Quantifying the co-evolution of economic activities locations with geo-historical data: Paris, 19th century

Juste Raimbault[1,2,3,4], and Julien Perret[1]

[1]LASTIG, Univ Gustave Eiffel, IGN-ENSG, Saint-Mandé, France
[2]CASA, UCL, London, UK
[3]UPS CNRS 3611 ISC-PIF, Paris, France
[4]UMR CNRS 8504 Géographie-cités, Paris, France

Urban systems are highly complex, what poses issues to ensure their resilience and sustainability and plan future cities [1]. One aspect of this complexity is their multidimensionality, even within subsystems such as the intra-urban location of economic activities. Indeed, different types of activities have specific location processes, yielding various patterns for their accessibility for example [2]. Understanding past dynamics of such social processes is crucial to build better urban models, theories and in practical terms insights for sustainable planning [3].

While recent and current urban dynamics are more and more easily tracked and quantified through the emergence of large urban data [4], historical dynamics on long time scales or on timeframes in a distant past are difficult to quantify due to the sparsity and heterogeneity of geo-historical data. At the macroscopic scale of urban systems, major transitions of past settlements systems have been modeled from an interdisciplinary perspective [5], and simulation models capturing various dimensions of systems of cities have been developed [6]. At the mesoscopic and microsocpic intra-urban scales, several issues are encountered when trying to build consistent dataset, such as geocoding [7] or vectorisation to reconstruct the dynamics of road networks [8].

This contribution builds on data produced by the Soduco research project [9] to investigate co-evolutionary dynamics in the location of economic activities. More precisely, we focus on the case of Paris in the middle of the 19th century. Using public domain scans for main economic activity repertoires ("*Didot-Bottin*"), an other work package of the project focused on building geo-referenced data containing activities of various professionals, using Optical Character Recognition techniques. We work on a sample of this data currently available, spanning 4 years between 1841 and 1844.

Starting from an initial dataset of 415,976 entries, we keep the ones with correct geographical coordinates (33%), and apply natural language processing (stemming and stop-word removal) to raw strings describing activities. From there, we focus on the stems with more than 100 occurences (578 stems), and code manually a broad category of economic activity for each (obtaining the coarse grain classification within: food, craftsmanship, art and literature, health, law and governance, service, teaching). We end at this stage with 36,072 entries with identified coordinates and broad activity. Source code and results are openly available on the git repository of this work at `https://github.com/JusteRaimbault/HistoricalData`.

To study co-evolutionary dynamics, we use the definition and characterisation method proposed by [10], which is based on weak Granger causality: two urban attributes will be co-evolving if they statistically exhibit a circular causal relationship in this context. Aggregating spatially on raster cells (10x10 grid for the whole Paris), we thus compute lagged correlations between the variations of activity counts between successive dates for each cell, for each pair of activity. We find several significant correlations (using Fisher confidence intervals), mostly negative for lagged correlations. This corresponds to a dynamic of sustitution of activities, with clusters rapidly replaced. Food and health are positively correlated and in co-evolution, while the other significantly co-evolving couple of activities is health and art but in a negative manner: in some districts medical professions replace artists and the contrary in others. We also consider simultaneous correlations, and find for example that food and craftsmanship have joint dynamics in the last time interval, but not during the first considered. Altogether, these results first confirm the existence of a co-evolution between some activities, unveil a precise characterisation of intra-urban socio-economic dynamics, and open the path towards more advanced thematic interpretations within an interdisciplinary context, such as with historians.

Current and future work also include (i) the extension of this study on longer time spans; (ii) the combination of Granger causality with geographically weighted regression, to optimise spatial neighbourhood considered in regressions [11], and (iii) the benchmark of methods to characterise co-evolution (including e.g. instrumental variables or causal machine learning methods) on this particular dataset.

[1] Michael Batty. *Inventing future cities*. MIT press, 2018.

[2] Antonio Paez. Network accessibility and the spatial distribution of economic activity in eastern asia. *Urban Studies*, 41(11):2211–2230, 2004.

[3] Céline Rozenblat, Denise Pumain, and Elkin Velasquez. *International and transnational perspectives on urban systems*. Springer, 2018.

[4] Jens Kandt and Michael Batty. Smart cities, big data and urban policy: Towards urban analytics for the long run. *Cities*, 109:102992, 2021.

[5] Lena Sanders. *Peupler la terre: De la préhistoire à l'ère des métropoles*. Presses universitaires François-Rabelais, 2018.

[6] Denise Pumain and Romain Reuillon. *Urban dynamics and simulation models*. Springer, 2017.

[7] Rémi Cura, Bertrand Dumenieu, Nathalie Abadie, Benoit Costes, Julien Perret, and Maurizio Gribaudi. Historical collaborative geocoding. *ISPRS International Journal of Geo-Information*, 7(7):262, 2018.

[8] Hanae El Gouj, Christian Rincón-Acosta, and Claire Lagesse. Urban morphogenesis analysis based on geohistorical road data. *Applied Network Science*, 7(1):1–26, 2022.

[9] Social dynamics in urban context. Open tools, models and data - Paris and its suburbs, 1789-1950. https://soduco.github.io/.

[10] Juste Raimbault. Characterising and modeling the co-evolution of transportation networks and territories. *arXiv preprint arXiv:2110.15950*, 2021.

[11] Chris Brunsdon, Stewart Fotheringham, and Martin Charlton. Geographically weighted regression. *Journal of the Royal Statistical Society: Series D (The Statistician)*, 47(3):431–443, 1998.