Cornell University
PLSCS/NTRES 6200
Spatial Modelling and Analysis
for agronomic, natural resources, environmental and regional studies

Spring semester 2021
Course information

D G Rossiter
d.g.rossiter@cornell.edu
http://www.css.cornell.edu/faculty/dgr2/

November 30, 2020
Orientation to course

1. for whom?

2. objectives

3. instructor

4. method

5. assignments & grading

6. project

7. schedule

8. some key references
For whom?

- **Graduate** and **senior undergraduate** students who want to advance their understanding of geographic information science and technology applied to problems in the agronomic, natural resources, environmental or regional studies sciences;
  
  - Student should have followed an undergraduate-level or introductory graduate-level courses in inferential statistics (e.g., STSCI 3200/BTRY 3020)
  - Helpful but not necessary: remote sensing, computer programming, GIS theory and practice (e.g., PLSCS 4200)

- Especially useful for students starting a **graduate research project** that uses spatial information in agronomy, soil science, natural resource management, regional science, hydrology, ecology . . .
  
  - “spatial”: information is from **known locations** with **known spatial extent**
Learning outcomes

• Student is able to analyze complex spatial problems with appropriate theory, statistical methods and computational tools
  – No “silver bullet”, each problem must be analyzed on its own terms
  – Always use domain knowledge and consider application requirements

• Student is able to organize data analysis as reproducible research (“literate data analysis”)

• Student is competent to process and visualize spatial data; with emphasis on open-source computer programs and publically-available data
  – Most labs use the the R environment for data analysis, (spatial) statistical computing and visualization,
  – One lab uses the interactive GeoDA spatial analysis program

[https://spatial.uchicago.edu/geoda]
Instructor

David G. Rossiter

- Adjunct Associate Professor CALS (10th year)
  - Graduate faculty *Soil & Crop Sciences* ‘Environmental Information Systems’ concentration
  - Graduate faculty *Regional Science* ‘Environmental Studies’ concentration

- Guest Researcher (*Gastmedewerker*), ISRIC–World Soil Information, Wageningen (NL) (6th year)

- Visiting Professor (客座教授), Chinese Academy of Sciences, Soil Science Research Institute Nanjing (中国科学院南京土壤研究所) (8th year – postponed due to COVID-19)

- Visiting Professor (讲座教授), Nanjing Normal University, School of Geography (南京师范大学地理学学院) (6th year – postponed due to COVID-19)

- Retired from University of Twente (NL), Faculty of Geoinformation Science & Earth Observation (after 17 years)
Instruction method

- Graduate-level course, emphasis is on **guided self-instruction**, including critical reading of primary literature, lab. self-paced tutorials

- Lectures are **overviews / orientations** to introduce **fundamental spatial analysis concepts and methods**
  - Each lecture period begins with \( \approx 20 \) minutes discussion of an **assigned pre-class reading**, and with Q&A.

- Emphasise ability to read, understand and apply **methods in journal papers**, advanced **textbooks** and **reference** books; necessary in graduate career

- Student **project** allows you to apply ideas and techniques of spatial analysis from this course or elsewhere to your research

- First nine (of 14) lab. periods are tutorial exercises on **methods** and **computation**; example applications to illustrate these; other lab. periods for project work and presentation.
Assignments

• Weeks 1-10: nine (9) set computer lab. exercises, small hand-in assignment; due before Tues. of following week

• Weeks 11-13 three (3) short question sets based on reading; due before Tues. of following week

• Week 14: no assignment
Introduction to course

Grading

• Course is 4 credits

• Nine lab. exercise assignments: graded 0–3; 45% of final
  – 0 = not submitted; 1 = weak attempt; 2 = some incorrect answers or procedures; 3 = (close to) perfect

• Three question sets: graded 0–3; 15% of final

• Project: 40% of final (breakdown: 80% project, 10% reporting, 10% presentation/discussion)

• Grading on absolute scale; Letter equivalents:

<table>
<thead>
<tr>
<th>≥ 97.5</th>
<th>A+</th>
<th>≥ 92.5</th>
<th>A</th>
<th>≥ 90.0</th>
<th>A-</th>
</tr>
</thead>
<tbody>
<tr>
<td>≥ 87.5</td>
<td>B+</td>
<td>≥ 82.5</td>
<td>B</td>
<td>≥ 80.0</td>
<td>B-</td>
</tr>
<tr>
<td>≥ 77.5</td>
<td>C+</td>
<td>≥ 72.5</td>
<td>C</td>
<td>≥ 70.0</td>
<td>C-</td>
</tr>
<tr>
<td>≥ 67.5</td>
<td>D+</td>
<td>≥ 62.5</td>
<td>D</td>
<td>≥ 60.0</td>
<td>D-</td>
</tr>
<tr>
<td>&lt; 60.0</td>
<td>F</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
**Project**

• A small project in which the student applies **spatial analysis** to a **problem of interest** relevant to student’s field of study;
  - Usually part of students’ graduate research or senior thesis.
  - May work in pairs if independent contributions can be evaluated.

• **Spatial** or **Spatio-temporal** analysis must be prominent;

• Prefer use of **own** data, **similar** obtained from colleague/advisor, or **publically-available** data; little credit for extensive data manipulation (not the main purpose of this course);

• Should be able to complete during allocated lab. time and related self-study time; graded as such;

• Milestones: (1) **brief proposal**: 16-March (week 6 Tuesday); (2) **proposal presentation** (literature review / detailed project plan): 6 April (week 9 Tuesday); (3) presentation of preliminary results (11 or 12 May, week 14 Tu/Wed); (4) **report**: ≈24-May (on scheduled exam date, TBD)
Week schedule

**Lectures:** Tuesday, Thursday 0805–0920. In person Plant Science 404, also streaming on-line; lectures recorded and subtitled on Canvas

- Content is equivalent to a 50-minute lecture, this gives more time for review and interaction.

- Start with questions and discussion of the key points in the assigned papers, as motivation for the lecture.

- On some Thursdays discuss a paper which uses techniques covered during the week.

...
Computer laboratory: Tuesday or Wednesday 1330-1420 or 1440-1530, Mann Library 409 (first 50 minutes) and on-line, with on-line consultation for remainder of week.

- you should select one labs, because of room occupation restrictions (8 maximum), but if there is space you can come to an additional lab.

- brief orientation; most time is self-paced with instructor available to answer questions; a few breaks to review key points

- use your own computer
  - any standard Mac OS/X, Windows, Linux system is suitable
  - most labs. use the R Project for Statistical Computing in the RStudio environment, one lab with GeoDA.

- discussion forum for asynchronous problem-solving and suggestions
Schedule – lectures

Week 1  Naïve analysis; The R environment; Literate data analysis; Spatial concepts; Coördinate Reference Systems

Week 2  The Universal Model of Spatial Variation; Local spatial dependence, Ordinary Kriging (OK);

Week 3  Model-based methods: Feature-space regression, trend surfaces

Week 4  Data-driven methods (“machine learning”)

Week 5  Remote sensing as a data source; PCA (Th only)

Week 6  Big data, open data

Week 7  Area-Based Spatial Data Analysis

...
Week 8  Time series analysis; Spatio-temporal analysis

Week 9  Point Pattern analysis

Week 10  Spatial sampling; clustering

Week 11  Uncertainty, data quality, metadata;

Week 12  Geospatial simulation

Week 13  Bayesian methods for spatial analysis

Week 14  Summary: newest developments in spatial analysis
Schedule – labs

**Week 1**  The R environment; Literate Data Analysis

**Week 2**  Geostatistics with R/gstat

**Week 3**  Trend surfaces, spatial regression

**Week 4**  Data-driven methods: classification & regression trees; random forests; optimization

**Week 5**  (no lab)

**Week 6**  Google Earth Engine

**Week 7**  Area spatial data analysis: GeoDa

**Week 8**  Time series analysis; Spatio-temporal analysis

...
Week 9  Point-pattern analysis

Week 10  Spatial sampling; clustering

Weeks 11–14  Individual project work (instructor available for consultation)

Week 15  Student project (ongoing) presentations/discussion
Key texts/references

Cornell has access to most of the e-books.

**Spatial modelling in general**


**Ecological modelling**


Statistics review

Electronic Statistics Textbook
https://docs.tibco.com/data-science/textbook


Spatial analysis in R

Datasets etc. at http://www.asdar-book.org/

https://doi.org/10.32614/RJ-2018-009

https://rspatial.org/
General GIS


Geostatistics


Statistical modelling


Machine learning


Land surface modelling

Spatial sampling


Time series, spatio-temporal


Data science

https://r4ds.had.co.nz

Ahmed, Zia (2020). Geospatial Data Science in R.
https://zia207.github.io/geospatial-r-github.io/index.html