1.3.5. Data Analysis in R – Advanced Course (Cologne University of Applied Sciences)

Name of Module/Course	Data Analysis in R - Advanced Course
Short description	In this course you will build on the knowledge you have obtained in the Beginners Course by moving into more advanced statistical analyses of data and the application of machine learning algorithms. Not only will these skills increase the possibilities for what you can achieve in your studies, they are increasingly becoming a requirement to gain employment in many fields. They provide you numerous opportunities for the modern day professional work. This course teaches how to use The R Project for Statistical Computing (commonly known as "R") for advanced data analysis, focusing on advanced statistical analyses and the implementation of machine learning. The intensive course starts at an intermediate level and moves to an advanced level. Please note that the course uses examples and data analysis techniques in the fields of **climate, geography** and **hydrology,** and it is therefore recommend that students in master's courses related to these topics attend.
Name of Programme	Natural Resources Management
Name of University	Cologne University of Applied Sciences
Name of Lecturer	Oscar Manuel Baez Villanueva & Joschka Thurner
Responsible University lecturer	Prof. Lars Ribbe

SWS	**Attendance (h)**	**Self-study (h)**	**Total workload (h)**
-	17.5	22.5	40

Start & end dates	**timeslot:**
Please check the website	7 interactive lectures and 2 sessions allocated for students to work on exercises.
Registration until	**Number of possible AGEP participants**
Please check the website	35
Content and goals of qualification	Objective: for students to build on their knowledge in using the R Project for Statistical Computing and use R for advanced statistical analyses and the application of machine learning algorithms in the fields of climate, geography and hydrology. **Module 1: Functions and parallel computation** 1. Writing and using functions in R 2. Parallel computation in R 3. Exercise – Optimising codes using parallel computation and functions **Module 2: Advanced spatial and temporal statistics** 1. Review of basic spatial and temporal statistics

	2. Methods for downscaling and upscaling raster products 3. Data extraction over areas of interest 4. Statistical trends analyses **Module 3: Machine learning algorithms in R (Part 1)** 1. Introduction to machine learning algorithms 2. Machine learning packages in R 3. Using machine learning for prediction 4. Example - machine learning for streamflow prediction **Module 4: Machine learning algorithms in R (Part 2)** 1. Using machine learning for classification 2. Example - machine learning for land cover classification 3. Example – machine learning for spatial prediction of variables **Module 5: Working with R Markdown to create reports** 1. Introduction to R Markdown 2. Introduction to Leaflet 3. Creating and compiling reports with code in PDF and HTML 4. Example: Creation of a report of spatial results in HTML
Preconditions for participation	Completion of the AGEP course “Data Analysis in R - Beginners Course” (or equivalent knowledge) Knowledge of statistics
Teaching Methods	Online, via Zoom
lesson format (online/face-to-face)	The format will be online and includes: 1. Lectures 2. Exercises 3. Data and scripts to reproduce examples and solve the exercises 4. Questions and Answer (Q&A) sessions
Assessment method	None
language	English
Inscription external student	www.agep-info.de