
Une méthode d'identification de causalités dans des données spatio-temporelles

Juste Raimbault^{1,2}

1. UMR CNRS 8504 Géographie-cités

2. UMR-T 9403 IFSTTAR LVMT

juste.raimbault@polytechnique.edu

RÉSUMÉ. Résumé.

ABSTRACT. Abstract in English

MOTS-CLÉS : Quelques mots clés

KEYWORDS: En anglais

1. Introduction

L'étude des processus spatio-temporels fortement couplés implique la prise en compte d'intrications entre ceux-ci généralement difficiles à isoler. Essence même des approches par la complexité, ces interactions qui sont à l'origine du comportement émergent d'un système font sens comme objet d'étude en lui-même, et une séparation des processus paraît alors contradictoire avec une vision intégrée du système. Dans le cas des systèmes territoriaux, l'exemple des interactions entre réseaux de transport et territoires est une excellente allégorie de ce phénomène : des méthodes isolant les "effets structurants" d'une infrastructure développées dans les années 70 (Bonnafous, Plassard, 1974) se sont révélées par la suite de l'instrumentation politique et sans fondement empirique (Offner, 1993). Le débat est toujours d'actualité puisque la question se pose toujours par exemple pour la construction de lignes à grande vitesse (Crozet, Dumont, 2011). La réalité des processus territoriaux est en fait bien plus compliqué qu'une simple relation causale entre la mise en place d'une infrastructure et les retombées sur le développement local, mais correspond bien d'une *co-évolution* complexe (Bretagnolle, 2009). Sur le temps long et à grande échelle, certains effets de renforcement des dynamiques dans les systèmes de villes par l'insertion dans les réseaux, ont été mis en valeur par l'application de la Théorie Evolutive des Villes (L'Espace géographique, 2014), montrant que le démêlage est toutefois possible dans certains cas par une compréhension globale du système.

A d'autres échelles,

Les régimes sous lesquels des identifications de causalité sont cohérentes ne sont pas non plus identifiés de manière évidente. La définition de la causalité

(Liu *et al.*, 2011) propose la détection de relations spatio-temporelles entre perturbations des flots de trafic, introduisant une définition particulière de la causalité basé sur une correspondance de points extrêmes. Les algorithmes associés sont toutefois spécifiques et difficilement applicables à des types de systèmes différents.

Les neurosciences ont développé de nombreuses méthodes répondant à des problématiques similaires. (Luo *et al.*, 2013) définit une causalité de Granger généralisée prenant en compte la non-stationnarité et s'appliquant à des régions abstraites issues d'imagerie fonctionnelle.

2. Méthode

Nous décrivons ici une méthode générique, basée sur un test similaire à la causalité de Granger (), pour tenter d'identifier des relations causales dans des systèmes spatiaux. Soit $X_j(\vec{x}, t)$ des processus aléatoires spatiaux unidimensionnels. Une réalisation d'un sous-système territorial est donnée par des ensembles de trajectoires pour chaque processus $x_{i,j,t}$. On suppose l'existence de fonctions

de correspondance Φ_{j_1,j_2} permettant de faire correspondre les réalisations de chaque composantes à un index unique (dans le cas le plus simple, on associera les variables sur les mêmes patches). Si $\text{argmax}_{\tau} \hat{\rho}[x_{j_1}, x_{j_2}]$ est clairement défini, son signe donnera alors le sens de la causalité entre les composantes j_1 et j_2 .

3. Résultats

3.1. Données Synthétiques

Cette méthode doit dans un premier temps être testée et partiellement validée, ce que nous proposons de faire sur des données synthétiques. (Raimbault *et al.*, 2014) est un modèle simple de morphogenèse urbaine (modèle RBD) faisant un candidat intéressant pour notre test. En effet, les variables explicatives de la croissance urbaine, les processus d'extension du réseau et le couplage entre densité urbaine et réseau sont assez élémentaires. Cependant, hormis dans des cas extrêmes (distance au centre détermine valeur foncière uniquement, le réseau dépendra de manière causale de la densité, ou distance au réseau seule, la causalité devrait être inversée), les régimes mixtes n'exhibent pas de causalités évidentes : c'est donc un parfait cas pour tester si la méthode est capable d'en détecter.

Nous utilisons une implémentation adaptée¹ du modèle initial, permettant de capturer les valeurs des variables étudiées pour chaque patch et à chaque pas de temps.

Nous explorons une grille de l'espace des paramètres du modèle RDB, faisant varier les paramètres de poids de la densité, de la distance au centre et de la distance au réseau, que l'on note respectivement (w_d, w_c, w_r) , dans $[0; 1]$ avec un pas de 0.1. Les autres paramètres sont fixés à leur valeurs par défaut données dans l'article. Pour chaque valeur des paramètres, nous procédons à $N = 100$ répétitions ce qui est suffisant pour une bonne convergence des indicateurs. Les explorations sont effectuées via le logiciel OpenMole (Reuillon *et al.*, 2013).

Nous calculons sur l'ensemble des patches les corrélations retardées entre les variables suivantes : densité locale, distance au centre et distance au réseau. Des regroupements par régimes peuvent se faire, lorsqu'un paramètre domine et impose le régime

3.2. Cas d'étude

3.2.0.1. Contexte

(Damm *et al.*, 1980)

1. disponible sur le dépôt ouvert du projet à

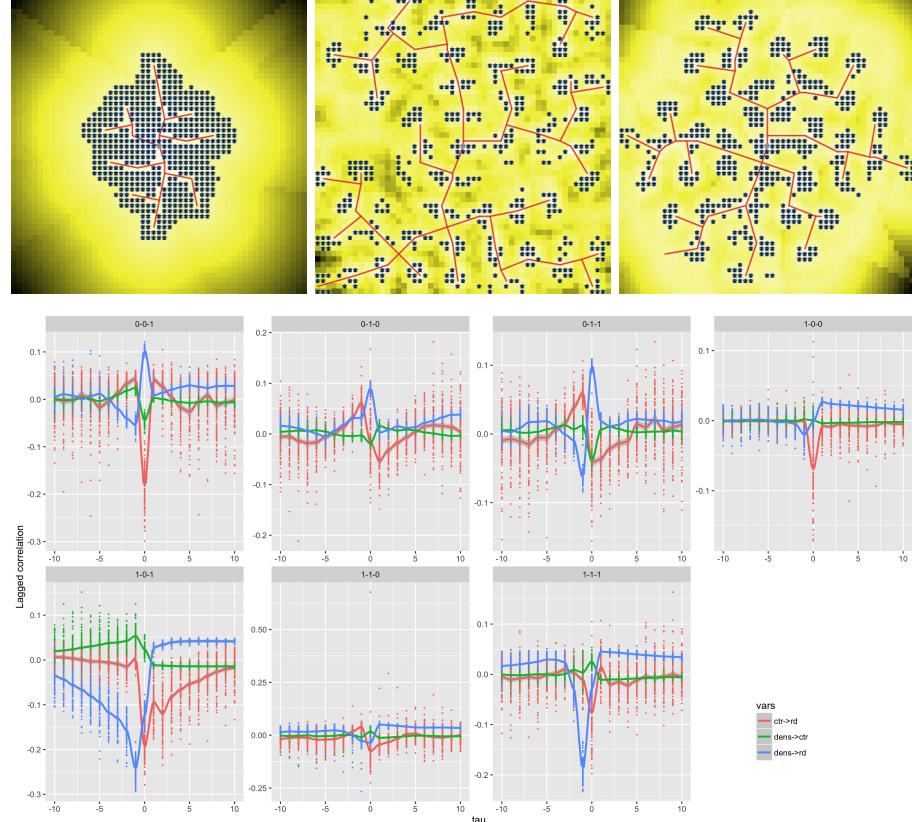


FIGURE 1. Correlations dans le modèle RDB (Première ligne) Exemples de configurations finales, obtenues avec (w_d, w_c, w_r) valant respectivement $(0, 1, 1), (1, 0, 1)$, et $(1, 1, 1)$. (Deuxième ligne) Corrélation retardées, pour chaque combinaison des paramètres, en fonction du retard τ . Les différentes couleurs correspondent

(STIF, 2010) projet arc express

(Beaucire, Drevelle, 2013) étude équilibrage est-ouest

(SDRIF, 2013) : SDRIF 2013

(Desjardins, 2010) bataille institutionnelle état-région

3.2.0.2. Données

Le nombre de transactions après nettoyage est de 632746, se répartissant sur l'ensemble des IRIS.

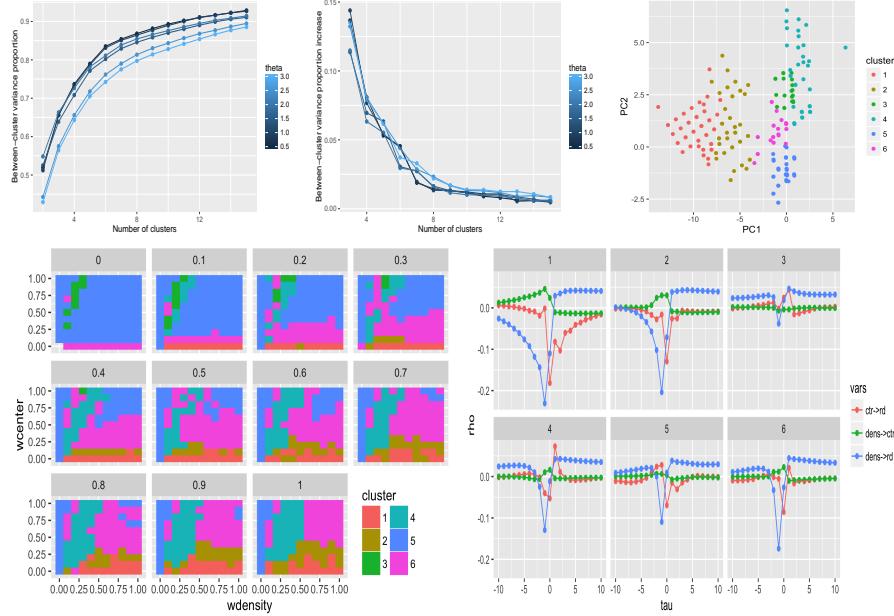


FIGURE 2. Identification de régimes d'interactions (Haut Gauche)
 Variance inter-cluster comme fonction du nombre de clusters. (Haut Milieu)
 Dérivée de la variance inter-cluster. (Haut Droite) Features dans un plan
 principal. (Bas Gauche) Diagramme de phase des régimes dans l'espace
 (w_d, w_c, w_r) . (Bas Droite) Trajectoires correspondantes des centroides.

4. Discussion

5. Conclusion

Bibliographie

- Beaucire F., Drevelle M. (2013). «grand paris express»: un projet au service de la réduction des inégalités d'accessibilité entre l'ouest et l'est de la région urbaine de paris? *Revue d'Économie Régionale & Urbaine*, n° 3, p. 437–460.
- Bonafous A., Plassard F. (1974). Les méthodologies usuelles de l'étude des effets structurants de l'offre de transport. *Revue économique*, p. 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. Consulté sur <http://tel.archives-ouvertes.fr/tel-00459720>
- Crozet Y., Dumont F. (2011). Retour sur les effets économiques du TGV. Les effets structurants sont un mythe (interview). *Ville, Rail et Transports*, n° 525, p. 48-51.

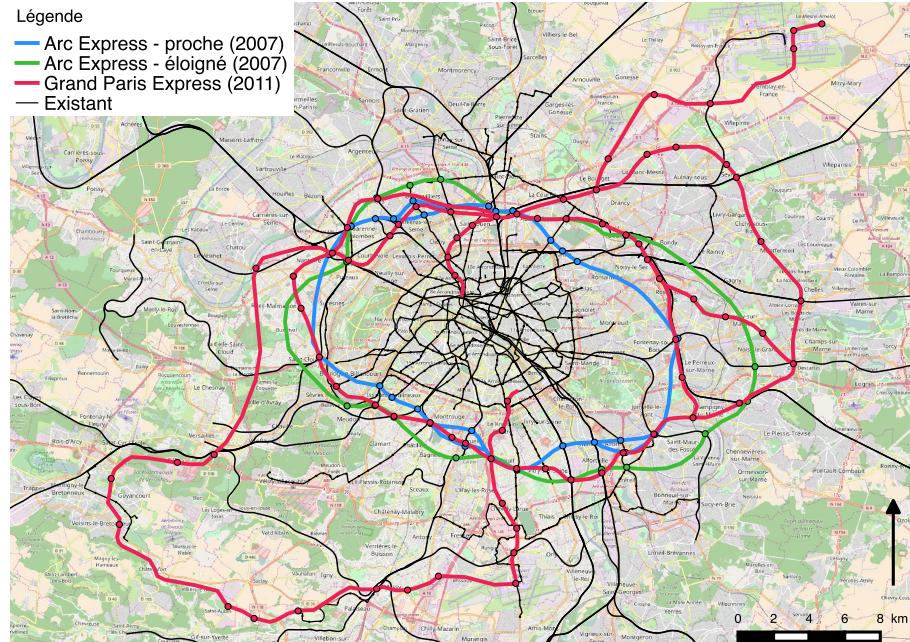


FIGURE 3. Projets de transport de la métropole du Grand Paris

Consulté sur <https://halshs.archives-ouvertes.fr/halshs-01094554>

- Damm D., Lerman S. R., Lerner-Lam E., Young J. (1980). Response of urban real estate values in anticipation of the washington metro. *Journal of Transport Economics and Policy*, p. 315–336.
- Desjardins X. (2010). la bataille du grand paris. *L'Information géographique*, vol. 74, n° 4, p. 29–46.
- L'Espace géographique. (2014). *Les effets structurants des infrastructures de transport, l'espace géographique 2014/1 (tome 43)*, p. 51-67.
- Liu W., Zheng Y., Chawla S., Yuan J., 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*, p. 1010–1018.
- Luo Q., Lu W., Cheng W., Valdes-Sosa P. A., Wen X., Ding M. et al. (2013). Spatio-temporal granger causality: A new framework. *NeuroImage*, vol. 79, p. 241–263.
- Offner J.-M. (1993). Les "effets structurants" du transport: mythe politique, mystification scientifique. *Espace géographique*, vol. 22, n° 3, p. 233–242.
- Raimbault J., Banos A., 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, uni-

versité de normandie, le havre, france; m. a. aziz-alaoui, c. bertelle, x. z. liu, d. olivier, eds.: pp. 51-60.

Reuillon R., Leclaire M., Rey-Coyrehourcq S. (2013). Openmole, a workflow engine specifically tailored for the distributed exploration of simulation models. *Future Generation Computer Systems*, vol. 29, n° 8, p. 1981–1990.

SDRIF. (2013). *Île-de-france 2030. orientations réglementaires et carte de destination générale des différentes parties du territoire.*

STIF. (2010). *Arcexpress, débat public sur le métro de rocade. dossier du maître d'ouvrage.*

archived at http://archive.wikiwix.com/cache/?url=http%3A%2F%2Fwww.debatpublic-arcexpress.org%2F_script%2Fntsp-document-file_download.php%3Fdocument_id%3D92%26document_file_id%3D106.