Simulation models in quantitative geography: towards integrated approaches

J. Raimbault[1,2,3]*
juste.raimbault@polytechnique.edu

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

Centrale Casablanca
June 20th 2019

Outline

1. Introduction

2. Model exploration methods

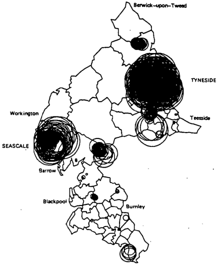
3. Interactions between networks and territories

1 Introduction

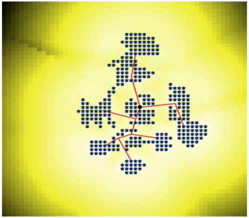
2 Model exploration methods

3 Interactions between networks and territories

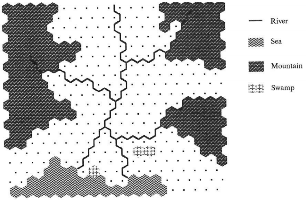
A long history of modeling and simulation in Geography



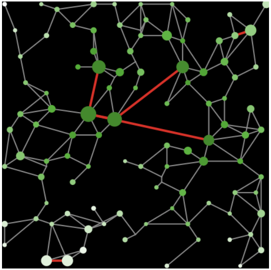
Geographical analysis machine [Openshaw et al., 1987]



Hybrid urban morphogenesis [Raimbault et al., 2014]



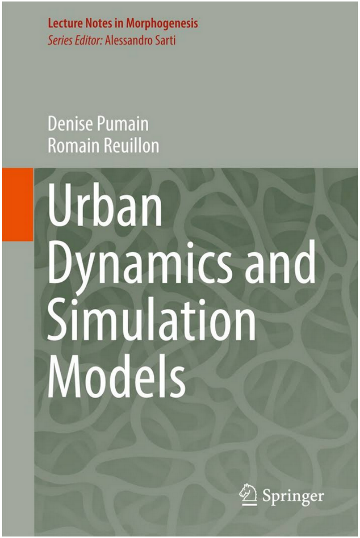
Simpop 1 model[Sanders et al., 1997]



SimpopNet model [Schmitt, 2014]

Necessity of simulation models in geography induced by complexities of these systems ?

- Ontological complexity [Pumain, 2003]
- Dynamical complexity: non-ergodicity and path-dependancy [Pumain, 2012]
- Complexity and co-evolution
- Complexity and emergence

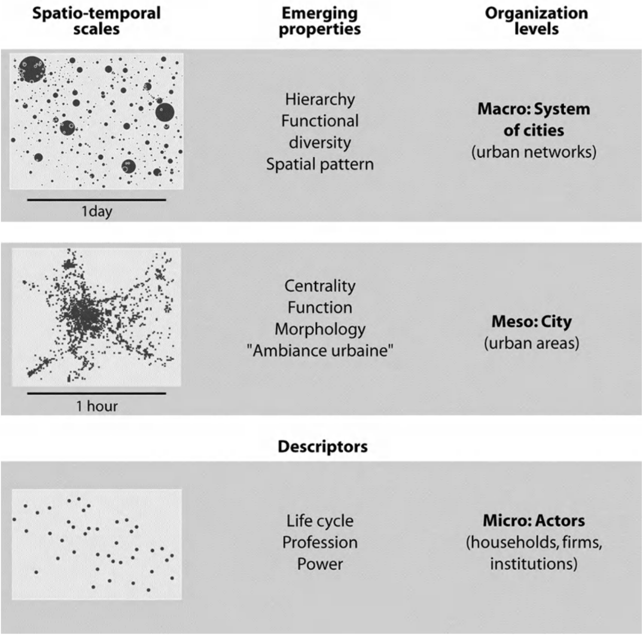


Development of evolutive urban theory [Pumain, 2018]

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

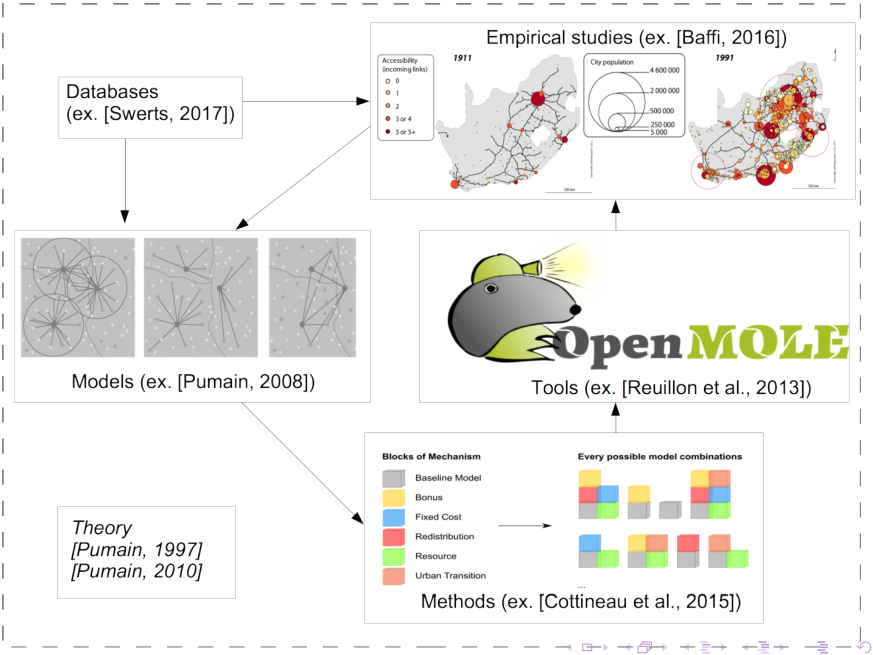
OpenMOLE

Evolutive Urban Theory

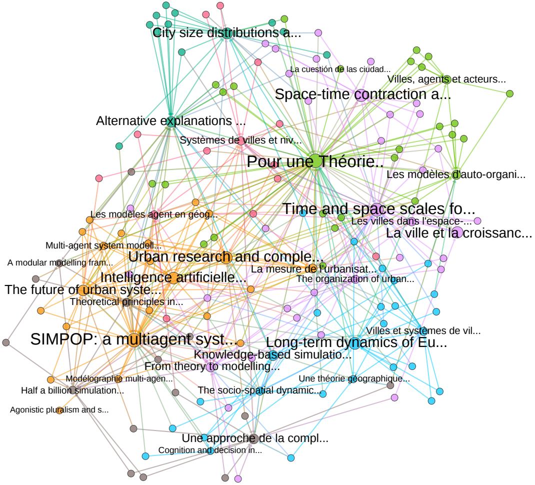


Systems of cities as co-evolutive systems in which interactions are crucial
[Pumain, 1997]
[Pumain, 2008]
[Pumain, 2018]

Iterative Construction of Knowledge across Domains



Evolutive urban theory



[Raimbault, 2017] Citation network analysis of key publications in the evolutive urban theory

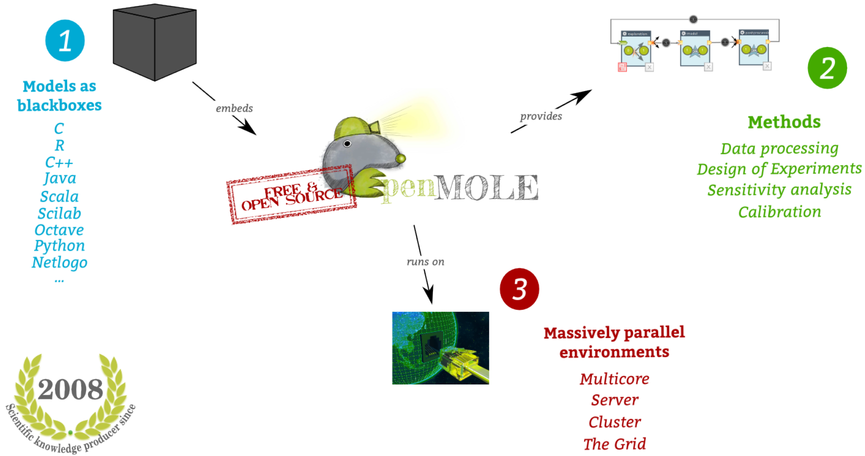
1 Introduction

2 Model exploration methods

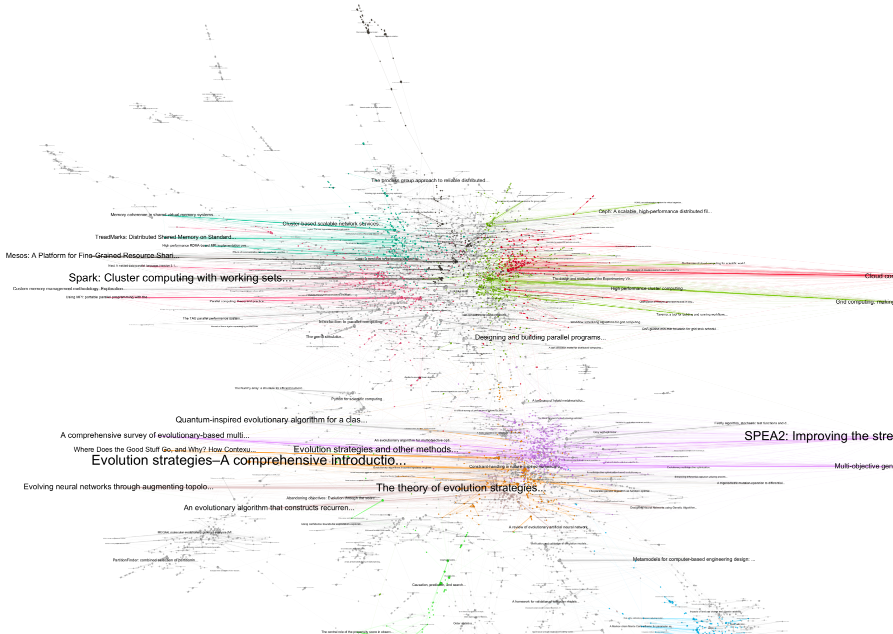
3 Interactions between networks and territories

The OpenMOLE manifesto

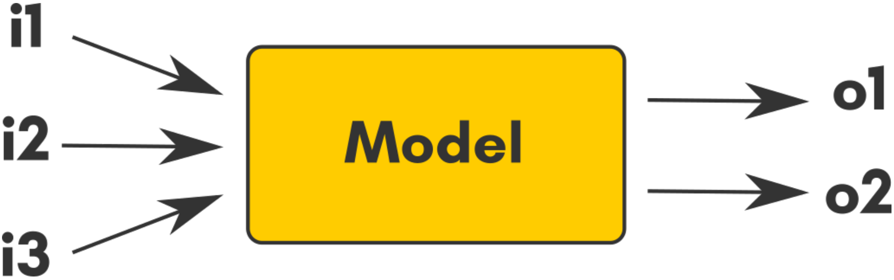
(i) Innovative exploration methods; (ii) Scaling of methods on high performance computing environments; (iii) No interference with the model.



Scientific environment of OpenMOLE



Simulation models



Included methods

- Parameter estimation
- Sensitivity analysis
- Robustness assessment
- Optimization

Designed in a scalable manner, handle stochasticity, usable on any models and environments.

Supported environments

Local prototyping, transparent passage to scale: zero déploiement, pas de connaissance technique requise, pas d'installation préalable.

Environnements pris en charge à l'heure actuelle:

- Multi-thread
- Delegation through SSH
- PBS
- SLURM
- Condor
- SGE
- OAR
- EGI Grid

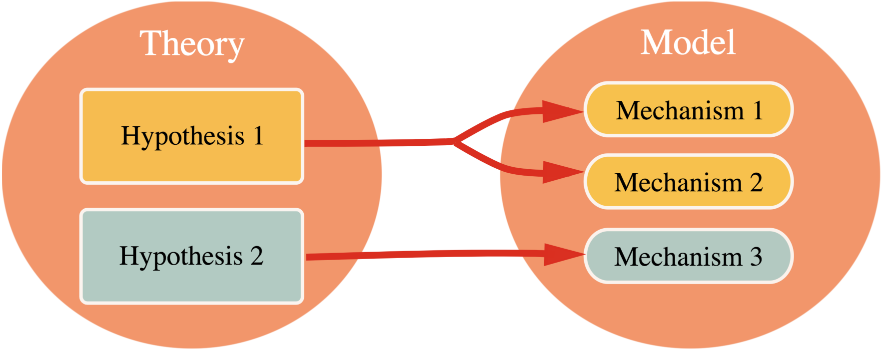
Towards computer-aided modeling

Theoretical framework and methods (algorithms) complementary to the modeling process

A modeling process which is:

1. Tractable: understand the choices made
2. Reproducible: verify the reasons of these
3. Reusable: study alternative choices

Descriptive models



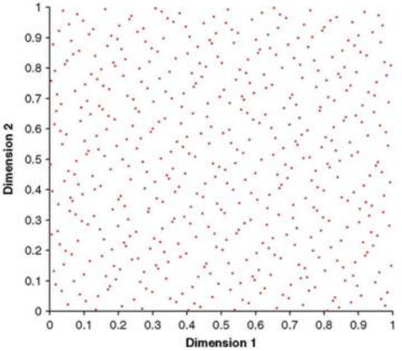
Design/evaluate a theory involving causal effects through its capacity to (re-)produce some patterns/data..

Model evaluation: How to assess

1. the sufficiency of mechanisms ?
2. the necessity of mechanisms ?
3. the uniqueness of the mechanisms ?

Sufficiency

Classical approach: Design of Experiments

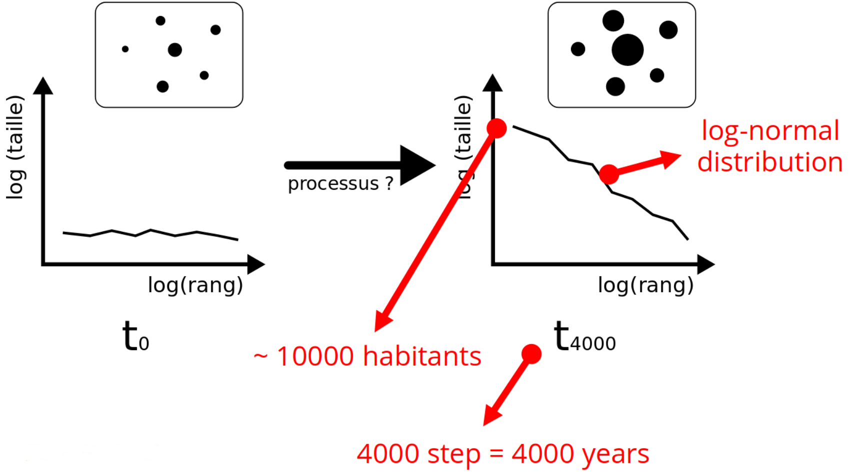


$\rightarrow$ produces a huge quantity of data; transform a problem into a data-mining problem; parameter space is mostly left unexplored due to the curse of dimensionality.

Method to assess sufficiency [Schmitt et al., 2015]

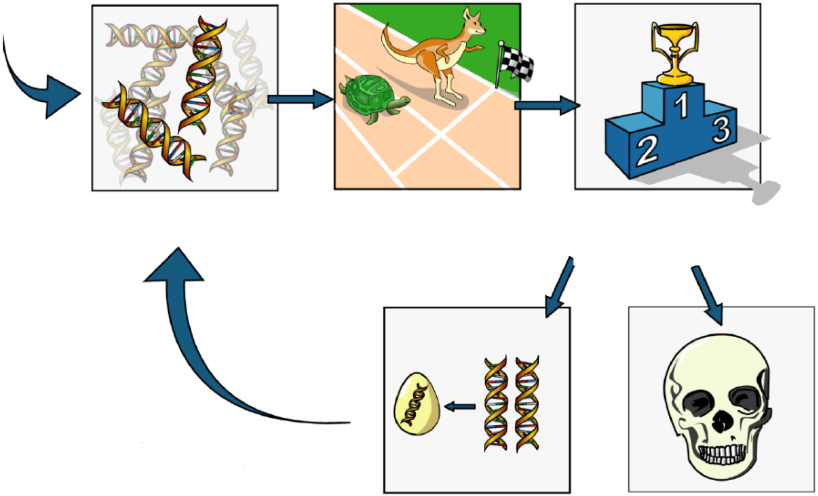
Inverse approach: from outputs to parameters

Formalising the expectations as indicators:



Calibration

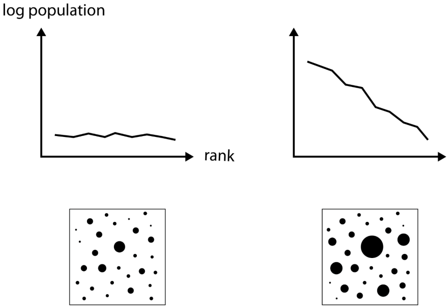
Genetic algorithm for calibration



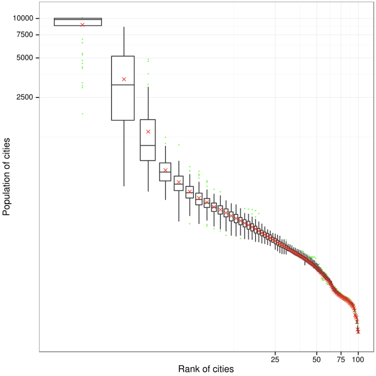
Calibration results

No compromise between the 3 objectives.

Searched pattern



Produced pattern



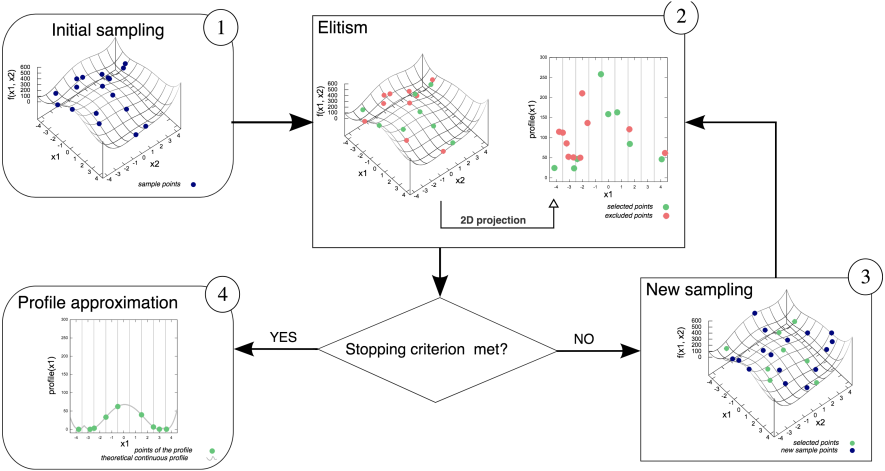
Performances: handles stochasticity: 100x gain, support for distributed computing: 1000x gain.

Necessity [Reuillon et al., 2015]

A new algorithm

1. Detects if a parameter is necessary
2. Better constraints the parameter range
3. As an indirect way to detect if some of the mechanisms are expandable

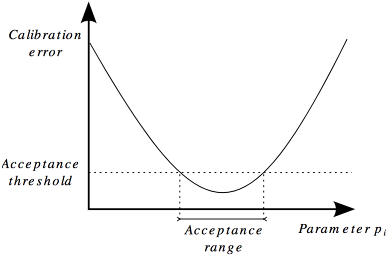
The profile algorithm



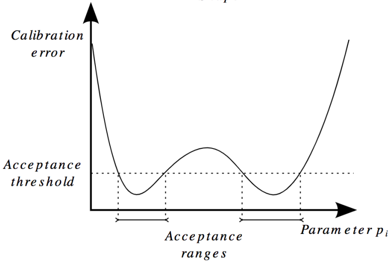
Computes the best calibration at fixed steps along one dimension.

The profile algorithm

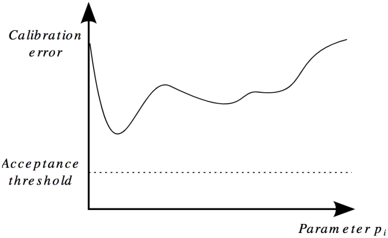
Shape 1



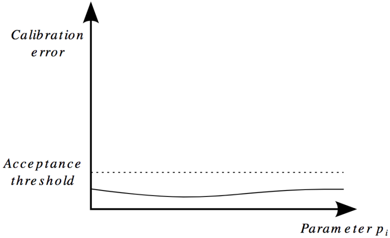
Shape 2



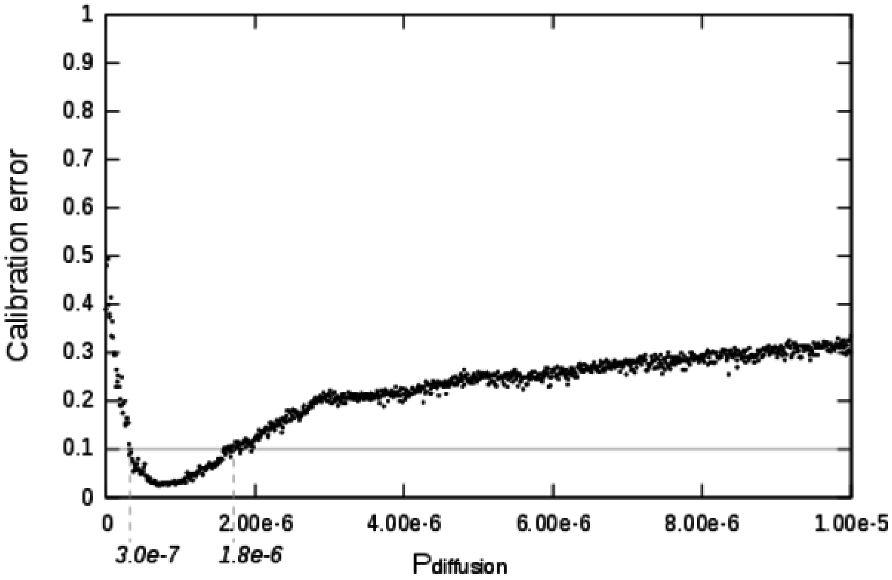
Shape 3



Shape 4

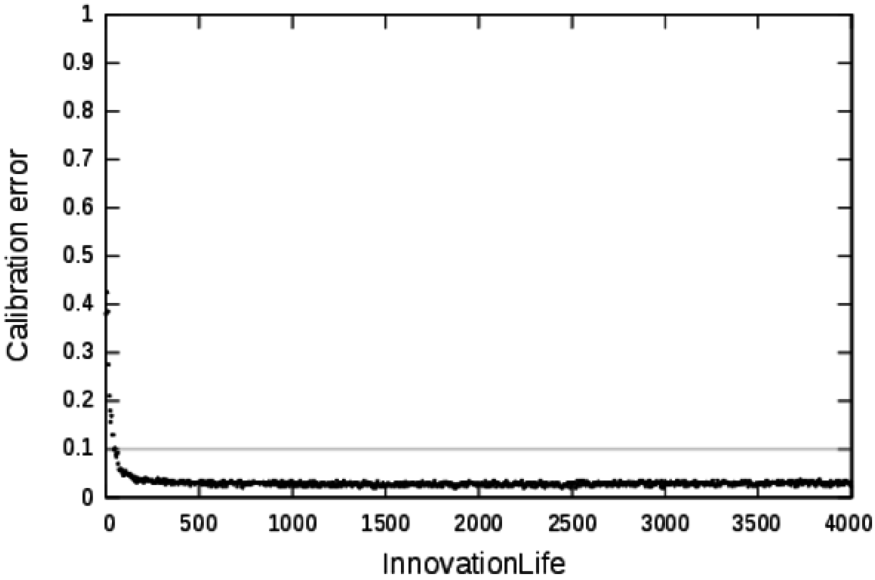


Results



Calibration error
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
0
2.00e-6
4.00e-6
6.00e-6
8.00e-6
1.00e-5
3.0e-7
1.8e-6
Pdiffusion

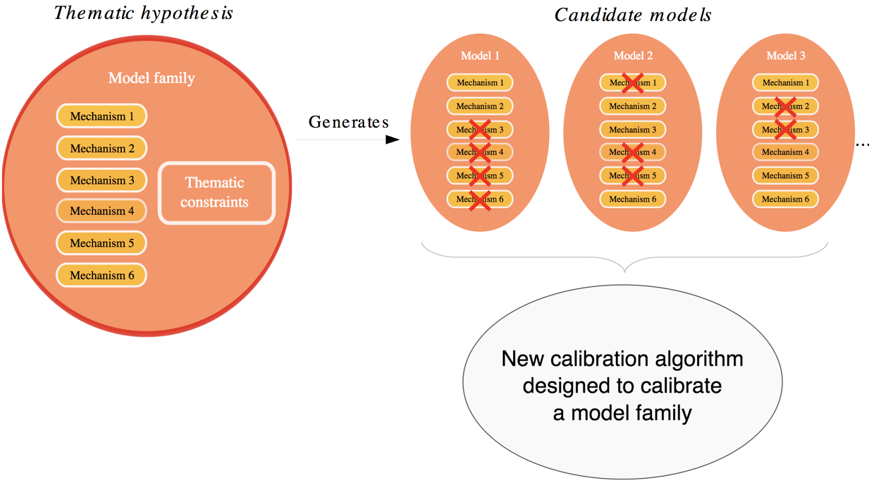
Results



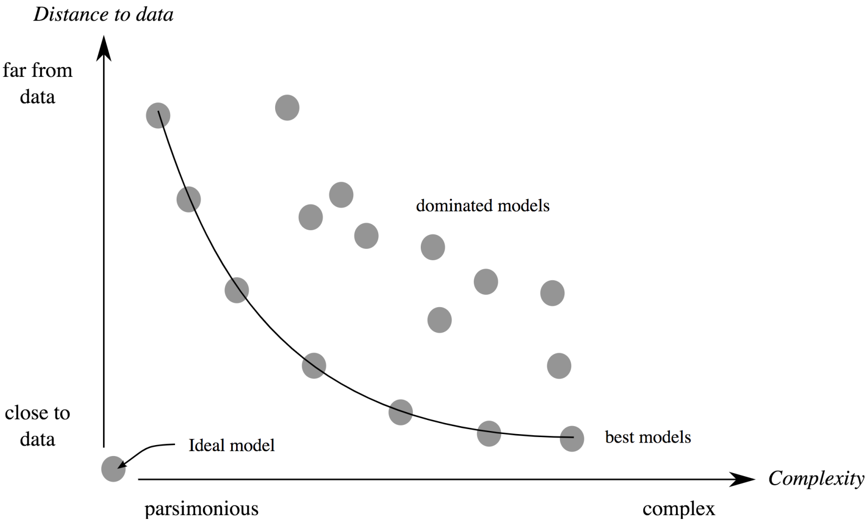
Calibration error
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
0
500
1000
1500
2000
2500
3000
3500
4000
InnovationLife

Unicity [Cottineau, 2014]

Automate the confrontation of alternative hypothesis / mechanisms.

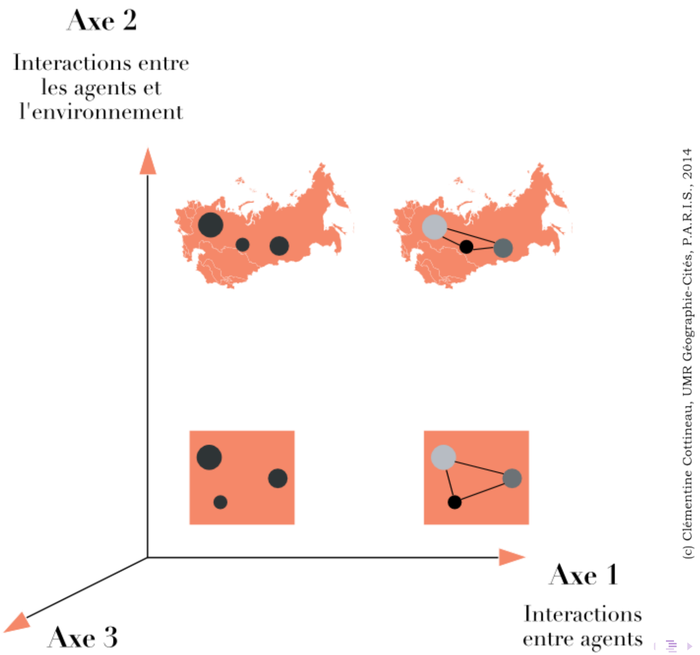


Objective



Distance to data
far from data
dominated models
close to data
Ideal model
best models
Complexity
parsimonious
complex

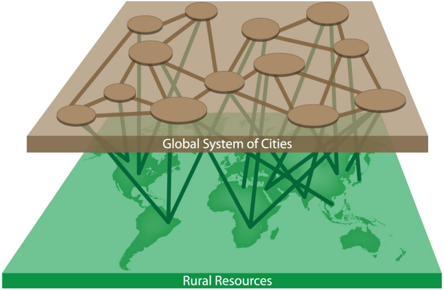
Multi-modeling (64 models)



Exemple of concurrent hypothesis

Exchange mechanism: market vs centralized

City growth: interurban interactions vs environmental situation



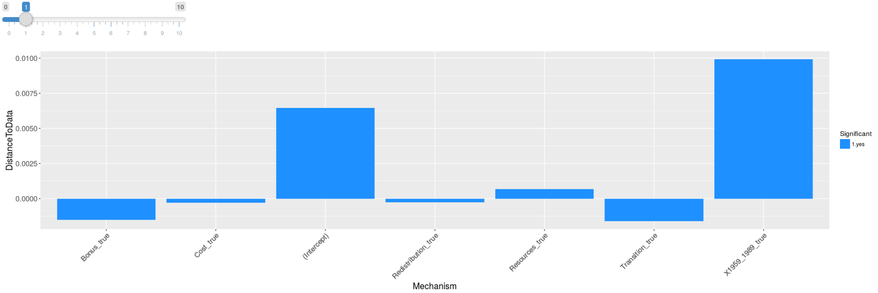
Calibration of model family

Compute the best set of parameters for all 64 models.

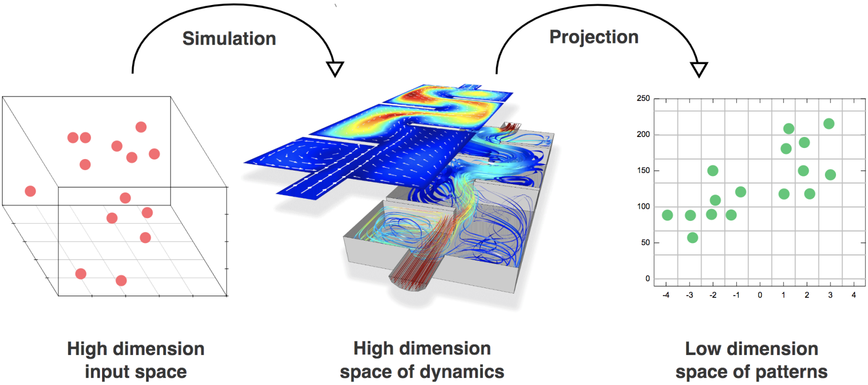
Contribution of mechanisms to the quality of simulation (closeness to data)

Models with different combination of mechanisms have been calibrated intensively against empirical data, using generic algorithms for more than 100000 generations. This plot shows the results of a regression explaining one measure of the quality of models (a small difference between simulated and empirical urban trajectories) by their mechanisms composition (the fact that any of the supplementary mechanisms is activated or not" Each bar represents the value of the estimated coefficient for each activated mechanism, in comparison with the same model structure without this mechanism, everything else being equal.

Statistical Significance (% of error)

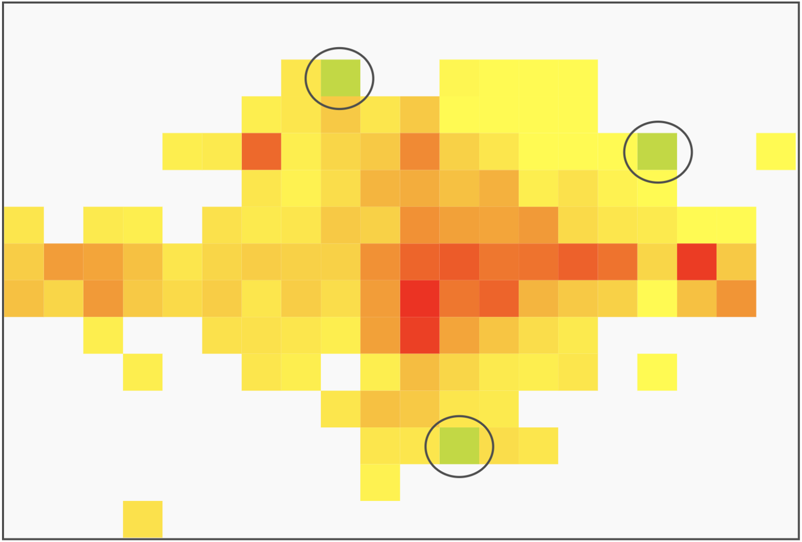


Alternative approach: pattern search [Chérel et al., 2015]

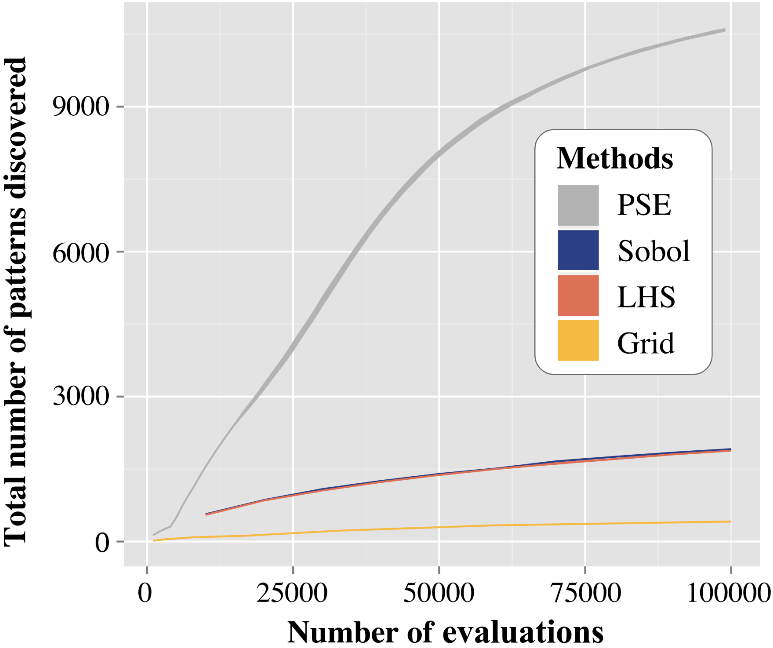


Novelty search

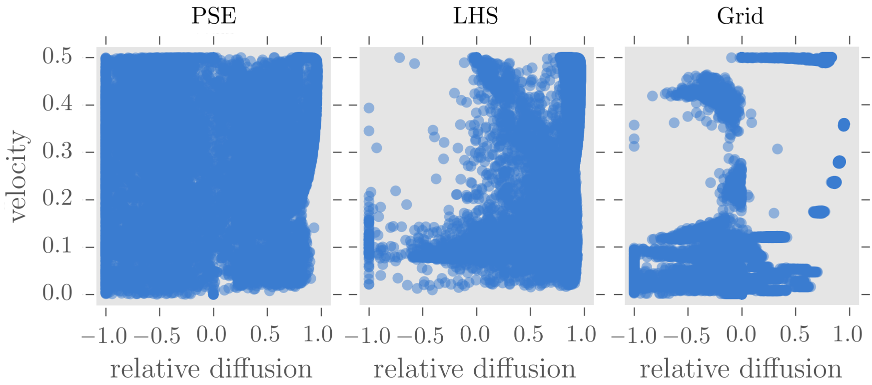
The inputs producing rare patterns have high fitness values.



Results

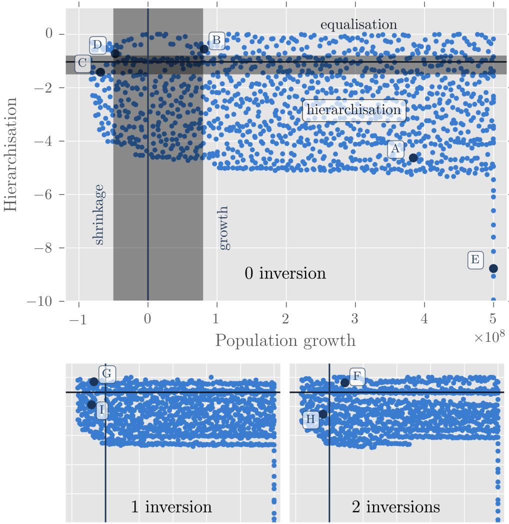

Methods
PSE
Sobol
LHS
Grid
Total number of patterns discovered
9000
6000
3000
0
0
25000
50000
75000
100000
Number of evaluations

Results



PSE
LHS
Grid
velocity
0.5
0.4
0.3
0.2
0.1
0.0
−1.0
−0.5
0.0
0.5
1.0
relative diffusion

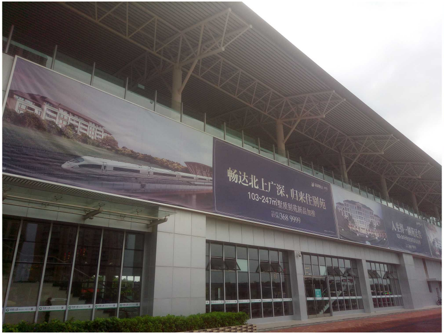
Results



equalisation
hierarchisation
shrinkage
growth
0 inversion
A
B
C
D
E
Hierarchisation
Population growth
$\times 10^8$
0
−2
−4
−6
−8
−10
−1
0
1
2
3
4
5
G
I
1 inversion
F
H
2 inversions

1. Introduction
2. Model exploration methods
3. Interactions between networks and territories

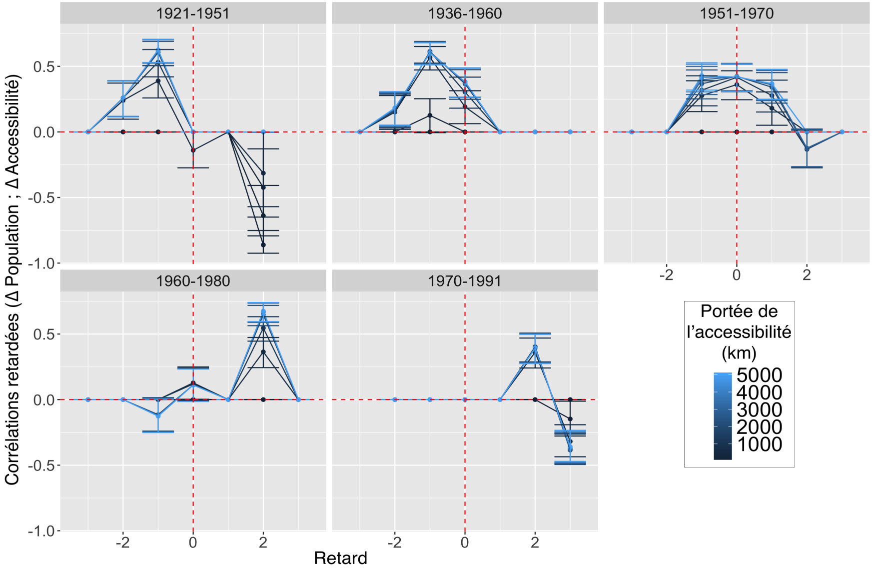
Interactions between networks and territories



Fieldwork observation of interactions between transportation and urban environment in Pearl River Delta: promotion of high speed, targeted urban development around stations.

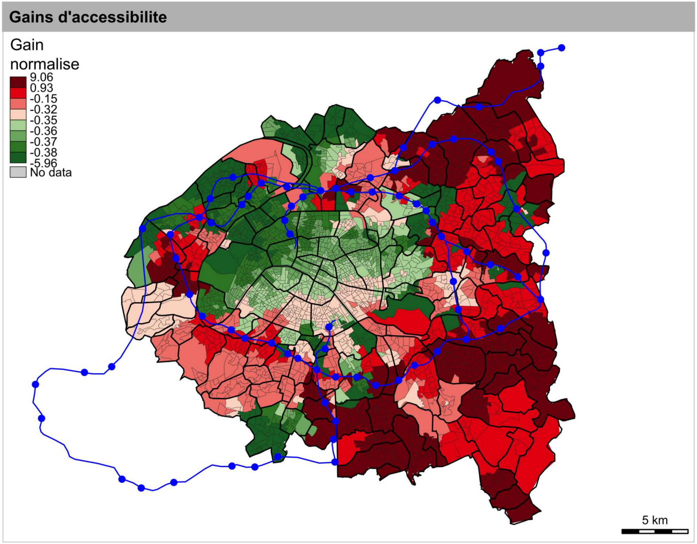
Raimbault, J. (2019). Evolving accessibility landscapes: mutations of transportation networks in China. In Aveline-Dubach, N., ed. *Pathways of sustainable urban development across China - the cases of Hangzhou, Datong and Zhuhai*, pp 89-108. Imago. ISBN:978-88-94384-71-0

Contrasted empirical observations



Inversion of the sens of causality between population growth and railway accessibility increase in South Africa during the 20th century.

Contrasted empirical observations



Relations plus complexes dans le cas du gain d'accessibilité permis par le Grand Paris Express et les dynamiques socio-économiques des territoires

Macroscopic scale:

- Interaction models between cities including networks → *Demonstration of network effects; exploration of interaction regimes*

Mesoscopic scale:

- Urban morphogenesis model coupling urban form and network growth → *Complementarity of multiple processes; calibration at the first and second order*
- Exploration of a model including transportation governance

Macroscopic scale:

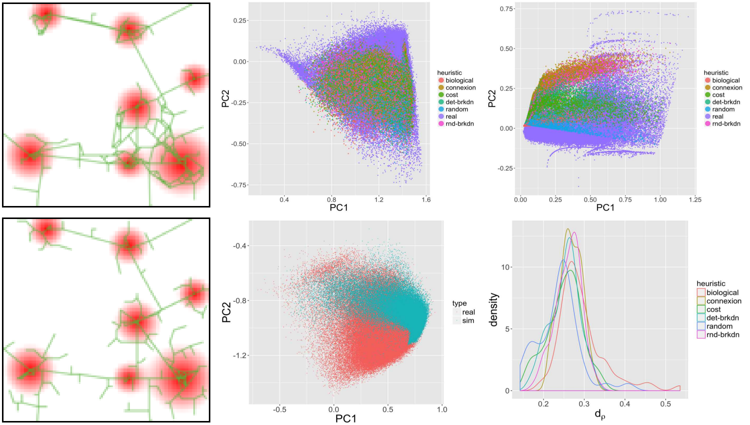
- Interaction models between cities including networks → *Demonstration of network effects; exploration of interaction regimes*

Mesoscopic scale:

- Urban morphogenesis model coupling urban form and network growth
→ *Complementarity of multiple processes; calibration at the first and second order*
- Exploration of a model including transportation governance

Mesoscopic models: morphogenesis

A morphogenesis model with reaction-diffusion and multi-modeling of network growth: complementarity of heuristics, calibration on forms and their correlations



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.

Raimbault, J. (2018). Multi-modeling the morphogenesis of transportation networks. In Artificial Life Conference Proceedings (pp. 382-383).

A simple Reaction-diffusion model

$\rightarrow$ Crucial role of the interplay between concentration forces and dispersion forces [Fujita and Thisse, 1996] in keeping Urban Systems at the border of chaos

$\rightarrow$ Potentiality of aggregation mechanisms (such as Simon model) to produce power laws

$\rightarrow$ Link with Reaction-diffusion approaches in Morphogenesis [Turing, 1952]

$\rightarrow$ Extension of a DLA-type model introduced by [Batty, 1991], with simple abstract processes of population aggregation and diffusion

Model Formalization

→ Grid world with cell populations $(P_i(t))_{1\leq i\leq N^2}$.

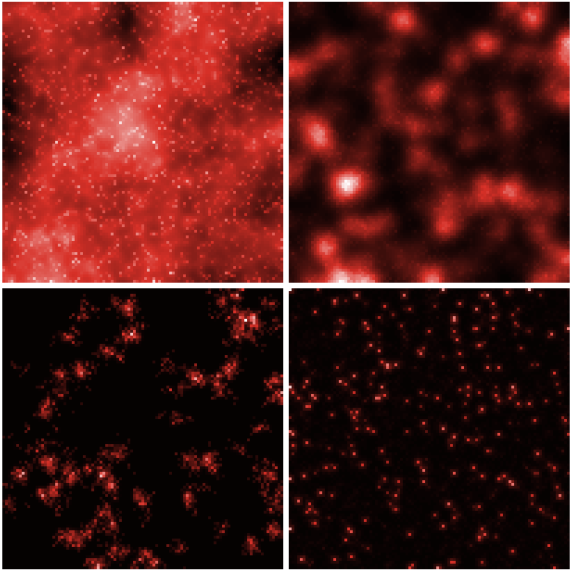
→ At each time step:

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.

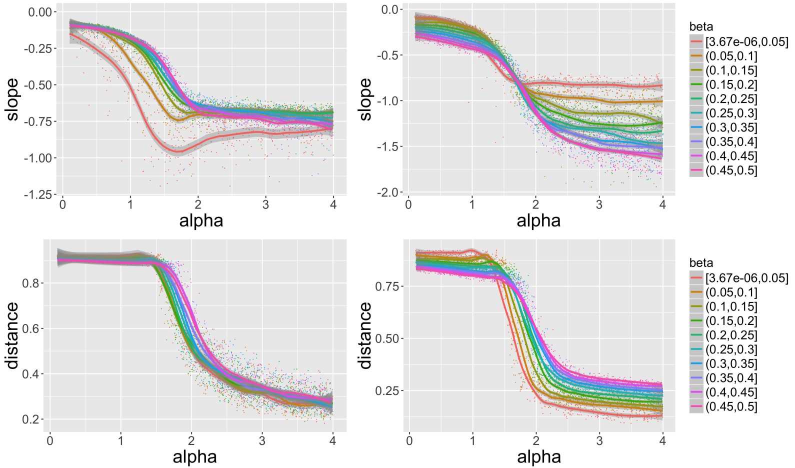
→ Output measured by morphological indicators: Moran index, average distance, rank-size hierarchy, entropy.

Generating Population Distributions



Examples of generated territorial shapes

Model behavior



Phase transitions of indicators unveiled by exploration of the parameter space (80000 parameter points, 10 repetitions each)

Path-dependence and frozen accidents

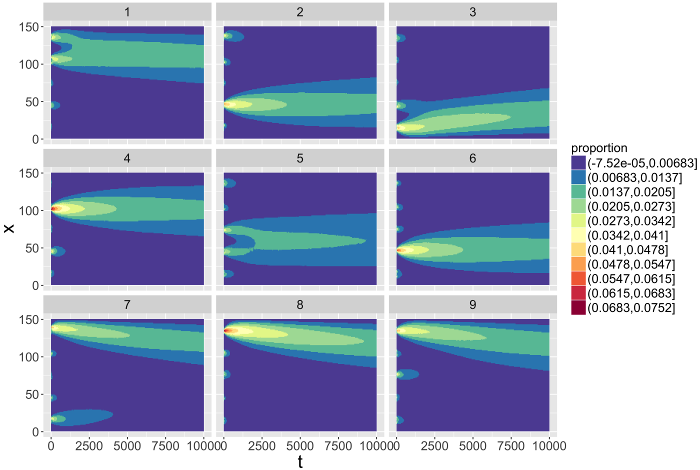
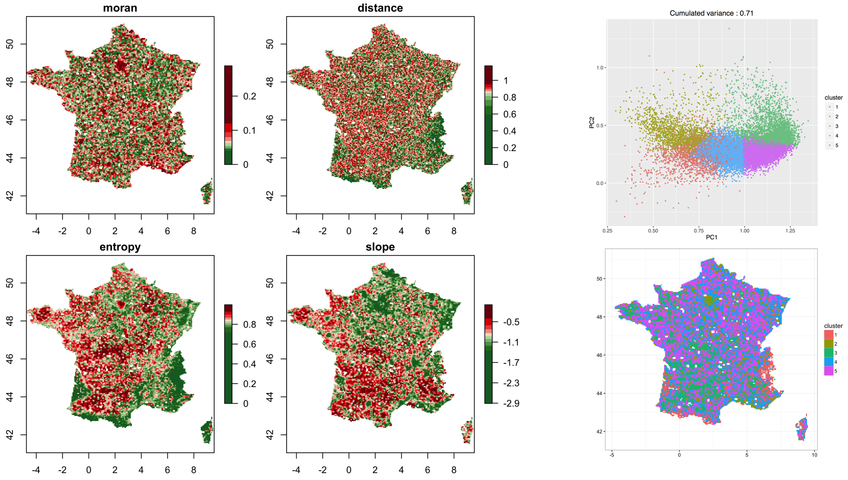


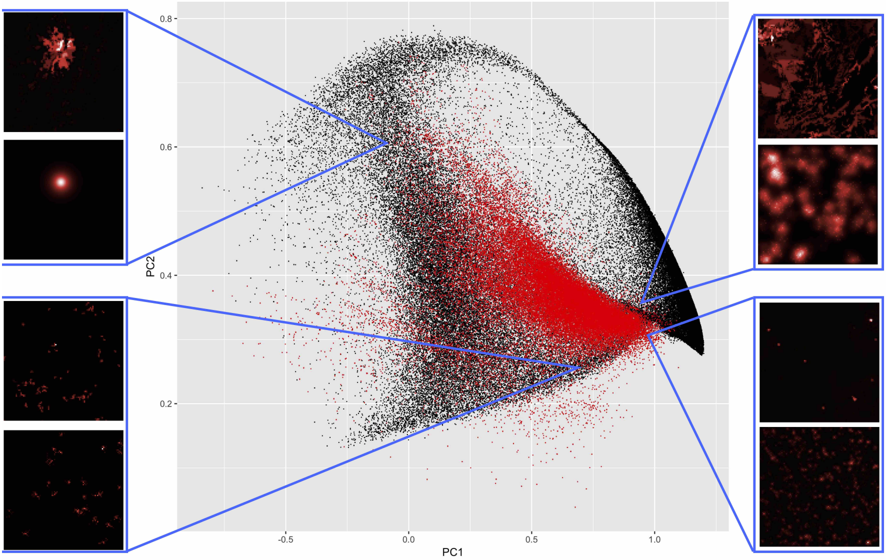
Illustration of path-dependence in a simplified one-dimensional version of the model: cell trajectories in time for 9 independent repetitions from the same initial configuration.

Empirical Data for Calibration



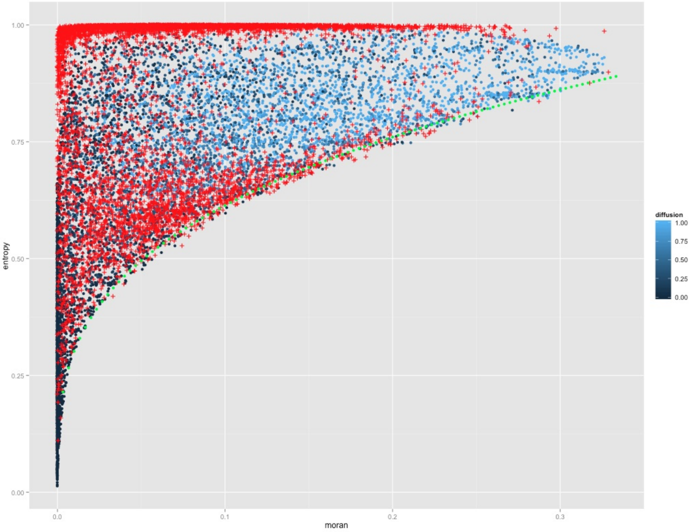
Computation of morphological indicators on population density data for Europe (shown here on France), morphological classification.

Model Calibration



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

Model Targeted Exploration



Potentialities of targeted model explorations: here feasible space using Pattern Space Exploration algorithm [Chérel et al., 2015].

Including more complex processes ?

Which ontology to include more complex functional properties ?

$\rightarrow$ Territorial systems as the strong coupling between territories and (potential and realized) networks [Dupuy, 1987].

$\rightarrow$ Networks convey functional notions of centralities and accessibility, among others ; have furthermore proper topological properties.

A Morphogenesis Model of co-evolution

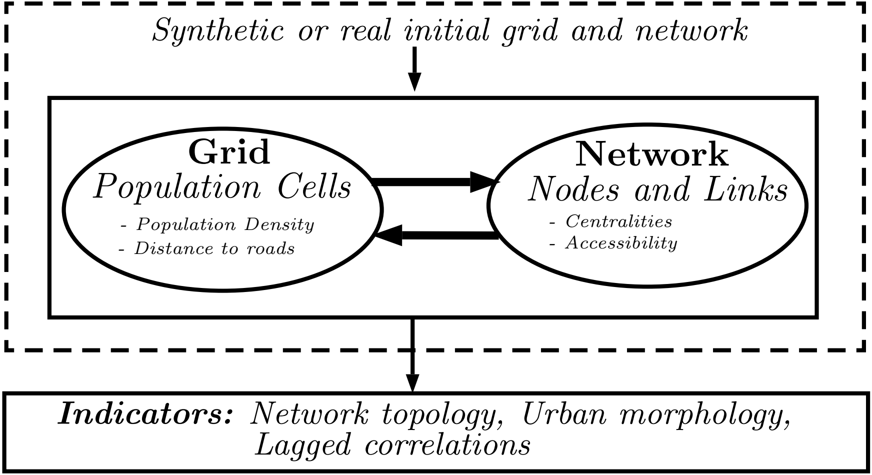
→ Coupled grid population distribution and vector transportation network, following the core of [Raimbault et al., 2014]

→ Local morphological and functional variables determine a patch-value, driving new population attribution through preferential attachment ; combined to population diffusion (reaction-diffusion processes studied before)

→ Network growth is also driven by morphological, functional and local network measures, following diverse heuristics corresponding to different processes (multi-modeling)

*Local variables and network properties induce feedback on both, thus a strong coupling capturing the **co-evolution***

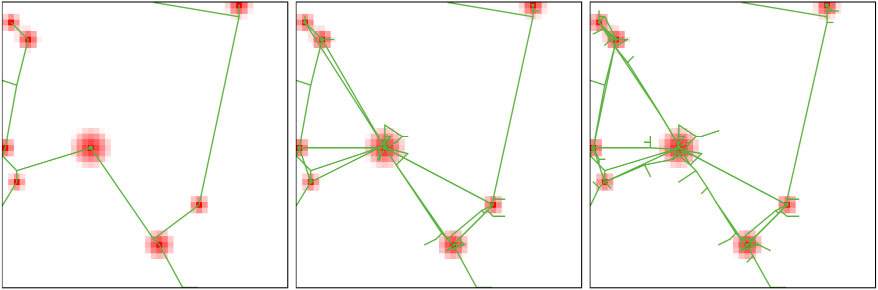
Model : Specification



Network Generation

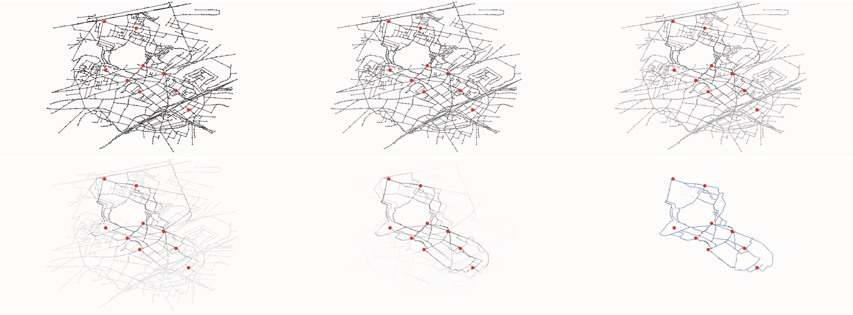
At fixed time steps :

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])



Biological network generation

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

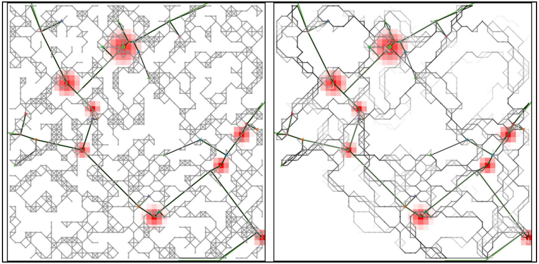


Application to the design of optimal bus routes

Biological Network generation

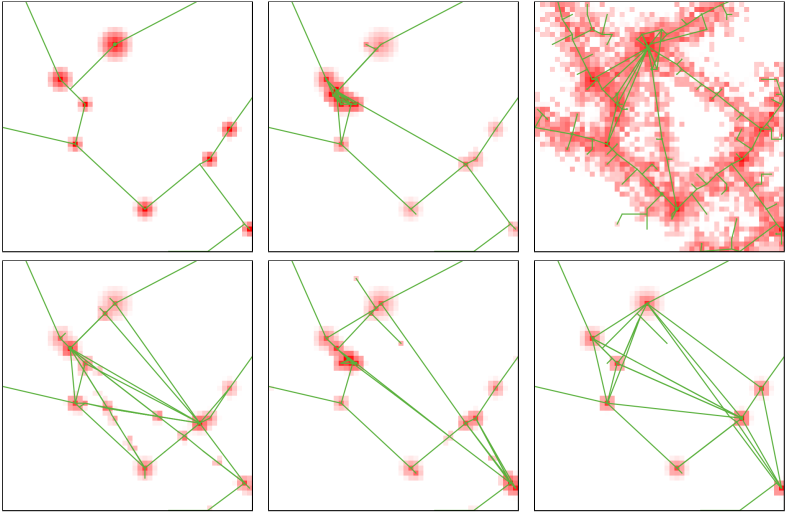
Adding new links with biological heuristic:

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



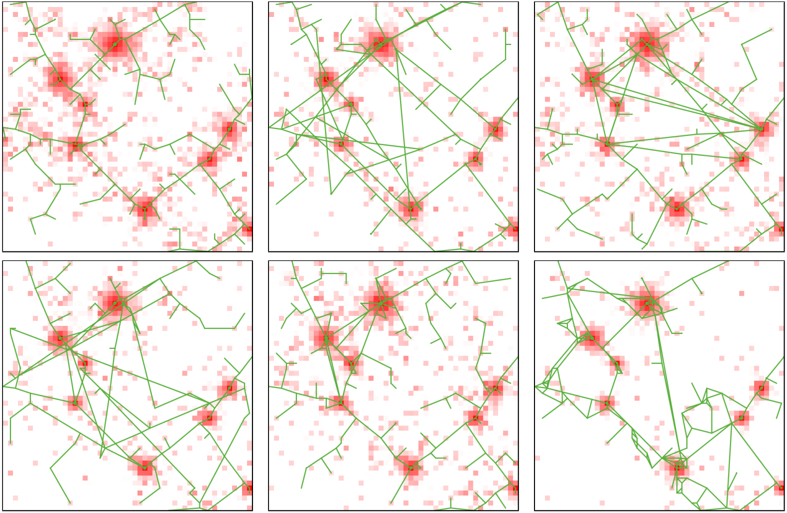
Intermediate stage for biological network generation

Generated Urban Shapes: Urban Form



In order: setup; accessibility driven; road distance driven; betweenness driven; closeness driven; population driven.

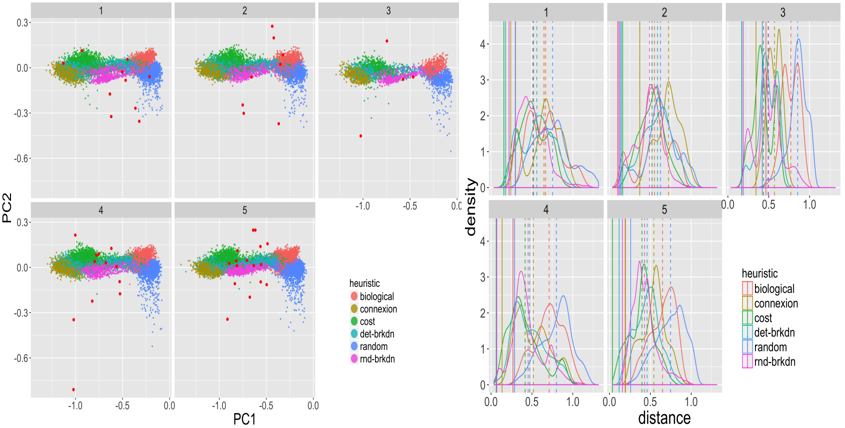
Generated Urban Shapes: Network



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

Results : Network Heuristics

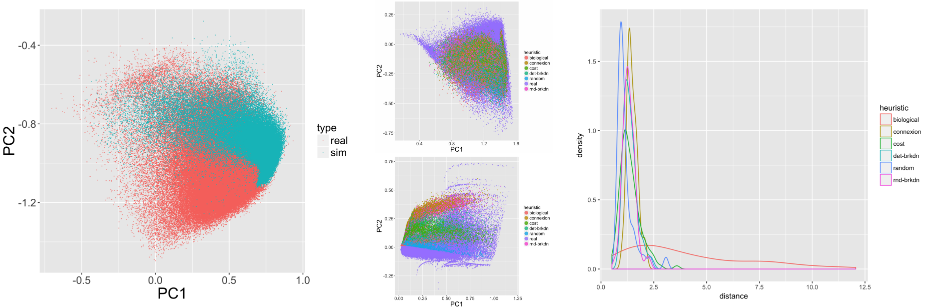
Comparison of feasible space for network indicators with fixed density



(Left) Feasible spaces by morphological class and network heuristic; (Right) Distribution of distances to topologies of real networks

Results : Calibration

Calibration (model explored with OpenMole [Reuillon et al., 2013], $\sim 10^6$ model runs) at the first order on morphological and topological objectives, and on correlations matrices.

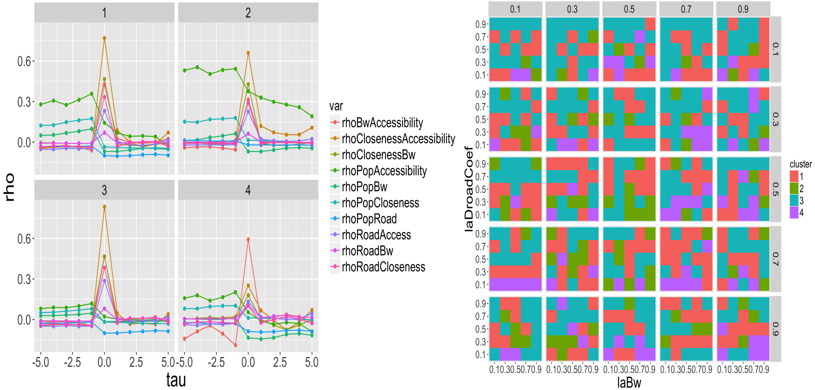


(Left) Full indicator space; (Middle) Morphological and Topology, by network heuristic; (Right) Distance distribution for cumulated distance for indicators and correlations.

Results : Causality Regimes

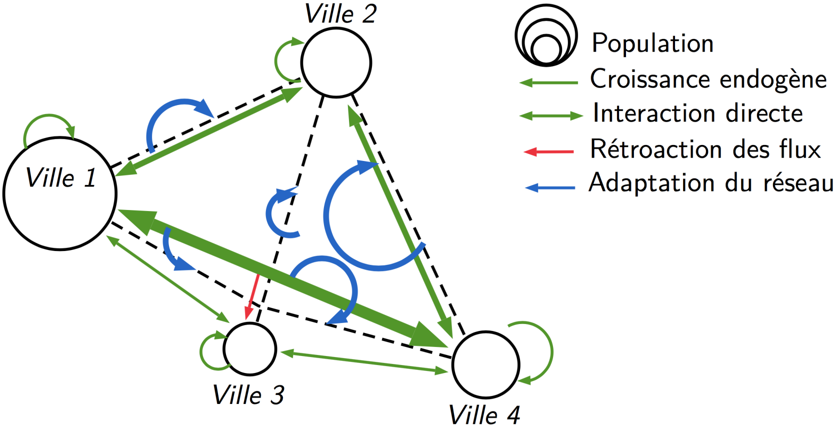
Unsupervised learning on lagged correlations between local variables unveils a diversity of causality regimes

$\rightarrow$ Link between *co-evolution regime* and morphogenetic properties of the urban system



(Left) Lagged correlation profiles of cluster centers; (Right) Distribution of regimes across parameter space

Macroscopic interaction model



Raimbault, J. (2018). Indirect evidence of network effects in a system of cities. Environment and Planning B: Urban Analytics and City Science, 2399808318774335.

Raimbault, J. (2019). Modeling the co-evolution of cities and networks. In Niel, Z., Rozenblat, C., eds. *Handbook of Cities and Network*, Edwar Elgar Publishing, *in press*.

Rationale : extend an interaction model for system of cities by including physical network as an additional carrier of spatial interactions

$\rightarrow$ Work under Gibrat independence assumptions, i.e. $\mathrm{Cov}[P_i(t), P_j(t)] = 0$. If $\vec{P}(t+1) = \mathbf{R} \cdot \vec{P}(t)$ where $\mathbf{R}$ is also independent, then $\mathbb{E}\left[\vec{P}(t+1)\right] = \mathbb{E}[\mathbf{R}] \cdot \mathbb{E}\left[\vec{P}\right](t)$. Consider expectancies only (higher moments computable similarly)

$\rightarrow$ With $\vec{\mu}(t) = \mathbb{E}\left[\vec{P}(t)\right]$, we generalize this approach by taking $\vec{\mu}(t+1) = f(\vec{\mu}(t))$

Macroscopic Model Formulation

Let $\vec{\mu}(t) = \mathbb{E}\left[\vec{P}(t)\right]$ cities population and (d_{ij}) distance matrix

Model specified by

$$f(\vec{\mu}) = r_0 \cdot \mathbf{Id} \cdot \vec{\mu} + \mathbf{G} \cdot \mathbf{1} + \mathbf{N}$$

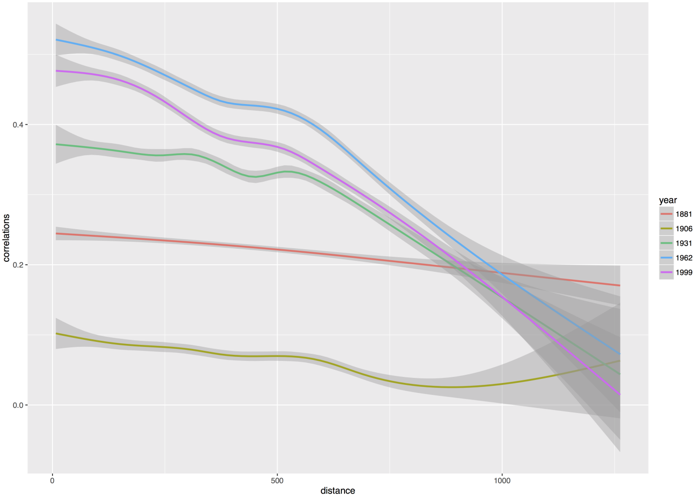
with

- $G_{ij} = w_G \cdot \frac{V_{ij}}{<V_{ij}>}$ and $V_{ij} = \left(\frac{\mu_i \mu_j}{\sum {\mu_k}^2}\right)^{\gamma_G} \exp\left(-d_{ij}/d_G\right)$
- $N_i = w_N \cdot \sum_{kl} \left(\frac{\mu_k \mu_l}{\sum \mu}\right)^{\gamma_N} \exp\left(-d_{kl,i}\right)/d_N$ where $d_{kl,i}$ is distance to shortest path between k,l computed with slope impedance $(Z = (1+\alpha/\alpha_0)^{n_0}$ with $\alpha_0 \simeq 3)$

Data : stylized facts

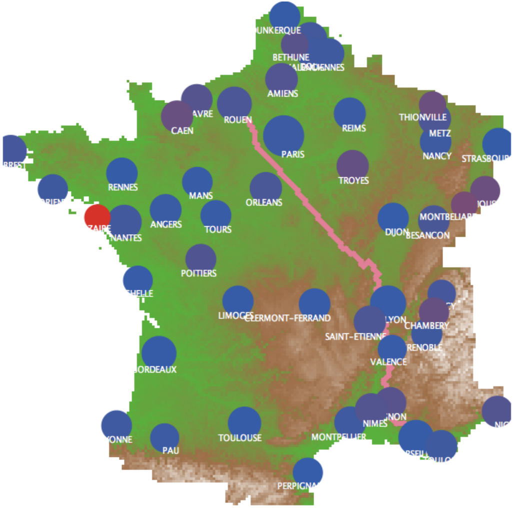
Population data for French-cities (Pumain-INED database : 1831-1999)

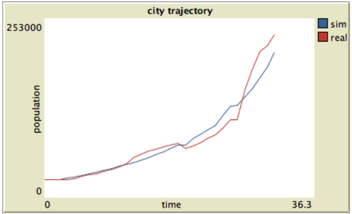
Non-stationarity of log-returns correlations function of distance

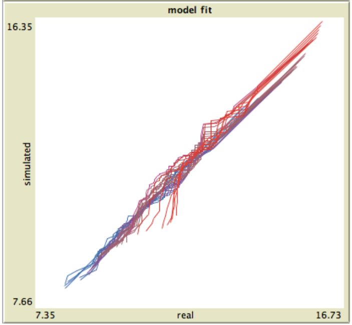


Geographic abstract network

Physical transportation network abstracted through a geographical shortest path network

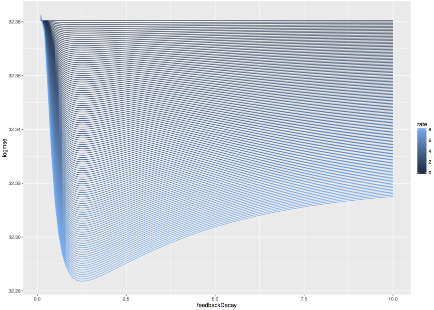


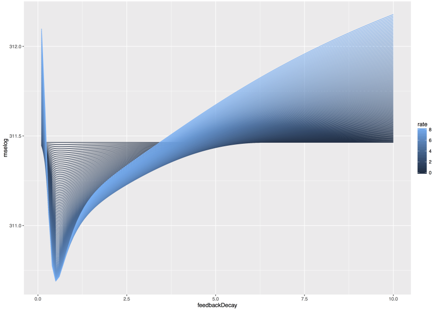




Results : model exploration

Evidence of physical network effects : fit improve through feedback at fixed gravity

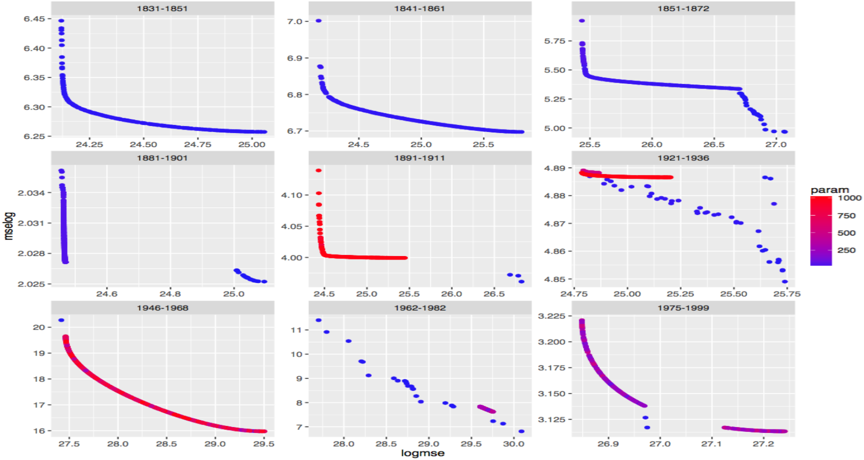




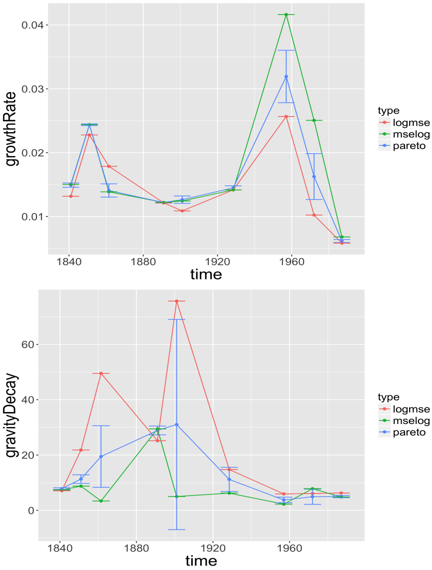
Results : model calibration

Model calibration using GA on computation grid, with software `OpenMole` [Reuillon et al., 2013]

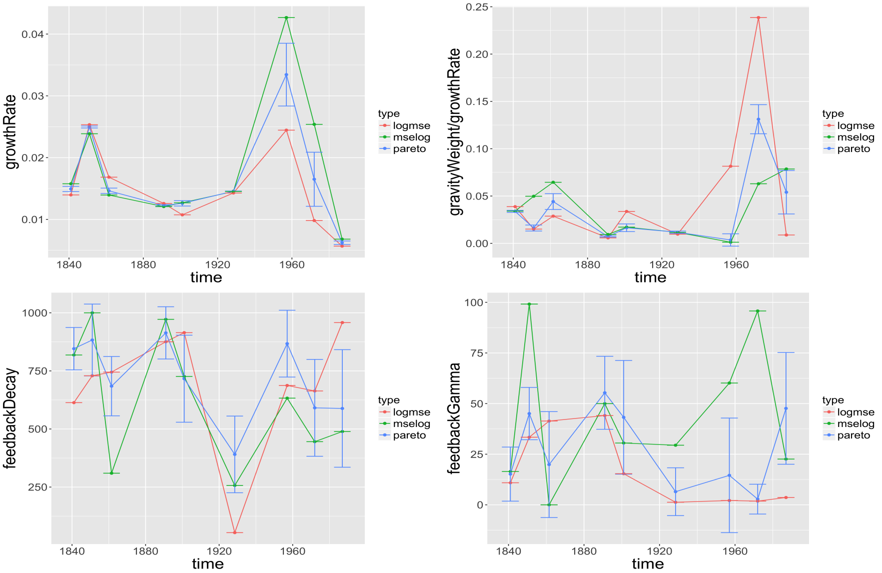
Pareto front for full model calibration, objectives MSE and MSE on logs

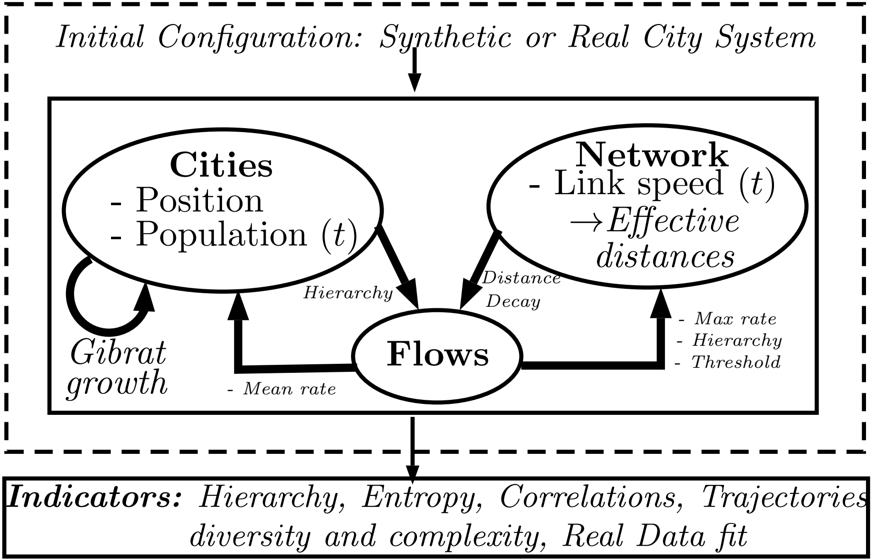


Results : non-stationary gravity model calibration



Results : non-stationary full model calibration




Initial Configuration: Synthetic or Real City System
Cities
- Position
- Population (t)
Network
- Link speed (t)
→Effective distances
Gibrat growth
Hierarchy
Distance Decay
Flows
- Mean rate
- Max rate
- Hierarchy
- Threshold
Indicators: Hierarchy, Entropy, Correlations, Trajectories diversity and complexity, Real Data fit

Model Formalization : Network Growth

Given the flow ϕ in a link, its effective distance is updated following

1. For the thresholded case

$$d(t+1) = d(t) \cdot \left(1 + g_{max} \cdot \left[\frac{1 - \left(\frac{\phi}{\phi_0}\right)^{\gamma_s}}{1 + \left(\frac{\phi}{\phi_0}\right)^{\gamma_s}} \right] \right)$$

2. For the full growth case

$$d(t+1) = d(t) \cdot \left(1 + g_{max} \cdot \left[\frac{\phi}{\max \phi} \right]^{\gamma_s} \right)$$

where γ_s is a hierarchy parameter, ϕ_0 a threshold parameter and g_{max} the maximal growth rate easily adjustable to realistic values by computing $(1 + g_{max})^{t_f}$

Model Formalization : Indicators

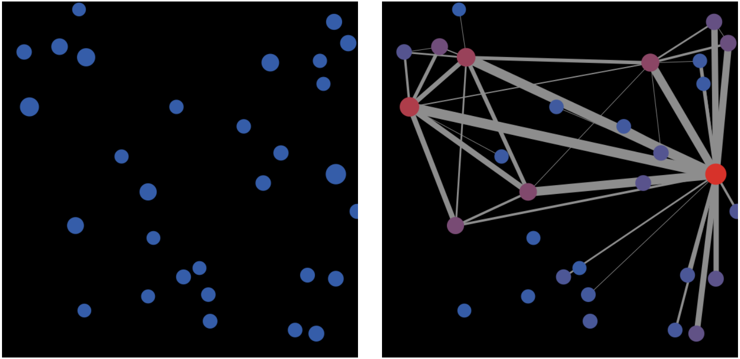
- Hierarchy, Entropy, Summary statistics in time
- Initial-final rank correlation (changes in the hierarchy) for variable X : $\rho\left[X_i(t=0), X_i(t=t_f)\right]$
- Trajectory diversity for variable X : with $\tilde{X}_i(t) \in [0;1]$ rescaled trajectories,

$$\frac{2}{N\cdot(N-1)}\sum_{i<j}\left(\frac{1}{T}\int_t\left(\tilde{X}_i(t)-\tilde{X}_j(t)\right)^2\right)^{\frac{1}{2}}$$

- Average trajectory complexity (number of inflexion points)
- Pearson correlations conditionally to distance $\hat{\rho}_d\left[(X(\vec{x}_1, Y(\vec{x}_2))|||\vec{x}_1-\vec{x}_2|| \sim d\right]$
- Lagged return correlations $\hat{\rho}_\tau\left[\Delta X(t), \Delta Y(t-\tau)\right]$ (Granger causality)

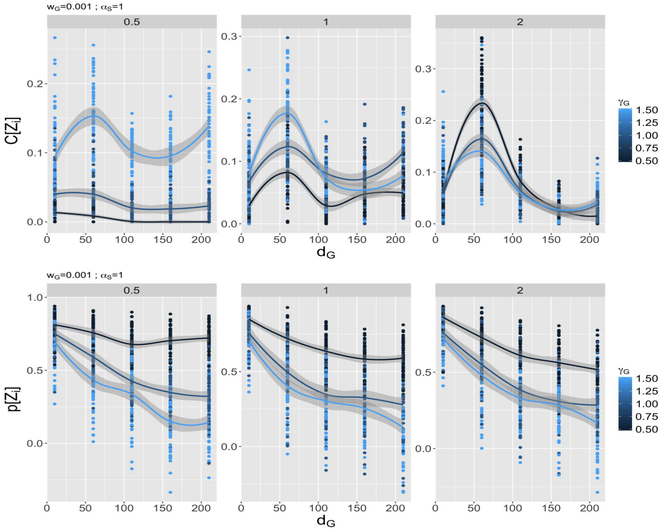
Model Specification : Abstract Network

Complete virtual network between cities, initialized with euclidian distances ; thresholded reinforcement of speeds as a function of flows.



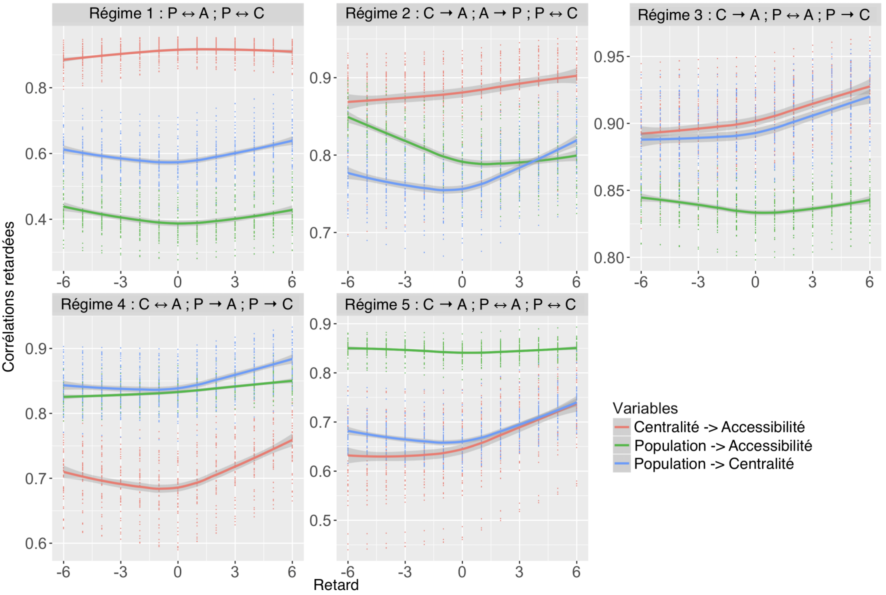
Exemple of run ($t_f = 30$). Level of red gives overall growth and link width flows.

Results: emergence of co-evolution niches



Trajectories of maximal complexity for intermediate values of interaction range.

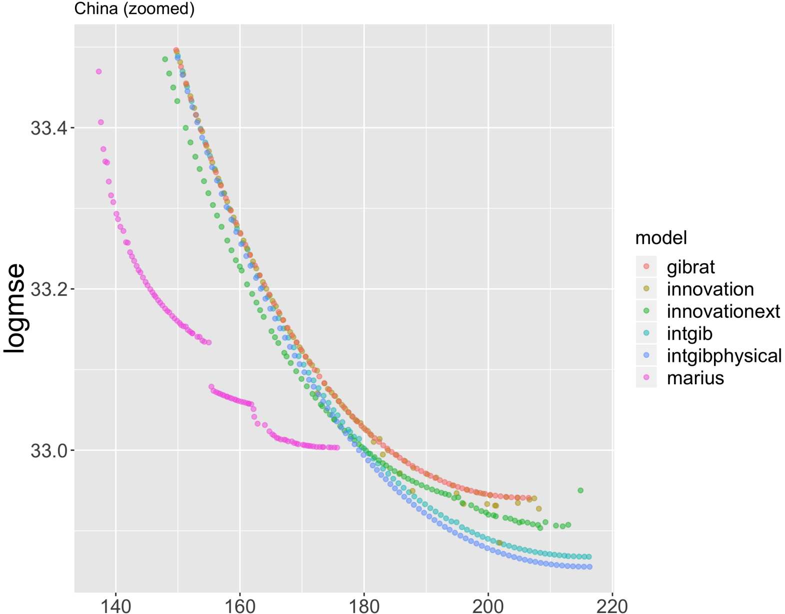
Results: co-evolution regimes



Multiple co-evolution regimes highlighted for synthetic urban systems,

Other example of multi-modeling

Benchmark of growth models for systems of cities [Raimbault, 2018]



Conclusion

→ Multiple ways to model urban systems and to extract knowledge from these: **towards more coupling and comparison of models.**

→ At which scale ? **Need for multi-scale models.**

→ With more refined urban characteristics and other dimensions? **Need for more interdisciplinarity.**

To use OpenMOLE (free and open software) and contribute:
`next.openmole.org`

Some references

Raimbault, J. (2018). Caractérisation et modélisation de la co-évolution des réseaux de transport et des territoires (Doctoral dissertation, Université Paris 7 Denis Diderot). https://halshs.archives-ouvertes.fr/tel-01857741

Raimbault, J. (2018). Indirect evidence of network effects in a system of cities. Environment and Planning B: Urban Analytics and City Science, 2399808318774335.

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.

OpenMOLE summer school

Next week is our summer school, soon receiving applications for next year

`exmodelo.org`

Submit to special session at CCS

Satellite session on methods and epistemology in modeling and simulation, at Conference on Complex Systems, 2nd October 2019

Submit your abstract before June 30th !

`https://iscpif.fr/ccs-satelllite-session-2019-new-methods/`

Submission link:

`https://easychair.org/conferences/?conf=simexplo2019`

Reserve Slides

Baffi, S. (2016).
Railways and city in territorialization processes in South Africa : from separation to integration ?
PhD thesis, Université Paris 1 - Panthéon Sorbonne.

Banos, A. and Genre-Grandpierre, C. (2012).
Towards new metrics for urban road networks: Some preliminary evidence from agent-based simulations.
In *Agent-based models of geographical systems*, pages 627–641. Springer.

Batty, M. (1991).
Generating urban forms from diffusive growth.
Environment and Planning A, 23(4):511–544.

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

Cottineau, C. (2014).
L'évolution des villes dans l'espace post-soviétique. Observation et modélisations.
PhD thesis, Université Paris 1 Panthéon-Sorbonne.

Cottineau, C., Reuillon, R., Chapron, P., Rey-Coyrehourcq, S., and Pumain, D. (2015).
A modular modelling framework for hypotheses testing in the simulation of urbanisation.
Systems, 3(4):348–377.

References III

Dupuy, G. (1987).
Vers une théorie territoriale des réseaux: une application au transport urbain.
In *Annales de Géographie*, pages 658–679. JSTOR.

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.

Openshaw, S., Charlton, M., Wymer, C., and Craft, A. (1987).
A mark 1 geographical analysis machine for the automated analysis of point data sets.
International Journal of Geographical Information System, 1(4):335–358.

Pumain, D. (1997).
Pour une théorie évolutive des villes.
Espace géographique, 26(2):119–134.

Pumain, D. (2003).
Une approche de la complexité en géographie.
Géocarrefour, 78(1):25–31.

Pumain, D. (2008).
The socio-spatial dynamics of systems of cities and innovation processes: a multi-level model.
The Dynamics of Complex Urban Systems, pages 373–389.

Pumain, D. (2010).
Une théorie géographique des villes.
Bulletin de la Société géographie de Liège, (55):5–15.

Pumain, D. (2012).
Urban systems dynamics, urban growth and scaling laws: The question of ergodicity.
In *Complexity Theories of Cities Have Come of Age*, pages 91–103. Springer.

Pumain, D. (2018).
An evolutionary theory of urban systems.
In *International and Transnational Perspectives on Urban Systems*, pages 3–18. Springer.

Raimbault, J. (2017).
An applied knowledge framework to study complex systems.
In *Complex Systems Design & Management*, pages 31–45.

Raimbault, J. (2018).
A systematic comparison of interaction models for systems of cities.
In *Conference on Complex Systems 2018*.

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.

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 VII

Reuillon, R., Schmitt, C., De Aldama, R., and Mouret, J.-B. (2015).
A new method to evaluate simulation models: The calibration profile (cp) algorithm.
Journal of Artificial Societies and Social Simulation, 18(1):12.

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.

Schmitt, C. (2014).
Modélisation de la dynamique des systèmes de peuplement: de SimpopLocal à SimpopNet.
PhD thesis, Paris 1.

Schmitt, C., Rey-Coyrehourcq, S., Reuillon, R., and Pumain, D. (2015).
Half a billion simulations: Evolutionary algorithms and distributed computing for calibrating the simpoplocal geographical model.
Environment and Planning B: Planning and Design, 42(2):300–315.

Swerts, E. (2017).
A data base on chinese urbanization: Chinacities.
Cybergeo: European Journal of Geography.

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.

References IX

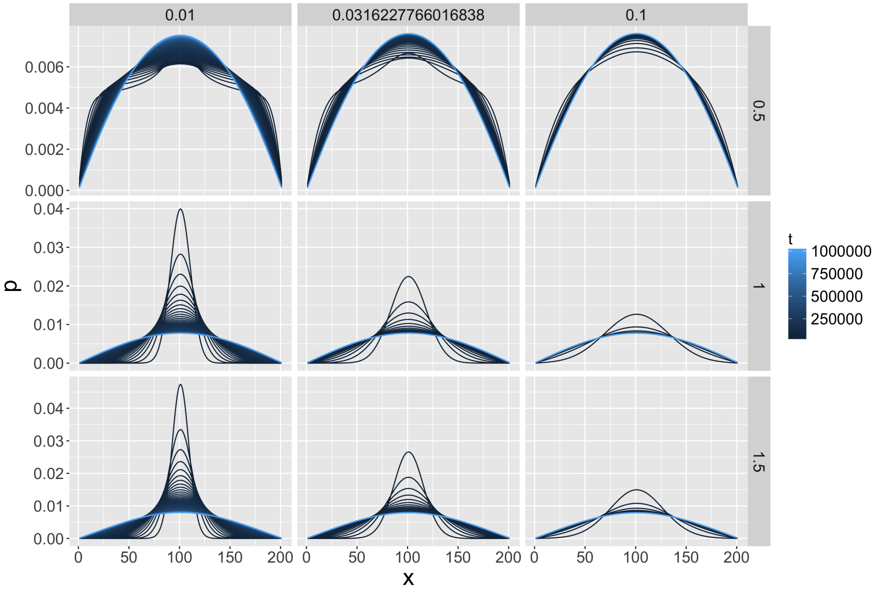
Turing, A. M. (1952).
The chemical basis of morphogenesis.
Philosophical Transactions of the Royal Society of London B: Biological Sciences, 237(641):37–72.

Model classification : PDE

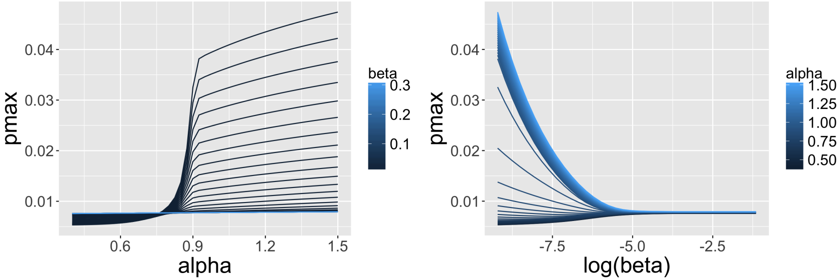
The one-dimensional model verifies the PDE :

$$\delta t \cdot \frac{\partial p}{\partial t} = \frac{N_G \cdot p^{\alpha}}{P_{\alpha}(t)} + \frac{\alpha\beta(\alpha-1)\delta x^2}{2} \cdot \frac{N_G \cdot p^{\alpha-2}}{P_{\alpha}(t)} \cdot \left(\frac{\partial p}{\partial x}\right)^2 + \frac{\beta\delta x^2}{2} \cdot \frac{\partial^2 p}{\partial x^2} \cdot \left[1 + \alpha \frac{N_G p^{\alpha-1}}{P_{\alpha(t)}}\right] \tag{1}$$

Stationary behavior of 1D model



Stationary behavior of 1D model



Morphological indicators

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

$$\mathscr{E} = \sum_{i=1}^{M} \frac{P_i}{P} \cdot \ln \frac{P_i}{P} \quad (2)$$

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

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

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

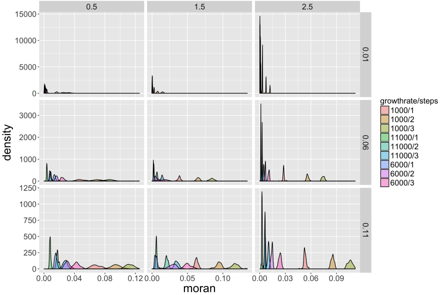
4. Mean distance between individuals

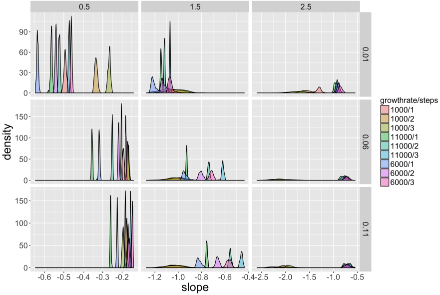
$$\bar{d} = \frac{1}{d_M} \cdot \sum_{i<j} \frac{P_i P_j}{P^2} \cdot d_{ij}$$

where d_M is a normalisation constant

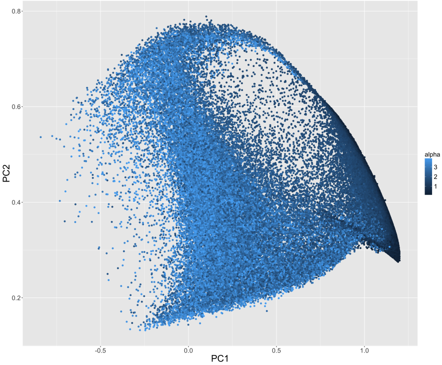
Model behavior : Convergence

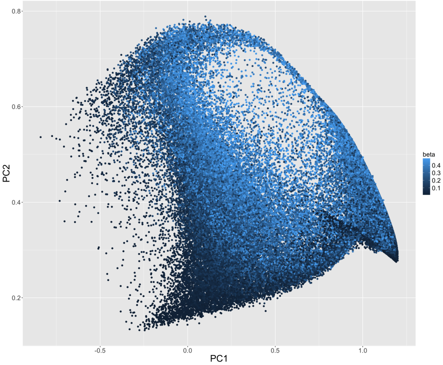
Large number of repetitions show good convergence properties





Model behavior





Empirical indicators computation

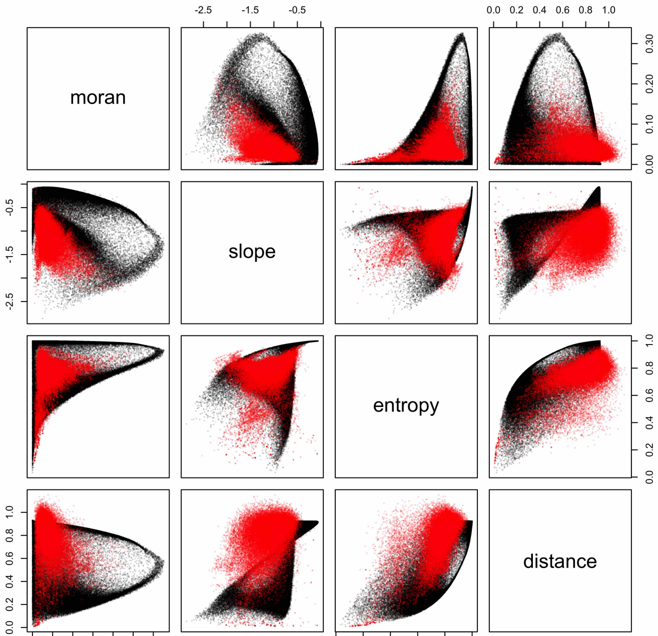
→ Eurostat population density raster (100m, simplified at 500m resolution)

→ Overlapping (10km offset) squares of 50km side : equivalent to smoothing, removes window shape effect. Not very sensitive to window size (tested with 30km and 100km)

→ Indicators computed using Fast Fourier Transform Convolution

→ Classification using repeated k-means ; number of clusters taken at transition in clustering coefficient.

Model calibration: all indicators



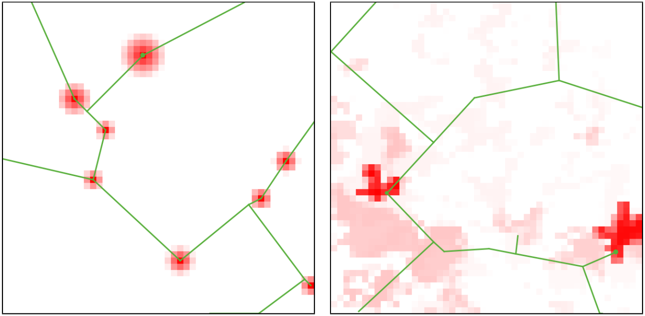
Model parameters

Heuristic	Param.	Name	Process	Domain	Default
Base	l_m	added links	growth	$[0;100]$	10
	d_G	gravity distance	potential	$]0;5000]$	500
	d_0	gravity shape	potential	$]0;10]$	2
	k_h	gravity weight	potential	$[0;1]$	0.5
	γ_G	gravity hierarchy	potential	$[0.1;4]$	1.5
Random	γ_R	random selection	hierarchy	$[0.1;4]$	1.5
	θ_R	random threshold	breakdown	$[1;5]$	2
Cost-benefits	λ	compromise	compromise	$[0;0.1]$	0.05
Biological	n_b	iterations	convergence	$[40;100]$	50
	θ_b	biological th.	threshold	$[0.1;1.0]$	0.5

Model setup

Synthetic setup: rank-sized monocentric cities, simple connection with bord nodes to avoid bord effects

Real setup: Population density raster at 500m resolution (European Union, from Eurostat)



Stopping conditions: fixed final time; fixed total population; fixed network size.

Network Indicators

Network Topology measured by:

- Betweenness and Closeness centralities: average and hierarchy
- Accessibility (weighted closeness)
- Efficiency (network pace relative to euclidian distance)
- Mean path length, diameter

Network nodes

Network baseline extension:
Adding a fixed number n_N of new nodes : for patches such that $d_r < d_0$, probability to receive a node is

$$p = P/P_{max} \cdot (\delta_M - \delta)/\delta_M$$

Nodes connected the shortest way to existing network.

Biological network morphogenesis model

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

Settings :

- Initial homogeneous network of tubes ij of length L_{ij}, variable diameter D_{ij}, carrying a flow Q_{ij}.
- Nodes i with a pressure p_i.
- N nodes are origin/destination points : randomly at each step one becomes source $p_{i_+} = I_0$ and one other sink $p_{i_-} = -I_0$

Biological network evolution

At each iteration :

1. Determination of flows with Kirchoff's law (electrostatic analogy) : Ohm's law $Q_{ij} = \frac{D_{ij}}{L_{ij}} \cdot (p_i - p_j)$ and conservation of flows $\sum_{j \to i} Q_{ij} = 0, \sum_{j \to i_\pm} Q_{i_\pm j} = \pm I_0$
2. Evolution of diameters (γ reinforcement parameter) by

$$\frac{dD_{ij}}{dt} = \frac{|Q_{ij}|^\gamma}{1 + |Q_{ij}|^\gamma} - D_{ij}$$

$\rightarrow$ Extraction of the final network after convergence given a threshold parameter for diameters

$\rightarrow$ Multi-scale model : diameters are constant during an iteration to obtain equilibrium flows

Biological network: indicators

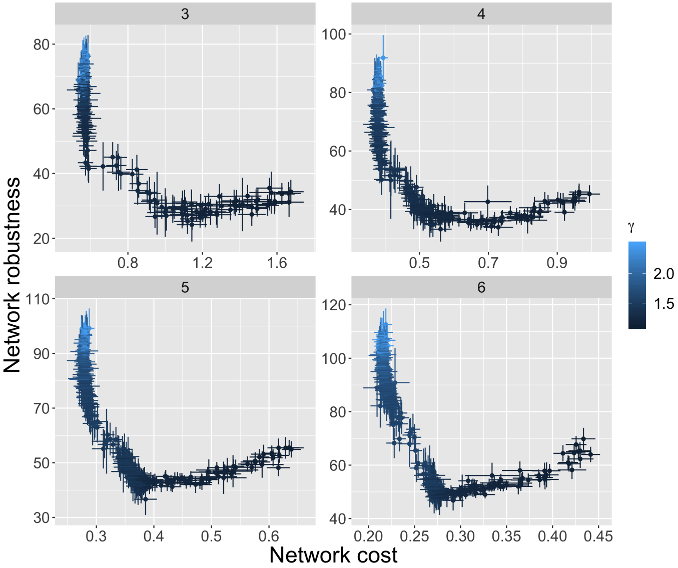
Behavior of the model evaluated with performance indicators for generated network (V_f, E_f), that are contradictory objectives :

- Construction costs $c = \sum_{ij \in E_f} D_{ij}(t_f)$
- Average performance [Banos and Genre-Grandpierre, 2012]

$$v = \frac{1}{|V_f|^2} \sum_{i,j \in V_f} \frac{d_{i \to j}}{||\vec{i} - \vec{j}||}$$

- Robustness (*Network Trip Robustness* index [?])

Biological network: optimal networks



Exploration of parameter space for synthetic network generation

Deterministic breakdown Network generation

1. Gravity potential given by

$$V_{ij}(d) = \left[(1-k_h) + k_h \cdot \left(\frac{P_i P_j}{P^2}\right)^{\gamma}\right] \cdot \exp\left(-\frac{d}{r_g(1+d/d_0)}\right)$$

2. $k \cdot N_L$ links are selected with lowest $V_{ij}(d_N)/V_{ij}(d_{ij})$, among which N_L links with highest (lest costly) are realized
3. Network is planarized

Results: components

With average betweenness centrality $\bar{bw}$ and average closeness centrality $\bar{cl}$, diameter r, average path length $\bar{l}$, relative speed v_0

Simulated point cloud:
$PC1 = -0.51\bar{bw} - 0.45\bar{l} + 0.57v_0 - 0.43r + 0.05\bar{cl}$ and
$PC2 = -0.45\bar{bw} + 0.17\bar{l} + 0.33v_0 + 0.8r + 0.1\bar{cl}$

Herfindhal index (20 width grid): first quartile at 0.54, a median at 0.76 and a third quartile at 1

Distance to real configurations:
$d(1,2) = \sqrt{(\bar{bw}_1 - \bar{bw}_2)^2 + (\bar{cl}_1 - \bar{cl}_2)^2 + (\bar{l}_1 - \bar{l}_2)^2}$, we use $d_{min} = \min_j d(S, R_j)$

Real point cloud: $PC1 = 0.12\bar{bw} - 0.09\bar{cl} + 0.98\bar{l}$ and
$PC2 = -0.20\bar{bw} - 0.97\bar{cl} - 0.06\bar{l}$

Population morphology classes

With 10 grids per class:

- Class 5: lowest Moran, high distance, hierarchy and entropy; numerous population centers that are localized and dispersed.
- Class 4: highest entropy and hierarchy; a small number of localized centers.
- Class 3: lowest distance and entropy; diffuse population.
- Class 2: highest Moran; one or a few centers with consequent size.
- Class 1: intermediate values for all indicators; a certain number of centers of intermediate size.

Population morphology classes

Class	Moran I	Distance $\bar{d}$	Entropy $\mathscr{E}$	Hierarchy γ
1	0.23	0.66	0.76	0.62
2	0.47	0.50	0.75	0.53
3	0.21	0.42	0.57	0.65
4	0.24	0.75	0.90	0.87
5	0.15	0.76	0.84	0.72