

Models for the Co-evolution of Cities and Networks

JUSTE RAIMBAULT^{1,2}

¹ UPS CNRS 3611 ISC-PIF

² UMR CNRS 8504 Géographie-cités

Keywords : *Urban System; Co-evolution; Transportation Network*

1 Introduction

2 Specification of co-evolution models

This section extends the logic of integrating a system of cities with a transportation network, which has been pursued in a static way for network behavior in the interaction model developed and explored in section ??, to propose a *macroscopic model of co-evolution for systems of cities*.

2.1 Rationale

This first approach relies in a direct extension of the interaction model within a system of cities described in chapter ??, at a macroscopic scale with an ontology typical to systems of cities. For the sake of simplicity, we still stick to an unidimensional description of cities by their population.

Concerning network growth, we propose also to stay at a relatively aggregated and simplified level, allowing to test growth heuristics at different levels of abstraction. In order to be flexible on model mechanisms, diverse processes can be taken into account, such as direct interactions between cities, intermediate interactions through the network, the feedback of network flows and a demand-induced network growth.

Empirical characteristics emphasized by [Thévenin et al., 2013] for the French railway network suggest the existence of feedbacks of network use, or of flows traversing it, on its persistence and its development, whose properties have evolved in time: a first phase of strong development would correspond to an answer to a high need of coverage, followed by a reinforcement of main link and the disappearance of weakest links.

The coupling between cities and the network will be achieved by the intermediate of flows between cities in the network: these capture the interactions between cities and have simultaneously an influence on the network in which they flow.

2.2 Model description

The urban system is characterized by populations $\mu_i(t)$ and the network $\mathbf{G}(t)$, to which can be associated a distance matrix $d_{ij}^G(t)$. Flows between cities ϕ_{ij} follow the expression given in ?? with network distance. The same way, the evolution of populations follows the specifications of the base model. The Fig. 1 shows the structure of the model.

Concerning the network, we assume that it evolves following the equation

$$\mathbf{G}(t+1) = F(\mathbf{G}(t), \phi_{ij}(t)) \quad (1)$$

such that the assignment of flows within the network and a local variation of its elements is possible. We propose in a first time to consider patterns linked to distance only, and to specify a relation on an abstract

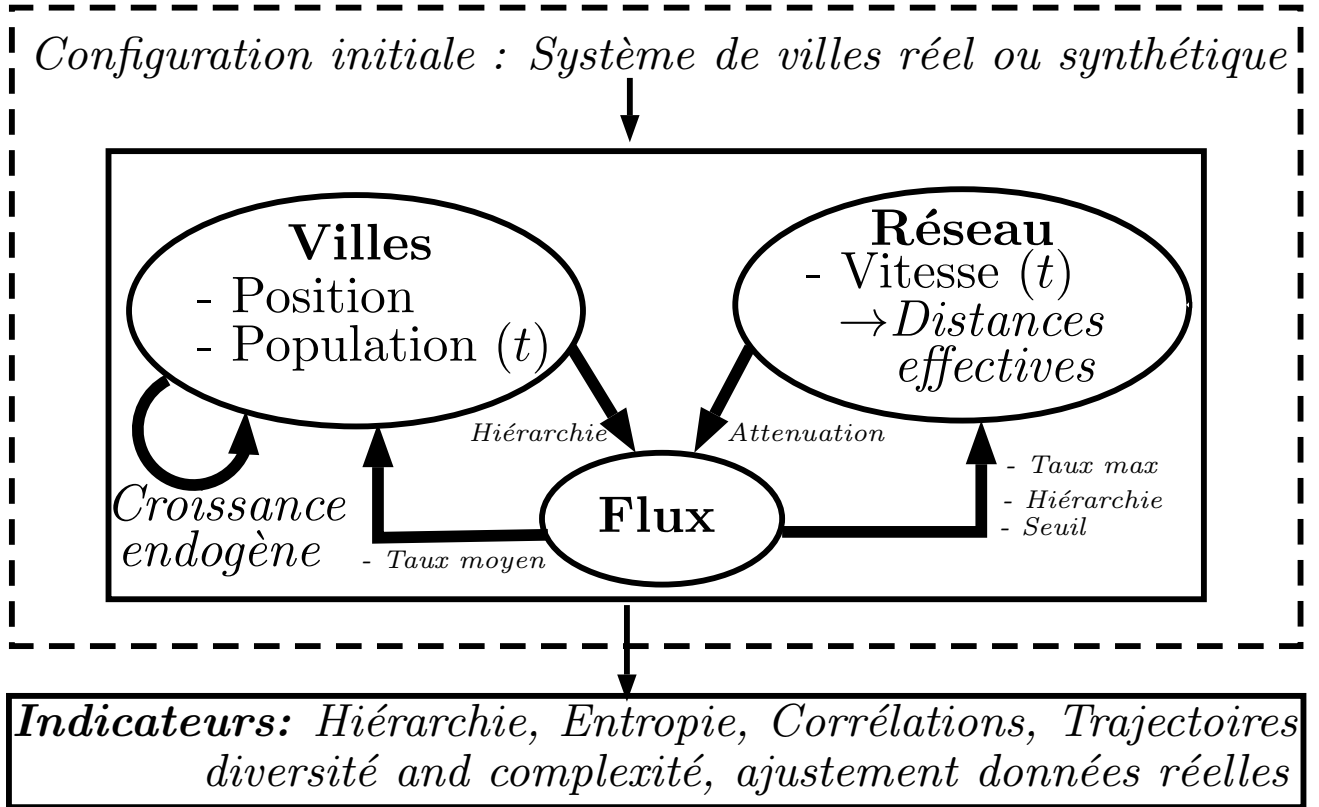


Figure 1: **Abstract representation of the model.** Ellipses correspond to main ontological elements (cities, network, flows), whereas arrows translate processes for which associated parameters are given. The model is described in its broader ecosystem of initialisation and output indicators.

network as

$$d_{ij}^G(t+1) = F(d_{ij}^G(t), \phi_{ij}(t)) \quad (2)$$

i.e. an evolution of the distance matrix only. In this spirit, we keep an interaction model strictly at a macroscopic scale, since a precise spatialization of the network would imply to take into account a finer scale that includes the local shape of the network which determines shortest paths.

Following a thresholded feedback heuristic, given a flow ϕ in a link, we assume its effective distance to be updated by:

$$d(t+1) = d(t) \cdot \left(1 + g_{max} \cdot \left[\frac{1 - \left(\frac{\phi}{\phi_0}\right)^{\gamma_s}}{1 + \left(\frac{\phi}{\phi_0}\right)^{\gamma_s}} \right] \right) \quad (3)$$

with γ_s a hierarchy parameter, ϕ_0 the threshold parameter and g_{max} the maximal growth rate at each step. This auto-reinforcement function can be interpreted the following way: above a limit flow ϕ_0 , the travel conditions improve, whereas they deteriorate below. The hierarchy of gain is given by γ_s , and since $\frac{1 - \left(\frac{\phi}{\phi_0}\right)^{\gamma_s}}{1 + \left(\frac{\phi}{\phi_0}\right)^{\gamma_s}} \rightarrow_{\phi \rightarrow \infty} -1$, g_{max} is the maximal distance gain. This function is similar to the one used by [Tero et al., 2007]¹.

2.3 Indicators

3 Results

3.1 Implementation

The coupling of the interaction model to a finer representation of the network (for example an encoding of the whole network structure) makes the full integration into an OpenMole plugin more difficult, as it was done for the model studied in ???. We need here an *ad hoc* implementation. The use of a workflow as a mediator for coupling is an interesting solution but which is realistic only for a weak coupling as in ???. One of the issues that the meta-modeling library for OpenMole that is currently being developed around OpenMole will have to tackle is the possibility to allow strong coupling (for example in the sense of a dynamical coupling during the evolution of the simulation) of heterogeneous components in a transparent way, in order to benefit from the advantages of different languages or of already existing implementations.

We choose here a full implementation with NetLogo, for the simplicity of coupling between components. A particular care is taken for the duality of network representation, both as a distance matrix and as a physical network, in order to facilitate the extension to physical network heuristics.

3.2 Exploration on a synthetic system of cities

The model is first tested and explored on synthetic city systems, in order to understand some of its intrinsic properties. In this case, we consider the model with an abstract network as specified above, i.e. without spatial description of the network and with evolution rules acting directly on d_{ij}^G given the previous specifications.

A synthetic city system is generated following the heuristic used in the previous section: (i) N_S cities are randomly distributed in the euclidian plan; (ii) populations are attributed to cities following an inverse power law, with a hierarchy parameter α_S and such that the largest city has a population equal to P_{max} , i.e. following $P_i = P_{max} \cdot i^{-\alpha_S}$.

¹Which uses $\Delta d = \Delta t \left[\frac{\phi^\gamma}{1+\phi^\gamma} - d \right]$. This function yield similarly a threshold effect, since the derivative vanishes at $\phi^* = \left(\frac{d}{1-d} \right)^{1/\gamma}$, but it can not be adjusted.

To simplify, several meta-parameters are fixed: the number of cities is fixed at $N_S = 30$, the maximal population at $P_{max} = 100000$ and the maximal network growth to $g_{max} = 0.005$. Final time is fixed at $t_f = 30$, what corresponds to distances divided approximatively by 5^2 , in order to comply to an empirical constraint: this corresponds to the evolution of the travel time between Paris and Lyon from around ten hours at the beginning of the century to two hours today, showed for example by [Thévenin et al., 2013]. We also neglect network effects at the second order by taking $w_N = 0$.

We explore a grid in the parameter space $\alpha_S, \phi_0, \gamma_s, w_G, d_G, \gamma_G$. We use the indicators introduced in ?? to quantify model behavior in the parameter space. We describe the results for $\alpha_S = 1$, what is the closest to existing city systems (in comparison to 0.5 and 1.5, see the systematic review of the rank-size law estimations done by [Cottineau, 2017]).

L'évolution de la centralité de proximité moyenne dans le temps est visualisée en Fig. ?? (haut) pour $w_G = 0.001$, et à (γ_G, ϕ_0) variables. Le comportement n'est pas sensible à d_G (voir graphique complet en ??). Cette évolution témoigne d'une transition en fonction du niveau de hiérarchie : lorsque celui-ci décroît, on observe l'émergence de trajectoires où la centralité moyenne croît dans le temps, ce qui correspond à des situations où l'ensemble des villes bénéficie en moyenne d'accroissements d'accessibilité.

Concerning the entropy of populations, for which the temporal trajectory is shown in Fig. ?? (bottom), all parameters give a decreasing entropy, i.e. a behavior of convergence of cities trajectories in time³.

Looking at the complexity of accessibility trajectories, we observe for values of $\phi_0 > 1.5$ a maximum of complexity as a function of interaction distance d_G , stable when w_G and γ_G vary (see also the exhaustive plots in Fig. ??, Appendix ??). This intermediate scale can be interpreted as producing regional subsystems, large enough for each to develop a certain level of complexity, et isolated enough to avoid the convergence of trajectories over the whole system. We reconstruct therein a spatial non-stationarity, typically observed in ??, and rejoin the concept of the ecological niche⁴ localized in space: the emergent subsystems that are relatively independent, are good candidates to contain processes of co-evolution. The emergence of this intermediate scale can be compared to the modularity of the French urban system showed by [Berroir et al., 2017].

Finally, the behavior of rank correlations for accessibility reveals that the interaction distance systematically increases the number of hierarchy inversions, what corresponds in a sense to an increase in overall system complexity. The hierarchy parameter diminishes this correlation, what means that a more hierarchical organization will impact a larger number of cities in the qualitative aspects of their trajectories. This effect is similar to the "first mover advantage" showed by [Levinson and Xie, 2011], which unveils a path dependency and an advantage to be rapidly connected to the network: in our case, the modifications in the hierarchy correspond to cities that benefit from their positioning in the network.

3.3 Application to the French system of cities

The model is then applied to the French system of cities on long time dynamical data: the Pumain-INED database for populations, spanning from 1831 to 1999 [?], with the evolving railway network from 1840 to 2000 [Thévenin et al., 2013]. Such a time span can be associated with structural effect on long time, as developed in ??. This application aims on the one hand at testing the ability of the model to reproduce a real dynamic of co-evolution, and on the other hand at extracting thematic information on processes through calibrated parameter values.

²Indeed, we can compute that the minimal multiplicative factor for distance is $(1 - g_{max})^{t_f}$, what gives for these values $(1 - 0.05)^{30} \simeq 0.214$, i.e. a division by 5 of the travel time.

³Indeed, the entropy for the population variable gives the dispersion of the distribution of populations, and thus its decrease translate a trend to concentrate in time.

⁴As it was already described in ??, an ecological niche in the sense of [Holland, 2012] corresponds to the relatively independent ecosystem in which there is co-evolution between the species.

3.3.1 Data

We work on railway network data constructed by [Thévenin et al., 2013]. The French railway network is particularly interesting jointly with population data already presented, since the covered time span is relatively close, and as [Thévenin et al., 2013] recalls, this transportation mode has at any times materialized the implication of public and private actors. It corresponds to different processes depending on the period, from a more decentralized management to a more centralized recently, and different technological materializations with for example the recent emergence of high speed trains [Zembri, 1997]. For each date in the population database, we extract the simplified abstract network in which all stations and intersections with a degree larger than two are linked with abstract links which speed and length attributes correspond to real values, at a granularity of 1km⁵. This yields the time-distance matrices between the cities included in the model.

3.3.2 Stylized facts

Before calibrating the model, we can observe the lagged correlation patterns in the dataset, by applying the causality regimes method. This empirical study should on the one hand allow us to verify well known stylized facts, and on the other hand to produce a preliminary knowledge of empirical system behavior. We compute as detailed above the closeness centrality through the network, given by $T_i = \sum_j \exp -d_{ij}/d_0$, and we study the lagged correlation between its derivative ΔT_i and the derivative of the population ΔP_i , given by $\hat{\rho}_\tau = \hat{\rho}[\Delta P_i(t), \Delta T_i(t - \tau)]$ estimated on a moving window containing T_w successive dates. We show in Fig. 2 the results obtained.

These results are important for at least two reasons. First, the behavior of the number of significant correlations as a function of T_w and d_0 allows us to find stationarity scales in the system. We observe on the one hand a specific spatial scale that gives a maximum for all temporal windows, at $d_0 = 100km$, what suggests the existence of consistent regional subsystems, which existence is stable in time: indeed, this value corresponds to the interaction distance. It remarkably coincides with the intermediate scale isolated in the synthetic model. On the other hand, long spatial ranges induce an optimal temporal scale, for $T_w = 4$ what corresponds to around twenty years: we identify it as the overall temporal stationarity scale of the system and study the lagged correlations for this value.

Secondly, the behavior of lagged correlations does not seem to comply to the existing literature. At the intermediate spatial scale, the values of ρ_+ , ρ_- exhibit no regularity. On the whole system, there is until 1946 close to no significant effect, then no causality between 1946 and 1975 (maximum at $\tau = 0$, non-significant minimum), and a 5 years shift of accessibility causing population after 1968 (the effect staying however doubtful). We do not reproduce the correlation effect between network centrality and place in the urban hierarchy advocated by [Bretagnolle, 2003]⁶, what lead us to question the existence of the “structural co-evolution” on long time described by BRETAGNOLLE in [Offner et al., 2014]. What [Bretagnolle, 2003] obtains is a simultaneous correspondence between growth rate and level of connectivity to the network (and not with network dynamic), but not in our sense a co-evolution, since no statistical relation is furthermore exhibited.

We rejoin the recent results of [Mimeur et al., 2017] that show the statistical non-significance of the correlation between growth rate and evolution of network coverage and accessibility, at a zero delay. Our results are less precise on the class of cities studied (they differentiate large and small cities, and work

⁵This processing is achieved thanks to the R package for transportation network analysis specifically developed for this thesis, see ??.

⁶As [Lemoy and Caruso, 2017] is not able to reproduce, for density profiles as a function of the distance to the center of European metropolis, the transition that allows [Guérois and Pumain, 2008] to define the peri-urban. These more or less recent works are not reproducible, producing neither code nor data, and giving only a superficial description of the methods, and it is thus impossible to know the origin of the qualitative divergence obtained. A good reproducibility together with the construction of systematic comparisons (*benchmarks*) of models, empirical analysis, that are recent but also to validate old studies, seems to be a reasonable solution to this kind of issue.

on a larger panel), but more general as they study variable delays and accessibility ranges, and are thus complementary.

3.3.3 Calibration of the abstract model

Expected results of the calibration on real data concern both the more or less accurate reproduction of real city population growth dynamics, i.e. to what extent the inclusion of a dynamical network can increase the explanatory power for trajectories, and also how realistic the evolution of network distance is. We still work with the abstract model.

Model evaluation We can add to the indicators used before a calibration indicator for distance. The particular property of adjustment for populations, that resides in the existence of a power law for the sizes of cities that made negligible the performance on medium and small cities in the case of a cumulated error, and suggested the addition of the indicator on the error on logarithms, is not present for distances that follow a distribution concentrated on a single order of magnitude. We use therefore a standard measure of fit, given by

$$\varepsilon_D = \log \left[\sum_t \sum_{i,j} \left(d_{ij}(t) - \tilde{d}_{ij}(t) \right)^2 \right]$$

where $d_{ij}(t)$ are observed distances and $\tilde{d}_{ij}(t)$ the simulated distances. It is simply a cumulated squared-error, as used for the comparison of origin-destination matrices in a similar case of simulation of a transportation network in [Jacobs-Crisioni and Koopmans, 2016].

Results We proceed to a non-stationary calibration, on the $(\varepsilon_P, \varepsilon_D)$ objectives, i.e. the squared-error on populations and on distances. The estimation is done with a moving window with the periods already used in ???. In order to have a limited dimension to explore, we take a fixed $w_N = 0$ to study the interactions only at the first order, knowing that the abstract network parameters $(g_{max}, \gamma_S, \varphi_0)$ are taken into account in the calibration. The calibration is done with a genetic algorithm in a way similar as in ???. The Fig. 3 shows the obtained Pareto fronts, and the Fig. 4 the evolution in time of parameter values for the optimal solutions.

We observe a large variability of the shape of Pareto fronts for the bi-objective calibration on population and distance, what witnesses more or less difficulty to simultaneously adjust population and distance. Some periods, such as 1891-1911 and 1921-1936, are close to have a simultaneous objective point for the two objectives, what would correspond to a good correspondence of the model to both trajectories of cities and trajectory of the network on these periods.

In comparison with calibration results of the model with static network of ??, when comparing the performances for the objective ε_G , we find periods where the static is clearly better (1831 and 1841 for example) and others where the co-evolutive model is better (1946 and 1962): thus, taking into account the co-evolution helps in some cases to have a better reproduction of population trajectories.

The values of optimal parameters in time, shown in Fig. 4, seem to contain some signal. The evolution of w_G and γ_G are coherent with the evolutions observed for the static model. For d_G , the model principally saturates on the maximal distance and the evolution is difficult to interpret.

However, the evolution of ϕ_0 could be a sign of a “TGV effect” in recent periods, through the secondary peak for population after 1960. Indeed, the construction of high speed lines has shortened distances between cities on top of the hierarchy, and an increase of the threshold ϕ_0 corresponds to an increase of the selectivity for a potential diminution of distances.

The calibrated g_{max} can finally be interpreted according to the history of the railway network (at least of all points in the Pareto front): a significant secondary peak in the first years, a minimum in the years corresponding to the stabilization of the network (1900), and an increase until today linked to the increase of train speeds and the opening of high speed lines.

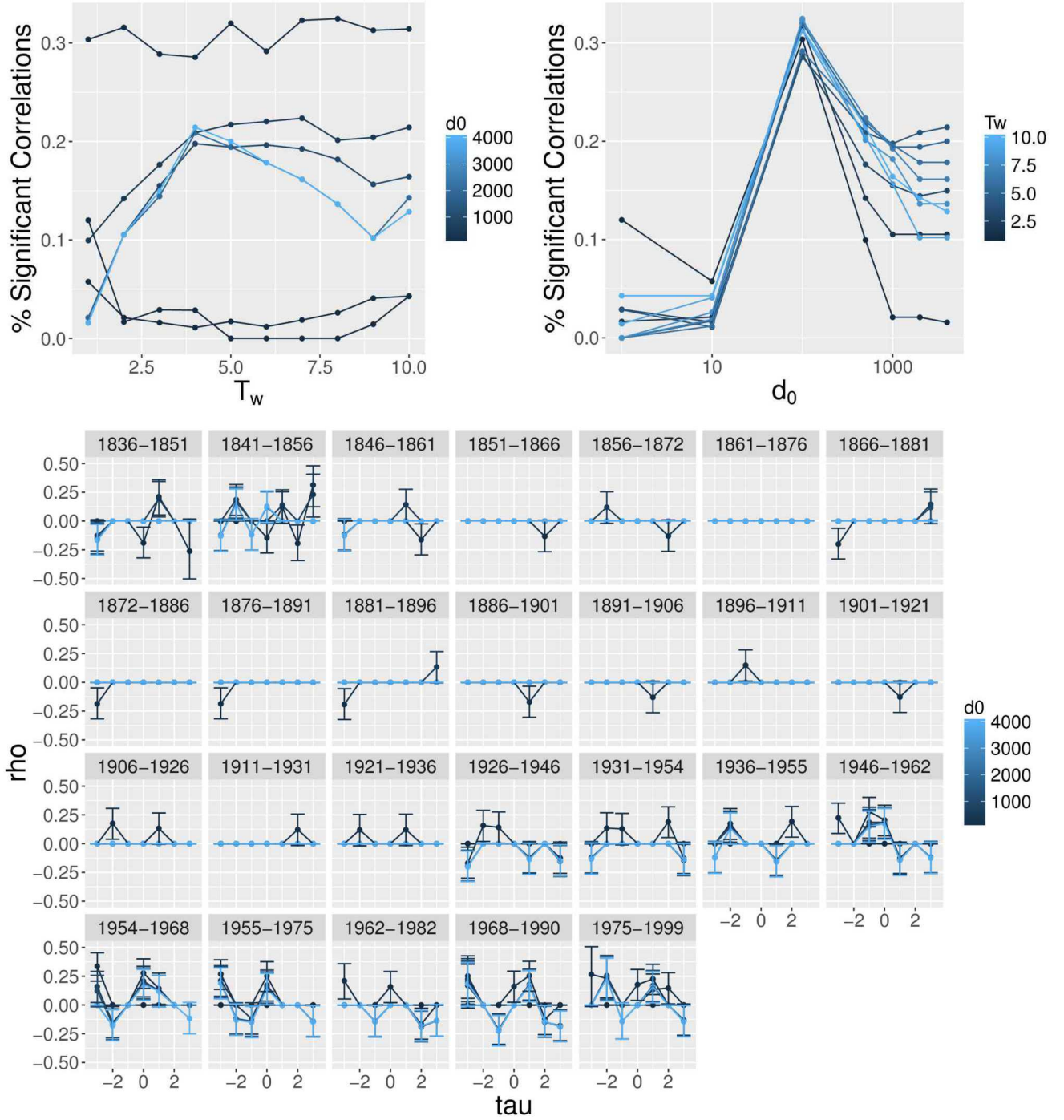


Figure 2: **Empirical lagged correlations for the French system of cities.** Correlations are estimated on a window of duration $5 \cdot T_w$, between population growth rates and the variations of closeness centrality with a decay parameter d_0 (see text). (*Top left*) Number of significant correlations (taken such that $p < 0.1$ at 95%) as a function of T_w for d_0 variable; (*Top right*) Number of significant correlations as a function of d_0 for T_w variable; (*Bottom*) For the “optimal” window $T_w = 4$, value of ρ_τ as a function of τ , for all successive periods.

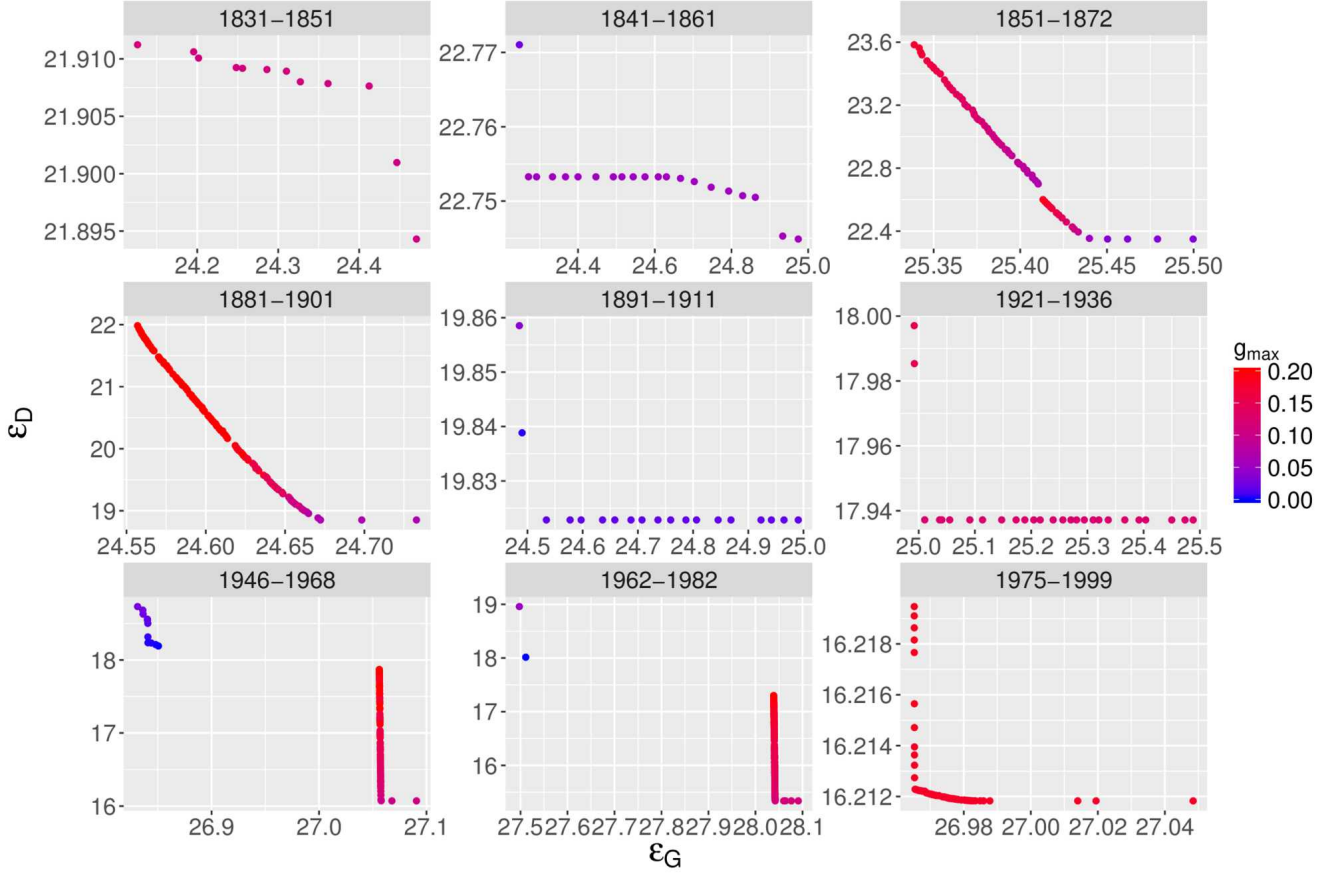


Figure 3: **Pareto fronts for the bi-objective calibration between population and distance.** Fronts are given for each calibration period and are colored according to g_{max} .

We have this way in a certain extent indirectly quantify interaction processes through the network and the processes of network adaptation to flows, in the case of a real system.

3.3.4 Model with a physical network

We now sketch the outline of a specification of the model with a physical network, what would in a sense correspond to an hybrid model combining different scales. The objective of such a specification would be on the one hand to study the difference in trajectories compared to the abstract network, i.e. to quantify the importance of economies of scale (due to common links), of congestion and also the possible compromises to take in order to spatialize the network. On the other hand, it would help to understand to what extent it is possible to produce realistic networks in comparison to autonomous network growth models for example. These issues are tackled at an other scale and for other ontological specifications in chapter ??.

Such a specification follows the frame of [Li et al., 2014], which model the co-evolution between transportation corridors and the growth of main poles at a regional scale.

The physical network we implement aims at satisfying a greedy criteria of local time gain. More precisely, we assume a self-reinforcement similar to [Tero et al., 2010] A specification analog to the one used before assumes a growth for each link, given also in a logic of self-reinforcement by:

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

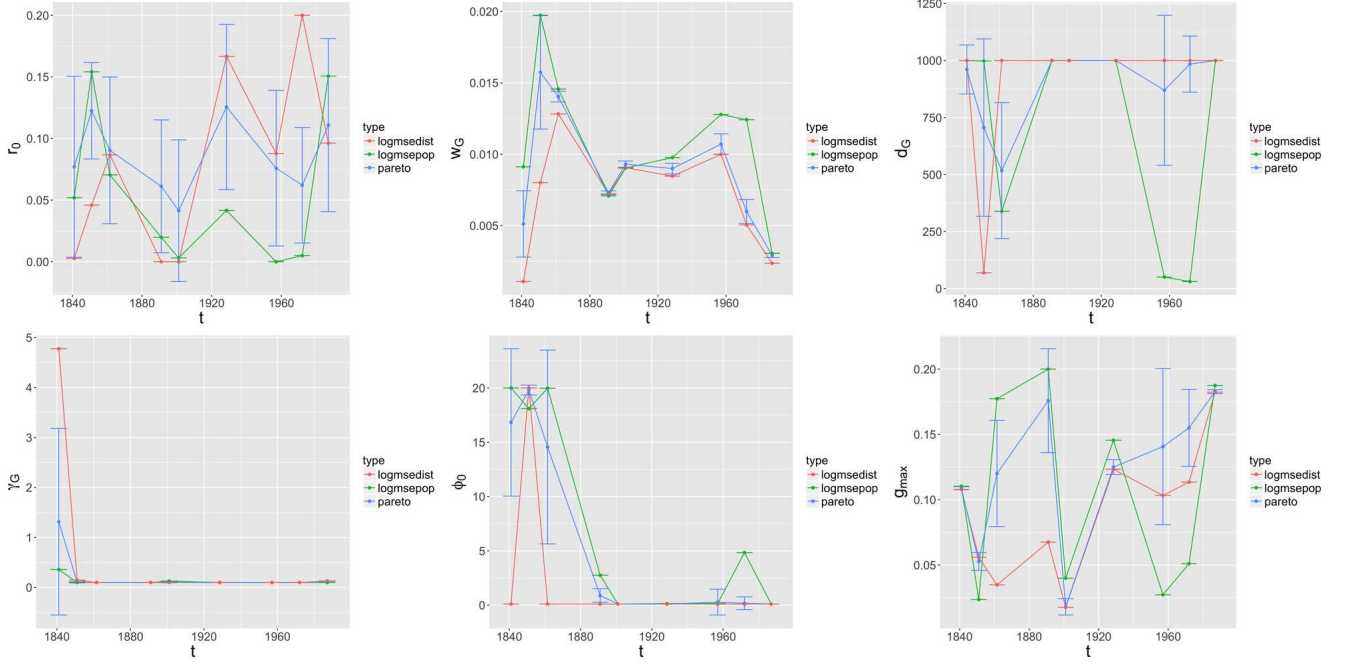


Figure 4: **Temporal evolution of optimal parameters.** From left to right and top to bottom, values of parameters $(r_0, w_G, d_G, \gamma_G, \phi_0, g_{max})$, respectively for the full Pareto front (blue), for the optimal point in the sense of the distance (red) and the optimal point in the sense of the population (green).

if ϕ is the flow in the link and $d(t)$ its effective distance. The threshold specification used before does indeed not allow a good convergence in time, in particular with the emergence of local oscillation phenomena.

We generate a random initial network, by perturbing the position of vertices of a grid for which a fixed proportion of links has been removed (40%) and by linking cities to the network through the shortest path. Links have all the same impedance, which then evolves according to the equation above. An example of a configuration obtained with this specification is given in Fig. 5. The good convergence properties (visual stabilization of network structure during restricted experiments) suggest the potentialities offered by this specification, which systematic exploration is out of the scope of this work.

4 Discussion

4.1 Application

The study of particular trajectories within a system of cities can allow to answer to specific thematic questions: for example, the influence of medium-sized cities on the global trajectory of the system, or the drivers of a more or less “successful” trajectory for this type of profile. In the case of the application to a real system, the mapping of deviation to the model in time can suggest regional particularities.

4.2 Developments

We also finally expect to be able through the model to compare urban systems in different geographical and political contexts, and at different scales. This should foster the understanding the implications of planning actions on the interactions between networks and territories. For example, French railway network has emerged through multiple operators, on the contrary to the Chinese high speed railway network, for which a more precise development could be considered.

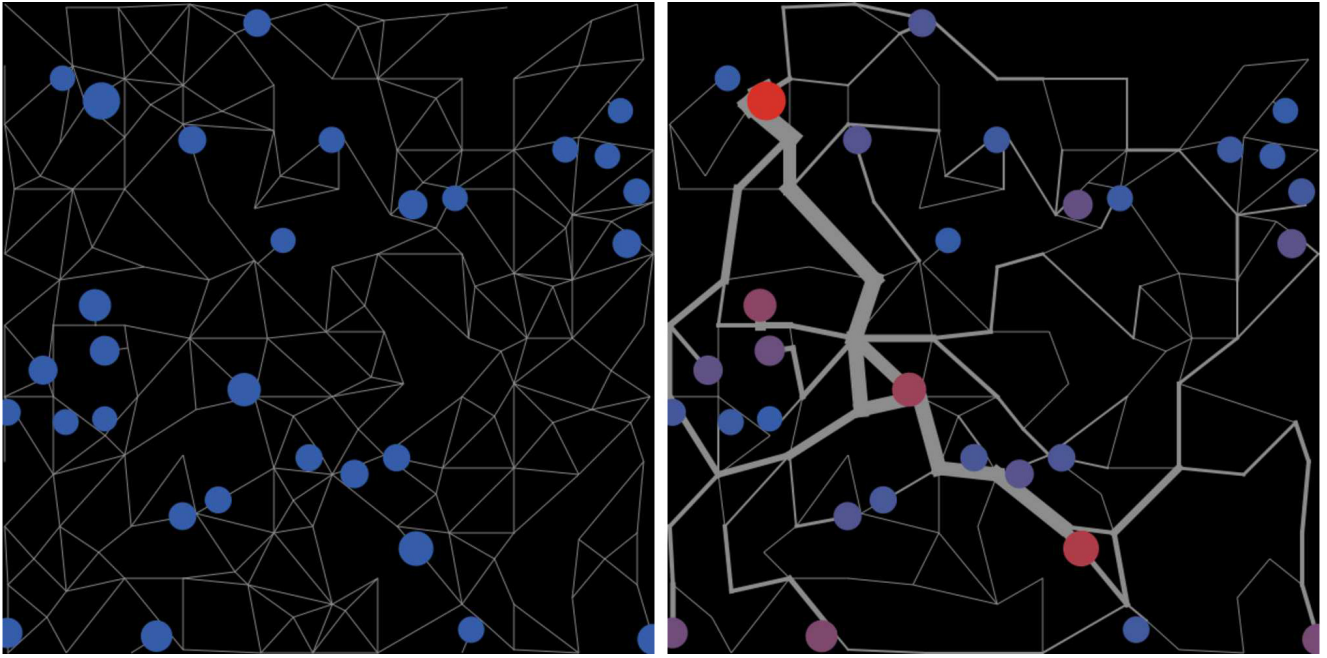


Figure 5: **Example of configuration obtained with a self-reinforcing network.** (*Left*) Initial random configuration, with uniform impedances; (*Right*) Final configuration obtained after 100 iterations.

Conclusion

References

- [Berroir et al., 2017] Berroir, S., Cattani, N., Dobruszkes, F., Guérois, M., Paulus, F., and Vacchiani-Marcuzzo, C. (2017). Les systèmes urbains français: une approche relationnelle. *Cybergeog: European Journal of Geography*.
- [Bretagnolle, 2003] Bretagnolle, A. (2003). Vitesse et processus de sélection hiérarchique dans le système des villes françaises. *Données urbaines*, 4.
- [Cottineau, 2017] Cottineau, C. (2017). Metazipf. a dynamic meta-analysis of city size distributions. *PLOS ONE*, 12(8):1–22.
- [Guérois and Pumain, 2008] Guérois, M. and Pumain, D. (2008). Built-up encroachment and the urban field: a comparison of forty european cities. *Environment and Planning A*, 40(9):2186–2203.
- [Holland, 2012] Holland, J. H. (2012). *Signals and boundaries: Building blocks for complex adaptive systems*. MIT Press.
- [Jacobs-Crisioni and Koopmans, 2016] Jacobs-Crisioni, C. and Koopmans, C. C. (2016). Transport link scanner: simulating geographic transport network expansion through individual investments. *Journal of Geographical Systems*, 18(3):265–301.
- [Lemoy and Caruso, 2017] Lemoy, R. and Caruso, G. (2017). Scaling evidence of the homothetic nature of cities. *arXiv preprint arXiv:1704.06508*.
- [Levinson and Xie, 2011] Levinson, D. and Xie, F. (2011). Does first last? the existence and extent of first mover advantages on spatial networks. *Journal of Transport and Land Use*.

- [Li et al., 2014] Li, Y., Lu, D., and Tian, Y. (2014). Modeling corridor and growth pole coevolution in regional transportation network. *Transportation Research Record: Journal of the Transportation Research Board*, (2466):144–152.
- [Mimeur et al., 2017] Mimeur, C., Queyroi, F., Banos, A., and Thévenin, T. (2017). Revisiting the structuring effect of transportation infrastructure: an empirical approach with the French Railway Network from 1860 to 1910. *Historical Methods: A Journal of Quantitative and Interdisciplinary History*.
- [Offner et al., 2014] Offner, J.-M., Beaucire, F., Delaplace, M., Frémont, A., Ninot, O., Bretagnolle, A., and Pumain, D. (2014). Les effets structurants des infrastructures de transport. *Espace Géographique*, (42):p–51.
- [Tero et al., 2007] Tero, A., Kobayashi, R., and Nakagaki, T. (2007). A mathematical model for adaptive transport network in path finding by true slime mold. *Journal of theoretical biology*, 244(4):553–564.
- [Tero et al., 2010] 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.
- [Thévenin et al., 2013] Thévenin, T., Schwartz, R., and Sapet, L. (2013). Mapping the distortions in time and space: The french railway network 1830–1930. *Historical Methods: A Journal of Quantitative and Interdisciplinary History*, 46(3):134–143.
- [Zembri, 1997] Zembri, P. (1997). Les fondements de la remise en cause du schéma directeur des liaisons ferroviaires à grande vitesse: des faiblesses avant tout structurelles. In *Annales de géographie*, pages 183–194. JSTOR.