

A scala library for spatial sensitivity analysis

J. Raimbault^{1,2,3,*}, J. Perret^{4,2} and R. Reuillon^{2,3}

¹ Center for Advanced Spatial Analysis, UCL

² UPS CNRS 3611 ISC-PIF

³ UMR CNRS 8504 Géographie-cités

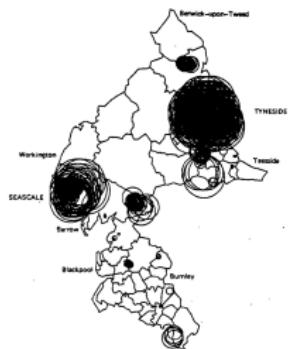
⁴ LaSTIG STRUDEL, IGN, ENSG, Univ. Paris-Est

* juste.raimbault@polytechnique.edu

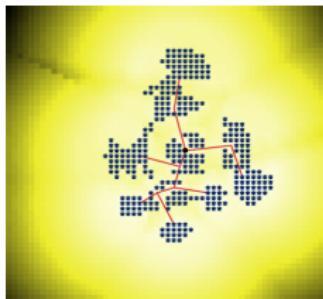


GISRUK 2020
July 21st, 2020

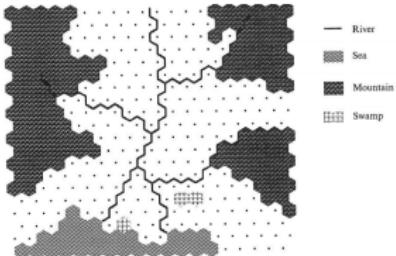
Modeling and simulation in geography



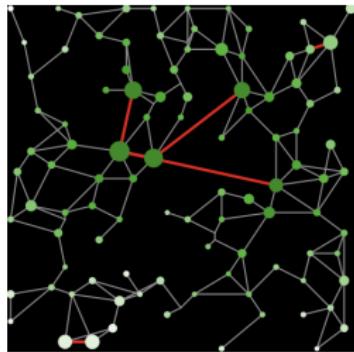
Geographical analysis machine
[Openshaw et al., 1987]



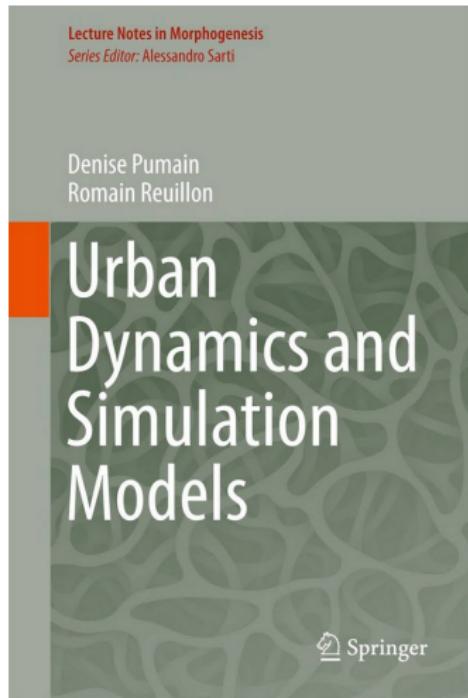
Hybrid urban morphogenesis
[Raimbault et al., 2014]



Simpop 1 model [Sanders et al., 1997]



SimpopNet model [Schmitt, 2014]



Development of an evolutionary urban theory [Pumain, 2018]

- Recurrent stylized facts on main systems of cities
- Construction of simulation models (with an explicative purpose)
- Tools and methods to explore simulation models

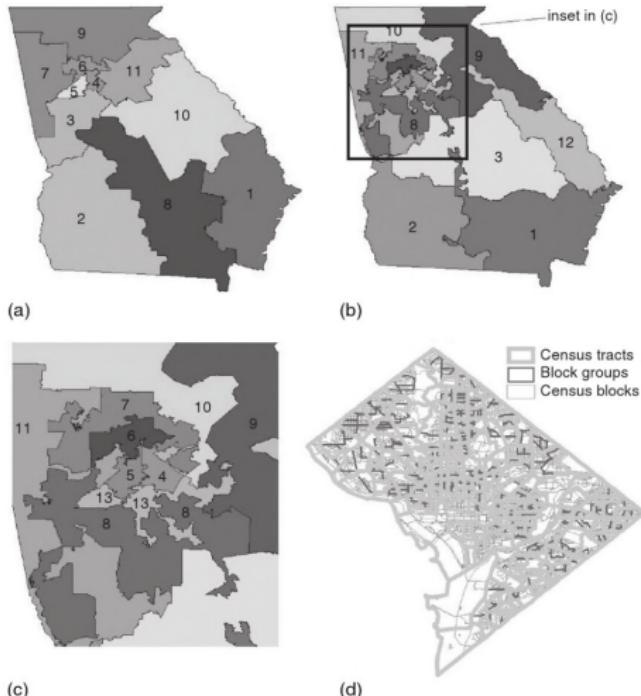


OpenMOLE

Classical problems in geography and spatial sciences:

- ▶ Modifiable Areal Unit Problem
- ▶ Dependancy of processes to scale
- ▶ Spatial non-stationarity
- ▶ Fuzzy and noisy data
- ▶ Genericity/particularity

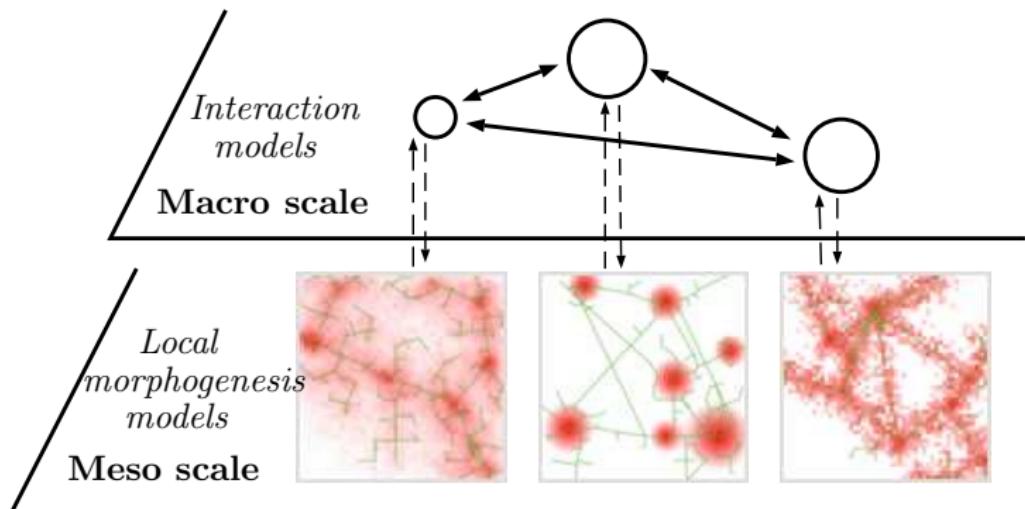
The Modifiable Areal Unit Problem



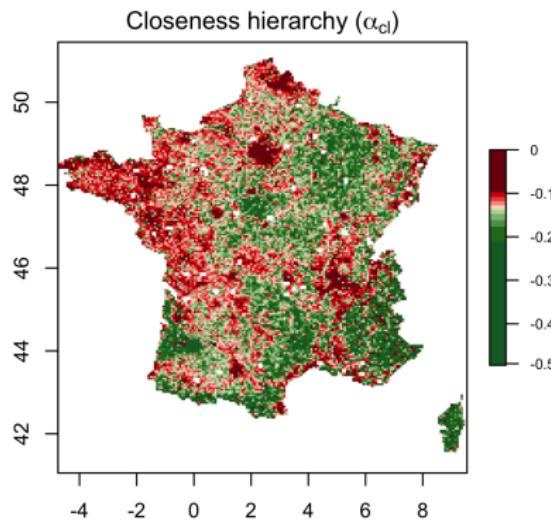
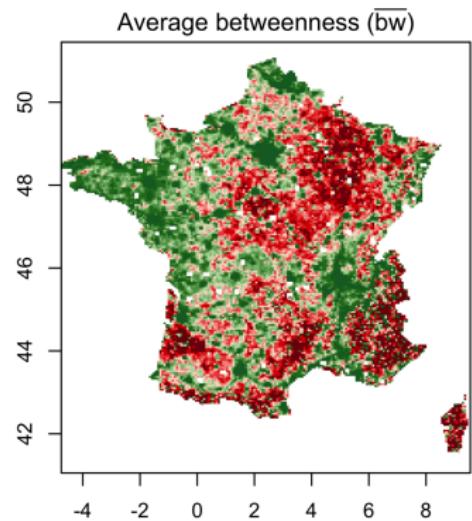
Wong, D. (2009). The modifiable areal unit problem (MAUP). The SAGE handbook of spatial analysis, 105, 23.

Multiscalar systems

Processes specific to scales, coupling implies dedicated ontologies
[Raimbault, 2019a]



Spatial non-stationarity of road network properties
[Raimbault, 2019b]

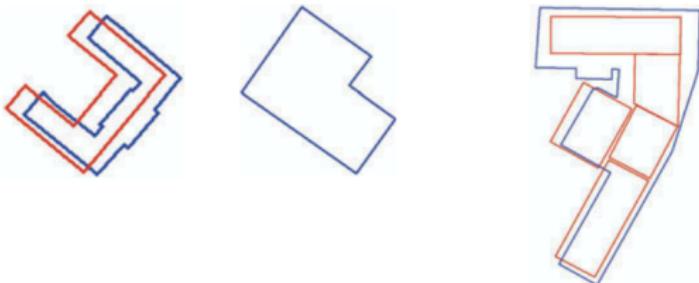


Noisy spatial data

Assessment of data quality in OpenStreetMap [Fan et al., 2014]

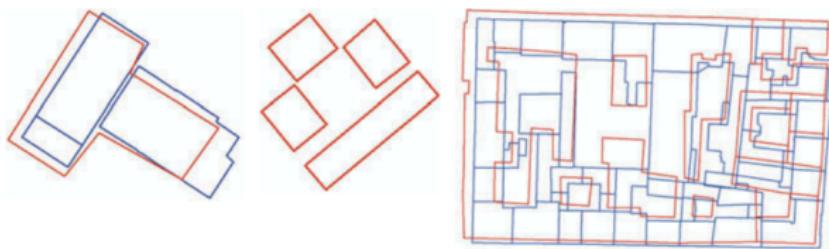
Relation	1:1	1:0	1: n
----------	-----	-----	--------

Illustration



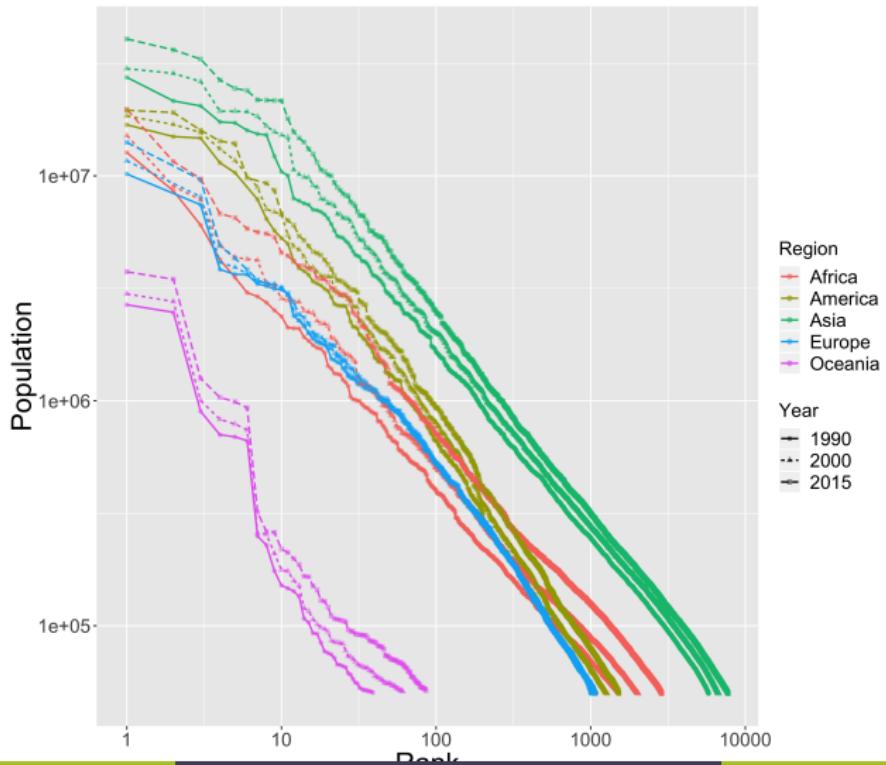
Relation	$n:1$	$0:1$	$n:m$
----------	-------	-------	-------

Illustration



Genericity and specificity

Urban systems are simultaneously universal and particular
[Raimbault et al., 2020]



Approach proposed

⇒ **spatial configuration are parameters too**

- ▶ “*Space matters*”: impact of spatial configuration on model behavior
- ▶ Model behaviors which are robust to spatial configuration
- ▶ Model behaviors which are robust to noise in parametrization real datasets

- coupling models with spatial configuration generators (spatial synthetic data) gives model sensitivity to space through sensitivity analysis of the coupled model
- synthetic urban forms resembling real configurations
- at different scales: microscopic (buildings), mesoscopic (population distribution), macroscopic (system of cities)

At the microscopic scale (district): generating building layouts

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

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

Urban form indicators for building layouts:

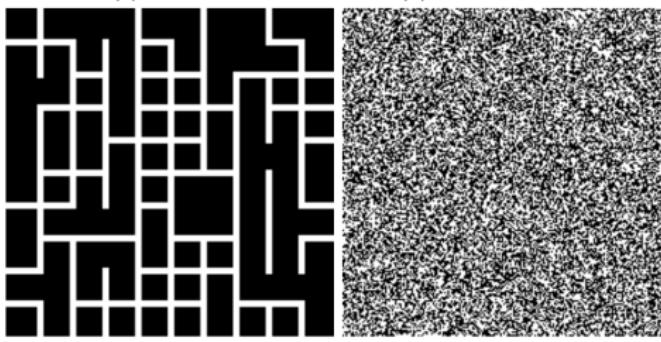
- ▶ density, number of buildings, average area
- ▶ Moran index and average distance on rasterized representation
- ▶ average detour in the free space
- ▶ mathematical morphology indicators (steps for erosion and dilation) [Serra, 1994]

Complementary generators



(a) Blocks

(b) Kernel mixture



(c) Network percolation

(d) Random

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

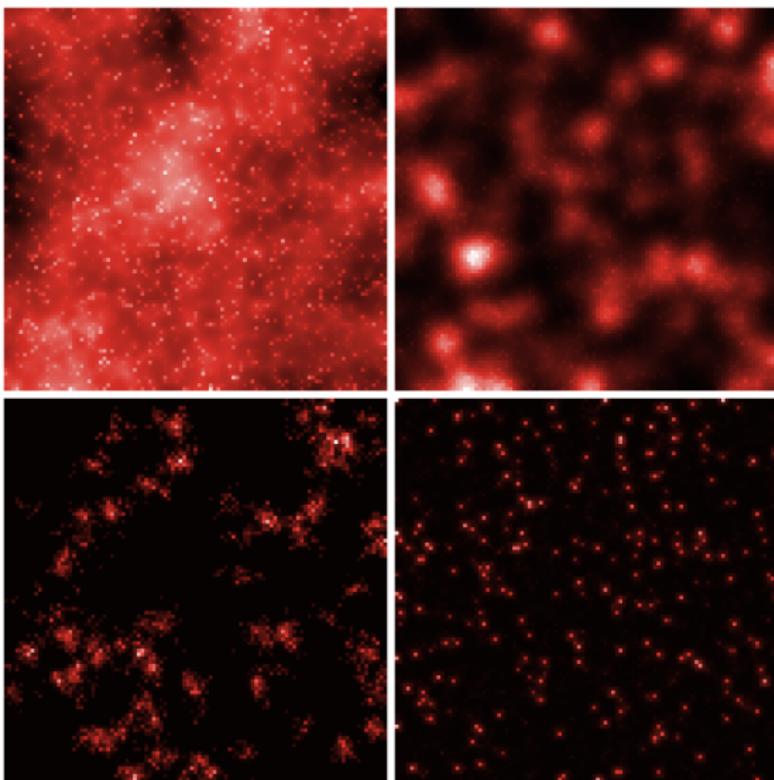
Raimbault, J. (2018). Calibration of a density-based model of urban morphogenesis. *PLoS one*, 13(9), e0203516.

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

- Crucial role of the interplay between concentration forces and dispersion forces [Fujita and Thisse, 1996] in keeping Urban Systems at the border of chaos
- Potentially of aggregation mechanisms (such as Simon model) to produce power laws
- Link with Reaction-diffusion approaches in Morphogenesis
[?]
- Extension of a DLA-type model introduced by [?], with simple abstract processes of population aggregation and diffusion

- Grid world with cell populations $(P_i(t))_{1 \leq i \leq N^2}$.
- At each time step:
 1. Population growth with exogenous rate N_G , attributed independently to a cell following a preferential attachment of strength α
 2. Population is diffused n_d times with strength β
- Stopping criterion: fixed maximal population P_m .
- Output measured by morphological indicators: Moran index, average distance, rank-size hierarchy, entropy.

Generating Population Distributions



Examples of generated territorial shapes

Morphological indicators

- Rank-size slope γ , given by $\ln(P_{\tilde{i}}/P_0) \sim k + \gamma \cdot \ln(\tilde{i}/i_0)$ where \tilde{i} are the indexes of the distribution sorted in decreasing order.
- Entropy of the distribution:

$$\mathcal{E} = \sum_{i=1}^M \frac{P_i}{P} \cdot \ln \frac{P_i}{P} \quad (1)$$

$\mathcal{E} = 0$ means that all the population is in one cell whereas
 $\mathcal{E} = M$ means that the population is uniformly distributed.

- Spatial-autocorrelation given by Moran index, with simple spatial weights given by $w_{ij} = 1/d_{ij}$

$$I = M \cdot \frac{\sum_{i \neq j} w_{ij} (P_i - \bar{P}) \cdot (P_j - \bar{P})}{\sum_{i \neq j} w_{ij} \sum_i (P_i - \bar{P})^2}$$

- Mean distance between individuals

$$\bar{d} = \frac{1}{d_M} \cdot \sum_{i < j} \frac{P_i P_j}{P^2} \cdot d_{ij}$$

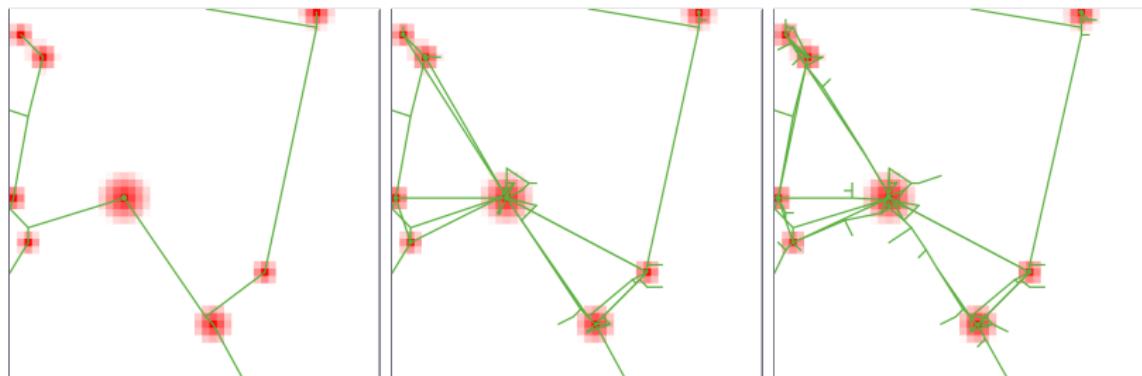
where d_M is a normalization constant

Network Generation models

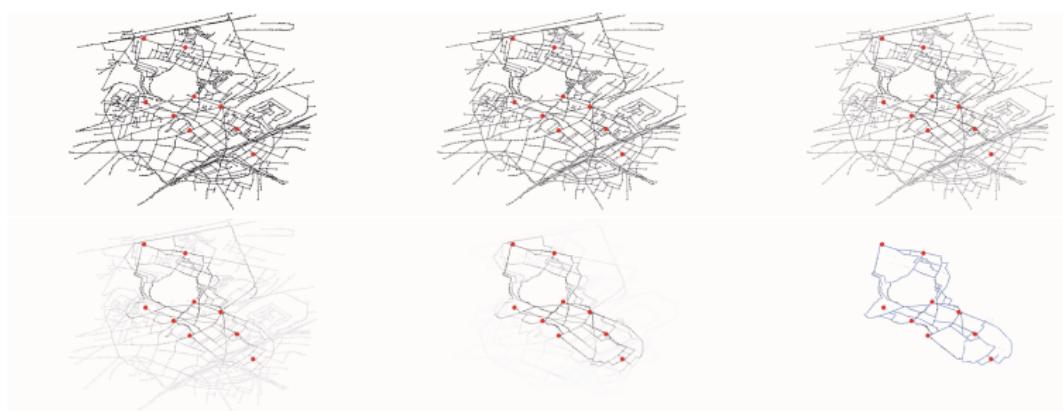
currently being implemented

Network generated conditionally to population; at fixed time steps:

1. Add new nodes preferentially to new population and connect them
2. Variable heuristic for new links, among: nothing, random, gravity-based deterministic breakdown, gravity-based random breakdown (from [Schmitt, 2014]), cost-benefits (from [Louf et al., 2013]), biological network generation (based on [Tero et al., 2010])



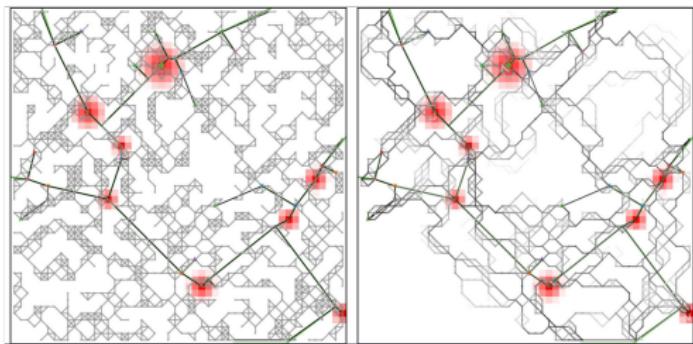
Model studied by [Tero et al., 2010] : exploration and reinforcement by a slime mould searching for resources



Application to the design of optimal bus routes

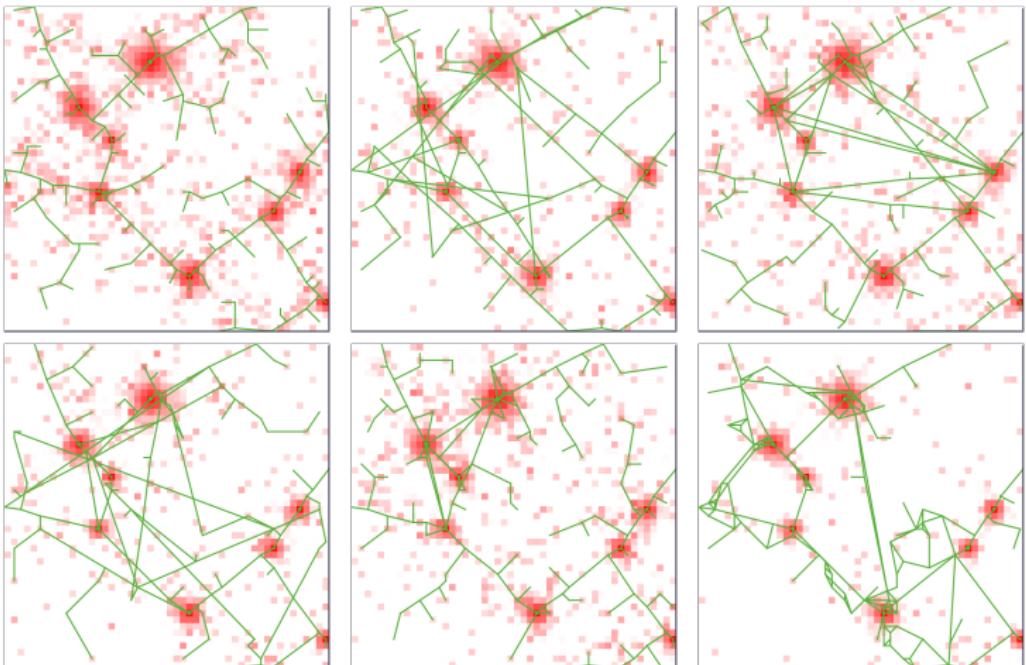
Adding new links with biological heuristic:

1. Create network of potential new links, with existing network and randomly sampled diagonal lattice
2. Iterate for k increasing ($k \in \{1, 2, 4\}$ in practice) :
 - ▶ Using population distribution, iterate $k \cdot n_B$ times the slime mould model to compute new link capacities
 - ▶ Delete links with capacity under θ_d
 - ▶ Keep the largest connected component
3. Planarize and simplify final network



Intermediate stage for biological network generation

Generated Urban Shapes: Network

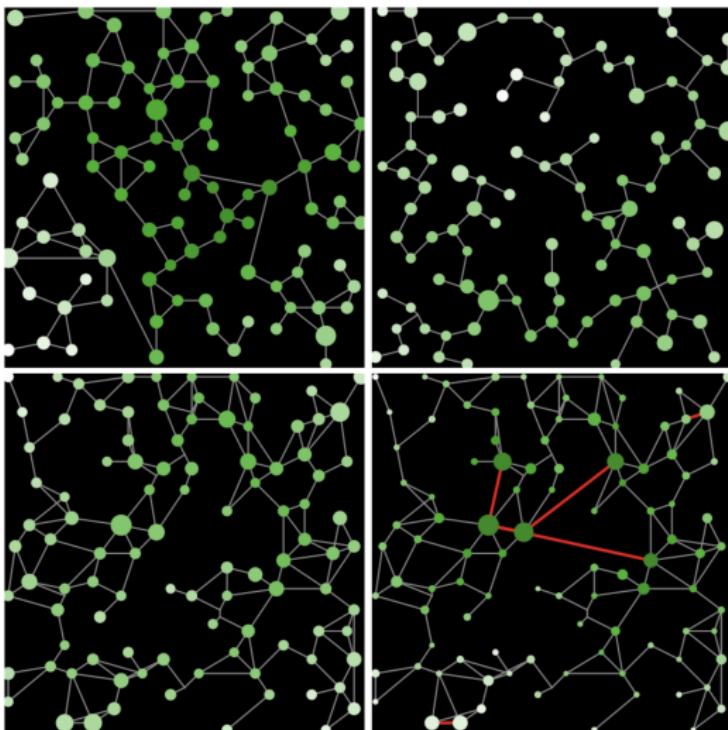


In order: connection; random; deterministic breakdown; random breakdown; cost-driven; biological.

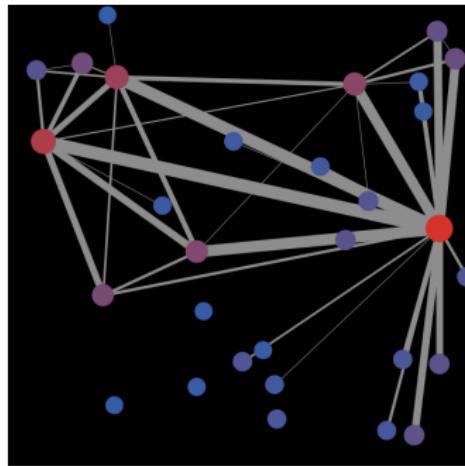
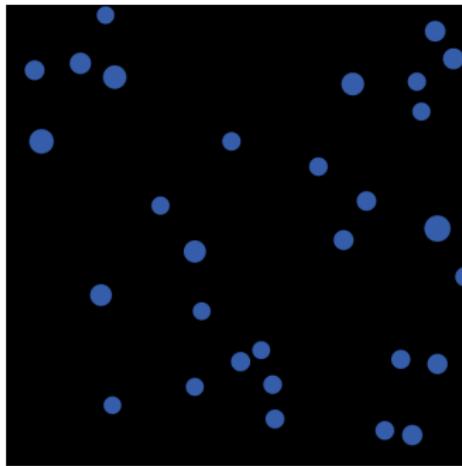
At the macroscopic scale: synthetic systems of cities

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

Synthetic system of cities and network for the SimpopNet model
[?]



Cities-network co-evolution model explored on synthetic systems of cities [?]



currently being implemented

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

- ▶ Impact of missing elements
- ▶ Impact of imprecise coordinates or topology
- ▶ Optimal matching between spatial datasets

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

Examples:

- ▶ simulating urban projects by modifying population of areas with a given spatial correlation structure
- ▶ simulating network disruptions or new transportation lines

In the spatial approach, spatial model indicators are also important: what kind of spatial structure does the model produce ?

- ▶ previous form indicators at different scales, applied on any spatialized variable or event: quantify level of aggregation, hierarchy, clustering
- ▶ spatial statistics indicators and methods
- ▶ more complicated approaches: fractals and multifractals, spatial datamining

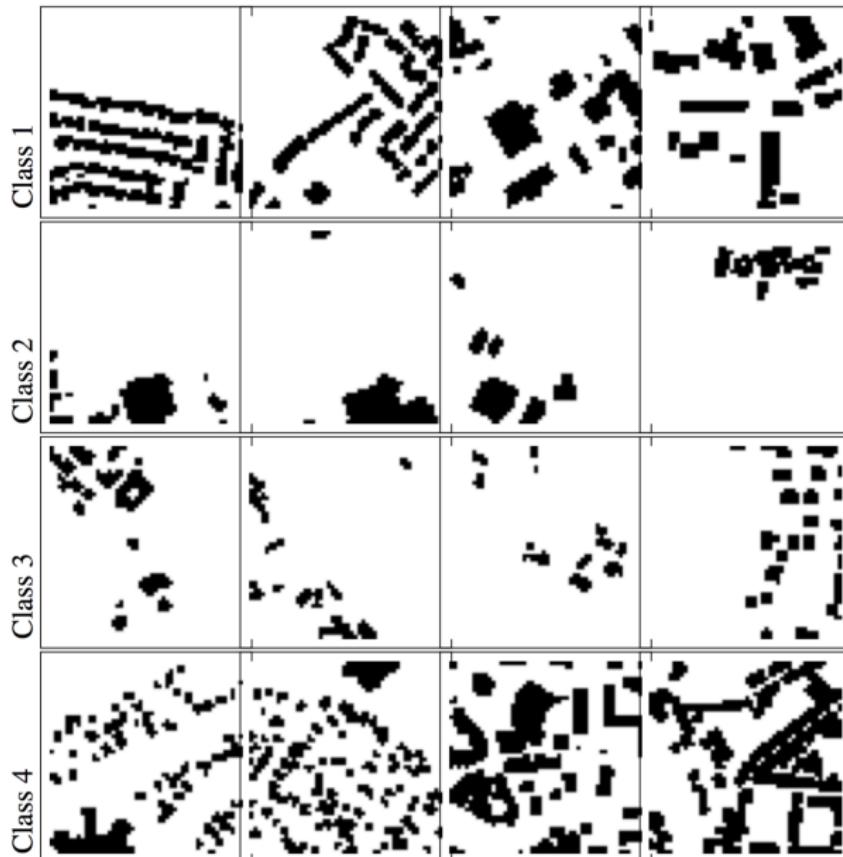
Integration into OpenMOLE

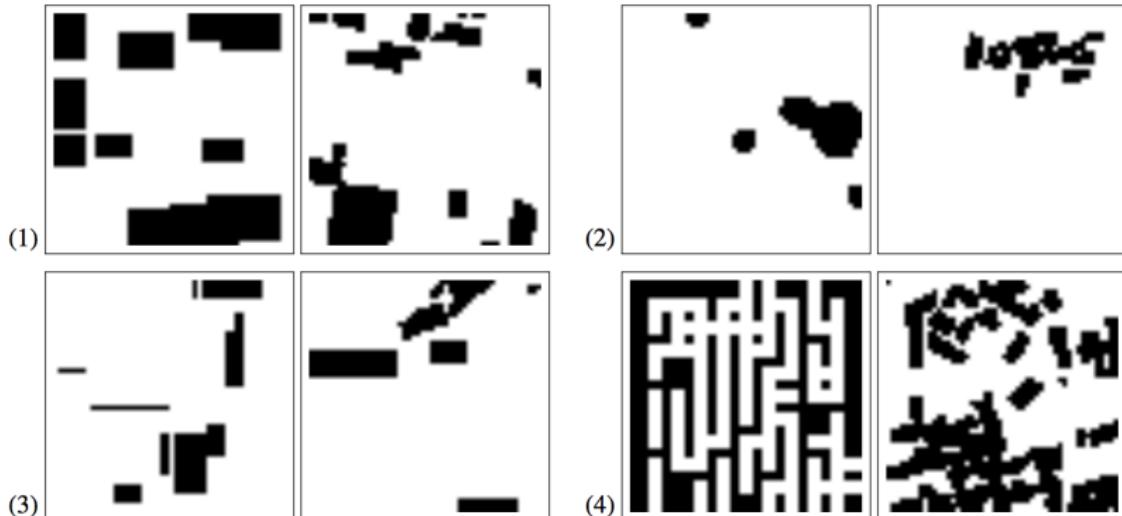


Sampled districts from OpenStreetMap

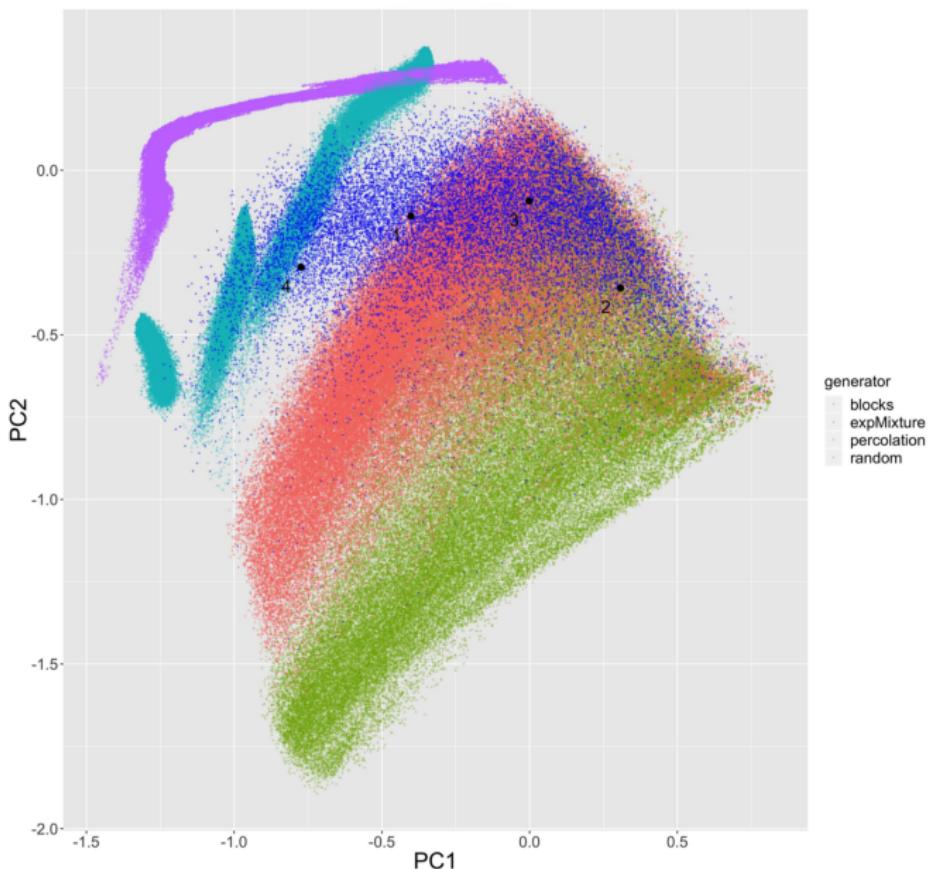


Classification of urban forms





Point cloud



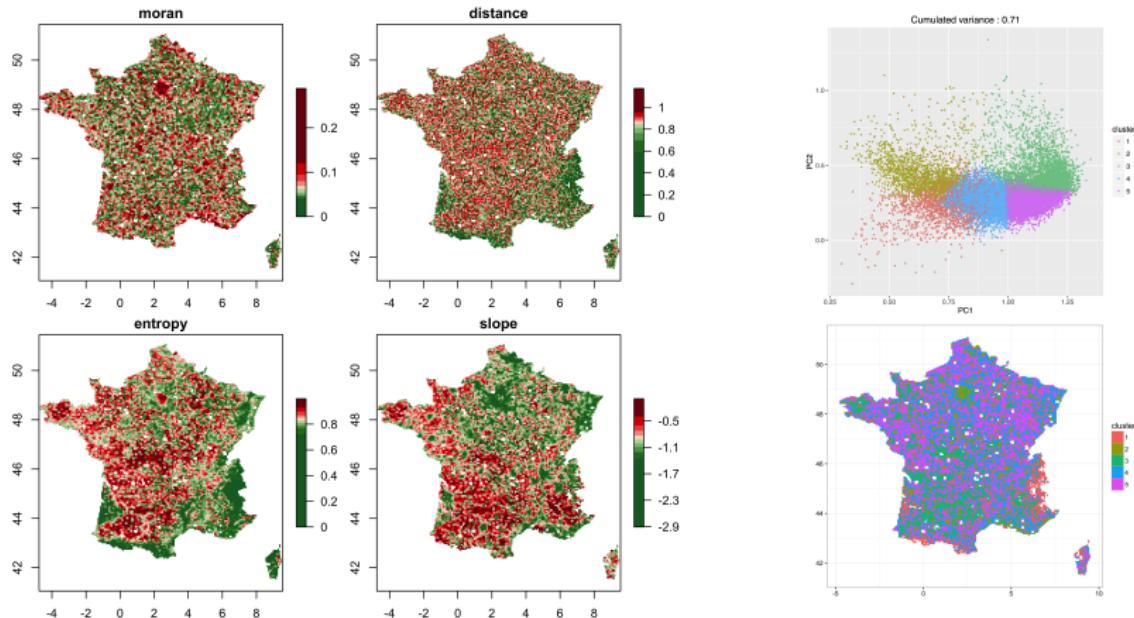
	Random	Blocks	Exp. Mixture	Percolation
Centroid 1	0.424 ± 0.011	0.106 ± 0.063	0.303 ± 0.101	0.325 ± 0.019
Centroid 2	0.809 ± 0.022	0.164 ± 0.099	0.184 ± 0.141	0.947 ± 0.019
Centroid 3	0.428 ± 0.019	0.095 ± 0.054	0.109 ± 0.064	0.541 ± 0.019
Centroid 4	0.515 ± 0.005	0.311 ± 0.077	0.589 ± 0.149	0.083 ± 0.025

Why not use calibration heuristics? Open question of fitting a point cloud; issue of projecting in a reduced dimension space

```
sampling =  
  BlocksGridSpatialSampling(  
    samples = 2,  
    gridSize = 10,  
    blocks = (p1gen in Range(1.0,10.0)),  
    blockMinSize = (p2gen in Range(2.0,4.0)),  
    blockMaxSize = (p3gen in Range(2.0,4.0)),  
    prototype = w  
)
```

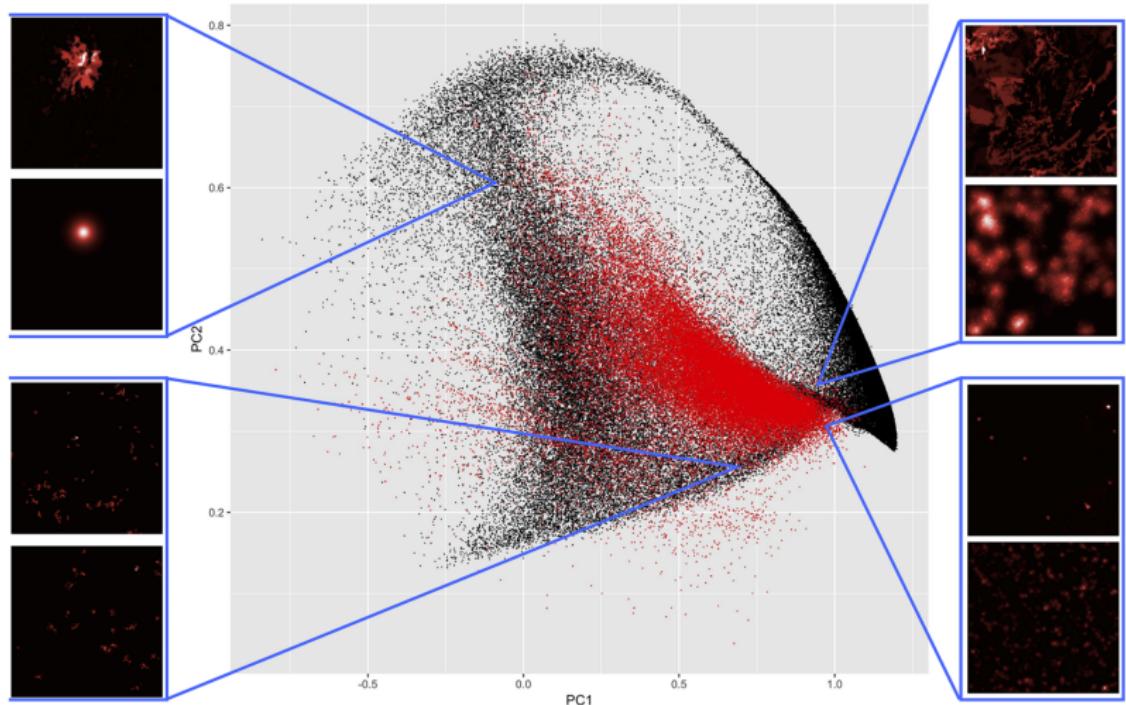
→ other generators have their own primitives
(ExpMixtureThresholdSpatialSampling,
PercolationGridSpatialSampling) and arguments (**see the documentation**)

Population grids: empirical Data for Calibration



Computation of morphological indicators on population density data for Europe (shown here on France), morphological classification.

Population grids: model Calibration

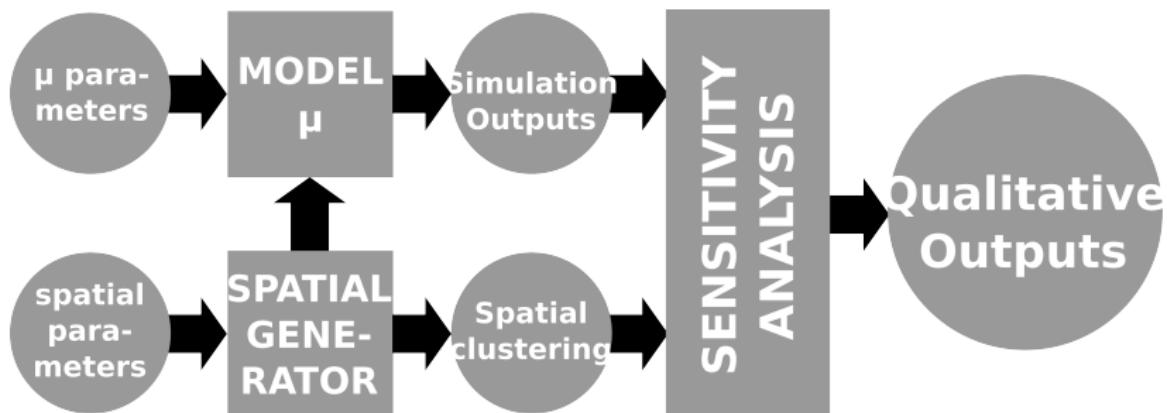


Brute force calibration by exploring the parameter space. Reproduction of most existing configuration in the morphological sense (here in principal plan).

Method flowchart

General workflow to test the spatial sensitivity of simulation models

Raimbault, J., Cottineau, C., Texier, M. L., Néchet, F. L., & Reuillon, R. (2018). Space Matters: extending sensitivity analysis to initial spatial conditions in geosimulation models. arXiv preprint arXiv:1812.06008.

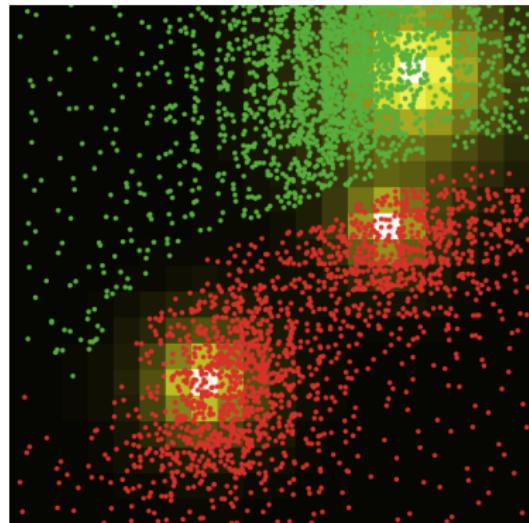
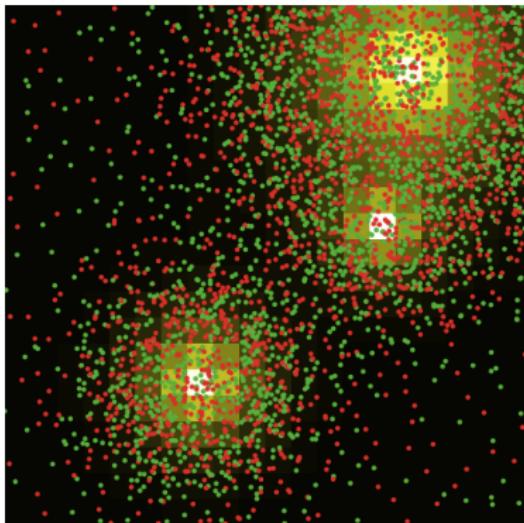


Relative distance of phase diagrams to compare global model behavior when meta-parameters change

$$d_r(\mu_{\vec{\alpha}_1}, \mu_{\vec{\alpha}_2}) = 2 \cdot \frac{d(\mu_{\vec{\alpha}_1}, \mu_{\vec{\alpha}_2})^2}{Var[\mu_{\vec{\alpha}_1}] + Var[\mu_{\vec{\alpha}_2}]}$$

Why could the Schelling model be sensitive to space ?

[?] network effects in Schelling model



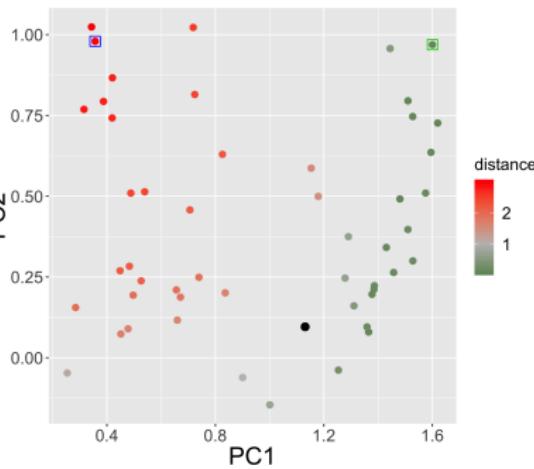
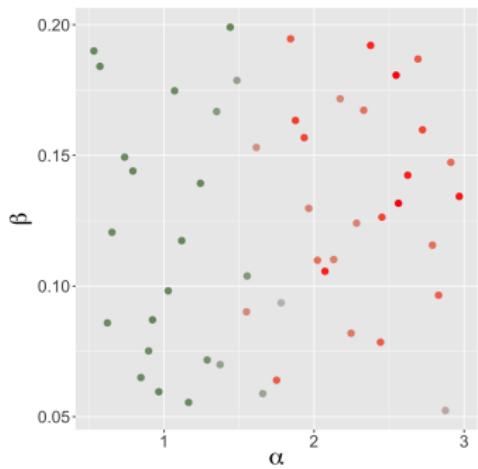
Influence of spatial generator parameters on model outputs

Simulation outcome by segregation index:	Dissimilarity		Entropy		Moran's I	
Intercept	-0.212 ***	-0.141 ***	-0.254 ***	-0.208 ***	-0.036 ***	-0.061 ***
Similarity Wanted (S)	1.212 ***	1.212 ***	1.250 ***	1.250 ***	0.550 ***	0.550 ***
quadratic term (S^2)	-0.942 ***	-0.942 ***	-0.963 ***	-0.963 ***	-0.428 ***	-0.438 ***
Vacancy Rate (V)	0.602 ***	0.602 ***	0.453 ***	0.453 ***	-0.027 ***	-0.027 ***
Minority Index (%Maj - %Min)	0.307 ***	0.307 ***	0.130 ***	0.130 ***	-0.067 ***	-0.067 ***
Density Grid = Polycentric		0.087 ***		0.052 ***		0.001 ***
Density Grid = Discontinuous		0.111 ***		0.068 ***		0.00
Attraction generator parameter α		-0.083 ***		-0.053 ***		0.014 ***
Diffusion generator parameter β		0.323 ***		0.218 ***		0.017 ***
R2 (%)	30.6	34.7	24.1	25.6	23.9	24.0
# of observations (sim. runs)	2,106,000	2,106,000	2,106,000	2,106,000	2,106,000	2,106,000
AIC	-70717.68	-198748.2	208213.8	166048.8	-4385990	-4387816

A model of resource collection

- ▶ agents collect a spatial resource
- ▶ the resource regrows at a certain rate only

Relative distances between phase diagrams



Spatial synthetic data in other disciplines ?

- ▶ spatial network generative models
- ▶ other disciplines, ecology, geosciences [?]?
- ▶ interaction 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
(not much in spatial statistics ?)

	OpenMOLE	spatialdata	planned
Micro grid spatial samplings	✓	✓	✓
Meso grid spatial samplings	✓	✓	✓
Macro spatial samplings	✗	✓	✓
Spatial network generation	✗	✓	✓
Real data import	✗	✗	✓
Real data perturbations	✗	✗	✓
Spatial statistics	✓	✓	✓
Hybrid methods	✗	✗	✓
Domain models (transportation, land-use)	✗	✓	✓

- **Space matters:** relevance of spatially-explicit models and spatial sensitivity analysis.
- **Synthetic data:** first experimental samplings included in OpenMOLE, soon more to come.
- **Disciplinary context:** strong contingency on included models and forms, please provide feedbacks, suggestions, needs, ideas from your viewpoint.

Open issues at <https://github.com/openmole/spatialdata/issues>

References I

-  Fan, H., Zipf, A., Fu, Q., and Neis, P. (2014). Quality assessment for building footprints data on openstreetmap.
International Journal of Geographical Information Science, 28(4):700–719.
-  Fujita, M. and Thisse, J.-F. (1996). Economics of agglomeration.
Journal of the Japanese and international economies, 10(4):339–378.
-  Louf, R., Jensen, P., and Barthelemy, M. (2013). Emergence of hierarchy in cost-driven growth of spatial networks.
Proceedings of the National Academy of Sciences, 110(22):8824–8829.

References II

-  Openshaw, S., Charlton, M., Wymer, C., and Craft, A. (1987).
A geographical analysis machine for the automated analysis of
point data sets.
International Journal of Geographical Information System,
1(4):335–358.
-  Pumain, D. (2018).
An evolutionary theory of urban systems.
In *International and Transnational Perspectives on Urban Systems*, pages 3–18. Springer.
-  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 III

-  Raimbault, J. (2019a).
A multi-scalar model for system of cities.
In *Conference on Complex Systems 2019*, Singapore, Singapore.
-  Raimbault, J. (2019b).
An urban morphogenesis model capturing interactions between networks and territories.
In *The Mathematics of Urban Morphology*, pages 383–409. Springer.
-  Raimbault, J., Banos, A., and Doursat, R. (2014).
A hybrid network/grid model of urban morphogenesis and optimization.
In *4th International Conference on Complex Systems and Applications*, pages 51–60.

References IV

-  Raimbault, J., Denis, E., and Pumain, D. (2020). Empowering urban governance through urban science: Multi-scale dynamics of urban systems worldwide. *Sustainability*.
-  Raimbault, J. and Perret, J. (2019). Generating urban morphologies at large scales. In *Artificial Life Conference Proceedings*, pages 179–186. MIT Press.
-  Sanders, L., Pumain, D., Mathian, H., Guérin-Pace, F., and Bura, S. (1997). Simpop: a multiagent system for the study of urbanism. *Environment and Planning B*, 24:287–306.

References V

-  Schmitt, C. (2014).
Modélisation de la dynamique des systèmes de peuplement: de SimpopLocal à SimpopNet.
PhD thesis, Paris 1.
-  Serra, J. (1994).
Morphological filtering: an overview.
Signal processing, 38(1):3–11.
-  Tero, A., Takagi, S., Saigusa, T., Ito, K., Bebber, D. P., Fricker, M. D., Yumiki, K., Kobayashi, R., and Nakagaki, T. (2010).
Rules for biologically inspired adaptive network design.
Science, 327(5964):439–442.