

Environmental Data management, analysis and visualisation with R

1. Introduction to R

Theory

- o Brief history of R
- o Key features of R (advantages, peculiarities, limitations)

Laboratory

- o Introduction to the environment of R (GUIs)
- o R and RSTUDIO installation
- o Introduction to the RSTUDIO environment

2. Objects in R

Theory

- o R Object categories
- o Use and utilization of the R objects
- o Type and classes of the R objects

Laboratory

- o Introduction to the R objects
- o Object conversions
- o Utilization of objects in real data

3. Vectors and operators in R

Theory

- o Operators
- o Numerical functions
- o Vectors

Laboratory

- o Operations with operators and vectors
- o Expression of numerical functions in R
- o Problem solving with user defined functions

4. Tables, databases and lists

Theory

- o Two-dimensional / multidimensional arrays
- o Types of data vectors
- o Data box and lists

Laboratory

- o Data entry
- o Converts between vectors / frames and lists
- o Data export

5. Introduction to the concept of project in RStudio

Theory

- o Analysis of the “project” in RStudio and the benefit of working in such a context
- o The parts and limitations of the RStudio project

Laboratory

- o Creation and work of a real RStudio project

6. Receive and manage data from remote databases with R

Theory

- o Introduction to Big Environmental Databases
- o Procedure and restrictions on receiving data from remote databases
- o Data management process

Laboratory

- o Example of receiving data from a remote database
- o Downloading of data from WEB databases and APIs
- o Data management problems

7. Specialized packages, commenting and naming in R

Theory

- o What are the packages of R
- o The comments in writing R scripts
- o Good practices regarding comments and object names in R

Laboratory

- o Download and install packages
- o Write code with comments and use them
- o Complete script writing with emphasis on its good organisation for interdisciplinary communication

8. Data processing and analysis using the dplyr package

Theory

- o The basic commands and symbolism of dplyr
- o Working procedure with dplyr
- o Comparison with other base R approaches

Laboratory

- o Perform analysis of real environmental data with dplyr
- o Export results to table formats
- o Create graphs with the basic functions of R

9. Visualization of data analysis using ggplot2

Theory

- o Basic structure of ggplot2 commands and arguments
- o Grammar of ggplot2

Laboratory

- o Creating graphics from environmental data analysis
- o Create graphics with multiple panels

- o Export graphs in various dimensions, resolutions and formats

10. Basic statistical analysis of data with R

Theory

- o The basic commands for performing statistical analysis
- o Commands for descriptive statistics results' exporting
- o Data exporting process in a variety of formats

Laboratory

- o Perform descriptive analysis of environmental data
- o Perform regression analysis
- o Formatting and exporting analysis results

11. Creating user defined functions in R

Theory

- o Introduction to the functions of R
- o Basic characteristics and structure of functions

Laboratory

- o Solve a problem using a user defined function

12. Versioning in R

Theory

- o Introduction to version control
- o Introduction to the basic code of the control tools

Laboratory

- o Install git
- o Git functions

13. Good practices in writing and correcting code in R

Theory

- o Basic code structures
- o Code search and debugging procedures
- o Introduction to the basic knowledge repositories for resolving code errors

Laboratory

- o Execute code and search for errors
- o Find and fix errors
- o Navigate and search for solutions in knowledge repositories

Learning Objectives

The course is an introduction to the environmental data analysis with R programming language. On completion the student will be capable of:

- Working the R language environment (native GUI and RStudio).

- Installing and operating R language on a personal computer.
- Entering its own research data locally or from remote databases.
- Modifying the dataset to the appropriate format for further analysis.
- Performing the basic data analysis (descriptive statistics, regression, etc.).
- Visualising the data and the analysis' results.
- Exporting and sharing the analysis results and the visualisations.
- Understanding and operating the versioning control frameworks (e.g. Git, GitHub).