Spatialdata: a scala library for spatial sensitivity analysis

J. Raimbault*[1,2,3], R. Reuillon[2] and J. Perret[4,2]

[1]Center for Advanced Spatial Analysis, University College London
[2]UPS CNRS 3611 ISC-PIF
[3]UMR CNRS 8504 Géographie-cités
[4]LaSTIG STRUDEL, IGN, ENSG, Univ. Paris-Est

Summary

The sensitivity analysis and validation of simulation models require specific approaches in the case of spatial models. We describe a scala library providing such tools, including synthetic generators for urban configurations at different scales, spatial networks, and spatial point processes. These can be used to parametrize geosimulation models on synthetic configurations, and evaluate the sensitivity of model outcomes to spatial configuration. The library also includes methods to perturbate real data, and spatial statistics indicators, urban form indicators, and network indicators. It is embedded into the OpenMOLE platform for model exploration, fostering the application of such methods without technical constraints.

KEYWORDS: Sensitivity analysis; Geosimulation; Spatial synthetic data; Model validation; Model exploration.

1 Introduction

The sensitivity of geographical analyses to the spatial structure of data is well known since the Modifiable Areal Unit Problem was put forward by Openshaw (1984). This type of issue has been generalized to various aspects since, including temporal granularity (Cheng and Adepeju, 2014) or the geographical context more generally (Kwan, 2012). When studying geosimulation models (Benenson and Torrens, 2004), similar issues must be taken into account, extending classical sensitivity analysis methods (Saltelli et al., 2004) to what can be understood as *Spatial Sensitivity Analysis* as proposed by Raimbault et al. (2019).

Several studies showed the importance of that approach. For example, in the case of Land-use Transport interaction models, Thomas et al. (2018) show how the delineation of the urban area can significantly impact simulation outcomes. Banos (2012) studies the Schelling segregation model on networks, and shows that network structure strongly influences model behavior. The spatial resolution in raster configurations can also change results (Singh et al., 2007).

*juste.raimbault@polytechnique.edu

Smith et al. (2009) spatial synthetic data microsimulation

2 Spatial sensitivity methods

Implementation and integration in OpenMOLE Reuillon et al. (2019) Reuillon et al. (2013) Passerat-Palmbach et al. (2017) workflow and reproducibility

3 Applications

Different applications of the library have already been described in the literature. Regarding the generation of synthetic data in itself,

Raimbault (2019)

Raimbault et al. (2019)

4 Discussion

References

Banos, A. (2012). Network effects in schelling's model of segregation: new evidence from agent-based simulation. *Environment and Planning B: Planning and Design*, 39(2):393–405.

Benenson, I. and Torrens, P. (2004). *Geosimulation: Automata-based modeling of urban phenomena*. John Wiley & Sons.

Cheng, T. and Adepeju, M. (2014). Modifiable temporal unit problem (mtup) and its effect on space-time cluster detection. *PloS one*, 9(6):e100465.

Kwan, M.-P. (2012). The uncertain geographic context problem. *Annals of the Association of American Geographers*, 102(5):958–968.

Openshaw, S. (1984). The modifiable areal unit problem. *Concepts and techniques in modern geography*.

Passerat-Palmbach, J., Reuillon, R., Leclaire, M., Makropoulos, A., Robinson, E. C., Parisot, S., and Rueckert, D. (2017). Reproducible large-scale neuroimaging studies with the openmole workflow management system. *Frontiers in neuroinformatics*, 11:21.

Raimbault, J. (2019). Second-order control of complex systems with correlated synthetic data. *Complex Adaptive Systems Modeling*, 7(1):1–19.

Raimbault, J., Cottineau, C., Le Texier, M., Le Nechet, F., and Reuillon, R. (2019). Space matters: Extending sensitivity analysis to initial spatial conditions in geosimulation models. *Journal of Artificial Societies and Social Simulation*, 22(4):10.

Reuillon, R., Leclaire, M., Raimbault, J., Arduin, H., Chapron, P., Chérel, G., Delay, E., Lavallée, P.-F., Passerat-Palmbach, J., Peigne, P., et al. (2019). Fostering the use of methods for geosimulation models sensitivity analysis and validation. In *European Colloquium on Theoretical and Quantitative Geography 2019*.

Reuillon, R., Leclaire, M., and Rey-Coyrehourcq, S. (2013). Openmole, a workflow engine specifically tailored for the distributed exploration of simulation models. *Future Generation Computer Systems*, 29(8):1981–1990.

Saltelli, A., Tarantola, S., Campolongo, F., and Ratto, M. (2004). Sensitivity analysis in practice: a guide to assessing scientific models. *Chichester, England*.

Singh, A., Vainchtein, D., and Weiss, H. (2007). Schelling's segregation model: Parameters, scaling, and aggregation. *arXiv preprint arXiv:0711.2212*.

Smith, D. M., Clarke, G. P., and Harland, K. (2009). Improving the synthetic data generation process in spatial microsimulation models. *Environment and Planning A*, 41(5):1251–1268.

Thomas, I., Jones, J., Caruso, G., and Gerber, P. (2018). City delineation in european applications of luti models: review and tests. *Transport Reviews*, 38(1):6–32.