DESCRIPTION OF THE CLASS

CODE	TITLE MULTIVARIATE ANALYSES FOR MARINE ECOLOGY
MU4MRM355 6 ECTS	**KEYWORDS:** Numerical ecology, data analysis, ordination, classification, machine learning.
M2	**PROFESSORS IN CHARGE:** Jean-Olivier IRISSON, Villefranche sur mer, Sorbonne Université Stéphane GASPARINI, Villefranche sur mer, Sorbonne Université **OTHER PROFESSORS:** Assistant professor

CLASS MODALITIES

TEACHING MODALITIES. 30h of theoretical lessons (including 9h of software training), 30h of tutorials including part of course application and part of personal project (processing multivariate data sets independently).

EVALUATION PROCEDURES. The theoretical courses are evaluated by a written exam (2 hours). The project is evaluated by an oral presentation (~10 min).

CLASS SUMMARY

This class focuses on statistical techniques used to analyse multivariate data and time series in oceanography or, more generally, in environmental sciences. We teach how to choose and associate appropriate statistical methods, the theoretical basis of these different methods and their practical implementation through the use of specialised software and programming, using R.

The topics covered are: distance measures relevant in ecology (Euclidean, Bray-Curtis, etc.), ordination methods in reduced space (ACP, AFC, MDS), multiple regression methods (linear model, RDA, CCA), unsupervised grouping methods (hierarchical, k- means), supervised classification and regression methods (i.e. Machine Learning; LDA, classification/regression trees, RandomForest, Multi Layer Perceptrons, Convolutional Neural Networks).

TEACHING GOALS

At the end of this course, you will know how to:
1. program analyzes in R and use specialized software;
2. choose an analysis method appropriate to the question at hand and the nature of the data;
3. implement multivariate analyzes and interpret their results;
4. conduct a complete analysis of a dataset already prepared.

PREREQUISITES

This class deals with multivariate analyses and assumes prior knowledge of classical inferential statistics (notion of variance, ANOVA, simple regression).

BIBLIOGRAPHY / WEBOGRAPHY

Numerical ecology. Legendre, Legendre, 2012. 3dr English edition. Elsevier.
Numerical ecology with R. Borcard, Gillet, Legendre, 2011. Springer
An Introduction to Applied Multivariate Analysis with R. Everitt, Hothorn, 2011. Springer.
Analyses factorielles simples et multiples. Escoffier, Pagès, 1998. Dunod.
Statistique exploratoire multidimentionelle. Lebart, Morineau, Piron, 1997. Dunod.

HOW THE CLASS WORKS

The theoretical basis of each method is detailed during a half-day course, focusing on the conditions of application of each method, their choice and the interpretation of their results. The mathematical underpinnings are present-ed to enable a complete understanding of the method but without requiring an advanced level of knowledge.

Then, each method is practiced during a half-day computer lab session, allowing to apply it on one or several datasets which highlight its characteristics well. These labs are all based on free and open source software; some methods are presented with software sporting a graphical interface (Tanagra and Past), others via programming in R. Training in these programs and in R is included at the beginning of the module.

Finally, the last two days are devoted to personal projects, during which students analyse a small dataset, which is either provided to them or that they brought with them. Each project is the subject of an oral presentation.

Practical information

This module can, in general, be accepted as a PhD class, sanctioned by an exam. At the very least, a certificate of participation can be issued.

The module is open to French and foreign students in a Master or PhD degree, assuming the prerequisites mentioned above are known. If foreign students register, the courses will be given in English.

The marine station of Villefranche offers accommodation and catering options.

Students are expected on the Saturday or Sunday preceding the start of classes. Classes begin at 09:00 on the first day and end at 12:00 on the last day.

Registration is through the e-Candidate platform.