Utrecht University
Department Human Geography & Spatial Planning
Department Physical Geography

INFOMSDASM
Spatial data analysis and simulation modelling

Applied Data Science (Master program)

Course manual
2020/2021

Simon Scheider
Derek Karssenberg

Spatial data analysis and simulation modelling

Contents

Aims of the course

The course teaches basic methods for comprehending, accessing and analyzing spatial information, as well as methods for model simulations of spatial systems integrated with spatial data. Students will learn theory and practice of accessing geodata sources, using geodata models, and how to transform geodata for spatial analysis purposes. This includes both raster and vector models as well as spatial networks. In the second half of the course, this will be extended towards spatio-temporal data, the analysis of this data, and simulation modelling which entails the encoding of mechanisms that steer temporal changes in the geographical domain. These models will be integrated with spatio-temporal data through calibration and model data assimilation techniques. The course includes theory related components as well as hands-on components teaching students how to use data science tools and how to program with geodata. The knowledge and experience acquired during the course will be applied in a case study executed at the end of the course. The learned methods can be applied in data science across various empirical domains, including Geoscience, Social Science, as well as environmental epidemiology and Life Science.

Requirements and organization

Students need to have successfully finished the course "Data Wrangling and Data Analysis" (INFOMDWR) within the Applied Data Science (ADS) Master program.

The course is given in period 2 (starting from week 46 2020 (10th November) until week 5 2021 (4th February)), in the time slots B (Tuesday morning 9:00 - 13:00 and Thursday afternoon 13:15 - 18:00). In addition to these supervised hours, students are expected to self-study with the course material, including reading, writing and finalizing assignments (20 hours in total per week).

Learning objectives

The learning goals of this course are as follows:

1. Understand geo-analytic concepts and models of geodata and apply basic computational methods of geo-spatial data analysis
2. Choose appropriate geo-analytic models and methods for spatial analytical tasks
3. Understand concepts, design and apply models for field-based and agent-based simulation
4. Understand spatial error models and apply them for data calibration
5. Analyze existing studies applying spatial data analysis and simulation modeling

Syllabus and mandatory readings

The first half of the course (week 1-4) is an introduction to geo-spatial data analysis. This includes the following *course modules* covered by corresponding lectures and labs (using Python 3.7 with QGIS and other libraries):

1. Geodata sources: Reference systems, data sources and geodata quality
    - Principles of reference systems, and how to transform them into each other
    - Geodata formats and retrieval possibilities, including OGC interface standards and geodata Web APIs
    - Assess the different dimensions of geodata quality, including spatial resolution, spatial accuracy, spatial extent and completeness
2. Geodata models: Core concepts, geodata types, measurement levels, basic geometry and spatial networks
    - Distinguish forms of geoinformation (objects, fields, networks, events and tracks) and their representation as geodata types (raster and vector models) and how they impact spatial analysis
    - Distinguish levels of measurement and how they impact cartographic mapping
    - Basic geometric models and manipulations
    - Spatial databases
3. Basic geo-analytical transformations:
    - Spatial overlay and map algebra
    - Distance-based analysis
    - Spatial network analysis

The second half (week 5-8) is an introduction to spatial simulation modeling. This includes the following content lectures and labs (using Python 3.7 with PCRaster, Campo, and other libraries):

4. Field-based simulation modeling
    - Map Algebra with raster maps
    - Local (point) models
    - Neighbourhood interaction
5. Agent-based simulation modeling
    - Decision making agents
    - Field-agents
    - Agent interaction
6. Integration of data and simulation models
    - Objective functions
    - Calibration algorithms and tools

To study these contents, **detailed mandatory readings will be given for each week on Blackboard**. In general, we will make use of the following books:

- *Chrisman 2002: Exploring Geographic Information systems, 2nd edition*
- *Burrough, McDonnell, Lloyd 2015: Principles of Geographical Information Systems, 3rd edition*
- *Laurini, Thompson 1992: Fundamentals of spatial information systems*
- *Monmonier 1996: How to lie with maps, 2nd edition*

Learning activities

The learning activities were designed around *pre-defined case studies* that will be picked by students (groups of 3) in the second half of the course, and which will be the basis for a *short paper assignment* and the *case study project* in the last two weeks. The case study is presented at the end of the course. The main course content is divided into 6 course modules (see syllabus) during the first 8 weeks. Each week there will be some *online lectures* (pre-recorded) about some module with corresponding *mandatory readings*. *Computer labs* are self-contained and apply the concepts introduced in the lectures/readings to some practical examples of spatial analysis and simulation modeling. Note that the course is held online, so that all contact moments are held on *Microsoft Teams (see Blackboard)*.

Online lectures and readings (every week)

Lectures consist of a number of pre-recorded videos, which need to be studied during a week by students on their own, together with mandatory readings. To discuss this content, there will be a corresponding *Q&A session* on Microsoft Teams (see Blackboard), in which students can ask questions about the presented material of both lectures and readings on the topic scheduled for the particular week. **For each week, students need to study the materials in advance of the Q&A session scheduled for that week.**

Labs with assignments (every week)

GIS and simulation modelling exercises need to be done for every lab based on some *lab manual*. The manuals also describe *an assignment* that needs to be submitted on Blackboard after the lab (at the end of the same week the lab is given). For a number of labs in the second part of the course students answer questions as part of the lab by entering their answer in Blackboard and no separate assignment document needs to be uploaded after the course - this will be explained at the start of these labs. The lab is not graded but is checked for sufficiency and a sufficient result is a requirement to pass the course. Students have time to work on labs during supervised hours (Tuesday morning and Thursday afternoon), as well as during unsupervised hours. **Submissions of assignments are obligatory and need to be uploaded to blackboard in the same week as the lab.**

Short paper assignment

Student groups of 3 need to write a *short paper* on some topic selected by students from a set of predefined topics (see last section). On the chosen topic, the group will write a *2-pager* (max. 1200 words excluding bibliography) based on self-chosen existing literature. This should include 2-3 articles on main concepts and *existing spatial data analysis or simulation study*. The short paper should include the following parts:

1. Problem (A methodological problem from data analysis and simulation modelling)
2. Concepts (Main concepts needed to understand the problem)
3. Approach (How concepts were applied in the chosen study)
4. Discussion (Remaining problems and improvements)

Submission is done via Blackboard. Two (written) feedback moments are planned for the short paper:

- Submission of *preliminary version* (Problem) in the week before Christmas (18th December), with *feedback by 11 January*
- Submission of *first version* (including Problem, Concepts, Approach) on the 14th of January, with *feedback* until 19th January
- Submission of *final version* on the 4th of February

Case study project

The case study is done (preferably) on the same chosen topic as the short paper and in the same **groups of 3** students in the last two weeks of the course. It consists of a small research/data science project with existing (provided) geodata sources. Students write a scientific report (max. 2400 words excluding bibliography) on the case study which is uploaded to Blackboard. There will be an intermediate presentation of the case study in the first case study week *on the 28th of January.* The final report is presented on the last day of the course (*4th February*).

Schedule and program

Week	Date	Lab B1-B2 (Tuesday 9:00-13:00)	Date	Lecture Q&A (Thursday 13:15-14:00)	Lab B3-B5 (Thursday 14:15-18:00)	Course module
46 2020	10.11.2020	Introduction (online) Lab 1.1 (retrieval and CRS)	12.11.2020	Lecture geodata sources	Lab 1.2 (data quality)	Geodata sources
47 2020	17.11.2020	Lab 2.1 (core concepts, data models and measurement levels)	19.11.2020	Lecture geodata models	Lab 2.2 (geometric manipulation)	Geodata models
48 2020	24.11.2020	Lab 2.3 (spatial databases)	26.11.2020	Lecture spatial databases	Lab 3.1 (overlay and map algebra)	
49 2020	01.12.2020	Lab 3.2 (distance based analysis)	03.12.2020	Lecture geo-analytical transformations	Lab 3.3 (network analysis)	Basic geo-analytical transformations
50 2020	08.12.2020	Lab 4 (field based model)	10.12.2020	Lecture field-based simulation modeling	Lab 4 (field based model)	Field-based simulation modeling
51 2020	15.12.2020	Lab 5 (agent based model)	17.12.2020	Lecture agent-based simulation modeling	Lab 5 (agent based model) preliminary version short paper	Agent-based simulation modeling
2 2021	12.01.2021	Lab 6 (model calibration)	14.01.2021	Lecture model calibration	Lab 6 (model calibration), first version short paper	
3 2021	19.01.2021	Lab 6 (model calibration)	21.01.2021	Exam		Model calibration
4 2021	26.01.2021	Case study introductions	28.01.2021	Case study		
5 2021	02.02.2021	Case study	04.02.2021	Case study presentations	submission report/ short paper	Case study

Note: The latest schedule is available on Blackboard.

Important dates and due dates:

- Short paper preliminary version (18 December 2020)
- Short paper; first full version January 14, 2021, final version February 04, 2021.
- Open book exam; January 21, 2021, (time 1 ½ h) covering questions for each module
- Case study proposal (Tuesday 26.01) in the form of a short presentation (5 min + discussion)
- Case study presentations (04.02) in two groups I and II
- Case study report; February 04, 2021

Assessments: what needs to be done to pass the course

1. *Computer lab assignments* (submission satisfying minimal requirements, for every lab)
2. *Short paper assignments*. Mark is given for the revised paper to be submitted by the end of the course, using criteria for academic research papers.
3. *Open book exam* on the theory (consisting of mandatory readings, e-lecture content, computer labs) is written online (on Blackboard) on the *21st of January 2020*.
4. *Case study report*. Mark is given using criteria for academic research papers, to be submitted by the end of the course.

Note: for the short paper and the case study, students form a *thematic group of 3 persons*. See Blackboard for details.

The *final mark M* is calculated as:

M = 0.15A + 0.5 B + 0.25C + 0.1 D

with the following marks: A, short paper assignment; B, exam; C, case study report; D, oral presentation on case study. A, B, C, and D are not rounded.

Absence (for instance as a result of illness or family circumstances) during the exam must be agreed with the coordinator of the course in advance by phone or email. You need to hand over a sick note (medical certificate from your doctor) afterwards to get access to a resit.

The course has been passed if the final grade is >=6 and all obligations have been fulfilled. If not, and only if 1) the final grade is 4.00 or higher and 2) all obligations have been fulfilled, a supplementary test (or repeat exam) can be offered. If the supplementary test has been successfully passed, the final grade of the course will be 6.0. However, students who didn't participate in the first exam due to unforeseen circumstances are exempted from this rule.

For details on the above and further information, see the OER (Education and Examination Regulations) of Natural Sciences.

Table 1: Matrix of how assessments test learning objectives

Learnin g obj. Assess -ments	1 Geo-analytic concepts	2 Spatial simulation concepts	3 Model calibration	4 Understand existing studies	5 Choose appropriate methods
Lectures					
Labs 1-3					
Labs 4-6					
Short paper					
Case study					
Exam					

Case study and short paper topics

The topics will be described on Blackboard in due time. Please keep an eye on the updates there.