

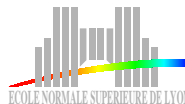
# Première approche :

## Analyse du nombre de locations de Vélo'v

Pierre BORGNAT

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

23 octobre 2009



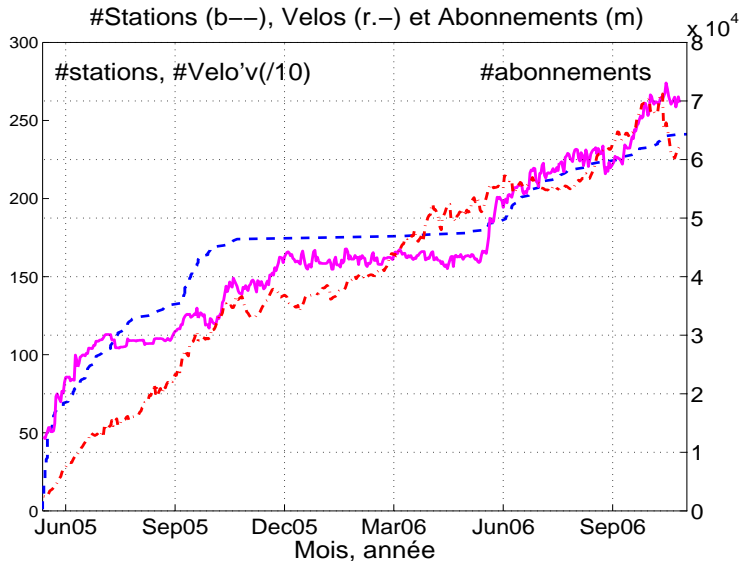
# Constation et objectif

- Évolution du système : nombre de stations, de vélo'v, d'usagers augmentent
- Rythmes de la semaine, des saisons
- Irrégularités : jours fériés, grèves,...
- **Objectif** : prédire la charge du système
- ⇒ analyse du nombre de trajets en vélo'v
- ⇒ modèle (statistique) des locations au fil du temps
- ⇒ méthode de prédiction
- Données globales, trajets vélo'v, météo

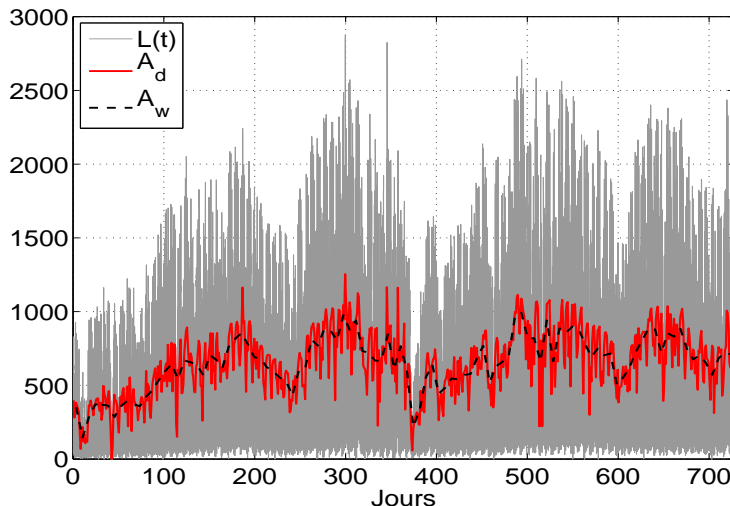
# Constation et objectif

- Évolution du système : nombre de stations, de vélo'v, d'usagers augmentent
- Rythmes de la semaine, des saisons
- Irrégularités : jours fériés, grèves,...
- **Objectif** : prédire la charge du système
- ⇒ analyse du nombre de trajets en vélo'v
- ⇒ modèle (statistique) des locations au fil du temps
- ⇒ méthode de prédiction
- Données globales, trajets vélo'v, météo

# Déploiement progressif du système



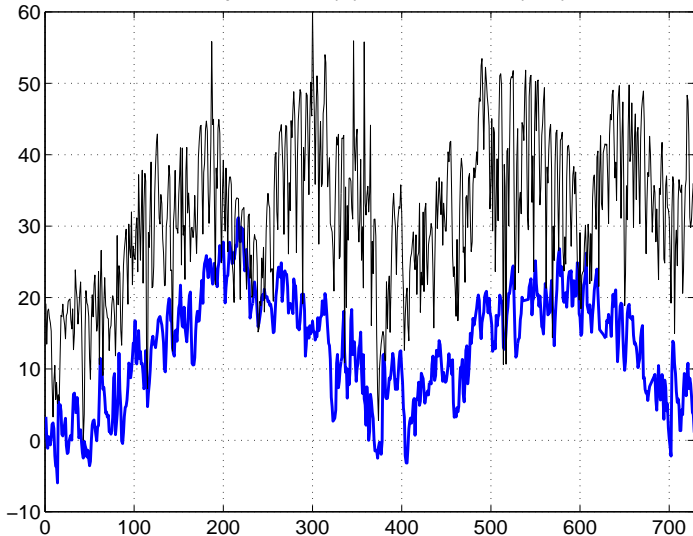
# Locations par heure, globalisées sur Lyon



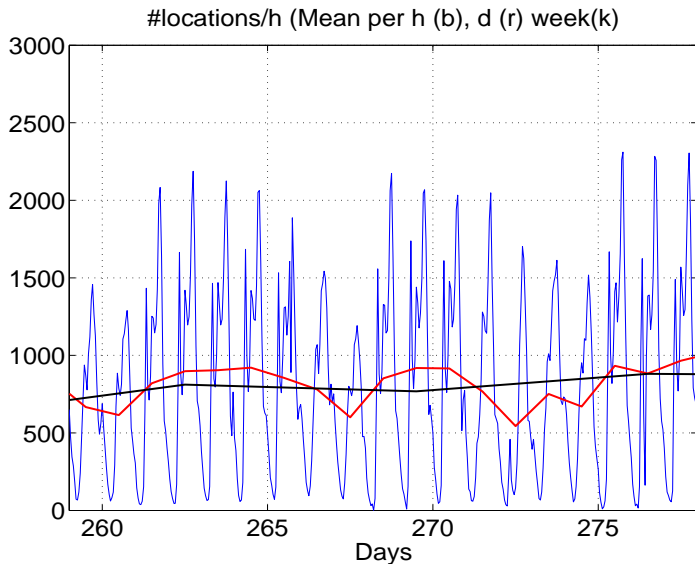
- agrégation sur  $\Delta = 1h$  (durée médiane des locations = 11')
- $t = 0$  au 17/12/2005, soit 2 ans de données

# Comparaison avec la température

Température (b) et #locations(/50)



## Zoom sur 3 semaine



# Étapes de l'analyse

- Utilisation totale sur chaque journée  $A_d(d)$
  - Moyenne répétée sur la semaine  $\langle L(t) \rangle_c$
  - Modèle pour un jour donné  $\sim A_d(d) \times \langle L(t) \rangle_c$
  - **Questions :** Prédiction de l'utilisation à la journée?
- Écriture mathématique :

$$A_d(d) = \frac{1}{24} \sum_{t \in (d)} L(t) \quad (1)$$

$$\langle L(t) \rangle_c = \frac{1}{N_w} \sum_{d=0}^{N_w-1} L(t + 168 d) \quad (2)$$



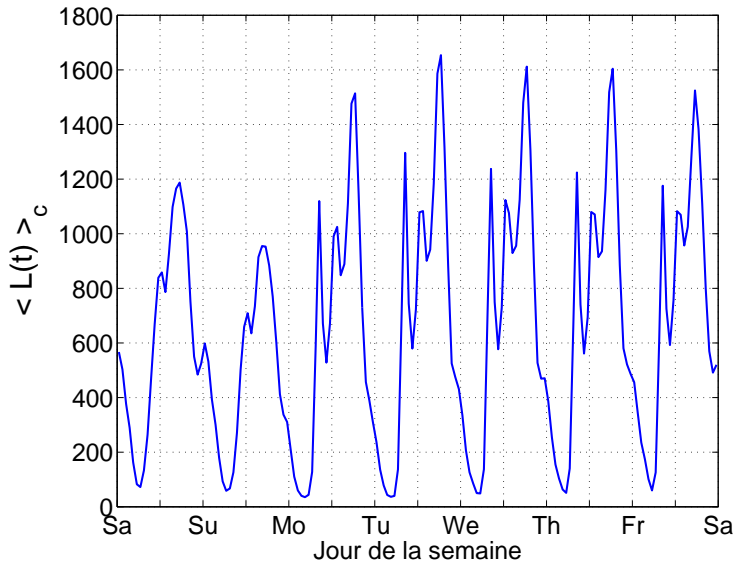
# Étapes de l'analyse

- Utilisation totale sur chaque journée  $A_d(d)$
- Moyenne répétée sur la semaine  $\langle L(t) \rangle_c$
- Modèle pour un jour donné  $\sim A_d(d) \times \langle L(t) \rangle_c$
- **Questions :** Prédiction de l'utilisation à la journée?
- Écriture mathématique :

$$A_d(d) = \frac{1}{24} \sum_{t \in (d)} L(t) \quad (1)$$

$$\langle L(t) \rangle_c = \frac{1}{N_w} \sum_{d=0}^{N_w-1} L(t + 168 d) \quad (2)$$

## Moyenne des locations sur la semaine



# Prédiction de l'utilisation à la journée

- Modèle statistique linéaire avec **différents facteurs** :
- Moyenne selon le jour de la semaine  $\alpha_0(d_7)$
- Nombre d'abonnés  $N_s(d)$
- Nombre de vélo'v  $N_v(d)$
- Température (en °C)  $\delta T(d) = T(d) - \langle T(d) \rangle$
- Pluie  $R(d)$  (en mm)
- Vacances ; grèves

$$\widehat{A_d(d)} = \alpha_0(d_7) + \alpha_1 \delta N_s(d) + \alpha_2 \delta N_v(d) + \alpha_3 \delta T(d) + \alpha_4 R(d) + \alpha_5 J_h(d) + \alpha_6 J_s(d) \quad (3)$$

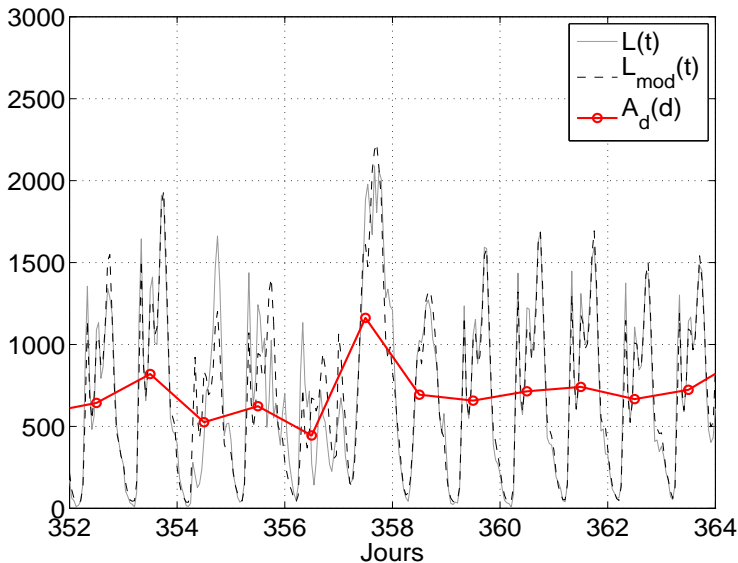
## Paramètres estimés

- Moyenne à la journée :  $A_0 = 17370 \pm 320$  vélo'v

Facteur	Abonnés	Vélo'v	Température	Pluie	Vacances	Grèves
	$\delta N_s(d)$	$\delta N_v(d)$	$\delta T(d)$	$R(d)$	$J_h(d)$	$J_s(d)$
Unité	abonnés	vélos	°C	mm		
Référ.	62 250	3 000	13.0	0.11		
std.	8 030	400	7.7	0.37		
coeff.	$\alpha_1$	$\alpha_2$	$\alpha_3$	$\alpha_4$	$\alpha_5$	$\alpha_6$
est.	1 860	-120	2270	-1280	-2900	20
IC <sub>-</sub>	1 210	-720	1980	-1520	-3700	-2900
IC <sub>+</sub>	2 560	+490	2560	-1030	-2100	+2900

- Sans ce modèle : erreur relative de 30%
- Avec modèle complet : erreur restante de 12%

# Comparaison modèle - données (locations horaires)

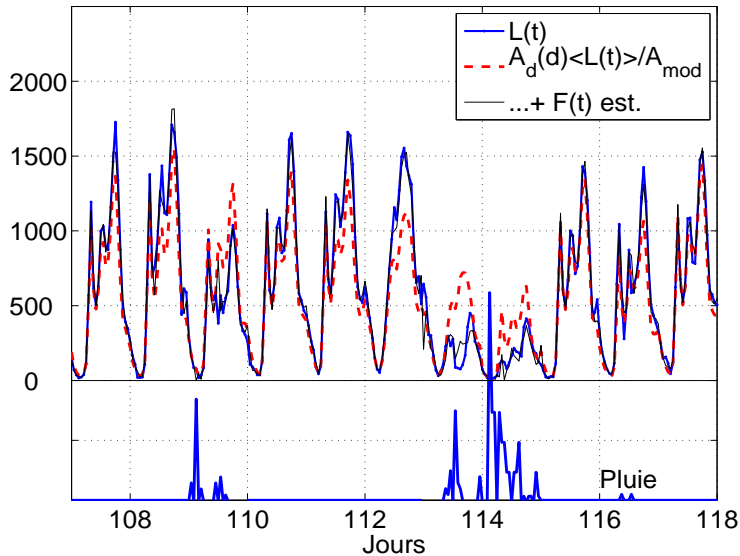


## Étude des locations une heure à l'avance

- **Objectif** : Prévoir le nombre d'utilisation de Vélo'v à l'heure prochaine en fonction de ce qui vient se passer dans l'heure.
- Introduire une **mémoire sur une heure** dans l'analyse statistique
- Ajouter des **corrections liées à la pluie**  
(effet potentiellement soudain)

$$\widehat{L(t+1)} = L_{\text{mod}}(t+1) + a_1 \left( L(t) - L_{\text{mod}}(t) \right) + \beta_1 \widehat{R}(t+1) \quad (4)$$

# Prévision des locations une heure à l'avance



# Conclusion de l'étude des locations horaires

- Outil pour prévoir l'utilisation globale des Vélo'vs :
  - → pour un jour donné, ou
  - → une heure à l'avance
- Facteurs dominants pour les locations Vélo'v :
  - répétition d'un même motif d'une semaine à l'autre, avec 3 pics journaliers
  - accroissement de la popularité du système
  - météo domine le comportement (!)
  - réduction de l'utilisation pendant les vacances

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