

Analyse statistique des données décrivant la dynamique et l'usage du système Vélo'V à Lyon

Pierre BORGNAT & Céline Robardet

CNRS – ENS Lyon, Laboratoire de Physique (UMR 5672), Université de Lyon
IXXI (Institut des Systèmes Complexes de Lyon)

9 mars 2010



Généralités sur Vélo'v

Problématiques de cet exposé

- quel usage des vélos ? heures, lieux, régimes de pointe ?
- évolution du système ? régulation possible ?
- ⇒ analyse quantitative des trajets et modèle prédictif
- visualisation spatiale et temporelle de l'activité
- ⇒ relier l'usage de Vélo'v aux populations ?

Les données disponibles

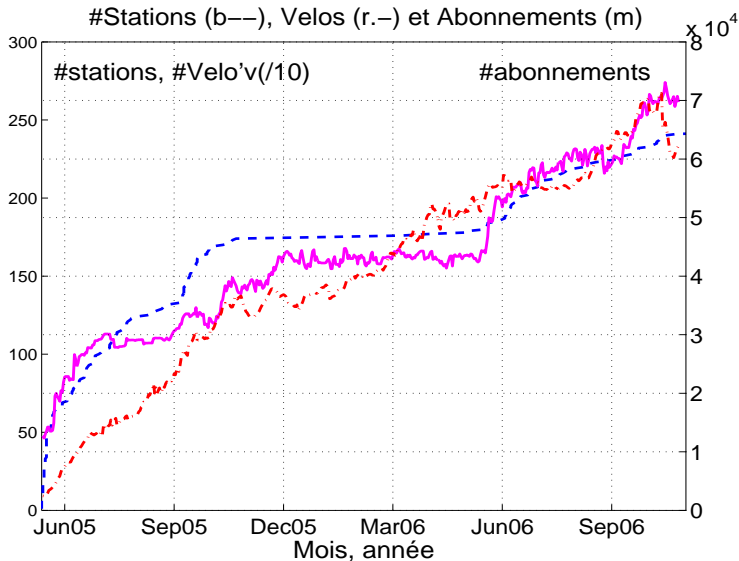
Tous les trajets du 25/05/2005 au 12/12/2007

- (13 millions d'événements)
 - heures (début, fin), durée
 - localisations (début, fin), longueur trajet
- informations anonymisées (pas d'information passager)
- nombre d'abonnés, de stations, de vélos
- données météo (température, pluie), calendrier
- taux de disponibilités en station

Comparaison à l'état de l'art

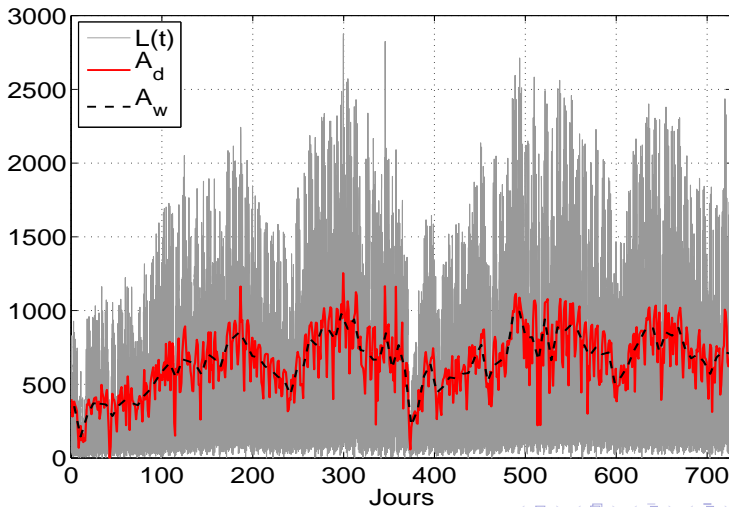
- enquête ménages ou terrain
 - échantillonnage ponctuels, jours ouvrés banalisés,
 - peu de détails sur 2 roues, puis vélos,...
 - définition du trajet ?
- Vélo'v : unité de vision (un trajet) ; exhaustif ; long terme

Déploiement progressif du système



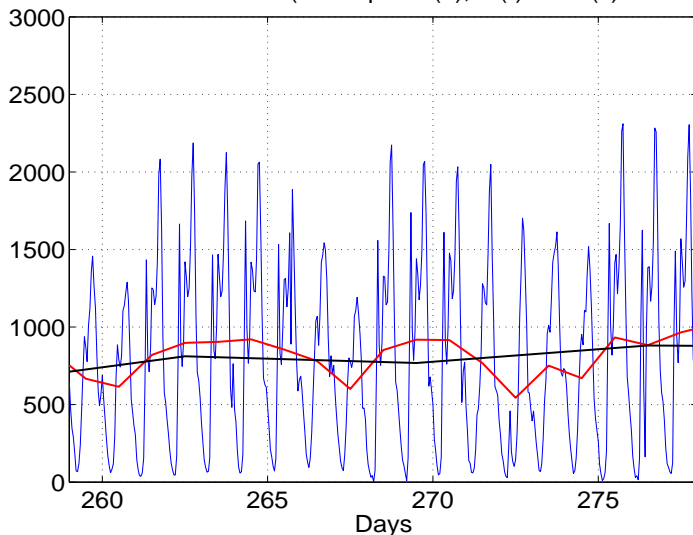
Nombre de locations par heure, globalisées sur Lyon

- agrégation sur $\Delta = 1\text{h}$ (durée médiane des locations = 11')
- $t = 0$ au 17/12/2005, 2 ans de données étudiées

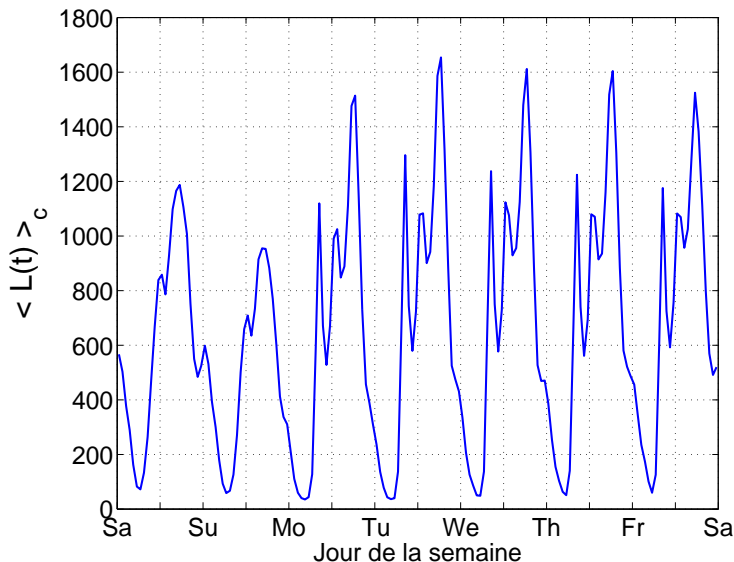


Zoom sur 3 semaines

#locations/h (Mean per h (b), d (r) week(k))

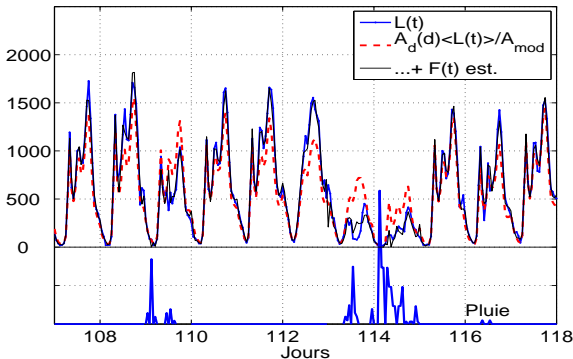


Moyenne cyclique des locations



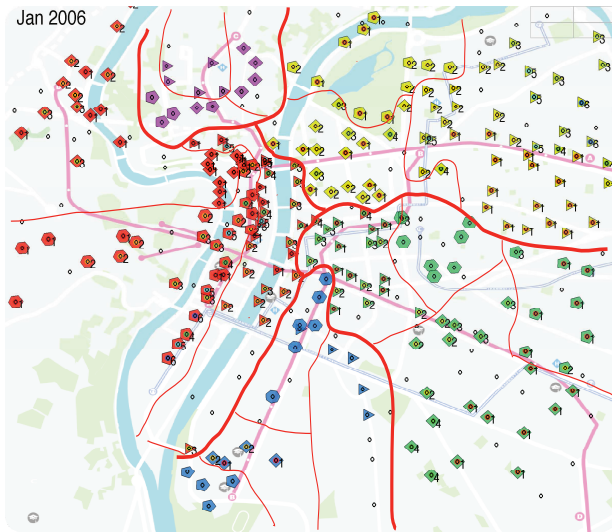
Modélisation statistique du nombre de locations

- # locations =
amplitude \times cycle hebdomadaire + fluctuations horaires
- amplitude (usage du jour) : régression linéaire (abonnés, vélos, température, pluie, vacances)



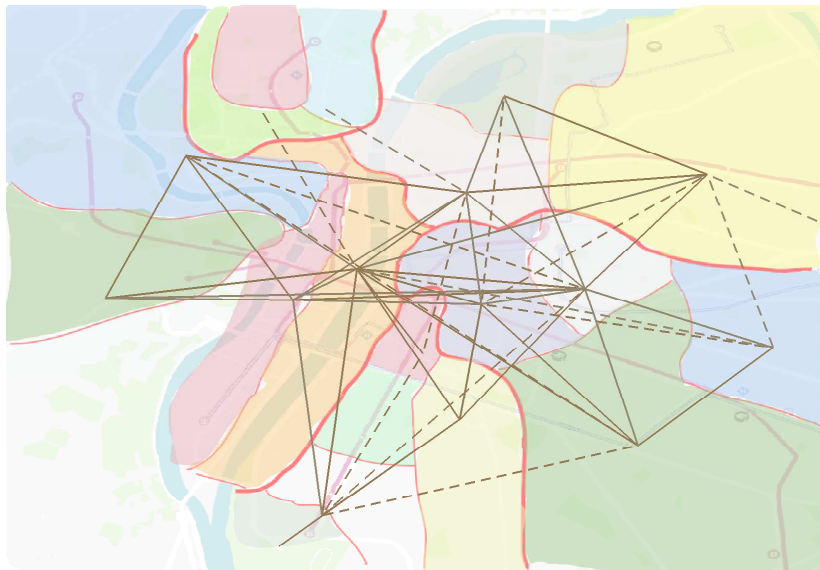
- fluctuations : modèle avec corrélation sur 1h et correction avec la pluie pendant l'heure (modèle ARX(1))

Groupes de stations échangeant les plus entre elles

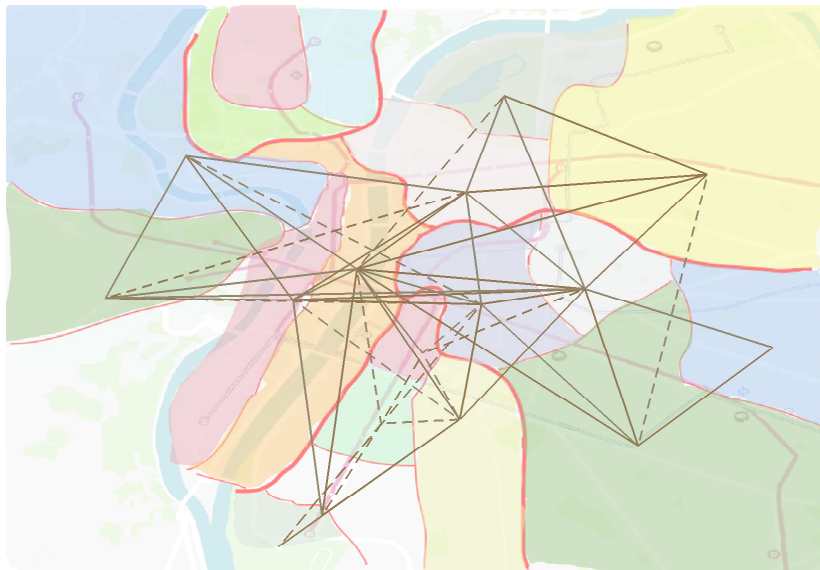


- Principe : détection automatique de *communautés hiérarchiques*

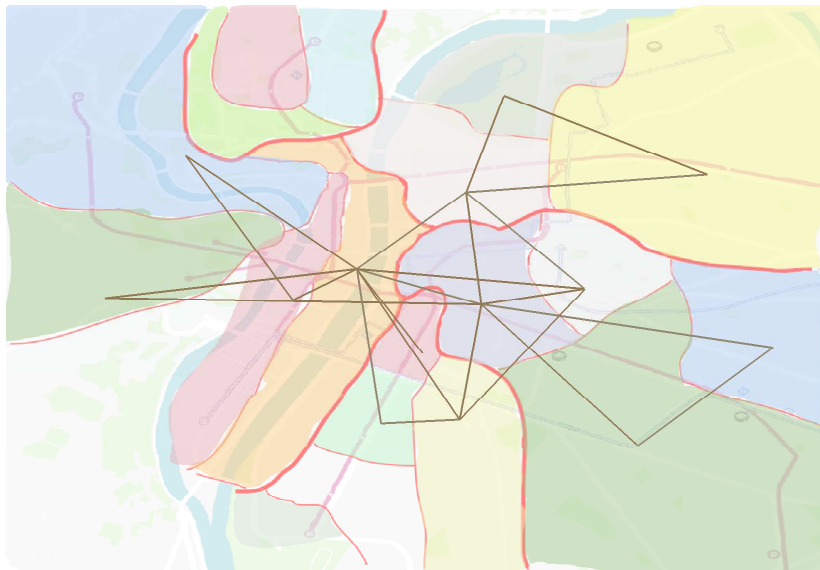
Groupes des flots actifs le matin



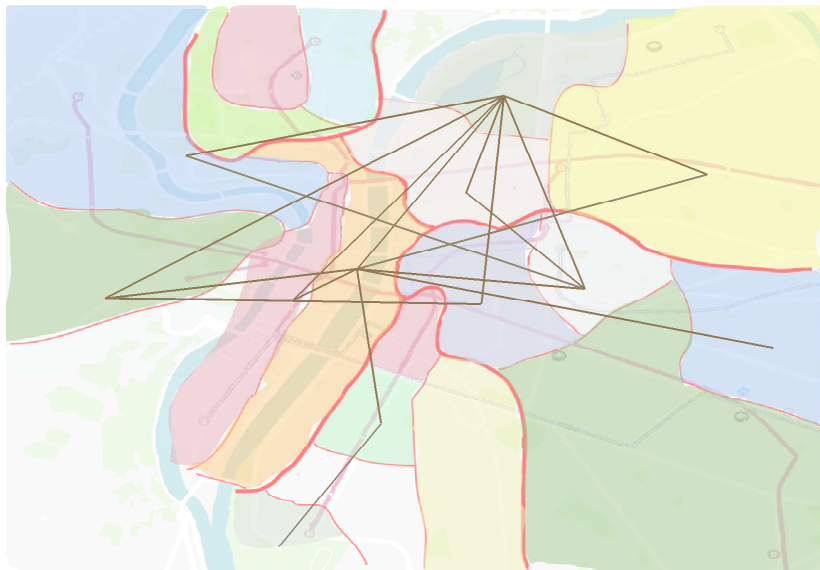
Groupes des flots actifs le soir



Groupes des flots actifs le midi



Groupes des flots actifs le week-end



Projet : relier usages des Vélo'v et populations ?

- Données supplémentaires :
 - recensements INSEE 1999-2005 et maintenant 2010,
 - découpage spatial en ÎLOTS ou IRIS,
 - données socio-économiques :
 - démographie (classe d'âge, pop. active, ménages, enfants, ...)
 - économie (revenus, emplois, commerces, infrastructures, ...)
 - sociologie (usage des transports, autres équipements, ...)
- Mise en relation : flux de Vélo'v / données socio-éco.
 - à travers la structure de corrélation
 - prévisions statistiques sous contraintes de parcimonie

Conclusion

- Enregistrements des locations Vélo'v : une base de données incomparable pour l'analyse des déplacements urbains en vélo
- Facteurs dominants l'emploi des Vélo'v :
 - **en temps** : gabarit cyclique, avec 3 pics journaliers
 - météo domine le comportement (!),
réduction de l'utilisation pendant les vacances
 - **en espace** : fort lien avec zone d'activités (commerces, universités) + population étudiante
- Perspectives : étude exhaustive des répartitions spatiales et dans le temps

<http://perso.ens-lyon.fr/pierre.borgnat>