Spatial sensitivity analysis of social simulation models

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

* juste.raimbault@ign.fr

[1]LASTIG, IGN-ENSG
[2]CASA, UCL
[3]UPS CNRS 3611 ISC-PIF
[4]UMR CNRS 8504 Géographie-cités

CCS 2023

Online session - Auditorium 2

October 17th 2023

Validation of simulation models

Typology of **validation methods for simulation models** from a systematic review [Raimbault, 2023]

- prediction
- sensitivity analysis
- uncertainty
- multiple/mixed methods
- benchmark
- calibration
- optimisation
- visualisation
- Pattern Oriented Modelling
- participatory
- exploration
- surrogate

$\rightarrow$ *Methods, standards and definitions strongly depend on disciplines and model functions [Raimbault, 2019c]*

Specificities of socio-spatial systems

→ Spatio-temporal non-stationarity and non-ergodicity [Raimbault, 2019b]

→ Fuzzy and noisy data [Olteanu-Raimond et al., 2015]

→ Genericity/specificity of patterns and processes [Raimbault et al., 2020a]

→ Modifiable Areal Unit Problem [Wong, 2004]

→ Multiscalar systems [Raimbault, 2021b]

Spatial sensitivity analysis

Spatial configurations are model parameters too

- "*Space matters*": impact of spatial configuration on model behavior
- Model behaviours which are robust to spatial configuration
- Model behaviours which are robust to noise in real datasets

$\Longrightarrow$ ***Construction of a generic methodology for spatial sensitivity analysis, including the generation of synthetic data, perturbation of real data and indicators***

Raimbault, J., Cottineau, C., Le Texier, M., Le Nechet, F., & Reuillon, R. (2019). Space Matters: Extending Sensitivity Analysis to Initial Spatial Conditions in Geosimulation Models. Journal of Artificial Societies and Social Simulation, 22(4).

Raimbault, J., Perret, J., & Reuillon, R. (2020). A scala library for spatial sensitivity analysis. GISRUK.

Synthetic data: general context

$\rightarrow$ coupling models with spatial configuration generators (spatial synthetic data) gives model sensitivity to space through sensitivity analysis of the coupled model

$\rightarrow$ synthetic urban forms resembling real configurations

$\rightarrow$ at different scales: microscopic (buildings), mesoscopic (population distribution), macroscopic (system of cities)

Synthetic building layouts

At the microscopic scale (district): generating building layouts

- systematic comparison of simple processual generators for building configurations
- introduction of morphological indicators
- calibration on sampled layouts from OpenStreetMap

Raimbault, J., & Perret, J. (2019, July). Generating urban morphologies at large scales. In Artificial Life Conference Proceedings (pp. 179-186). MIT Press.

Synthetic population grids

At the mesoscopic scale: population grids

- a reaction-diffusion model for population distributions
- urban form measures at the mesoscopic scale
- synthetic generators coupling population and road networks

Raimbault, J. (2018). Calibration of a density-based model of urban morphogenesis. PloS one, 13(9), e0203516.

Raimbault, J. (2019). An urban morphogenesis model capturing interactions between networks and territories. In The Mathematics of Urban Morphology (pp. 383-409). Birkhäuser, Cham.

Synthetic systems of cities

At the macroscopic scale: synthetic systems of cities

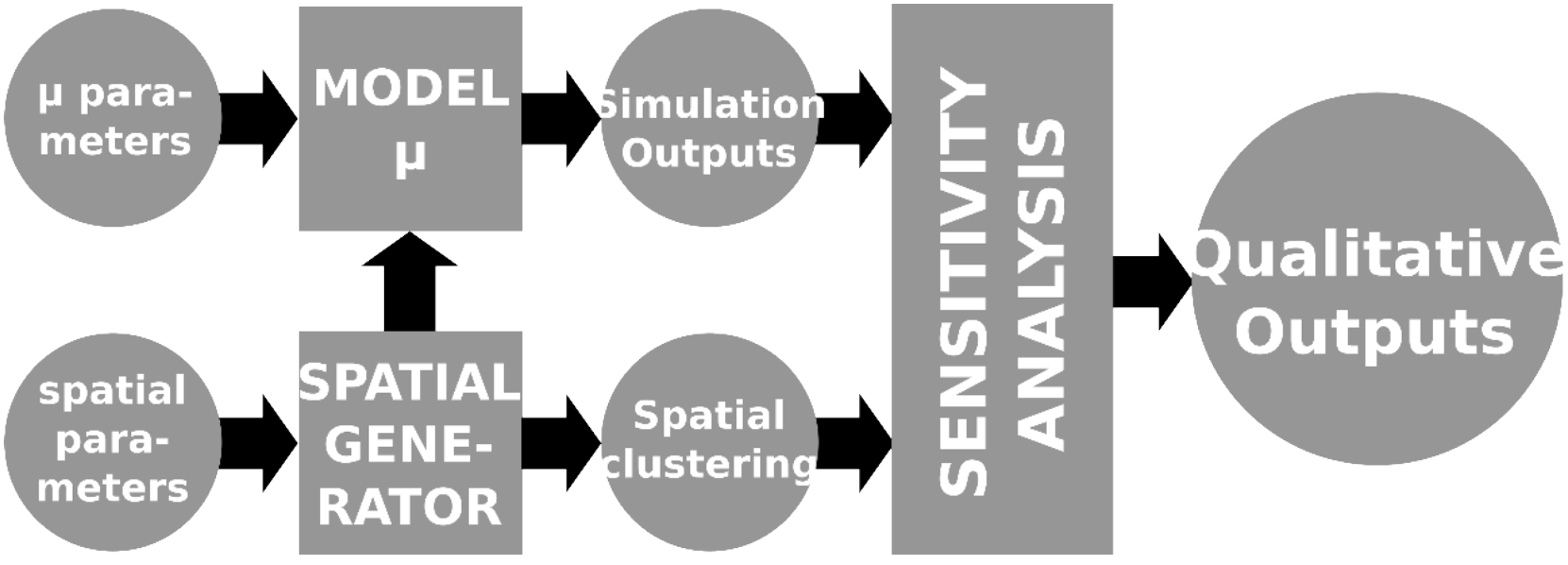
- Evolutive urban theory: systems of cities follow general stylised facts [Pumain, 2018]
- Rank-size law [Pumain et al., 2006]
- Central place theory

$\rightarrow$ *Cities-network co-evolution model explored on synthetic systems of cities*

Raimbault, J. (2021). Modeling the co-evolution of cities and networks. In Handbook of Cities and Networks (pp. 166-193). E. Elgar.

Spatial sensitivity analysis

Generic workflow to test the spatial sensitivity of simulation models



→ Application on the Schelling and Sugarscape models, showing important effects of the spatial configuration

Raimbault, J., Cottineau, C., Le Texier, M., Le Nechet, F., & Reuillon, R. (2019). Space Matters: Extending Sensitivity Analysis to Initial Spatial Conditions in Geosimulation Models. Journal of Artificial Societies and Social Simulation, 22(4).

Implementation: integration into OpenMOLE

Library implemented in `scala`: advantages of functional and object programming; Apache Spark; no widely used GIS library in scala.

`https://github.com/openmole/spatialdata`

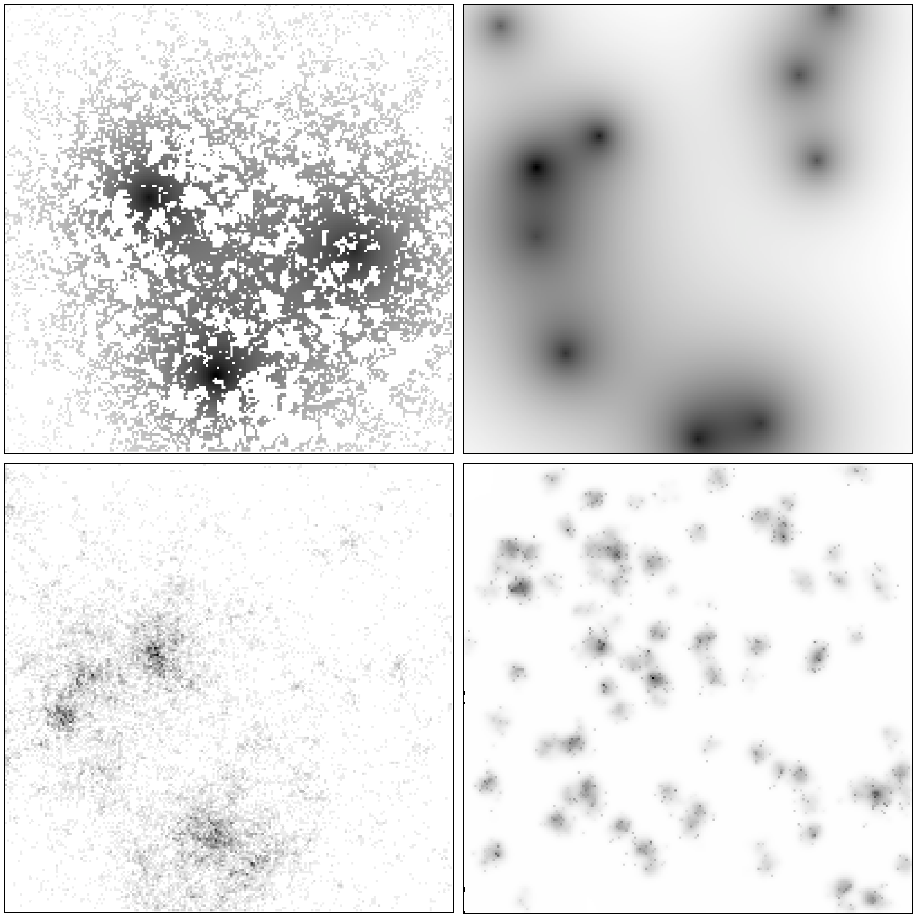
$\rightarrow$ integration into the OpenMOLE model exploration open source software [Reuillon et al., 2013]

Enables seamlessly (i) model embedding; (ii) access to HPC resources; (iii) exploration and optimization algorithms

`https://openmole.org/`

Benchmarking population generators

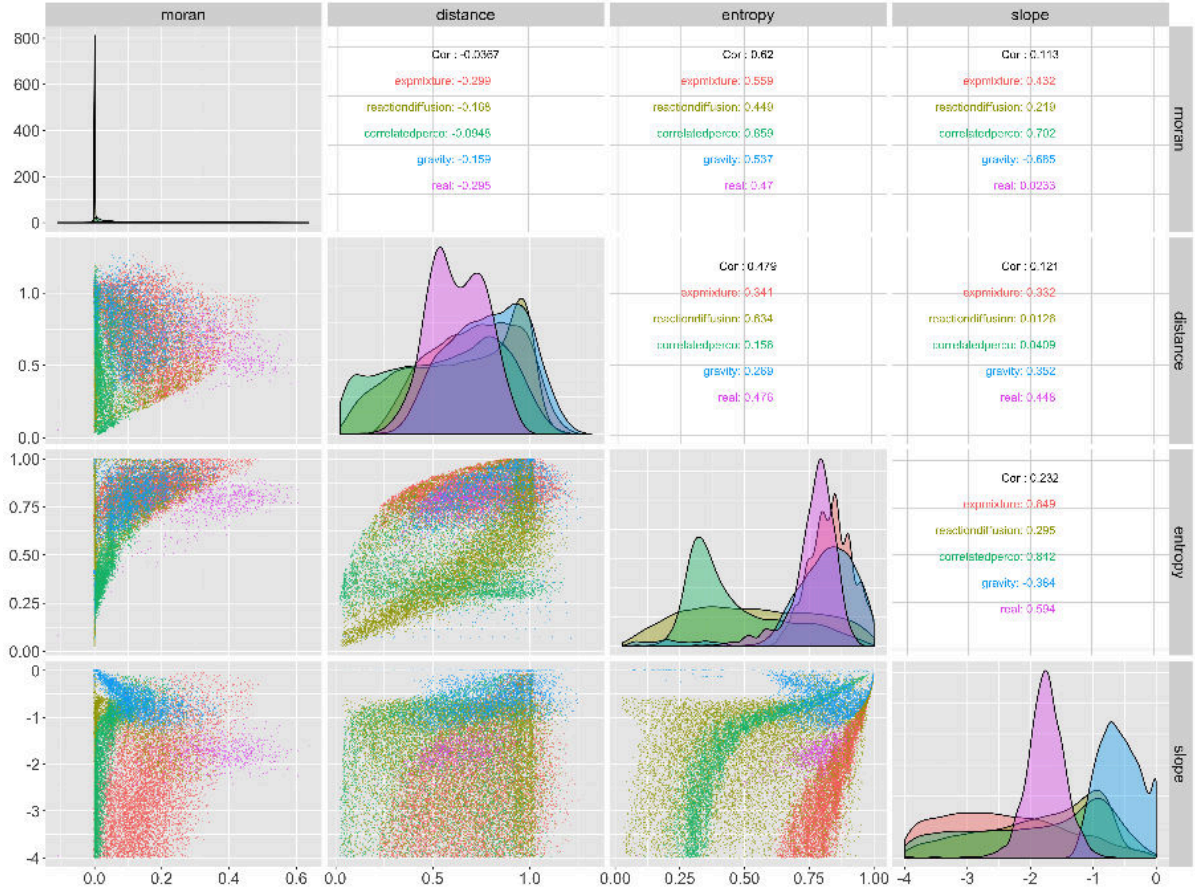
Which complementarity of different urban form generators at the mesoscopic scale (population distribution)?



Comparison of kernel mixture, reaction-diffusion [Raimbault, 2018a], gravity-based model [Li et al., 2021] and correlated percolation [Makse et al., 1998]

Raimbault, J. (2020). A comparison of simple models for urban morphogenesis. arXiv preprint arXiv:2008.13277.

Complementary feasible space of generators



A diversity-search algorithm [Chérel et al., 2015] provides feasible morphological spaces of the urban morphogenesis models, showing complementarity.

Real data perturbation

→ *How does noise in real data impacts the result ?*

- Impact of missing elements
- Impact of imprecise coordinates or topology
- Matching between spatial datasets to mitigate noise

→ *How does perturbation of real data allows to explore scenarios*

$\Longrightarrow$ spatial sensitivity analysis based on noise propagation between perturbed real input data and simulation models

Issue: "spatial noise" is specific to each type of spatial structure and does not follow standard distributions (for example road networks or building polygons)

→ **how to define small variations in the morphological space?**

Example of spatial noise: BDTopo

Map of Strasbourg building evolution (2012-2022) using IGN BDTopo: change in data specifications leads to split polygons, apparitions of artefacts, perturbations of polygons boundaries

(see Subdense project: `https://github.com/subdense`)

Towards variance-based indicators

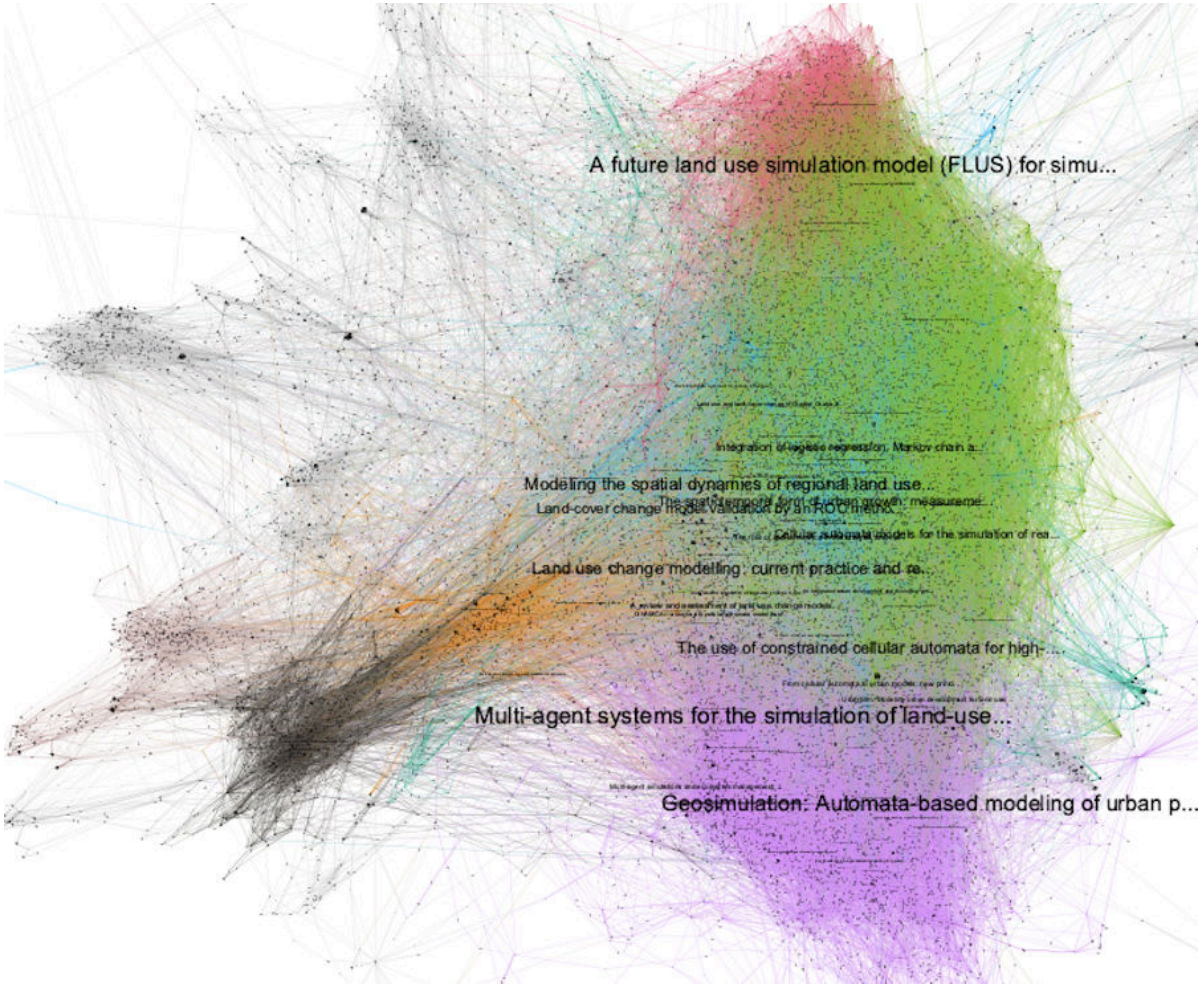
Global Sensitivity Analysis methods [Saltelli et al., 2008] introduce interpretable indicators based on conditional variances, with a robust theoretical statistical background, and are widely established for simulation models

→ *How to compute such indicators with spatial configuration as a parameter?*

→ *For the estimation to be correct, the full space must be sampled with a high discrepancy (using high-dimensional sequences for standard parameters).*

→ *Use diversity search to sample the feasible space of generators, compare with meta-parameter sampling; current application to the Schelling model to compare with the ad-hoc indicators of spatial sensitivity analysis based on distance between phase diagrams.*

Interdisciplinary systematic review



Systematic literature review and mapping of validation methods for simulation methods [Raimbault, 2023]

$\rightarrow$ link with similar approaches in other disciplines such as environmental science dealing with e.g. spatial uncertainty [Koo et al., 2020]

$\rightarrow$ link with approaches focusing more on uncertainty due to processes? (see presentation yesterday by C. De Bezenac)

Discussion: more open questions

- network generative models [Raimbault, 2018b], correlated synthetic data [Raimbault, 2019a]
- domain models: LUTI, urban dynamics
- other disciplines: ecology [Koo et al., 2020], geosciences [Mogheir et al., 2004]?
- link with data driven disciplines ? (planning, architecture, spatio-temporal datamining)
- genericity of some models? (reaction-diffusion)
- synthetic data generation methods (synthetic populations)
- synthetic data at the core of applied statistics methodology (less in spatial statistics?)
- port the library to more classic languages (python, R)

Conclusion

→ **Space matters:** relevance of spatially-explicit models and spatial sensitivity analysis for social simulation.

→ **Generic methodology:** generators included in OpenMOLE, and current and future developments of the `spatialdata` library will be continuously integrated.

→ **Disciplinary context:** need interdisciplinary feedback and more methods and models transfers.

Get the library at `https://github.com/openmole/spatialdata`

Use OpenMOLE `https://openmole.org`

References I

Chérel, G., Cottineau, C., and Reuillon, R. (2015).
Beyond corroboration: Strengthening model validation by looking for unexpected patterns.
PloS one, 10(9):e0138212.

Koo, H., Iwanaga, T., Croke, B. F., Jakeman, A. J., Yang, J., Wang, H.-H., Sun, X., Lü, G., Li, X., Yue, T., et al. (2020).
Position paper: Sensitivity analysis of spatially distributed environmental models-a pragmatic framework for the exploration of uncertainty sources.
Environmental modelling & software, 134:104857.

References II

Li, Y., Rybski, D., and Kropp, J. P. (2021).
Singularity cities.
Environment and Planning B: Urban Analytics and City Science, 48(1):43–59.

Makse, H. A., Andrade, J. S., Batty, M., Havlin, S., Stanley, H. E., et al. (1998).
Modeling urban growth patterns with correlated percolation.
Physical Review E, 58(6):7054.

Mogheir, Y., De Lima, J., and Singh, V. (2004).
Characterizing the spatial variability of groundwater quality using the entropy theory: I. synthetic data.
Hydrological processes, 18(11):2165–2179.

References III


Olteanu-Raimond, A.-M., Mustiere, S., and Ruas, A. (2015).
Knowledge formalization for vector data matching using belief theory.

Journal of Spatial Information Science, (10):21–46.


Pumain, D. (2018).
An evolutionary theory of urban systems.
International and transnational perspectives on urban systems, pages 3–18.


Pumain, D., Paulus, F., Vacchiani-Marcuzzo, C., and Lobo, J. (2006).
An evolutionary theory for interpreting urban scaling laws.
Cybergeo: European Journal of Geography.

References IV


Raimbault, J. (2018a).
Calibration of a density-based model of urban morphogenesis.
PloS one, 13(9):e0203516.


Raimbault, J. (2018b).
Multi-modeling the morphogenesis of transportation networks.
In *Artificial Life Conference Proceedings*, pages 382–383. MIT Press One Rogers Street, Cambridge, MA 02142-1209, USA journals-info ....


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

References V


Raimbault, J. (2019b).
An urban morphogenesis model capturing interactions between networks and territories.
The mathematics of urban morphology, pages 383–409.


Raimbault, J. (2019c).
Validation levels and standards depending on models types and functions.
In *CCS 2019, Singapore*. NTU Singapore.


Raimbault, J. (2020).
A comparison of simple models for urban morphogenesis.
arXiv preprint arXiv:2008.13277.

References VI


Raimbault, J. (2021a).
Modeling the co-evolution of cities and networks.
In *Handbook of Cities and Networks*, pages 166–193.


Raimbault, J. (2021b).
Strong coupling between scales in a multi-scalar model of urban dynamics.
arXiv preprint arXiv:2101.12725.


Raimbault, J. (2023).
Validation of geosimulation models: a systematic review.
In *ECTQG 2023, Braga*. Coimbra University.

References VII


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).


Raimbault, J., Denis, E., and Pumain, D. (2020a).
Empowering urban governance through urban science: Multi-scale dynamics of urban systems worldwide.
Sustainability, 12(15):5954.

References VIII


Raimbault, J. and Perret, J. (2019).
Generating urban morphologies at large scales.
In *Artificial Life Conference Proceedings*, pages 179–186. MIT Press One Rogers Street, Cambridge, MA 02142-1209, USA journals-info ....


Raimbault, J., Perret, J., and Reuillon, R. (2020b).
A scala library for spatial sensitivity analysis.
GISRUK.


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.

References IX


Saltelli, A., Ratto, M., Andres, T., Campolongo, F., Cariboni, J., Gatelli, D., Saisana, M., and Tarantola, S. (2008).
Global sensitivity analysis: the primer.
John Wiley & Sons.


Wong, D. W. (2004).
The modifiable areal unit problem (maup).
In *WorldMinds: geographical perspectives on 100 problems: commemorating the 100th anniversary of the association of American geographers 1904–2004*, pages 571–575. Springer.