New Course Code and Title

ES7023 Fundamentals of Data Science for Earth and Environmental Systems Science

Details of Course

Summary of course content
Modeling, simulation, statistical learning and data science methods are powerful tools for earth and environmental systems sciences. This course will cover the major concepts for building and evaluating models, including fundamentals of statistical and machine learning. Topics covered include (1) basic concepts and tools in data science, (2) statistical thinking, (3) goals and principles of scientific modeling, (4) model development, (5) model calibration and selection, (6) sensitivity analysis, (7) model evaluation, (8) model predictions, (9) results visualization and communication. Students will gain hands-on experience in developing models and simulations (using R programming language).

Rationale for introducing this course
This graduate-level course will provide the fundamental knowledge and skills for students to conduct PhD-level research using proper scientific modeling and data analysis processes (data exploration, model development, calibration, selection, sensitivity analysis, multi-model combination, visualization, etc). Through individual projects, it will further enable students to jump-start their PhD research through active analysis of their own data.

Aims and objectives
The aim of the course is to provide graduate students with the fundamental knowledge and skills to conduct scientific modeling in earth and environmental system sciences. Each student will conduct bi-weekly assignment, along with an individual project of their choosing. The project will be an opportunity to engage deeply with the modeling process including using methods from outside the course material, geared towards their individual research interest.

Syllabus
- Introduction to course, algorithms, programing, statistical learning
  Basic Review of statistics, introduction to R
- Goal and principles of scientific modeling; modeling examples
- Designing and programing models (statistical and machine learning)
- Model Calibration and Selection
- Sensitivity Analysis
- Model Evaluation
- Model Predictions
- Model Uncertainties
- Multi-model combination
- Visualization
- Ethical data science and AI applications

Assessment

Engagement with class and online forum 10%
Coding and report assignments 60%
Project 30%
(All assessments will be individually evaluated)
Total: 100%

Final Exam Duration
NA

Hours of Contact/Academic Units
39 hours / 3 AU

To be offered with effect from
(state Academic Year and Semester)
AY2018 Sem 2