

---

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

Juste Raimbault <sup>1,2</sup>

*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 des relations entre processus complexes à isoler. 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 detection 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  $\Phi_{j_1,j_2}$  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  $\operatorname{argmax}_\tau \hat{\rho}[x_{j_1}, x_{j_2}]$  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 adaptée<sup>1</sup> 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

## 4. Discussion

## 5. Conclusion

## Bibliographie

- 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.
- 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, 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., 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.

---

1. disponible sur le dépôt ouvert du projet à

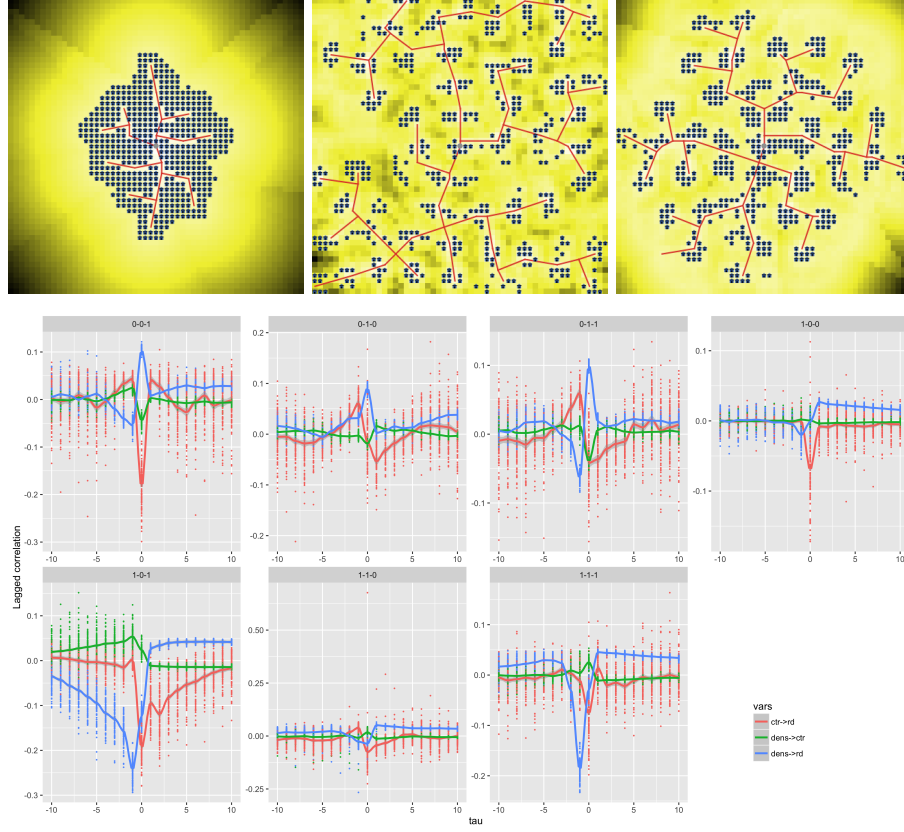


FIGURE 1. **Correlations dans le modèle RDB** (Première ligne) Exemples de configurations finales, obtenues avec  $(w_d, w_c, w_r)$  valant respectivement  $(0, 1, 1)$ ,  $(1, 0, 1)$ , et  $(1, 1, 1)$ . (Deuxième ligne) Corrélations retardées, pour chaque combinaison des paramètres, en fonction du retard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