KTH Architecture and
the Built Environment

Spatial data analysis in practice - SDAP

7,5 credits

Level: PhD course
Course code FAG3170
Department: Urban planning and Built Environment, ABE, KTH
Time: April-June 2017
Number of students: Min. 10, Max. 25
Cost: SEK 8 000 students
Sponsored by: Safeplaces network

1. Motivation

Spatial data analysis involves exploratory and confirmatory analyses that facilitate and enhance decision making with spatial data. This includes techniques that help visualize data, identify patterns and processes, detect outliers and anomalies, test hypotheses and theories, and generate new spatial data and knowledge. It is therefore not a surprise that spatial data analysis is an emerging field, with advancing technologies and evolving applications that are fundamental for a number of sciences, geography, architecture and urban planning, transportation, criminology, demography, epidemiology and economics, just to name a few. Mobile data and location based services, from apps, allied to Geographical Information Systems (GIS) provide the necessary setting for enhanced analysis of patterns and process over time and space. The KTH course *Spatial Data Analysis in Practice - SDAP* (course code FAG3170) offers examples of conceptual and applied research on spatial analysis capturing some of the most recent developments in this area.

2. Course aim and learning outcomes

Spatial data analysis in practice involves an encompassing set of skills that includes manipulation of spatial data, exploration of spatial statistics techniques, modelling in combination with Geographic Information Systems. In this course students are trained to become users of spatial data analysis techniques. Students will gain a broad knowledge of the diversity of current approaches, which methods are at hand and examples of applications using spatial data analysis in different fields (e.g.,). After completing the course the students should be able to:

- identify the appropriate approaches/techniques in spatial data analysis
- use relevant knowledge to solve spatial-related problems using real-life data sets and spatial statistical tools, including visualization, interpolation, pattern identification and modeling (spatial regression analysis)
- develop both technical and social skills by working in pairs to solve real-life problems using different statistical software

- to analyze results of practical exercises and be able to point out challenges and advantages with those tested techniques
- develop, interpret and critically reflect upon results of a case study using one (or more) spatial data analysis technique(s) learned during the course.
- be able to use their new skills in spatial data techniques and communicating them to an audience (written, graphically and orally).
- recognize and express the value of incorporating the spatial dimension of phenomena and processes in social sciences

2. Main content

The course is composed of 15 lectures divided in 3 parts. In the first part the nature of the geographical data is discussed whilst the identification of spatial patterns is the focus in the second part of the course together with the theoretical introduction to confirmatory spatial data analysis. The third part deals with examples of applications spatial data analysis/visualization of large share of data and development of the final project.

PART I - Thinking spatially: introduction to GIScience/The nature of spatial data, Data quality/Spatial structure of spatial data/Exploratory spatial data analysis

PART II – Area interpolation, ESDA cluster detection methods, Introduction to confirmatory analysis, OLS regression analysis and spatial modelling/autoregressive models.

PART III – Examples of applications spatial data analysis/ visualization of large share of data/Glimpses of future of spatial data analysis/Project (study case) and Project presentation.

3. The structure of the course and learning activities

The course is composed of **lectures** followed by **practical exercises**[1], about 1 hour long after the lecture. Practical exercises provide contextualized step-by-step instructions to students so they can apply the appropriate approaches to real-life datasets and obtain the results that can be later used for their analysis. Real-life examples make learning more interesting for the students simply by giving life to theory which is far too often dry and abstract. A 'forum time' when students may raise issues (after each practical exercise but before each new lecture) will be part of the daily schedule (about 15 minutes). These activities provide feedback to the students as they learn.

Reading is an essential part of the course. Reading the required literature before each lecture is highly encouraged to all students.

The final assessment is based on the development of a project. In groups of two, students are trained to "make a case" of why they want to test a certain type of technique using real-life data sets (<u>case study project</u>), with hypothesis testing. Students need to show that they are using the technique properly (according to intended learning outcomes) and provide results. Students have also to think critically about their results and chosen technique. If the technique or data were not adequate, why

[1] You are encouraged to bring your own laptop since two software packages used in this course are free-of-charge and can be downloaded from the web.

did they fail? Examples of alternative approaches (what could be the next step when results show 'unexpected patterns') are to be discussed as an integral part of the course. A written report should be handed in to the teacher. Students are encouraged to use your own pc to develop the final project.

Results of the case study should be shown by the students at the department as posters at the end of the course – text of the case studies should be condensed – students will stand next to the posters and be prepared to present their case study and answer questions.

4. Lecturers

Two top rank experts in this area will be lecturing in the course: Prof Robert Haining, Emeritus of the Department of Geography, University of Cambridge, UK and Prof Luc Anselin, Director, Center for Spatial Data Science, The University of Chicago, USA (http://spatial.uchicago.edu). The course will also include examples of applications in several disciplines, including urban planning, real estate and other social disciplines (e.g.. Manchester University, Uppsala University, Karolinska Institute).

The head teacher of the course is Vania Ceccato, Department of Urban Planning and Built Environment, ABE, at Royal Institute of Technology (KTH).

5. Literature

The readings required for the course are a book by Haining 2003 *Spatial data analysis: theory and practice,* a collection of articles, and excerpts from books. All lectures require pre-reading. Literature will be available at BILDA. The first chapters in Haining (2003) and Bivand et al. (2013) should be read by the students before the first class (available in BILDA three weeks before the course starts).

6. Language of instruction

The course will be taught in English.

7. Prerequisites and expectations

Anyone who is a PhD student in any relevant subject area (e.g., Urban and Regional Planning, Economics, Real Estate, Geography, Demography, Criminology, Environmental Sciences) is eligible to take this course. However, having knowledge in Geographical Information Systems and/or basic statistics is an advantage. The first week provides students with basic introduction to the course and tools. The first chapters in Haining (2003) should be read by the students before the first class (available in BILDA three weeks before the course starts).

8. Assessment

In order to pass the course, students need to:

- attend lectures and perform practical exercises
- participate in discussion in class ("forum-time" and hand in results on time)
- participate in and contribute to the work with the case study in group (maximum 2 individuals) that involves the development, the interpretation and critical reflection of results

using spatial data analysis techniques learned during the course. Use of own datasets is encouraged.

- Present their case study in a poster session at the department (written, graphically and orally). Department personnel (researchers, PhD students) and peers assess each group performance together with the content presented in the poster. The written report is handed in to the head teacher.

9. Grades

Pass/Fail.

10. References: a selection

Anselin, L. and Griffith, D. A. (1988) Do spatial effects really matter in regression analysis? *Papers of the Regional Science Association*, 65: 11–34.

Anselin, L., Florax, R.J.G.M., and Rey, S.J. (2004) *Advances in spatial econometrics eethodology, Tools and Applications*. Springer.

Anselin, L., Sridharan, S., and Gholston, S. (2007) Using exploratory spatial data analysis to leverage social indicator databases: the discovery of interesting patterns. *Social Indicators Research*, 82: 287–309.

Bailey, T.C., Gatrell, A.C. (1996*) Interactive spatial data analysis*. Harlow: Longman.

Bivand, R. S., Pebesma, E., & Gómez-Rubio, V. (2013). Hello World: Introducing Spatial Data Applied Spatial Data Analysis with R (pp. 1-16). New York, NY: Springer New York.

Bivand, R. (2007) Regional growth in Western Europe: an empirical exploration of interactions with agriculture and agricultural policy. In: *European Regional Growth*, Berlin: Springer.

Ceccato, V. (2009) Crime in a city in transition: the case of Tallinn, Estonia. *Urban Studies*, 46:1593-1610.

Ceccato, V., Oberwittler, D. (2008) Comparing spatial patterns of robbery: evidence from a Western and a Eastern European city. *Cities*, 25: 185–196.

Cressie, N (2015) Statistics for spatial data (Revised edition). New York, Chichester, Toronto, Brisbane:Wiley.

Fischer, M. M., Getis, A. (eds.) (2010) *Handbook of applied spatial analysis: software tools, methods and applications*. Berlin: Springer-Verlag.

Gatrell, A. and Bailey, T. (1996) Interactive spatial data analysis in medical geography. *Social Science & Medicine*, 42: 843–855.

Gatrell, A., Bailey, T., Diggle, P. and Rowlingson, B. (1996) Spatial point pattern analysis and its application in geographical epidemiology. *Transactions of the Institute of British Geographers*, 21: 256–274.

Haining, R. P. (1990) *Spatial data analysis in the social and environmental sciences*. Cambridge: Cambridge University Press.

Haining, R. (2003) *Spatial Data Analysis: Theory and Practice*. Cambridge: Cambridge University press.

Haining, R, and A Cliff. (2003) Using a Scan statistic to map the incidence of an infectious disease: measles in the USA 1962-1995. Paper read at Proceedings of the Geomed conference 2001, at Paris.

12. Number of students and commitment

The course is designed for maximum 25 students. The course will not be offered for less than 10 students.

This is an intensive course so you have to allocate time for lectures, labs and final project. Only register to the course if you feel you can commit yourself fully to the course.

You are welcome to send an email to asifai@kth.se expressing your interest course. We will get back to you latest late February.

---

13. Schedule *(preliminary, Updated 14 March 2017)*

Lecture 1	6th april	Nature spatial Data	Vania Ceccato
Lecture 2	6th april	Data quality	Vania Ceccato
Lecture 3	7th April	ESDA	Vania Ceccato
Lecture 4	18th April	ESDA-cluster	Bob Haining
Lecture 5	19th April	Cluster-Temporal	Bob Haining
Lecture 6	20th April	Ols	Bob Haining
Lecture 7	24th April	Spatial modelling	Bob Haining
Lecture 8	25Th April	Spatial modelling	Bob Haining
Lecture 9	11 May	Advanced topic I	Luc Anselin
Lecture 10	12 May	Advanced topic I	Luc Anselin
Lecture 11	18th May	Application 1	Vania Ceccato
Lecture 13	19th May	Application 2	Reka Solymosi
Lecture 14	25th May	Application 3	Mats/annat
Lecture 15	26th May	Application 5	John Östh
Project execution	29th May-11th June		
Project presentation	12th June		