Two-day statistics intensive course IV for soil scientists from March 24th until March 25th 2022 (in English)

IV. Multivariate statistics: principal component analysis, partial least squares regression, cluster analyses, regression trees and rule-based models using R for soil scientists

The media and scientific journals of different disciplines repeatedly address the topic of erroneous research due to insufficient statistical knowledge (see e.g., Ainsworth (2007, Nature 448, 849)). To some extent there may be similar problems in soil science.

Typical problematic fields in the area of multivariate statistics may be (I) a lack of knowledge of the opportunities and limitations of multivariate approaches; and (II) insufficient description of the multivariate analyses in publications.

The intensive course IV aims to improve soil scientists' statistical knowledge. A main objective is to reduce the occurrence of the above-mentioned problems in soil science research and publications.

Date: Intensive course IV: March, 19th 2020 10:00 a.m. until March 20th 6:00 p.m

Location: Kassel University, 37213 Witzenhausen, Nordbahnhofstr. 1a. In case of legal restrictions due to the Covid-19 pandemic, the intensive course will be held as Zoom session.

Costs: The costs for the intensive course IV including catering during coffee breaks (excluding accommodation and meals) are 166.00€. The price stated here is a final price and must be transferred before the start of the course.

Number of participants: The number of participants is restricted to maximal 25.

Laptops: Course participants should bring laptops with R and RStudio software already loaded onto them. If necessary, laptops can be borrowed by prior arrangement.

Materials: Lecture notes (80 pages), exercises and model solutions will be provided.

Recommended literature: Crawley (2012). The R Book. 2nd Ed., Wiley. Wehrens (2020). Chemometrics with R. 2nd Ed., Springer; Everitt & Hothorn (2011). An Introduction to Applied Multivariate Analysis with R. Springer.

Lecturer: Prof. Dr. Bernard Ludwig

Schedule of the intensive course:
Thursday, 24.03.2022

Time	Contents
10:00 - 13:30	Welcome, matrix operations, calculation of eigenvalues and eigenvectors, centering and z-transformation
13:30 - 14:30	Lunch break
14:30 - 16:00	Variance-covariance and correlation, principal component analysis (PCA), calculations, presentations and interpretations
16:00 - 16:10	Coffee break
16:10 - 18:00	Exercises on PCA using R

Friday, 25.03.2022

Time	Contents
10:00 - 13:30	Pretreatment of data (use of the Savitzky-Golay filter), principal component regression (PCR) and partial least squares regression (PLSR)
13:30 - 14:30	Lunch break
14:30 - 16:00	Exercises on PCR and PLSR using R
16:00 - 16:10	Coffee break
16:10 - 18:00	Introduction to partitional and hierarchical clustering. Presentation of regression trees and rule-based models.

Registration is open until March 1st 2022. Registration and general queries: Prof. Bernard Ludwig, Kassel University, bludwig@uni-kassel.de

Please note that the intensive course may be cancelled if not enough registrations have been received. In this case or in the event of cancellation due to illness of the lecturer or any other events beyond the control of the lecturer, there will be no claims possible, except for the reimbursement of participation costs (166.00€).

Additional information on statistical training courses: see

https://www.uni-kassel.de/fb11agrar/en/sections-/-facilities/environmental-chemistry/statistics-courses