

# Complementarity of generative models for road networks

J. Raimbault<sup>1,2,3,\*</sup>

<sup>1</sup> Center for Advanced Spatial Analysis, University College London

<sup>2</sup> UPS CNRS 3611 ISC-PIF

<sup>3</sup> UMR CNRS 8504 Géographie-cités

\* `juste.raimbault@polytechnique.edu`

## Abstract

Understanding the dynamics of road networks has theoretical implications for urban science and practical applications for sustainable long-term planning. Various generative models to explain road network growth have been introduced in the literature. We propose in this paper a systematic benchmark of such models integrating different paradigms (spatial interactions, cost-benefit compromises, biological network growth), focusing on the feasible space of generated network measures. We find a quantitatively high complementarity between the different models. This confirms the necessity of a plurality of urban models, in line with integrative approaches to urban systems.

## 1 Introduction

Territorial systems are complex in the sense that they span across multiple dimensions, and are produced by various agents at different spatial and temporal scales [Batty, 2007]. They are furthermore adaptive and self-organised, as witnesses the resilience of co-evolving city systems on long time scales [Pumain, 2021]. Among the built environment, environmental, and social components, one crucial driver of spatial structure and dynamics of these territorial systems are transportation infrastructure networks, and more particularly road networks. They irrigate territories by distributing accessibility, and their topology and hierarchy often shapes spatial interactions and thus urban and land-use dynamics [Wegener and Fürst, 2004]. Strong effects of path-dependance furthermore highlights the role of these networks [Ribeill, 1985].

Although accessibility is nowadays strongly multimodal [Cats and Birch, 2021], road networks have been the dominant transportation infrastructure in urban systems over long historical periods [Verdier and Bretagnolle, 2007], and they still are an important aspect of both urban form at the microscopic scale and accessibility at the mesoscopic and macroscopic scales. Understanding the link be-

tween underlying processes driving the growth of road networks and their shape is an important subject, towards potential application to sustainable planning.

Generative processes for road networks are multiple: for example a combination of self-organisation and top-down planning may leave a significant signature in topological features of these networks [Barthelemy et al., 2013]. In order to explain such evolution and in some cases reproduce existing networks, multiple simulation models have been introduced in the literature by different disciplines. These span across a broad spectrum, from data-driven models to parsimonious stylised models. [Xie and Levinson, 2009] propose a broad survey of such models, including economics, transport geography, transport planning and network science. The diversity of processes taken into account can range from economics of investment [Xie and Levinson, 2011], to negotiations between planning stakeholders [Raimbault and Le Néchet, 2021], full top-down planning [Szell et al., 2021], self-organisation through morphogenesis [Tirico et al., 2018], or local optimisation [Barthélemy and Flammini, 2009], among others.

Most of these models include empirically documented processes, and have been validated against stylised facts such as hierarchical organisation of the network or the distribution of network measures, and also often compared against existing networks with reasonable success. However, to what extent these processes and models are complementary remains an open question. We propose in this paper a systematic benchmark of several road network growth models, to clarify the necessity of considering such diverse models. We extend the multi-modeling approach of [Raimbault, 2018b] by focusing on the feasible space of network measures as model outputs. We consider in particular the following mechanisms for link addition, which were chosen to illustrate a diversity of underlying assumptions and frameworks: (i) gravity potential deterministic breakdown [Raimbault, 2019a]; (ii) random potential breakdown [Raimbault, 2020]; (iii) cost-benefits compromise [Louf et al., 2013]; (iv) biological network generation [Raimbault, 2018a].

The rest of this paper is organised as follows: we first describe in details each road network growth model included in the benchmark; we then describe the implementation and model setup, followed by numerical experiments aimed at comparing feasible spaces for each model; we finally discuss implications and possible extensions.

## 2 Road network growth models

### 2.1 Baseline dynamics

All models are included within the same baseline dynamics, for a better comparability. We use the co-evolution model between population density and road network described by [Raimbault, 2019b], but evolve the network only. At a mesoscopic spatial scale, given a population density distributed as a grid (typically with size 50x50), road network nodes and links are added iteratively. More precisely, starting from an empty network, at each time step:

1. new network nodes are added, with a preferential attachment to population and to existing roads to determine the cell, and a random position within the cell; a fixed number  $N_n = 20$  of nodes is considered;
2. added nodes are connected to the existing network with a connectivity algorithm: the closest network component is determined, and a shortest distance connection is created (either to the closest node, or with a perpendicular projection to the closest link);
3. a fixed number of links  $N_l$  (model parameter) is added, following one of the models included in the benchmark;
4. the resulting network is made planar, by adding nodes at intersections of crossing links.

The model can be run for a fixed number of time steps  $t_f$ , or until a total number of links  $N_f$  is reached.

## 2.2 Link addition models

**Null model** The null model corresponds to no additional link besides the ones added to keep the network connected. This yields tree networks with no loops.

**Random links** A second null model consists in the addition of random links. A random node is chosen, and connects to an other node to which it is not already connected. This procedure is done  $N_l$  times.

**Random gravity potential breakdown** The model explored in [Raimbault, 2020] adds random links according to spatial interactions between nodes. Given a population for each node  $P_i$  (determined here by the sum of the population within the catchment area of each node), a classical gravity model gives interaction forces between nodes. The origin city of the new link is chosen hierarchically to population, and the destination city following spatial interactions. The link is realised if the current network imposes a high detour (i.e. if the rate between euclidian distance and network distance is below a given threshold). This procedure is repeated until  $N_l$  links have been added.

**Deterministic gravity potential breakdown** The model formulated by [Raimbault, 2019a] also add links by considering spatial interactions between nodes. Still accounting for node populations, a spatial interaction model yields interaction forces between nodes (see [Raimbault, 2019a] for detailed equations). Potential links are taken as a fixed proportion among all couples, with the highest potential. A cost criteria at the last stage selects the links with the lowest detour among the potential links. Note that an other heuristic assuming strong top-down planning and investment could at this stage select links with the highest detour.

**Cost-benefit aggregation** The cost and benefits included in the previous heuristic are also considered through aggregation by [Louf et al., 2013]. For each node couple, a cost function aggregating interaction potential with construction cost (as the constructed link distance) normalised by an aggregation parameter  $\lambda$ , is computed. The  $N_l$  links with the lowest cost are constructed.

**Biological network generation** [Raimbault, 2018a] uses a slime-mould model to generate synthetic networks. We use the same model with the following process: (i) generate a background grid of local links between neighbour cells, on top of the existing network; (ii) run the slime-mould self-reinforcement model in this provisory network with origins/destinations as node couples with population, for a given number of steps  $t_b$ ; (iii) remove all provisory link with a capacity below a fixed threshold  $\theta_b$ ; (iv) simplify the resulting network (by removing nodes with degree exactly two) and keep the new links to add them to the evolving network.

## 2.3 Indicators

The generated road networks are quantified using the following measures: diameter, mean path length, mean betweenness centrality, mean closeness centrality, mean detour, cumulated length. More indicators could be taken into account but [Raimbault, 2019b] showed that these already provided a reasonable coverage.

# 3 Results

## 3.1 Data

We setup the model with real population grids. Therefore, the Eurostat population grid was used [Batista e Silva et al., 2013]. 50 configurations were randomly selected, equally distributed into morphological clusters provided by [Raimbault, 2018b] (measures of the spatial distribution of population). The initial population distribution is considered as a model parameter.

## 3.2 Implementation

The models are implemented into a single one in the NetLogo software, and integrated into the OpenMOLE software for model exploration and validation [Reuillon et al., 2013]. The models are open source and available at <https://github.com/JusteRaimbault/NetworkGrowth>

## 3.3 Exploration

We explore the parameter space of each model by using a Latin Hypercube Sampling, and sampling 40,000 parameter points for each model. The index of population configuration is included as a parameter in each case. As each

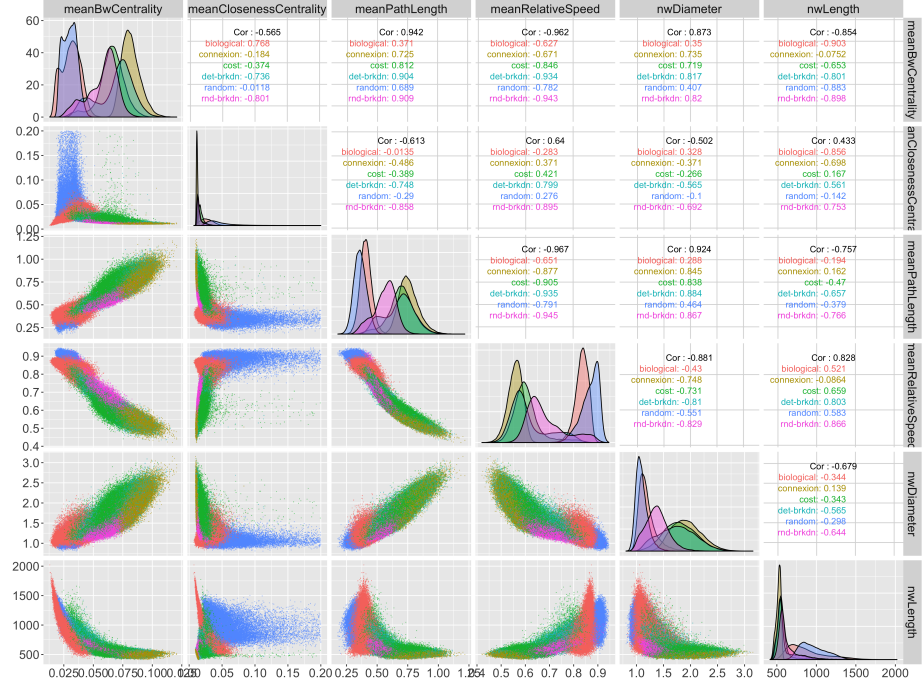


Figure 1: Scatterplot of feasible space for each model, between all indicators (rows and columns). The diagonal histograms give the statistical distribution of each indicators. Color give the model. Estimated correlations between indicators, stratified by model, are also given.

model has a low number of parameters (below 5, for the deterministic gravity breakdown), we assume that the obtained point cloud cover most of the feasible space. More robust results should be obtained with diversity search algorithms as detailed in the discussion below.

We show in Fig. 1 the scatterplot of indicators across models. We obtain overall important parts of the space which are occupied by a single model, suggesting a complementarity. Surprisingly, the closest model to the random null model seem to be the biological network, while the connection heuristic is close to the cost-benefit (which could be expected as only local links are added in the latest case). The random breakdown seems to be the model the most distinguished from others. We also find different performance patterns, as in terms of relative speed, connection model has the lowest while the random has the highest (this indicator is not weighted by population).

To further quantify the complementarity between models, we compute the overlap between point clouds. Therefore, we estimate hypervolumes and compute their relative intersection as  $I(O1 \rightarrow O2) = Vol(O1 \cap O2)/Vol(O2)$ . This non-symmetric measure gives the proportion of point cloud  $O2$  overlapping with

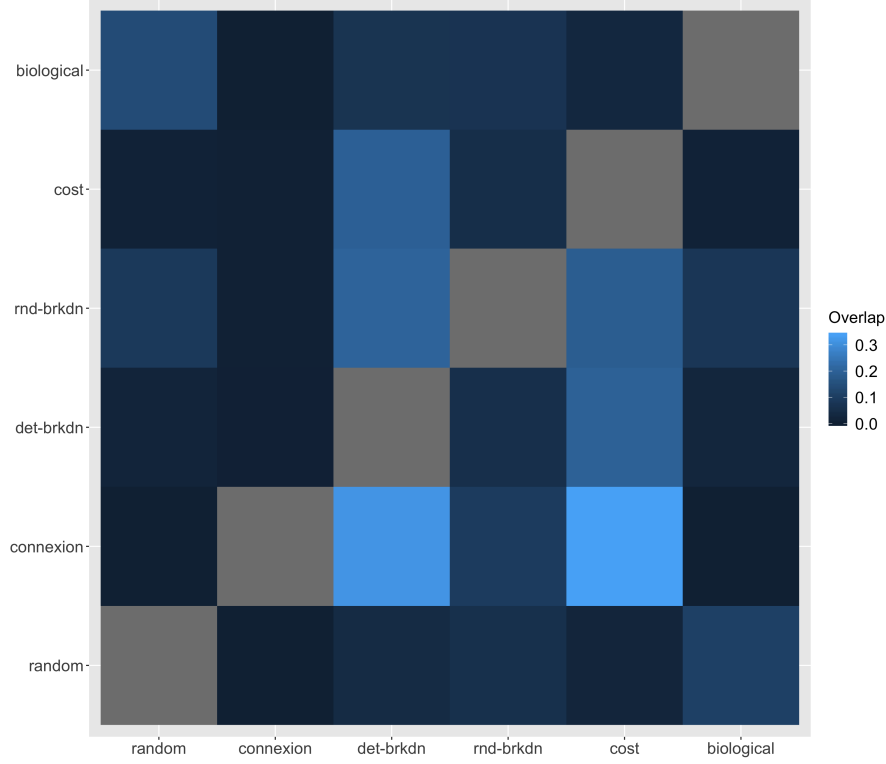


Figure 2: Estimated overlaps between hypervolumes of model outputs.

O1. The corresponding matrix is shown in Fig. 2. We find a relatively low level of overlapping, as the highest is just above 0.3, for the connection cloud in the deterministic breakdown and the cost. More than half of couples are below 0.1, confirming that point clouds are mostly distinct in the multidimensional indicator space.

Overall, this numerical exploration confirms the complementarity of the diverse heuristics included in our benchmark.

## 4 Discussion

Our result contribute to a general effort of urban models comparison and benchmarking, which is an essential step towards the construction of integrative models and theories. We have shown quantitatively a strong complementarity between diverse processes driving the growth of road networks. This confirms the necessity to include a plurality of models, views and disciplines in the understanding of road network dynamics.

Current work is focused on a more systematic determination of feasible spaces. To that end, we use the Pattern Space Exploration (PSE) algorithm introduced by [Chérel et al., 2015] and implemented into the OpenMOLE platform. This algorithm, which is a particular case of diversity search algorithms, has been shown to be efficient to discover a large number of patterns and effectively cover output space in low dimensions. This would allow obtaining more robust measures of the complementarity of the different models included.

An other important ongoing extension is the calibration of models on real network measures. In this work, real setups were used only for population distributions. Using databases with a broader span such as the GHSL database [Florczyk et al., 2019], more urban regions could be considered. Corresponding street networks could then be extracted from OpenStreetMap as it was done in [Raimbault, 2019b], and real network measures included in the point cloud. Depending on the obtained coverage, specific calibration algorithms may be needed to try approaching the points out of the simulated space.

Finally, this work would gain to include more generative models for road networks. [Molinero and Hernando, 2020] recently proposed a network model based on self-reinforcement processes. [Queyroi, 2018] compares biological with shortest-path routing procedures to simulate transport networks. [Mimeur, 2016] introduces a model for rail networks which includes the progressive hierarchicalisation of the network. The governance processes core to the model developed by [Raimbault and Le Néchet, 2021] or the investment processes described in [Cats et al., 2020], allow inferring results on other dimensions of the transport systems, but it is not clear to what extent they are able to produce fundamentally different networks than the ones obtained through simpler models. All these other approaches, among others, are also relevant candidates to be included in such systematic benchmark. One difficulty remains to be able to formalise, setup and implement all models within a single comparable context.

## References

- [Barthelemy et al., 2013] Barthelemy, M., Bordin, P., Berestycki, H., and Gribaudi, M. (2013). Self-organization versus top-down planning in the evolution of a city. *Scientific reports*, 3(1):1–8.
- [Barthélemy and Flammini, 2009] Barthélemy, M. and Flammini, A. (2009). Co-evolution of density and topology in a simple model of city formation. *Networks and spatial economics*, 9(3):401–425.
- [Batista e Silva et al., 2013] Batista e Silva, F., Gallego, J., and Lavalle, C. (2013). A high-resolution population grid map for europe. *Journal of Maps*, 9(1):16–28.
- [Batty, 2007] Batty, M. (2007). Complexity in city systems: Understanding, evolution, and design. *A planners encounter with complexity*, pages 99–122.

- [Cats and Birch, 2021] Cats, O. and Birch, N. (2021). Multi-modal network evolution in polycentric regions. *Journal of Transport Geography*, 96:103159.
- [Cats et al., 2020] Cats, O., Vermeulen, A., Warnier, M., and van Lint, H. (2020). Modelling growth principles of metropolitan public transport networks. *Journal of Transport Geography*, 82:102567.
- [Chérel et al., 2015] Chérel, G., Cottineau, C., and Reuillon, R. (2015). Beyond corroboration: Strengthening model validation by looking for unexpected patterns. *PloS one*, 10(9):e0138212.
- [Florczyk et al., 2019] Florczyk, A. J., Corbane, C., Ehrlich, D., Freire, S., Kemper, T., Maffenini, L., Melchiorri, M., Pesaresi, M., Politis, P., Schiavina, M., et al. (2019). Ghsl data package 2019. *Luxembourg, EUR*, 29788(10.2760):290498.
- [Louf et al., 2013] 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.
- [Mimeur, 2016] Mimeur, C. (2016). *Les traces de la vitesse entre réseau et territoire : approche géohistorique de la croissance du réseau ferroviaire français*. Theses, Université de Bourgogne.
- [Molinero and Hernando, 2020] Molinero, C. and Hernando, A. (2020). A model for the generation of road networks. *arXiv preprint arXiv:2001.08180*.
- [Pumain, 2021] Pumain, D. (2021). Co-evolution as the secret of urban complexity. In *Handbook on Cities and Complexity*. Edward Elgar Publishing.
- [Queyroi, 2018] Queyroi, F. (2018). Biological and shortest-path routing procedures for transportation network design. *arXiv preprint arXiv:1803.03528*.
- [Raimbault, 2018a] Raimbault, J. (2018a). Des systèmes naturels aux systèmes urbains: génération de réseaux de transport optimaux par modèle slime-mould. In *BioMim Expo 2018*.
- [Raimbault, 2018b] Raimbault, J. (2018b). Multi-modeling the morphogenesis of transportation networks. In *Artificial Life Conference Proceedings*, pages 382–383. MIT Press.
- [Raimbault, 2019a] Raimbault, J. (2019a). Second-order control of complex systems with correlated synthetic data. *Complex Adaptive Systems Modeling*, 7(1):1–19.
- [Raimbault, 2019b] Raimbault, J. (2019b). An urban morphogenesis model capturing interactions between networks and territories. In *The mathematics of urban morphology*, pages 383–409. Springer.



- [Raimbault, 2020] Raimbault, J. (2020). Unveiling co-evolutionary patterns in systems of cities: a systematic exploration of the simpopnet model. In *Theories and Models of Urbanization*, pages 261–278. Springer.
- [Raimbault and Le Néchet, 2021] Raimbault, J. and Le Néchet, F. (2021). Introducing endogenous transport provision in a luti model to explore polycentric governance systems. *Journal of Transport Geography*, 94:103115.
- [Reuillon et al., 2013] 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.
- [Ribeill, 1985] Ribeill, G. (1985). Aspects du développement du réseau ferré français sur la longue durée. l’approche historique. *FLUX Cahiers scientifiques internationaux Réseaux et Territoires*, 1(1):10–25.
- [Szell et al., 2021] Szell, M., Mimar, S., Perlman, T., Ghoshal, G., and Sinatra, R. (2021). Growing urban bicycle networks.
- [Tirico et al., 2018] Tirico, M., Balev, S., Dutot, A., and Olivier, D. (2018). Morphogenesis of complex networks: A reaction diffusion framework for spatial graphs. In *International Conference on Complex Networks and their Applications*, pages 769–781. Springer.
- [Verdier and Bretagnolle, 2007] Verdier, N. and Bretagnolle, A. (2007). L’extension du réseau des routes de poste en france, de 1708 à 1833. In *histoire des réseaux postaux en Europe du XVIIIe au XXIe siècle*, pages 155–193. Comité pour l’histoire de la Poste.
- [Wegener and Fürst, 2004] Wegener, M. and Fürst, F. (2004). Land-use transport interaction: State of the art. *Available at SSRN 1434678*.
- [Xie and Levinson, 2009] Xie, F. and Levinson, D. (2009). Modeling the growth of transportation networks: A comprehensive review. *Networks and Spatial Economics*, 9(3):291–307.
- [Xie and Levinson, 2011] Xie, F. and Levinson, D. (2011). *Evolving transportation networks*. Springer Science & Business Media.