



EDA-EMERGE Local Course 2

“Statistic course – Environmental statistics using R”

Venue: Institute for Environmental Research, RWTH Aachen University, Aachen, Germany

Organizer: Prof. Dr. H. Hollert, Dr. Richard Ottermanns, Jonas Hausen, Benjamin Daniels

Date: 28.10.2014

Time: 9h00 – 17h00

Course description:

The course focused on practical experience with statistical analysis of ecotoxicological data. The participants learned to effectively use R for their work including data documentation, exploratory analysis, statistical tests and data visualization. Furthermore, the course provided knowledge on the most important statistical approaches in ecotoxicology such as: NOEC/LOEC, linear regression, non-linear regression including dose-response analysis, ANOVA models, species sensitivity distributions (SSD) and multivariate statistics (classification, ordination, principal response curve). The training was practice orientated, involving lectures and exercises.

This amounted to a minimum total academic involvement of 7.5 hours equivalent to 0.25 ECTS points for the participants.

AGENDA

Statistic course Tuesday, 28.10.2014			
Time	Title	Lecturer	Supporting documents
9:00-9:10	Welcome by the host	Dr. Richard Ottermanns	None
9:15-9:50	Lecture Slot 1: Descriptive statistics	Dr. Richard Ottermanns	Slide-pdfs
9:50-11:00	Tutorials and Exercises Slot 1	Dr. Richard Ottermanns, Jonas Hausen, Benjamin Daniels	Exercises Solutions R-Howto
11:00-11:30	Lecture Slot 2: Statistical testing	Dr. Richard Ottermanns	Slide-pdfs
11:30-12:30	Tutorials and Exercises Slot 2	Dr. Richard Ottermanns, Jonas Hausen, Benjamin Daniels	Exercises Solutions R-Howto
12:30-13:30	Lunch		
13:30-14:00	Lecture Slot 3: Curve fitting	Dr. Richard Ottermanns	Slide-pdfs
14:00-15:10	Tutorials and Exercises Slot 3	Dr. Richard Ottermanns, Jonas Hausen, Benjamin Daniels	Exercises Solutions R-Howto
15:10-15:40	Lecture Slot 4: Multivariate statistics	Dr. Richard Ottermanns	Slide-pdfs
15:40-16:50	Tutorials and Exercises Slot 4	Dr. Richard Ottermanns, Jonas Hausen, Benjamin Daniels	Exercises Solutions R-Howto
16:50-17:00	Concluding remarks	Dr. Richard Ottermanns	None

Course content

In detail the statistic course covered the following topics:

- Descriptive statistics
 - Measure of statistical location and dispersion; types of scales, plots
 - Correlation matrix, covariance matrix
 - Statistical pre-tests (outlier test, test for normal distribution, test for homogeneity of variance)
- Statistical testing
 - Overview test procedures
 - Parametric tests (z-test, t-test)
 - Non-Parametric test procedures (Wilcoxon-test, U-Test)
 - Analysis of variance
 - Post-hoc tests
 - Multiple tests
 - χ^2 -tests
- Curve fitting
 - Regression
 - Questions for regression and correlation
 - Data for linear regressions
 - Linear regression
 - Correlation coefficient
 - The fitted regression line
 - The residual sum of squares
 - The residual standard deviation
 - The coefficient of determination r^2
 - Dose-response analysis
 - Species sensitivity distribution
 - Non-linear regression models
 - Non-linear regression models (2 parameters)
 - Comparison of 2 parameter models
 - Moving average for ECX determination
- Multivariate statistics
 - Multivariate statistics
 - Why & when to use it

- Optimal dimensionality of statistical models
- Dimensionality reduction
- Overview of statistical modelling approaches
- Classification
 - What it does
 - Classification methods
 - How it works
 - Based on distances
 - An example for cluster analysis
- Ordination
 - What it does
 - Ordination methods
 - Principal Component analysis (PCA): How it works
 - Based on covariance/correlation
 - An example for PCA
- Constrained ordination for community analysis: PRC
 - Redundancy Analysis (RDA) as basis
 - Principal response Curve (PRC): What it does
 - PRC: How it works
 - An example for PRC: Mesocosm experiment