

Identification of causalities in spatio-temporal data

J. Raimbault^{1,2}

juste.raimbault@parisgeo.cnrs.fr

¹UMR CNRS 8504 Géographie-cités

²UMR-T IFSTTAR 9403 LVMT

Sageo 2017 - Rouen

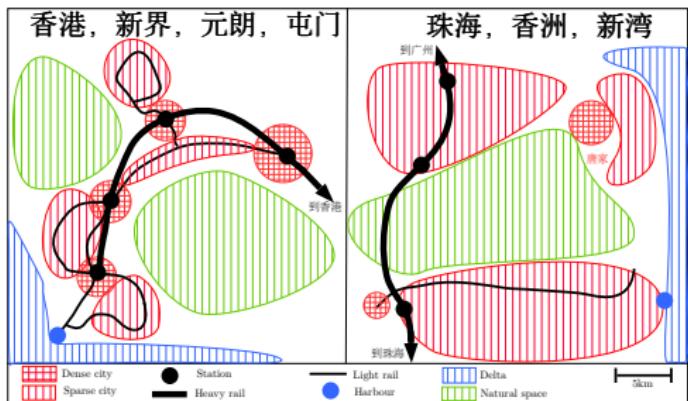
November 7th 2017

Spatio-temporal complexity



Source : [Arnaud et al., 2013]

Complex relations between networks and territories



Fieldwork in Pearl River Delta unveils local manifestations of the co-evolution of transportation networks and territories. Source : Author.

The myth of structuring effects

From [Bonnafous and Plassard, 1974] to [Offner, 1993] : do transportation infrastructures structure territories ?

- Existence of co-evolutive processes [Bretagnolle, 2009]
- At large scale, existence of structural urban system dynamics [Offner et al., 2011]
- The question of circular causalities arises at all scales (e.g. metropolitan scale and mobility [Cerqueira, 2017]) and in various fields (knowledge spillovers and innovation [Audretsch and Feldman, 1996])

Causality in Geography

- Classical geography already investigated causal links in space [Loi, 1985]
- [Claval, 1985] : beyond reductionist causality in systemic analysis
- Systemogenesis introduced by [Durand-Dastes, 2003] focuses on dynamics and path-dependency
- Towards a complex approach to causality ? [Morin, 1976]

Existing approaches in spatio-temporal causality

Transportation Networks and Territories

- Lagged correlations : [Levinson, 2008] London population and network connectivity ; [Gargi Chaudhuri and Keith C Clarke, 2015] historical data in North Italy
- Instrumental variables : [Duranton and Turner, 2012] US highways and employment ; [Berger and Enflo, 2017] clear effect of swedish railway on cities trajectories

Spatio-temporal correlations

- Matching method for traffic flows [Liu et al., 2011]
- Generalized granger causality in neuroscience [Ke et al., 2007]
- Spatio-temporal correlations in Computer Vision [Ke et al., 2007]

Research objective

- Genericity and operability of existing approaches ?
- Grasp complexity in a simple way ?
- At the interface of knowledge domains (methodology, modeling and empirical) [Raimbault, 2017]

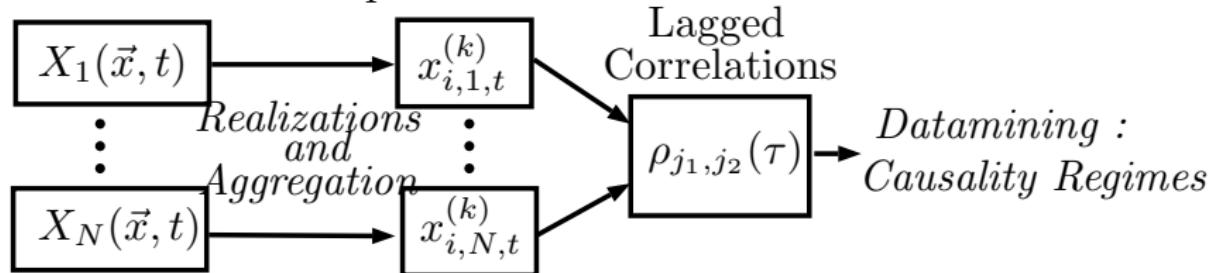
Research objective :

Explore a generic method based on patterns of spatio-temporal lagged correlations : notion of Granger causality ; validation on synthetic data and application to a case study.

Method: Rationale

Spatio-temporal
stochastic field

Trajectories of
spatial units



Method: Formalization

Correlation estimator $\hat{\rho}$ applying in time, space and repetitions, i.e. covariance is estimated by $\hat{\text{Cov}}[X, Y] = \hat{\mathbb{E}}_{i,t,k}[XY] - \hat{\mathbb{E}}_{i,t,k}[X]\hat{\mathbb{E}}_{i,t,k}[Y]$

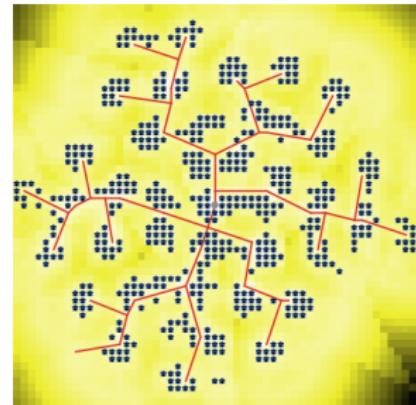
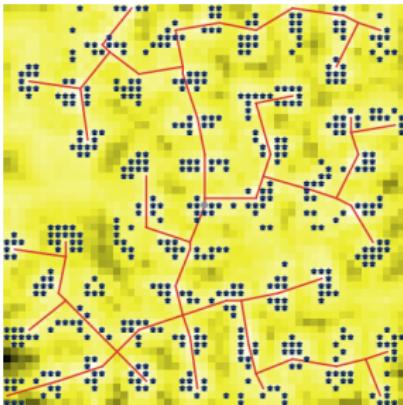
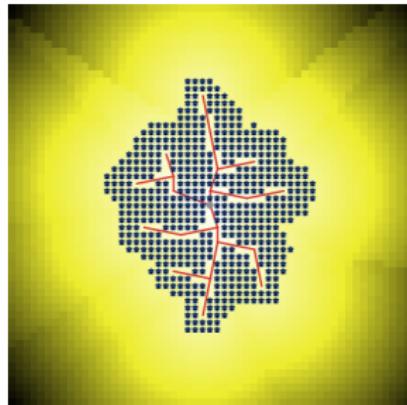
Lagged Correlation defined by

$$\rho_\tau[X_{j_1}, X_{j_2}] = \hat{\rho}\left[x_{i,j_1,t-\tau}^{(k)}, x_{i,j_2,t}^{(k)}\right] \quad (1)$$

Patterns of $\text{argmax}_\tau \rho_\tau[X_{j_1}, X_{j_2}]$ or $\text{argmin}_\tau \rho_\tau[X_{j_1}, X_{j_2}]$ (assumed clearly defined : e.g. statistical significance, minimal value) capture the sense of causality between j_1 and j_2

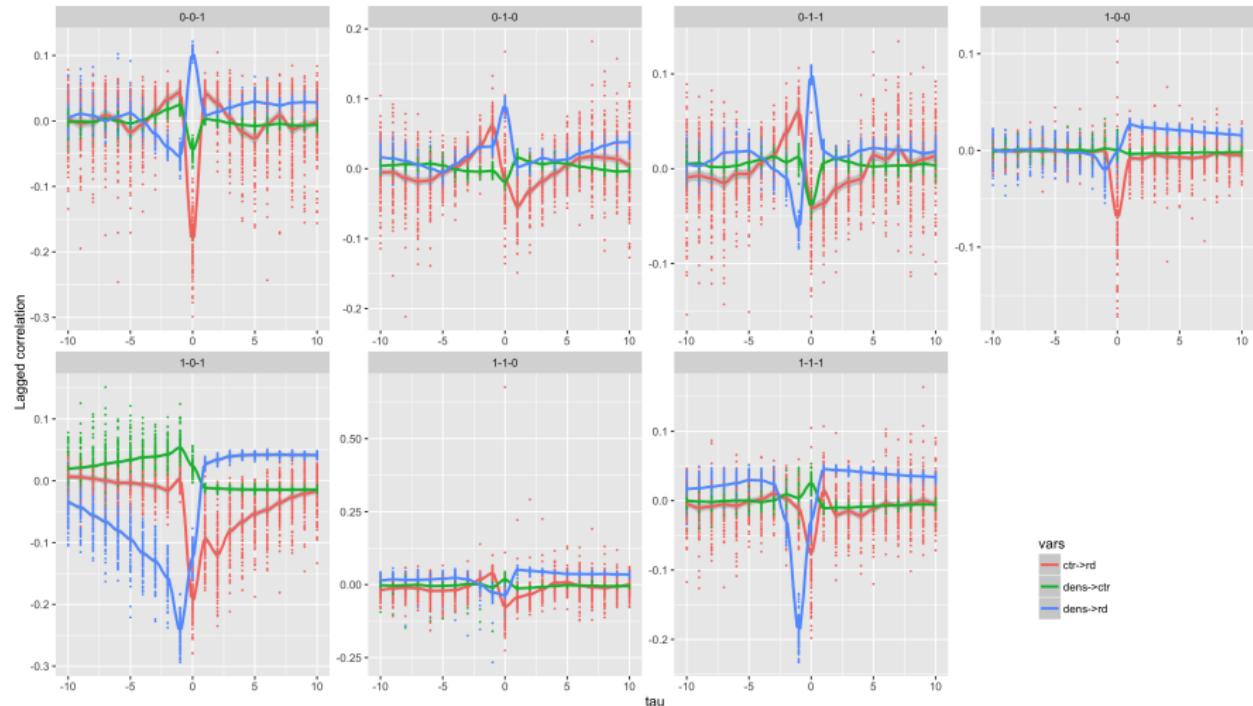
→ Datamining on ρ_τ (possibly parametrized values as $\rho_\tau^{(\omega)}$) to understand causality patterns.

Validation: Synthetic Data



Synthetic urban configurations generated by an hybrid morphogenesis model from [Raimbault et al., 2014]

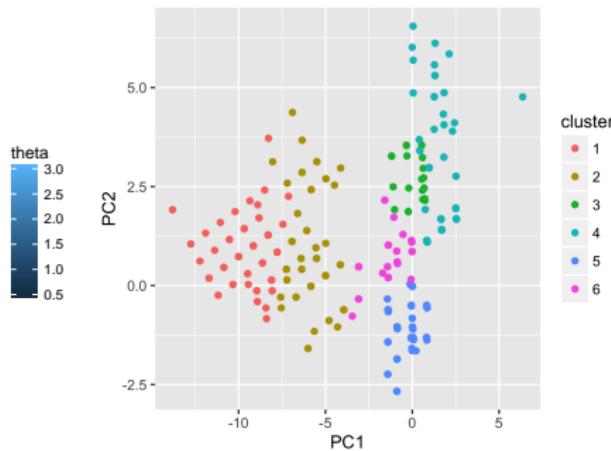
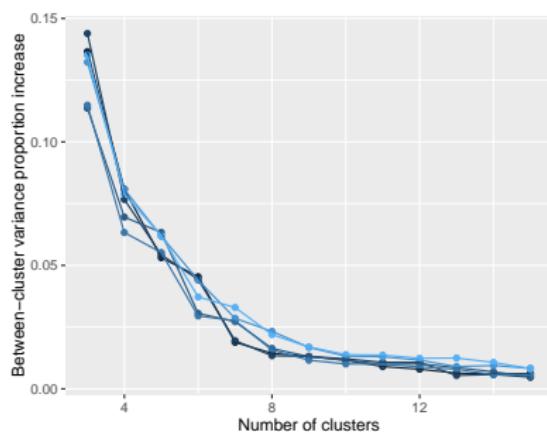
Profiles of lagged correlations



Values of ρ_τ for all couples of three explicative variables (density, distance to center, distance to roads), for 8 extreme parameter points

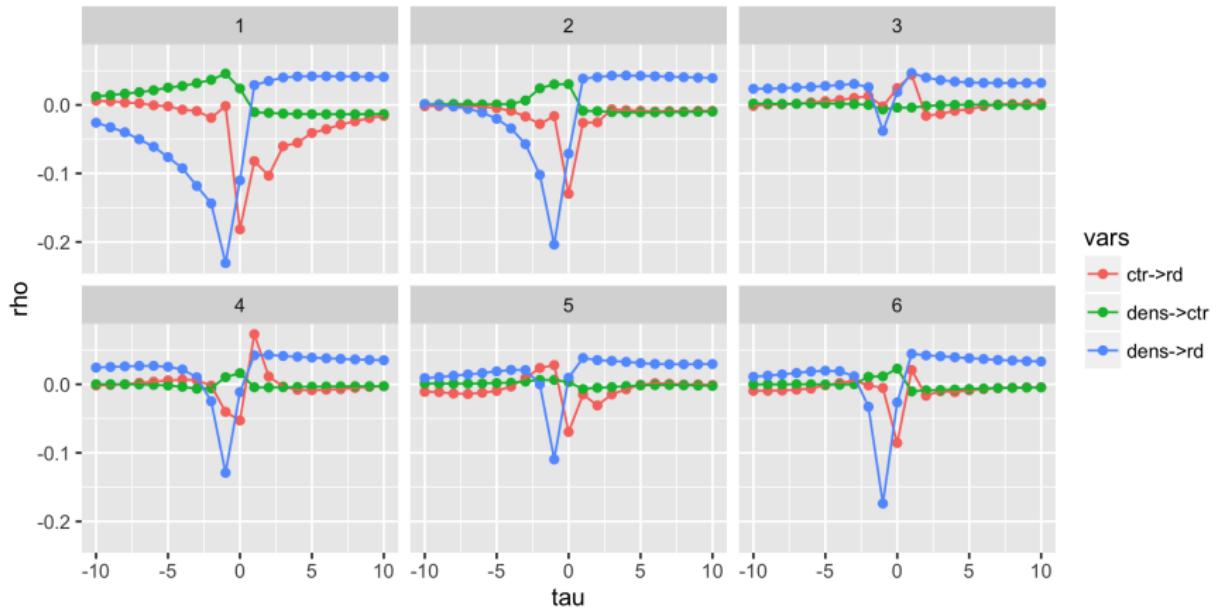
Unveiling Endogenous causality regimes

Intensive exploration of model parameter space (1000 parameters points x 100 repetitions) with OpenMole software [Reuillon et al., 2013]



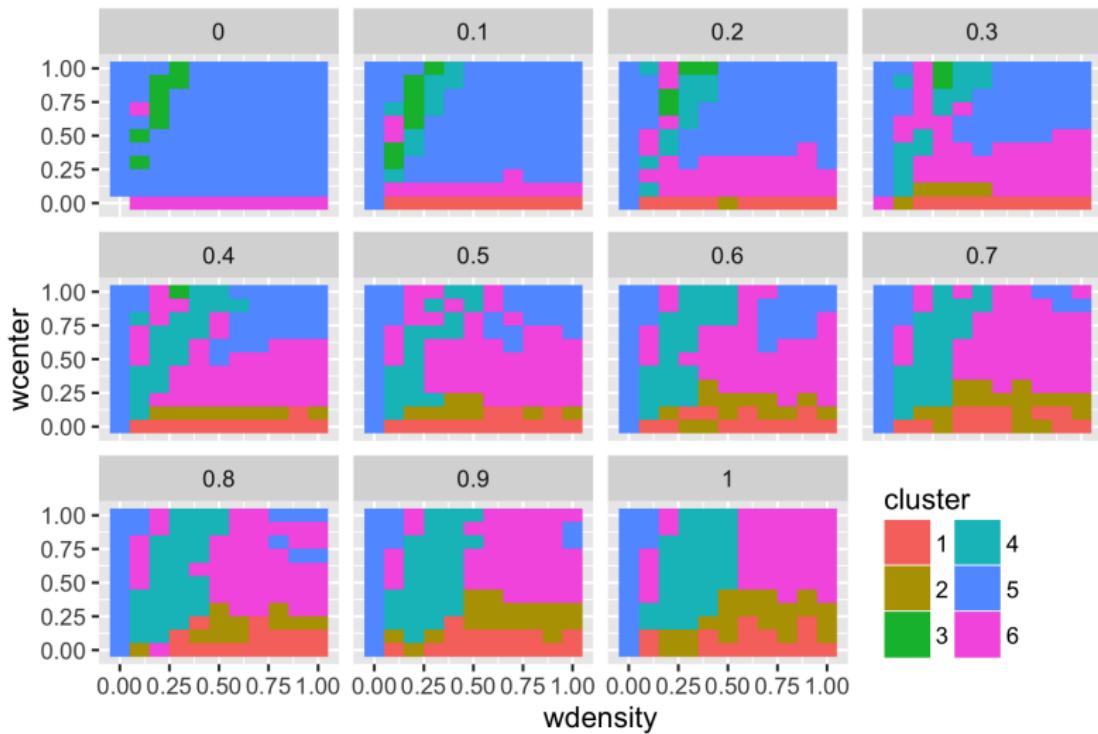
Unsupervised classification (robust k-means) on τ_{min}, τ_{max} features: (Left) Derivative of clustering coefficient for number of clusters k ; (Right) PCA visualisation of classification for “optimal” k

Consistence and interpretation of regimes



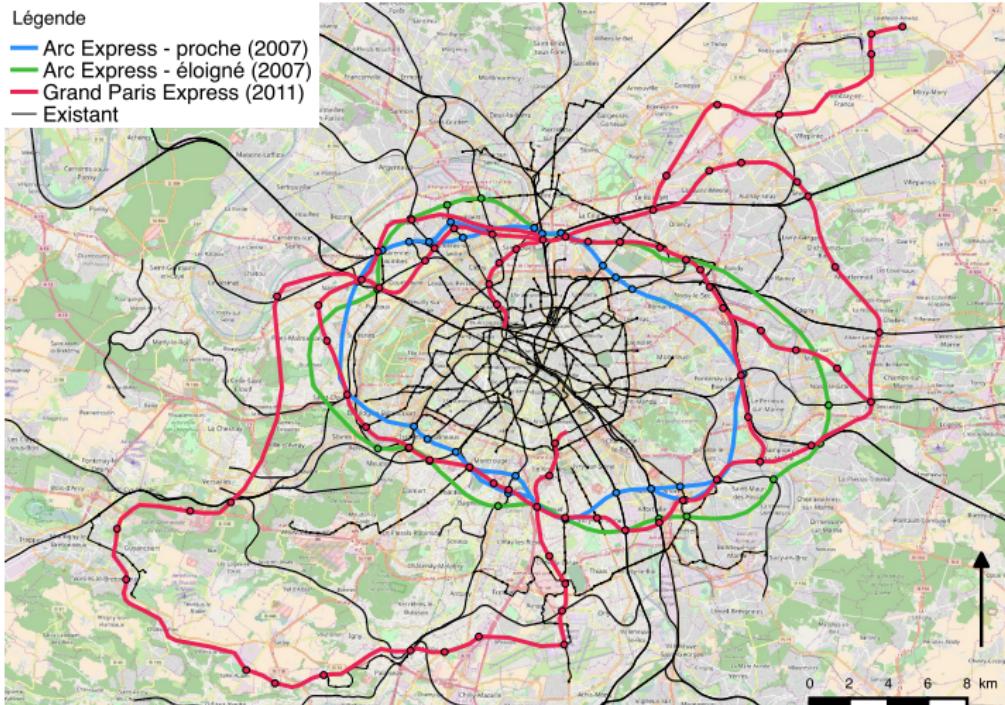
Values of cluster centers in terms of ρ_τ

Consistence and interpretation of regimes



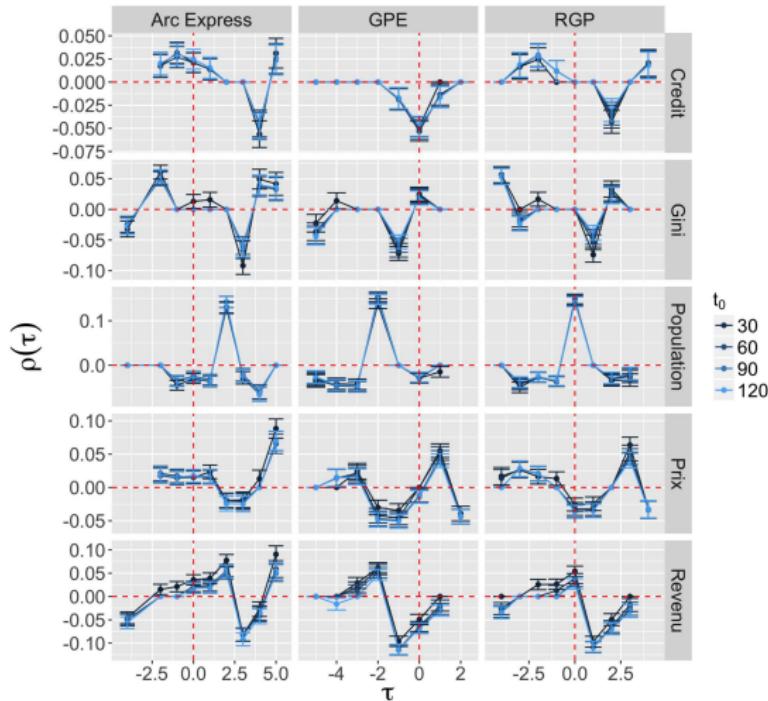
Position of clusters in the parameter space w_i

Application: Case study



Successive projects for the Grand Paris new transportation infrastructure

Application: Results



Values of ρ_τ for the different projects (columns) and different variables (rows), with accessibility differentials

Discussion

Implications

- Lagged correlation patterns on real data to investigate “structuring effects” in complex systems
- The operational concept of *Causality Regimes* introduces a novel way to look at co-evolution in models of simulation

Developments

- Characterisation of spatio-temporal diffusion : testing the spatial diffusion of innovation and the evolutive urban theory [Pumain, 2010]
- Optimal spatial scales for stationarity : link with GWR ?
[Brunsdon et al., 1998]

Conclusion

- A method validated on synthetic data and showed operational on a real system
 - At the interface of knowledge domains : theory, modeling, empirical, methodological
 - At the interface of disciplines : spatial analysis, statistics, datamining
-
- Code, data and results available at
<https://github.com/JusteRaimbault/CityNetwork>
 - Paper on arXiv at <https://arxiv.org/abs/1709.08684>
 - Acknowledgments : We thank the *European Grid Infrastructure* and its *National Grid Initiatives* (*France-Grilles* in particular) to give the technical support and the infrastructure.

Reserve slides

Reserve Slides

Granger causality

Granger causality test based on VAR processes :

$$X(t) = \sum_{0 \leq \tau \leq \tau_Y} b_\tau Y(t - \tau)$$

If there exists b_τ such that $|b_\tau| > 0$ significantly, then Y Granger-causes X .

We have then $\rho_\tau(Y, X) > 0$.

Morphogenesis

Morphogenesis (*Oxford dictionary*)

- ① *Biology* : The origin and development of morphological characteristics
- ② *Geology* : The formation of landforms or other structures.

History of the notion

- Started significantly with embryology around 1930 [Abercrombie, 1977]
- Turing's 1952 paper [Turing, 1952], linked to the development of Cybernetics
- first use in 1871, large peak in usage between 1907-1909, increase until 1990, decrease until today. *Scientific fashion* ?

Defining Morphogenesis

Meta-epistemological framework of imbricated notions:
Self-organization \supseteq Morphogenesis \supseteq Autopoiesis \supseteq Life

Properties:

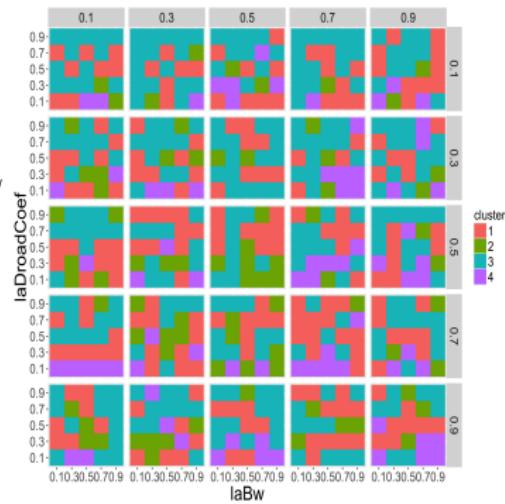
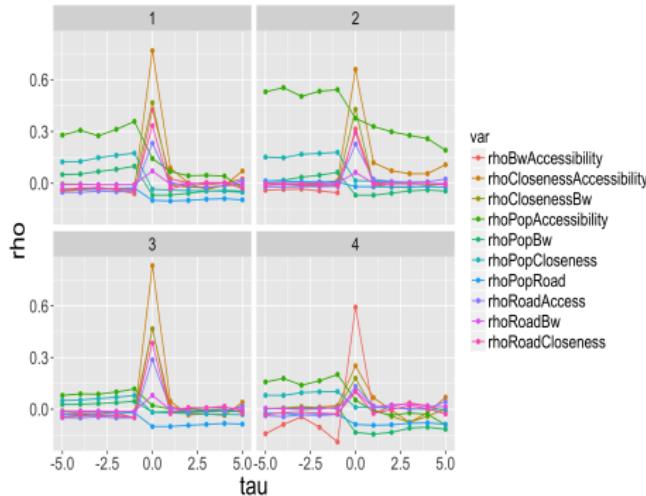
- Architecture links form and function
- Emergence strength [Bedau, 2002] increases with notion depth, as bifurcations [Thom, 1974]

Definition of Morphogenesis : *Emergence of the form and the function in a strongly coupled manner, producing an emergent architecture*
[Doursat et al., 2012]

Causality Regimes in a model of co-evolution

Unsupervised learning on lagged correlations between local variables unveils a diversity of causality regimes in a model co-evolving urban form and network topology

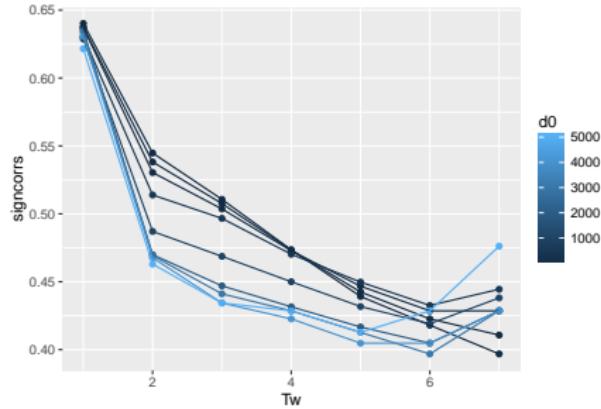
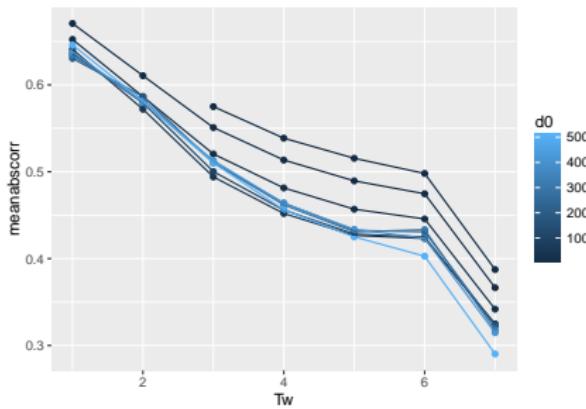
→ Link between *co-evolution regime* and morphogenetic properties of the urban system



(Left) Lagged correlation profiles of cluster centers; (Right) Distribution of regimes ↗ ↘ ↙ ↘

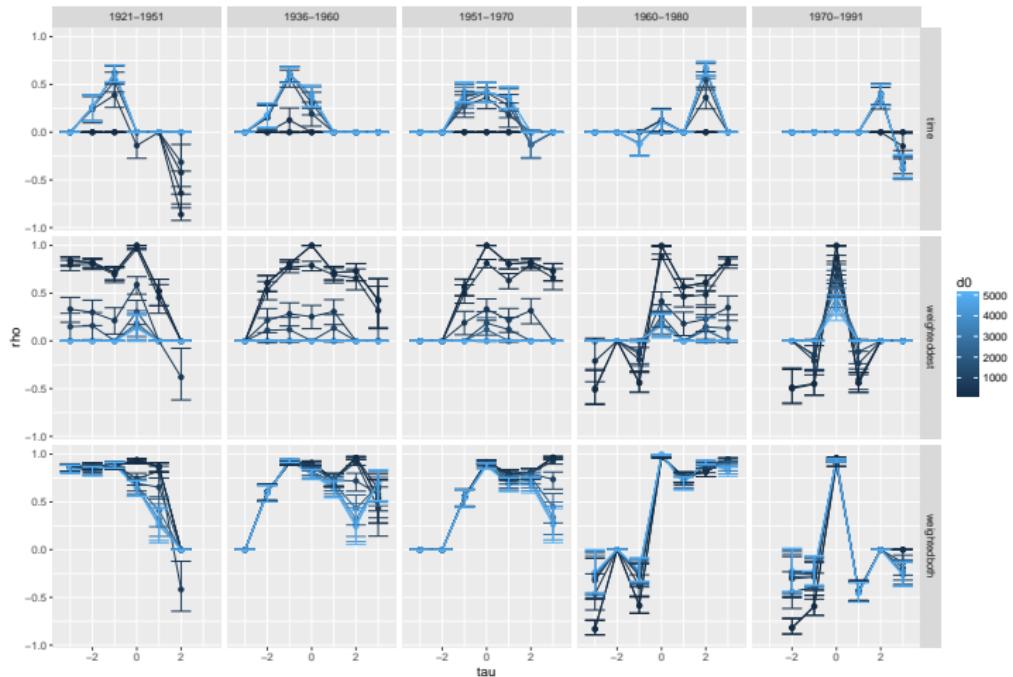
Application to South Africa : Stationarity scales

Optimal estimation time window and spatial range for accessibility



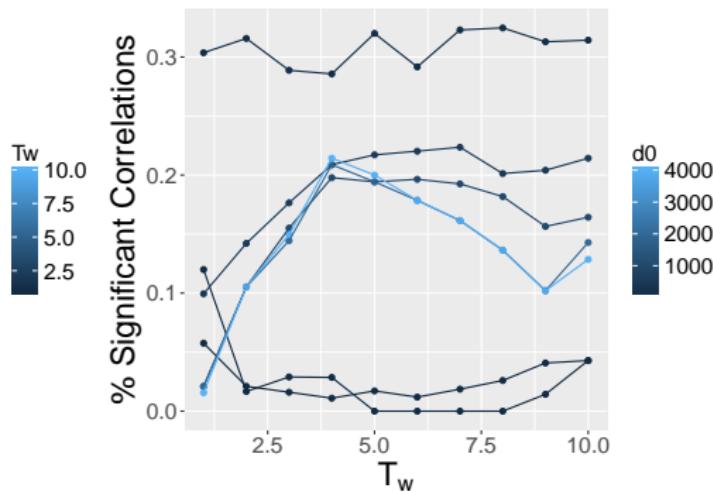
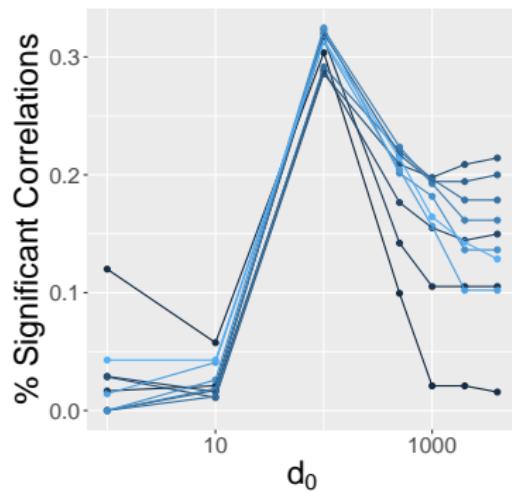
Application to South Africa : Causality Patterns

Clear inversion of the sense of Granger causality suggests a structural segregation effect of the apartheid laws

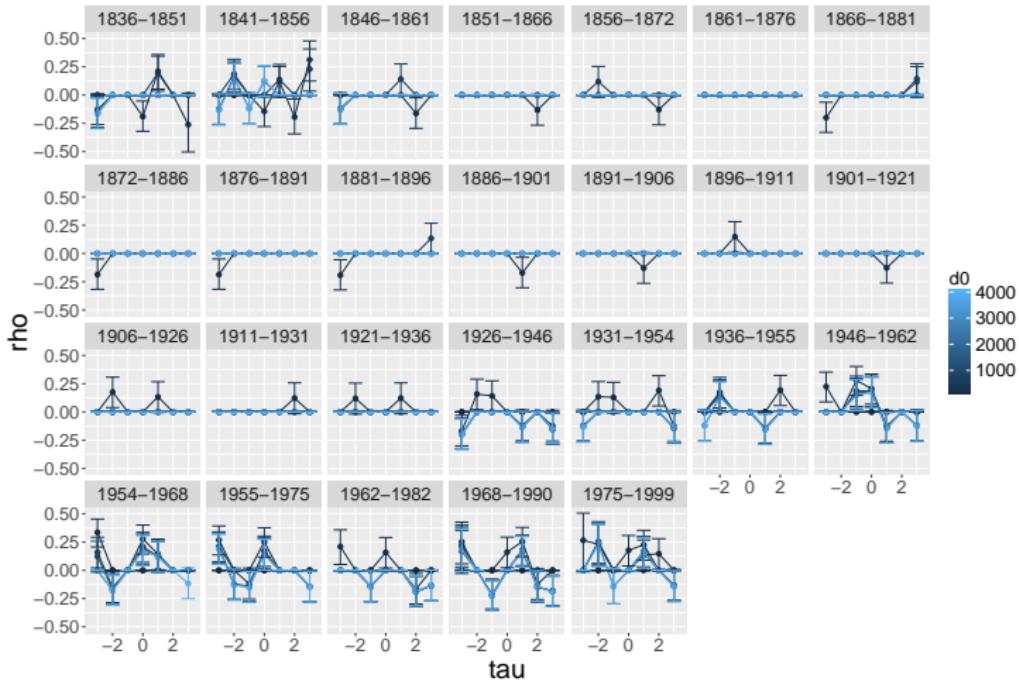


Application to France : Stationarity scales

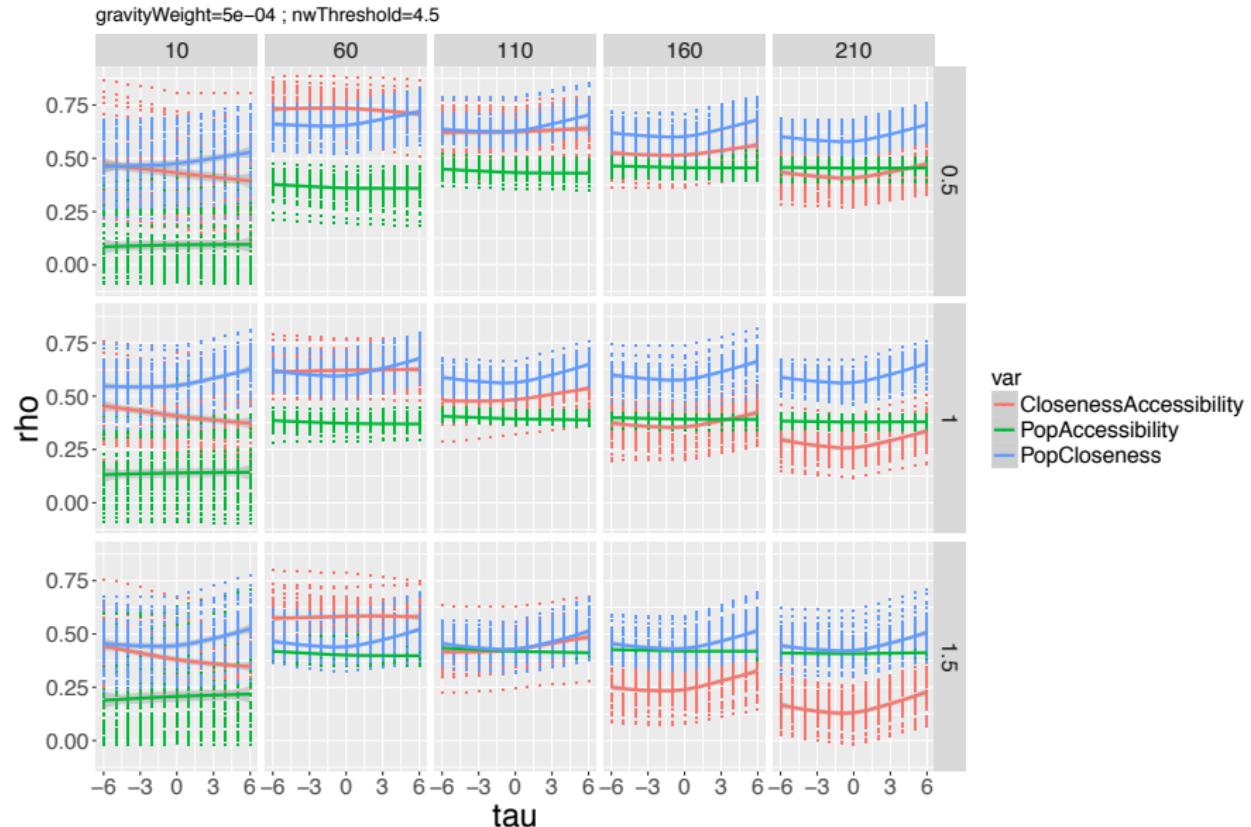
Optimal estimation time window and spatial range for accessibility



Application to France : No significant correlation



Macroscopic co-evolution : Correlation Patterns



References I

-  Abercrombie, M. (1977).
Concepts in morphogenesis.
Proceedings of the Royal Society of London B: Biological Sciences,
199(1136):337–344.
-  Arnaud, B., Bussi, M., and Guermond, Y. (2013).
Le grand paris et l'axe seine.
Métropoles, (13).
-  Audretsch, D. B. and Feldman, M. P. (1996).
R&d spillovers and the geography of innovation and production.
The American economic review, 86(3):630–640.
-  Bedau, M. (2002).
Downward causation and the autonomy of weak emergence.
Principia: an international journal of epistemology, 6(1):5–50.

References II

-  Berger, T. and Enflo, K. (2017).
Locomotives of local growth: The short-and long-term impact of railroads in sweden.
Journal of Urban Economics, 98:124–138.
-  Bonnafous, A. and Plassard, F. (1974).
Les méthodologies usuelles de l'étude des effets structurants de l'offre de transport.
Revue économique, pages 208–232.
-  Bretagnolle, A. (2009).
Villes et réseaux de transport : des interactions dans la longue durée, France, Europe, États-Unis.
Hdr, Université Panthéon-Sorbonne - Paris I.

References III

-  Brunsdon, C., Fotheringham, S., and Charlton, M. (1998). Geographically weighted regression. *Journal of the Royal Statistical Society: Series D (The Statistician)*, 47(3):431–443.
-  Cerqueira, E. V. (2017). Les inégalités d'accès aux ressources urbaines dans les franges périphériques de belo horizonte (brésil): quelles évolutions? *EchoGéo*, (39).
-  Claval, P. (1985). Causalité et géographie. *Espace géographique*, 14(2):109–115.

References IV

-  Doursat, R., Sayama, H., and Michel, O. (2012).
Morphogenetic engineering: toward programmable complex systems.
Springer.
-  Durand-Dastes, F. (2003).
Les géographes et la notion de causalité.
-  Duranton, G. and Turner, M. A. (2012).
Urban growth and transportation.
The Review of Economic Studies, 79(4):1407–1440.
-  Gargi Chaudhuri and Keith C Clarke (2015).
On the Spatiotemporal Dynamics of the Coupling between Land Use
and Road Networks: Does Political History Matter?
Environment and Planning B: Planning and Design, 42(1):133–156.

References V

-  Ke, Y., Sukthankar, R., and Hebert, M. (2007). Spatio-temporal shape and flow correlation for action recognition. In *Computer Vision and Pattern Recognition, 2007. CVPR'07. IEEE Conference on*, pages 1–8. IEEE.
-  Levinson, D. (2008). Density and dispersion: the co-development of land use and rail in london. *Journal of Economic Geography*, 8(1):55–77.
-  Liu, W., Zheng, Y., Chawla, S., Yuan, J., and Xing, X. (2011). Discovering spatio-temporal causal interactions in traffic data streams. In *Proceedings of the 17th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 1010–1018. ACM.

References VI

-  Loi, D. (1985).
Une étude de la causalité dans la géographie classique française.[l'exemple des premières thèses régionales].
Espace géographique, 14(2):121–125.
-  Morin, E. (1976).
La méthode: la nature de la nature.
Le Seuil.
-  Offner, J.-M. (1993).
Les "effets structurants" du transport: mythe politique, mystification scientifique.
Espace géographique, 22(3):233–242.

References VII

-  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.
-  Pumain, D. (2010).
Une théorie géographique des villes.
Bulletin de la Société géographie de Liège, (55):5–15.
-  Raimbault, J. (2017).
An applied knowledge framework to study complex systems.
ArXiv e-prints.

References VIII

-  Raimbault, J., Banos, A., and Doursat, R. (2014).
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.
-  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.
-  Thom, R. (1974).
Stabilité structurelle et morphogénèse.
Poetics, 3(2):7–19.

References IX



Turing, A. M. (1952).

The chemical basis of morphogenesis.

Philosophical Transactions of the Royal Society of London B: Biological Sciences, 237(641):37–72.