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

*** juste.raimbault@parisgeo.cnrs.fr**

Abstract

We propose a stochastic model of urban growth that generates 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 the whole spectrum 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.

1 Introduction

Urban Systems are complex socio-technical objects

[1] 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 [2] to modulate growth

patterns to ressemble real configurations. In the same spirit, our model situates at similar scales and can be qualified as a morphogenesis model.

2 Materials and Methods

2.1 The urban growth model

Rationale Our model is an extension of the diffusion-limited aggregation model studied in [3]. Indeed, the tension between antagonist aggregation and sprawl mechanisms may be an important process in urban morphogenesis. For example [4] 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.

Scale

Settings The model D proceeds iteratively the following way. An square grid of width N, initially empty, is represented by population $(P_i(t))_{1\leq i\leq N^2}$. 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

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

- a fraction β of population is diffused to four closest neighbors is operated n_d times

2.2 Indicators

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 [5]. The number of dimensions can be reduced to obtain a robust description with a few number of independent indicators [6]. For the choice of indicators, we follow the analysis done in [7] 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 a power law distribution by $\ln P_i/P_0 \sim k - \gamma \cdot \ln i/i_0$.
2. Distribution Entropy
$$\mathcal{E} = \sum_{i=1}^{N} \frac{P_i}{P} \cdot \ln \frac{P_i}{P}$$
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
$$\bar{d} = \sum_{i<j} \frac{P_i P_j}{P^2} \cdot d_{ij}$$

3 Results

3.1 Real Data

We use the population density grid provided openly by

Empirical morphological measures for calibration are the one described in the empirical chapter, i.e. the calibration is done on morphological objectives (entropy, hierarchy, spatial auto-correlation, mean distance) against real values computed on the set of 50km sized grid extracted from european density grid [12].

3.2 Generating Synthetic Patterns

The model was implemented in a first time in NetLogo [8] for exploration and visualization purposes, later in `Scala` for performance reasons and easy integration into OpenMole [9] for HPC model exploration. Computation of indicator values on geographical data was done in `R` using the `raster` package [10].

Generation of urban patterns.

The model as few parameters but is able to generate a very wide variety of shapes, extending beyond existing forms. In particular, its dynamical nature allows through P_m parameter to choose final regime that can be non-stationarity (generally chaotic shapes), semi-stationarity or total stationarity. Fig. **??** shows examples of generated shapes.

3.3 Model Behavior

In the study of such a computational model of simulation, the lack of analytical tractability must be balanced by an extensive knowledge

Convergence Indicators show good convergence property and bimodal statistical distribution for cumulated points in the parameter space confirm the existence of superposed regimes : gaussian distribution gives stationary configurations, whereas inverse log-normal distribution are close to real data shape and correspond to non-stationary regime. For one point and a large number of repetitions, we find that 50 repetitions are enough to obtain a 95% confidence interval smaller than σ around indicator mean.

Exploration of parameter space Parameter space is explored using a grid in first experiments, than a Latin Hypercube Sampling exploration. Parameter bounds are $\alpha \in [0.2, 2], \beta \in [0, 0.1], n_d \in \{0, \ldots, 4\}, N_G \in [500, 3000], P_m \in [2000, 100000]$. Fig.**??** shows the result. We also use the parameter space exploration algorithm [11] 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.

Figure 1. Precise calibration of the model. 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.

Statistical analysis. A statistical analysis (basic models) of indicator behaviors remains to be done and interpreted (one is done conjointly with network in paper corresponding to next section).

3.4 Model Calibration

Calibration Process 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. 1 shows the points we obtain for four different values of the threshold ranging from 10^{-6} to 10^{-3}.

Calibration refinement 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.

4 Discussion

4.1 Thematic interpretation of growth behavior

We still need to interpret the positions of typical shapes within parameter space in order to confirm the thematic interpretation of parameters. Depending on results of calibration refinement, we may obtain necessary and sufficient parameters to explain growth at this scale and a corresponding interpretation.

4.2 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.

5 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. We will use this model for other purposes in the following.

Supporting Information

S1 Figure

Bold the first sentence.

Acknowledgments

References

1. Andersson C, Lindgren K, Rasmussen S, White R. Urban growth simulation from "first principles". Physical Review E. 2002;66(2):026204.

2. 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.

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

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

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

6. 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.

7. 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;.

8. Wilensky U. NetLogo. 1999;.

9. 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.

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

11. 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.

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