Calibration of a Density-based Model of Urban Morphogenesis

Raimbault Juste[1,2,*]

1 UMR CNRS 8504 Géographie-cités, Paris, France

2 UMR-T 9403 IFSTTAR LVMT, Champs-sur-Marne, France

*** Corresponding Author**

Email : juste.raimbault@polytechnique.edu

Abstract

We study a stochastic model of urban growth generating spatial distributions of population densities, at an intermediate scale between economic models at the macro scale and land-use evolution models focusing on local relations. Integrating simply the two opposite key processes of aggregation ("preferential attachment") and diffusion (urban sprawl), we show that we can capture a significant part of existing urban forms in Europe. An extensive exploration and calibration of the proposed model allows determining the region of parameter space corresponding morphologically to observed european urban systems, providing an validated thematic interpretation to model parameters, and furthermore determining the effective dimension of the urban system at this scale regarding morphological objectives.

Introduction

The study of urban growth, and more particularly its quantification, is more than ever a crucial issue in a context where most of the world population live in cities which expansion has significant environmental impacts [1] and that have therefore to ensure an increased sustainability and resilience to climate change. The understanding of drivers for urban growth can lead to better integrated policies. It is however a question far from being solved in the diverse related disciplines: Urban Systems are complex socio-technical systems that can be studied from a large variety of viewpoints. Batty has advocated in that sense for the construction of a dedicated science defined by its objects of study more than the methods used [2], what would allow easier coupling of approaches and therefore Urban Growth models taking into account heterogeneous processes. The processes that a model can grasp are also linked to the choice of the scale of study. At a macroscopic scale, models of growth in system of cities are mainly the concern of economics and geography. [3] shows that in first approximation, the Gibrat's model postulating random growth rates not depending on city size, yield the well-know Zipf's law, or rank-size law, which is a typical stylized fact witnessing hierarchy in systems of cities. It was however shown empirically that systematic deviations to this law exist [4], and that spatial interactions may be responsible for it. Models integrating spatial interactions include for example [5] that introduces a growth model in which these interactions, that are function of distance and the geography, play a significant role in growth rates. More recently, [6] has extended this model by taking into account innovation waves between cities as a driver. The interplay of space, economic and population growth is studied by the Marius model [7] in the case of the former Soviet Union, on which model performance is shown improved compared to models without interactions.

At smaller scales, that can be understood as microscopic or mesoscopic depending on the resolution and extent of models, agents of models fundamentally differ. Space is generally taken into account in a finer way, through neighborhood effects for example. For example, [8] propose a micro-based model of urban growth, with the purpose to replace non-interpretable physical mechanisms with agent mechanisms, including interactions forces and mobility choices. Local correlations are used in [9], which develops the model introduced in [10], to modulate growth patterns to ressemble real configurations. The world of Cellular Automata (CA) models of Urban Growth [11] also offers numerous examples. [12] introduced a generic framework for CA with multiple land use, based on local evolution rules. A model with simpler states (occupied or not) but taking into account global constraints is studied by [13]. The Sleuth model, initially introduced by [14] for the San Francisco Bay area, and for which an overview of diverse applications is given in [15], was calibrated on areas all over the world, yielding comparative measures through the calibrated parameters.

Closely related to CA models but not exactly similar are Urban Morphogenesis models, which aim to simulate the growth of urban form from autonomous rules. [16] suggested that the fractal nature of cities is closely to the emergence of the form from the microscopic socio-economic interactions, namely urban morphogenesis. [17] develops a morphogenesis model for urban roads alone, with growth rules based on geometrical considerations. These are shown sufficient to produce a broad range of patterns analog to existing ones. Similarly, [18] couples a CA with an evolving network to reproduce stylized urban form, from concentrated monocentric cities to sprawled suburbs. The Diffusion-Limited-Aggregation model, coming from physics, and which was first studied for cities by [19], can also be seen as a morphogenesis model. These kind of models, that sometimes can be classified as CA, have generally the particularity of being parsimonious in their structure.

We study in this paper a morphogenesis model, at the mesoscopic scale, aimed at being simplistic in its rules and variables, but trying to be accurate in the reproduction of existing patterns. The underlying question is to explore the performance of simple mechanisms in reproducing complex urban patterns. We consider abstract processes, namely aggregation and diffusion, candidates as partially explanatory drivers of urban growth, based on population only, that will be detailed in model rationale below. An important aspect we introduce is the quantitative measure of urban form, based on a combination of morphological indicators, to quantify and compare model outputs and real urban patterns. Our contribution is significant on several points: (i) we compute local morphological characteristics on a large spatial extent (full European Union); (ii) we give significant insights into model behavior through extensive exploration of the parameter space; (iii) we show through calibration that the model is able to reproduce most of existing urban forms across Europe, and that these abstract processes are sufficient to explain urban form alone.

The rest of this paper is organized as follows: we first describe formally the model and the morphological indicators. We then detail values of morphological measures on real data, study the behavior of the model by exploring its parameter space and through a semi-analytical approach to a simplified case, and we describe results of model calibration.

Material and Methods

Urban growth model

Rationale Our model is based on widely accepted ideas of diffusion-aggregation processes for Urban Processes. The combination of attraction forces with repulsion, due for example to congestion, already yield a complex outcome that has been shown under some simplifying assumptions to be representative of urban growth processes. A model capturing these processes was introduced by [21], as a cell-based variation of the DLA model [19]. Indeed, the tension between antagonist aggregation and sprawl mechanisms may be an important process in urban morphogenesis. For example [22] opposes centrifugal forces with centripetal forces in the equilibrium view of urban spatial systems, what is easily transferable to non-equilibrium systems in the framework of self-organized complexity: a urban structure is a far-from-equilibrium system that has been driven to this point by these opposite forces. The two contradictory processes of urban concentration and urban sprawl are captured by the model, what allows to reproduce with a good precision a large number of existing morphologies. We can expect aggregation mechanisms such as preferential attachment to be good candidates in urban growth explanation, as it was shown that the Simon model based on them generates power-laws typical of urban systems (scaling laws for example) [20]. The question at which scale is it possible and relevant to define and try to simulate urban form is rather open, and will in fact depend on which issues are being tackled. Working in a typical setting of morphogenesis, the processes considered are local and our model must have a resolution at the micro-level. We however want to quantify urban form on consistent urban entities, and will work therefore on spatial extents of order 50 100km. We sum up these two aspects by stating that the model is at the *mesoscopic scale*.

Formalization We formalize now the model and its parameters. The world is a square grid of width N, in which each cell is characterized by its population $(P_i(t))_{1\leq i\leq N^2}$. We consider the grid initially empty, i.e. $P_i(0) = 0$, but the model can be easily generalized to any initial population distribution.

The simulation model proceeds iteratively the following way. At each time step, until total population reaches a fixed parameter P_m,

- total population is increased of a fixed number N_G (growth rate), following a preferential attachment such that

$$\mathbb{P}[P_i(t+1) = P_i(t)+1 | P(t+1) = P(t)+1] = \frac{(P_i(t)/P(t))^\alpha}{\sum (P_i(t)/P(t))^\alpha} \tag{1}$$

- a fraction β of population is diffused to cell neighborhood, this operation being repeated n_d times

Table 1. Summary of model parameters.

Parameter	Notation	Process	Range
Total population	P_m	Macro-scale growth	$[1e4, 1e6]$
Growth rate	N_G	Meso-scale growth	$[500, 30000]$
Aggregation strength	α	Aggregation	$[0.1, 4]$
Diffusion strength	β	Diffusion	$[0, 0.5]$
Diffusion steps	n_d	Diffusion	$\{1, \ldots, 5\}$

Indicators [23] : sort of morphological analysis

As our model is only density-based, we propose to quantify its outputs through spatial morphology, i.e. characteristics of density spatial distribution. We need therefore quantities having a certain level of robustness and invariance. For example, two polycentric cities should be classified as morphologically close whereas a direct comparison of distributions (Earth Mover Distance e.g.) could give a very high distance between configurations depending on center positions. To tackle this issue, we refer to the Urban Morphology Analysis literature which proposes an extensive set of indicators to describe urban form [24]. The number of dimensions can be reduced to obtain a robust description with a few number of independent indicators [25]. For the choice of indicators, we follow the analysis done in [26] where a typology of large european cities is obtained in consistence with qualitative knowledge. Let denote $(P_i)_{1 \leq i \leq N}$ the population of cells, sorted in decreasing order, d_{ij} the distance between cells i, j, and $P = \sum_{i=1}^{N} P_i$ total population. The indicators are the following :

1. Rank-size slope γ, expressing the degree of hierarchy in the distribution, computed by fitting with Ordinary Least Squares a power law distribution by $\ln P_i/P_0 \sim k - \gamma \cdot \ln i/i_0$.

2. Distribution Entropy

$$\mathcal{E} = \sum_{i=1}^{N} \frac{P_i}{P} \cdot \ln \frac{P_i}{P} \tag{2}$$

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

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

4. Mean distance between individuals, which captures population concentration

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

Real Data

We compute morphological measure on real urban density data, namely the population density grid of th European Union at 100m resolution provided openly by Eurostat [27]. The morphological measures used for calibration are the one described above that are the same used to classify model outputs. The calibration of the model is thus done on morphological objectives (entropy, hierarchy, spatial auto-correlation, mean distance). We show in Fig. 1 maps giving values of indicators for France, to ease readability. Maps for the full European union are available in S1 Text. The choice of the resolution, the space range, and the shape of the window on which indicators are computed, is made according to the thematic specifications of the model : We however tested the sensitivity to window size and shape

Results

Generation of urban patterns

Implementation The model is implemented both in NetLogo [28] for exploration and visualization purposes, and in `Scala` for performance reasons and easy integration into OpenMole [29], which allows a transparent access to High Performance Computing environments. Computation of indicator values on geographical data is done in `R` using the `raster` package [30]. Source code and results are available on the open repository of the project at `https://github.com/JusteRaimbault/Density`. Raw

figures/Fig1.png

Figure 1. Empirical values of morphological indicators.

figures/Fig2.png

Figure 2. Example of the variety of generated urban shapes.

datasets for real indicator values and simulation results are available on Dataverse at http://dx.doi.org/10.7910/DVN/WSUSBA.

Generated Shapes The model has a relatively small number of parameters but is able to generate a large variety of shapes, extending beyond existing forms. In particular, its dynamical nature allows through P_m parameter to choose between final configurations that can be non-stationarity or stationarity. Fig. 2 shows examples of generated shapes for different parameter values and different regimes.

Model Behavior

In the study of such a computational model of simulation, the lack of analytical tractability must be compensated by an extensive knowledge of model behavior in the parameter space [31]. This type of approach is typical of what Arthur calls the *Computational shift in modern science* [32].

Convergence First of all we need to assess the convergence of the model and its behavior regarding stochasticity. We run for a sparse grid of the parameter space consisting of 81 points, with 100 repetitions for each point. Corresponding histograms are shown in S1 Text. Indicators show good convergence properties: most of indicators are easily statistically discernable across parameter points, and these are distinguished without ambiguity when taking into account all indicators. We use this experiment to find a reasonable number of repetitions needed in larger experiments. For each point, we estimate the Sharpe ratios for each indicators as

Exploration of parameter space We sample the Parameter space using a Latin Hypercube Sampling, with parameter as $\alpha \in [0.1, 4], \beta \in [0, 0.5], n_d \in \{1, \ldots, 5\}, N_G \in [500, 30000], P_m \in [1e4, 1e6]$. This type of cribbing is a good compromise to have a reasonable sampling without being subject to the dimensionality curse within normal computation capabilities.

Semi-analytical Analysis

Our model can be understood as a type of reaction-diffusion model, that have been widely used in other fields such as biology: similar processes were used for example by Turing in its seminal paper on morphogenesis [33]. An other way to formulate the model typical to these approaches is by using Partial Differential Equations. We propose to gain insights into long-time dynamics by studying them on a simplified case. We consider the system in one dimension, such that $x \in [0; 1]$ with $1/\delta x$ cells of size δx. Each cell is characterized by its population as a random variable $P(x, t)$. We work on their expected values $p(x, t) = \mathbb{E}[P(x, t)]$, and assume that $n_d = 1$. As developed in Supplementary Material S2 Text, we show that this simplified process verifies the following PDE:

figures/Fig3.png

Figure 3. Behavior of indicators.

$$\delta t \cdot \frac{\partial p}{\partial t} = \frac{N_G \cdot p^{\alpha}}{P_{\alpha}(t)} + \frac{\alpha\beta(\alpha-1)\delta x^2}{2} \cdot \frac{N_G \cdot p^{\alpha-2}}{P_{\alpha}(t)} \cdot \left(\frac{\partial p}{\partial x}\right)^2 + \frac{\beta \delta x^2}{2} \cdot \frac{\partial^2 p}{\partial x^2} \cdot \left[1 + \alpha \frac{N_G p^{\alpha-1}}{P_{\alpha(t)}}\right] \quad (3)$$

[34]

This section allow us to learn the following properties, that are important to better understand the previous simulation results

1. Existence of a stationary solution for population proportion
2. Characteristic distance of the stationary distribution
3. Existence of bifurcations in the random case

Model Calibration

We use a specific calibration process: a principal component analysis allows to maximize the cumulated distance between generated points and reals points. We select then the point cloud that overlaps real points in the (PC1,PC2) plan, given a distance threshold. Fig. 5 shows the points we obtain for four different values of the threshold ranging from 10^{-6} to 10^{-3}.

figures/Fig4.png

Figure 4. Randomness and frozen accidents.

Discussion

more refined model : thresholded aggregation and diffusion (distance to center for diffusion, maximum density for aggregation)

Calibration refinement and Targeted Exploration We plan in further work to extract the exact parameter space covering all real situations and provide interpretation of its shape (correlations between parameters). Its volume in different directions should give the relative importance of parameters.

We also use the parameter space exploration algorithm [35] implemented in OpenMole, and obtain in Fig. **??** the lower bound in Moran-entropy plan, that unexpectedly exhibit a scaling relationship that we aim to explore further.

Integration into a multi-scale growth model

It could be possible to couple this model with a Gibrat (or Favaro-pumain) at Europa scale (macro) (with addition of consistence on migration constraints), where meso growth rates which were exogenous before are top-down determined, and bottom-up feedback is done through local aggregation level, influence importance of each area.

figures/Fig5.png

Figure 5. Model calibration. The principal component analysis is conducted to maximize the spread of the differences between real data and model output, i.e. on the set $\{|R_i - M_j|\}$ where R_i is the set of real points, M_j the set of model outputs. We select then the overlapping cloud at threshold θ, by taking models output closer to real point cloud than θ in the (PC1,PC2) plan.

figures/Fig6.png

Figure 6. PSE exploration. Scatterplotsof Moran against Entropy, with blue points obtained with LHS and red with PSE exploration. Lower bound is in green.

[36]

Conclusion

In conclusion, this first modeling step provide an accurately calibrated spatial urban growth model at the mesoscopic scale that can reproduce any European urban pattern in terms of urban form. Further work is needed for an interpretation of parameter influence and the determination of effective independent dimensions of the urban system at this scale.

Supporting Information

S1 Text

Extended Model Exploration. Extended figures for model exploration.

S2 Text

Semi-analytical Analysis. Analytical and numerical developments for the simplified model.

References

1. Seto KC, Güneralp B, Hutyra LR. Global forecasts of urban expansion to 2030 and direct impacts on biodiversity and carbon pools. Proceedings of the National Academy of Sciences. 2012;109(40):16083–16088.
2. Batty M. The new science of cities. Mit Press; 2013.
3. Gabaix X. Zipf's law for cities: an explanation. Quarterly journal of Economics. 1999;p. 739–767.
4. Rozenfeld HD, Rybski D, Andrade JS, Batty M, Stanley HE, Makse HA. Laws of population growth. Proceedings of the National Academy of Sciences. 2008;105(48):18702–18707.
5. Bretagnolle A, Mathian H, Pumain D, Rozenblat C. Long-term dynamics of European towns and cities: towards a spatial model of urban growth. Cybergeo: European Journal of Geography. 2000;.
6. Favaro JM, Pumain D. Gibrat Revisited: An Urban Growth Model Incorporating Spatial Interaction and Innovation Cycles. Geographical Analysis. 2011;43(3):261–286.
7. Cottineau C. L'évolution des villes dans l'espace post-soviétique. Observation et modélisations. Université Paris 1 Panthéon-Sorbonne; 2014.
8. Andersson C, Lindgren K, Rasmussen S, White R. Urban growth simulation from "first principles". Physical Review E. 2002;66(2):026204.
9. Makse HA, Andrade JS, Batty M, Havlin S, Stanley HE, et al. Modeling urban growth patterns with correlated percolation. Physical Review E. 1998;58(6):7054.
10. Makse HA, Havlin S, Stanley H. Modelling urban growth. Nature. 1995;377(1912):779–782.

11. Batty M, Xie Y. From cells to cities. Environment and planning B: Planning and design. 1994;21(7):S31–S48.

12. Xie Y. A Generalized Model for Cellular Urban Dynamics. Geographical Analysis. 1996;28(4):350–373. Available from: http://dx.doi.org/10.1111/j.1538-4632.1996.tb00940.x.

13. Ward DP, Murray AT, Phinn SR. A stochastically constrained cellular model of urban growth. Computers, Environment and Urban Systems. 2000;24(6):539–558.

14. Clarke KC, Gaydos LJ. Loose-coupling a cellular automaton model and GIS: long-term urban growth prediction for San Francisco and Washington/Baltimore. International journal of geographical information science. 1998;12(7):699–714.

15. Clarke KC, Gazulis N, Dietzel C, Goldstein NC. A decade of SLEUTHing: Lessons learned from applications of a cellular automaton land use change model. Classics in IJGIS: twenty years of the international journal of geographical information science and systems. 2007;p. 413–427.

16. Frankhauser P. Fractal geometry of urban patterns and their morphogenesis. Discrete Dynamics in Nature and Society. 1998;2(2):127–145.

17. Courtat T, Gloaguen C, Douady S. Mathematics and morphogenesis of cities: A geometrical approach. Physical Review E. 2011;83(3):036106.

18. Raimbault J, Banos A, Doursat R. A hybrid network/grid model of urban morphogenesis and optimization. In: Proceedings of the 4th International Conference on Complex Systems and Applications (ICCSA 2014), June 23-26, 2014, Université de Normandie, Le Havre, France; M. A. Aziz-Alaoui, C. Bertelle, X. Z. Liu, D. Olivier, eds.: pp. 51-60.; 2014. .

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

20. Sheridan Dodds P, Rushing Dewhurst D, Hazlehurst FF, Van Oort CM, Mitchell L, Reagan AJ, et al. Simon's fundamental rich-gets-richer model entails a dominant first-mover advantage. ArXiv e-prints. 2016 Aug;.

21. Batty M. Hierarchy in cities and city systems. In: Hierarchy in natural and social sciences. Springer; 2006. p. 143–168.

22. Fujita M, Thisse JF. Economics of agglomeration. Journal of the Japanese and international economies. 1996;10(4):339–378.

23. Guérois M, Pumain D. Built-up encroachment and the urban field: a comparison of forty European cities. Environment and Planning A. 2008;40(9):2186–2203.

24. Tsai YH. Quantifying urban form: compactness versus' sprawl'. Urban studies. 2005;42(1):141–161.

25. Schwarz N. Urban form revisited—Selecting indicators for characterising European cities. Landscape and Urban Planning. 2010;96(1):29 – 47. Available from: http://www.sciencedirect.com/science/article/pii/S0169204610000320.

26. Le Néchet F. De la forme urbaine à la structure métropolitaine: une typologie de la configuration interne des densités pour les principales métropoles européennes de l'Audit Urbain. Cybergeo: European Journal of Geography. 2015;.

27. EUROSTAT. Eurostat Geographical Data; 2014. Available from: http://ec.europa.eu/eurostat/web/gisco/geodata/reference-data/administrative-units-statistical-units.

28. Wilensky U. NetLogo. 1999;.

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

30. Hijmans RJ. Geographic data analysis and modeling. 2015;.

31. Banos A. Pour des pratiques de modélisation et de simulation libérées en Géographies et SHS. HDR Université Paris. 2013;1.

32. Arthur WB. Complexity and the Shift in Modern Science; 2015. Conference on Complex Systems, Tempe, Arizona.

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

34. Tadmor E. A review of numerical methods for nonlinear partial differential equations. Bulletin of the American Mathematical Society. 2012;49(4):507–554.

35. Chérel G, Cottineau C, Reuillon R. Beyond Corroboration: Strengthening Model Validation by Looking for Unexpected Patterns. PLoS ONE. 2015 09;10(9):e0138212. Available from: http://dx.doi.org/10.1371%2Fjournal.pone.0138212.

36. Zhang Z, Su S, Xiao R, Jiang D, Wu J. Identifying determinants of urban growth from a multi-scale perspective: A case study of the urban agglomeration around Hangzhou Bay, China. Applied Geography. 2013;45:193–202.