

Handbook of spatial analysis

The 11th European Forum for Geography and Statistics conference
Helsinki, Finland October 16 - October 18

Loonis Vincent

October 17, 2018



Introduction

- Insee has undertaken the writing of a handbook of spatial analysis.
- The objectives are :
 - to propose a list and the description of the statistical methods which could be implemented only with a point based system,
 - to disseminate, promote and share the results among NSIs and NMAs,
 - to promote the application of spatial statistics into the statistical production process.
- The handbook is supported by Eurostat and the EFGS,
 - A foreword co-signed by
 - Mariana Kotzeva: Director General of Eurostat,
 - Janusz Dygaszewicz, President of EFGS.

Insee Méthodes



N° 131
Octobre 2018

Manuel d'analyse spatiale

Théorie et mise en oeuvre pratique avec R
Insee - Eurostat
Sous la direction de Vincent LOONIS
Coordonné par Marie-Pierre de BELLEFON



eurostat



Insee Méthodes



N° 131
October 2018

Handbook of Spatial Analysis

Theory and Application with R
Insee - Eurostat
Directed by Vincent LOONIS - Coordinated by Marie-Pierre de BELLEFON



eurostat



- Why a new handbook ?
- Table of contents,
- Building bridges inside Insee, and with the academic community,
- Going further: dissemination, capacity building.

Why a new handbook ?

- A lot of initiatives have already been undertaken to
 - better integrate spatial and statistical information (GSGF, GEOSTAT 2, Geostat 3)
 - foster collaboration between NMCA and NSI (UN GGIM, EFGS)
 - better disseminate geocoded statistical information (grid system, data vizualisation...)
- Insee benefits from these initiatives to reshape its geocoding system, in collaboration with IGN.
- Insee wanted to be me more proactive while taking advantage of its specificity : the economic and social studies.
 - The starting point of the handbook is an existing point based system.
 - The handbook raises the question:
How can such a point based system be useful to improve the economic or social studies or the statistical production process ?

Why a new handbook ?

The existing handbooks are

- Theoretical: Cressie (1992), Anselin (2013), Gelfand et al. (2010), LeSage and Pace (2009);
- Practical: Bivand et al. (2008), Brunsdon and Comber (2015)
- Theoretical and practical: Fischer and Getis (2009), Haining and Haining (2003), Schabenberger and Gotway (2017).

Insee's Handbook

- is theoretical and practical (implementation with R),
- focuses on NSIs specific topics, spatial sampling, small area statistics, geographical confidentiality,
- does not claim to be comprehensive.

The handbook consists of 4 parts and 14 chapters

- 1 Describing geocoded data,
- 2 Measuring the importance of spatial effects,
- 3 Taking spatial effects into account,
- 4 Extensions,

Describing geocoded data

The first part aims at describing how to gather statistical units according to their location.

- Chapter 1 **Descriptive spatial analysis** explains how data can be taken up with R and introduces some concepts of maps semiotics.
- Chapter 2 **Codyfying the neighbourhood structure** helps define the actual spatial interaction between statistical units.

Measuring the importance of spatial effects

Everything is related to everything else, but near things are more related than distant thing (First Tobler's law of geography).

Part II introduces 3 categories according of to their **Data Generating Process (DGP)**:

- The location of the statistical units is fixed for **areal data**. Chapter 3 introduces then global or local **Spatial auto-correlation indices**.
- Chapter 4 deals with **Spatial distribution of points**, whose location is a random variable. The goal is to highlight attractiveness or repulsiveness between points.
- chapter 5 **Geostatistics** deals with data whose spatial support is continuous.

Taking spatial effects into account

Most of the data managed by NSIs are areal data. We consider such data with spatial interaction.

- Chapter 6 **Spatial econometrics**, chapter 7 **Spatial econometrics on panel data** models the spatial interaction related to the explanatory and unobserved variables.
- Heterogeneity means that the strength of relation varies with the location of the observation. Chapter 8 **spatial smoothing** and chapter 9 **Geographically-Weighted regression** deal with heterogeneity.

Taking spatial effects into account

So far the previous methods are meant to be used with exhaustive and comprehensive data. **What happens for survey data ?**

- Chapter 10 **Spatial sampling** explains how to benefit from a geocoded sample frame to select an efficient sample.
- Chapter 11 presents the potential pitfalls when using **Spatial econometrics on survey data**.
- Chapter 12 Explains how taking spatial correlation into account can improve **Estimation on small areas**.

- Chapter 13 introduces **Graph analysis**
- Chapter 14 **Confidentiality of spatial data** explains how to assess and manage the risk of disclosure while preserving spatial autocorrelation.

Building bridges

Around 40 people contributed to the handbook as authors or reviewers. The writing has been of very good opportunity to build bridges:

- Within Insee
 - 18 authors from various departments
- with the academic community
 - 7 authors from various universities and schools of statistics.
- with both of them, and even more
 - A workshop organized in Insee premises, to present the handbook, gathered more than 100 people.

Going further

The writing, the proofreading, the printing of the French and the English versions are now finished.

- 250 copies of the French version have already been sent to various French official bodies.
- 100 copies of the English have to be sent to various official non French bodies.
- The French version will be available for free on Insee's website next week.
- The English version will be available for free on Eurostat and EFGS websites, in the coming weeks.
- Insee is setting up a training course relying on the handbook.

Going further

If you are interested in receiving one of 100 English copies, feel free to send an email subject to the following constraints :

- Send it to vincent.loonis@insee.fr
- Use the word *handbook* as keyword in the subject of you email.
- Only one copy by organization (NSI, NMCA, international bodies)
- Describe shortly but precisely your post address
 - First name, Last name
 - Organization
 - Address
 - Country

References I

- Anselin, L. (2013). *Spatial econometrics: methods and models*, volume 4. Springer Science & Business Media.
- Bivand, R. S., Pebesma, E. J., Gomez-Rubio, V., and Pebesma, E. J. (2008). *Applied spatial data analysis with R*, volume 747248717. Springer.
- Brunsdon, C. and Comber, L. (2015). *An introduction to R for spatial analysis and mapping*. Sage.
- Cressie, N. (1992). Statistics for spatial data. *Terra Nova*, 4(5):613–617.
- Fischer, M. M. and Getis, A. (2009). *Handbook of applied spatial analysis: software tools, methods and applications*. Springer Science & Business Media.
- Gelfand, A. E., Diggle, P., Guttorp, P., and Fuentes, M. (2010). *Handbook of spatial statistics*. CRC press.

References II

- Haining, R. and Haining, R. P. (2003). *Spatial data analysis: theory and practice*. Cambridge University Press.
- LeSage, J. and Pace, R. K. (2009). *Introduction to spatial econometrics*. Chapman and Hall/CRC.
- Schabenberger, O. and Gotway, C. A. (2017). *Statistical methods for spatial data analysis*. CRC press.