

- exam (individual) - 50%
- oral presentation (group work) - 50%

Supervision

Office hours	Yes
Assistant.e.s	Yes
Forum	Yes
Others	moodle

Resources

Moodle Link

- https://go.epfl.ch/ENV-513