Proposed short course: “Improving statistical evaluations in the geosciences”
Convener: Bernard Ludwig

Almost all scientific studies rely to some extent on correct statistical analyses. While statistical software packages for scientists offer great opportunities and provide many powerful tools (e.g., in data mining and exploratory statistics), there are many pitfalls, which may result in wrong or non-reproducible manuscripts. This problem has been known for a long time and has been addressed explicitly in some research fields other than the geosciences. For instance, Ainsworth (2007, Nature 448, 849) stated: "The abysmal standard of statistical analysis in much of genetic epidemiology is little short of scandalous".

This short course aims to address potential problems in geoscientific studies and to reduce the number of non-reproducible studies.

A. Fundamental issues in design of experiments and statistical analyses
The following fundamental issues will be addressed

- Time spent for experimental designs. Advantages and disadvantages of selected experimental designs. Missing randomization. Observational study vs. controlled experiments
- Pseudo-replication vs. true replications and how to deal with it. Wrong model formulations
- “Obsession” with p values: Statistical significance and geoscientific relevance
- Statistical tests: conditions for the application of modelling and hypothesis testing
- Dealing with suspected outliers
- Logistic vs. linear regression
- Number of experimental treatments vs. power of tests. Number of replicates required for predictive modelling
- Use and misuse of correlation analyses
- Investigating and dealing with interactions between factors or predictors

B. Communication issues in geoscientific studies
A number of studies may be of great interest, but the information provided may be simply insufficient for reproducible studies. The following communication issues will be addressed

- Discrepancies between hypotheses stated in an introduction vs. actual hypotheses tested by the tests
- Lack of information on anova and/or regression models: response variables, factors and factor levels, fixed and random effects, predictors, error terms considered, interactions studied
- Lack of information on residual inspection. Missing information on conditions tested for the different tests
- Lack of information on model simplification. Missing information about significant predictors
- Lack of essential information for principal component analyses
- Lack of information on data transformation - e.g., “When necessary, data were transformed” is not that helpful

C. Selected additional issues in geoscientific studies
In some studies, improvements may be possible and the following fields will be addressed.

- Dealing with variance heterogeneity
- Use of contrasts instead of multiple mean testing
- Use of mixed regression and anova models
- Including squared and cubic contributions in models instead of solely relying on linear contributions. Lack of fit

- Box Cox transformation
- Validation or cross-validation instead of a sole focus on calibration.
- Model types

Examples will be shown using the software R