

Analyse de données de mobilité : Cas des systèmes de vélos en libre (VLS)

Etienne Côme, Latifa Oukhellou
IFSTTAR-COSYS, Laboratoire Grettia

11 Avril 2013



Plan

- 1 Introduction
 - Problématique
 - Le système Vélib' en quelques graphes
- 2 Regroupement de stations à partir de profils d'usage temporels
- 3 Clustering de matrices O-D dynamiques
- 4 Autres réalisations
 - Visualisation interactive
 - Recherche d'itinéraire
 - Projection de l'usage sur le réseau routier et cyclable
- 5 Travaux actuels & futurs

Vélos en Libre Service (VLS)

Nouveau mode doux de mobilité

- Vélib' : Système complexe de vélos partagés à grande échelle
- Données d'usage du système Vélib' :
 - ▶ Données d'occupation des stations
 - ▶ **Données des trajets** effectués sur 2 mois Sept. et Avril 2011 (Mairie de Paris et JC. Decaux)
 - ▶ Données **temporelles** : dates et heures de départ et d'arrivée
 - ▶ Données **spatiales** : stations de départ et d'arrivée
 - ▶ Type d'abonnement (longue ou courte durée)
- Nouvelles problématiques d'analyse des mobilités urbaines avec des challenges innovants

Objectifs

Objectifs "Opérationnels"

- Eviter le problème d'une station vide au départ et pleine au retour
⇒ Optimisation des politiques de re-déploiements des vélos
- Concevoir de nouveaux systèmes, ou étendre un système existant
⇒ Positionnement et dimensionnement des stations en fonction des données socio-économiques et géographiques, ...

Objectifs "Traitements"

- Rechercher des motifs d'usage récurrents du vélib', modèles spatio-temporels
- Construire des modèles de prédiction d'usage, estimation de la demande

Effets temporels

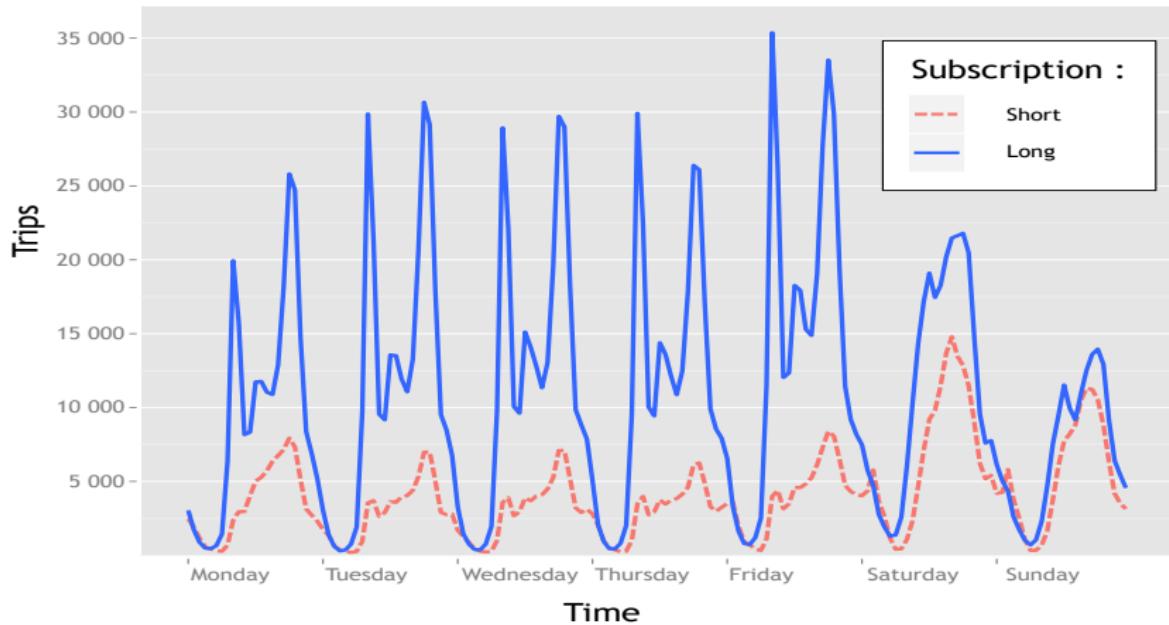


FIGURE 1: Nombre de trajets / heure (courte / longue abonnements)

Effets spatiaux



FIGURE 2: trajets entrants entre [6h,7h] en semaine

Effets spatiaux

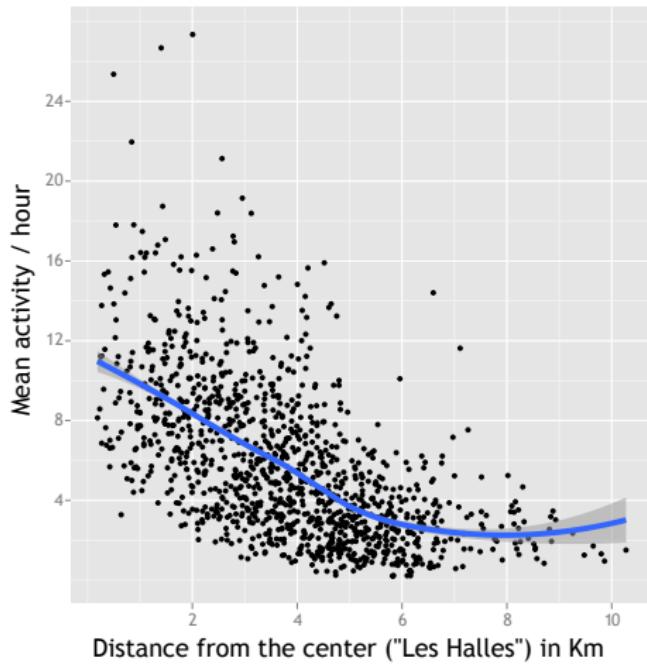


FIGURE 3: Activités stations / distance aux "Halles (centre de Paris)"

Analyse exploratoire des données

Méthodologie générale

- Utiliser des algorithmes de clustering (regroupement automatique) pour trouver des formes d'usage type du Vélib'
- Croiser les groupes ou clusters trouvés avec des données géographiques et socio-économiques de la ville
⇒ Facteurs influents sur l'usage du système VLS.

Deux pistes de travail :

- 1 Trouver des groupes de stations avec des profils d'usage similaires
- 2 Segmenter / Résumer la dynamique temporelle globale du système

Regroupement de stations à partir de profils d'usage temporels

Regroupement de stations à partir de leurs profils d'usage temporels

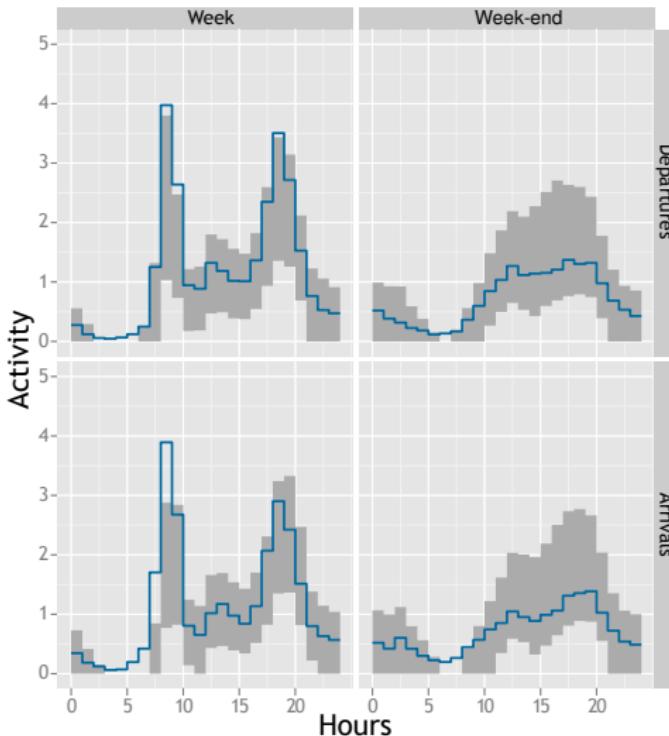
Objectifs :

- Stations décrites par les dynamiques de flux entrants et sortants
 - ▶ X_{sd}^{out} : # de vélos pris à la station s le jour d à l'heure t
 - ▶ X_{sd}^{in} : # de vélos déposés à la station s le jour d à l'heure t
- $\mathbf{X}_{sd} = (X_{sd1}^{in}, \dots, X_{sd24}^{in}, X_{sd1}^{out}, \dots, X_{sd24}^{out})$
- Prise en compte des jours de semaine / week end.
- Croiser les résultats avec d'autres variables explicatives : population, emplois, loisirs, ...
- Analyse réalisée avec 8 groupes
(bon compromis : interprétations/attache aux données)

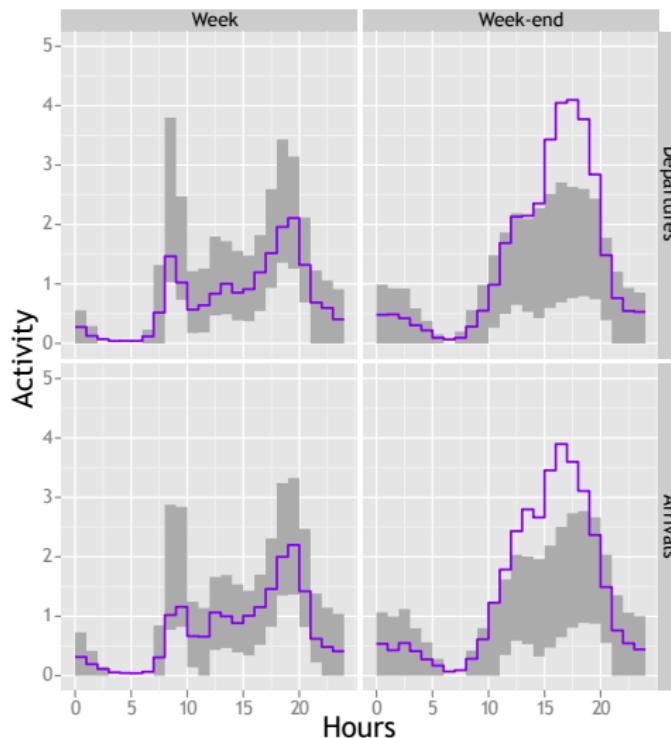
Pôles multimodaux



Pôles multimodaux



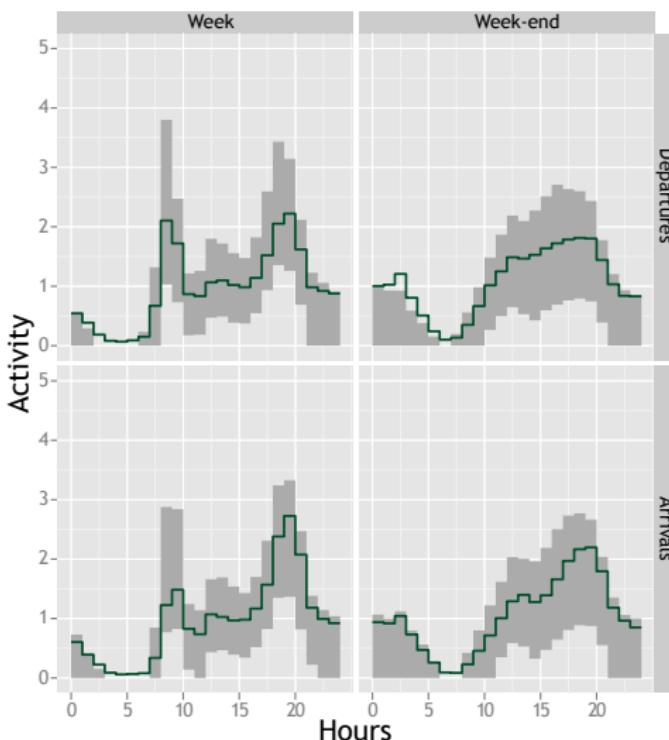
Parcs



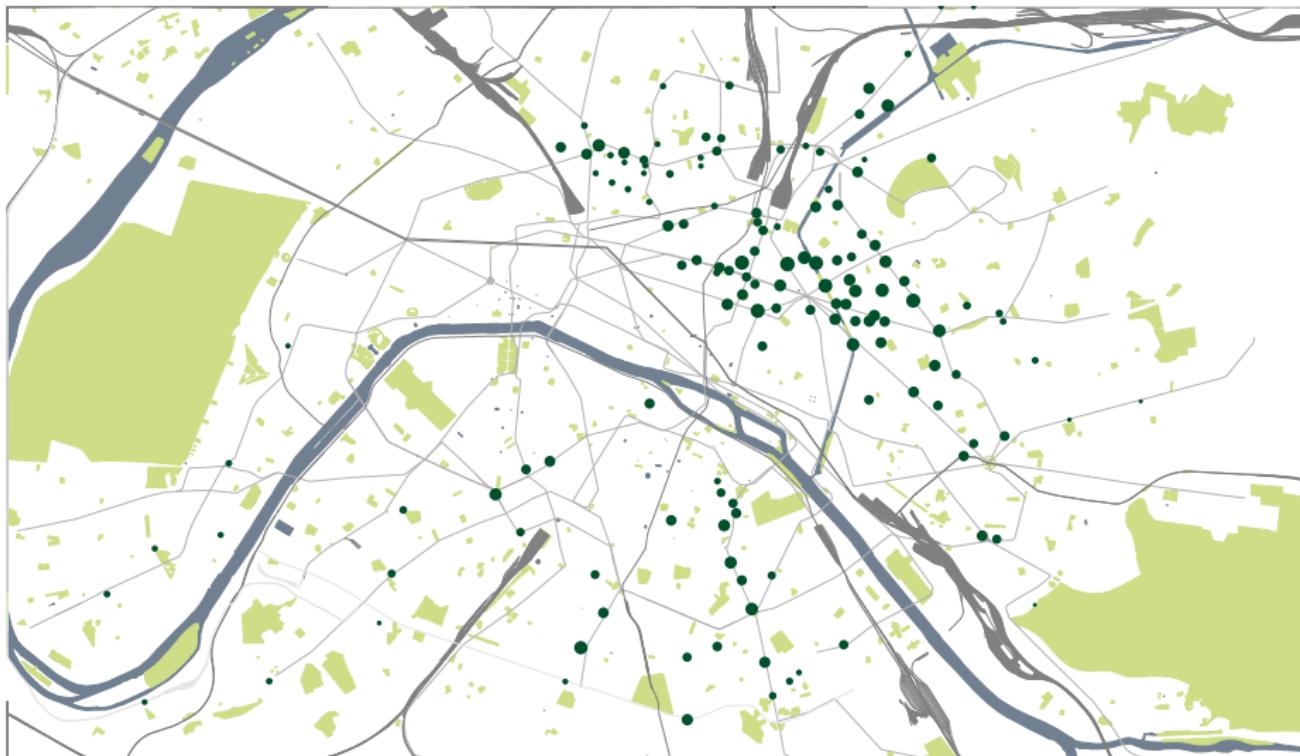
Parcs



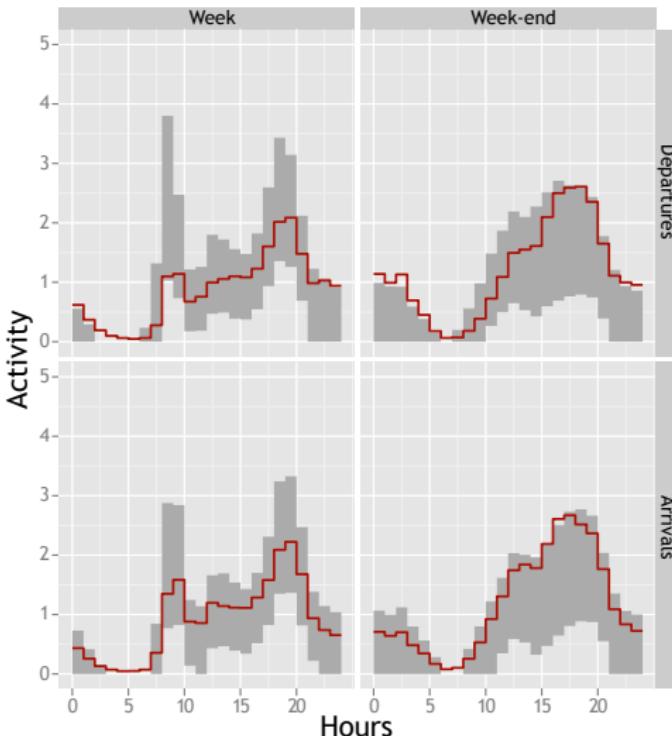
Sorties nocturnes



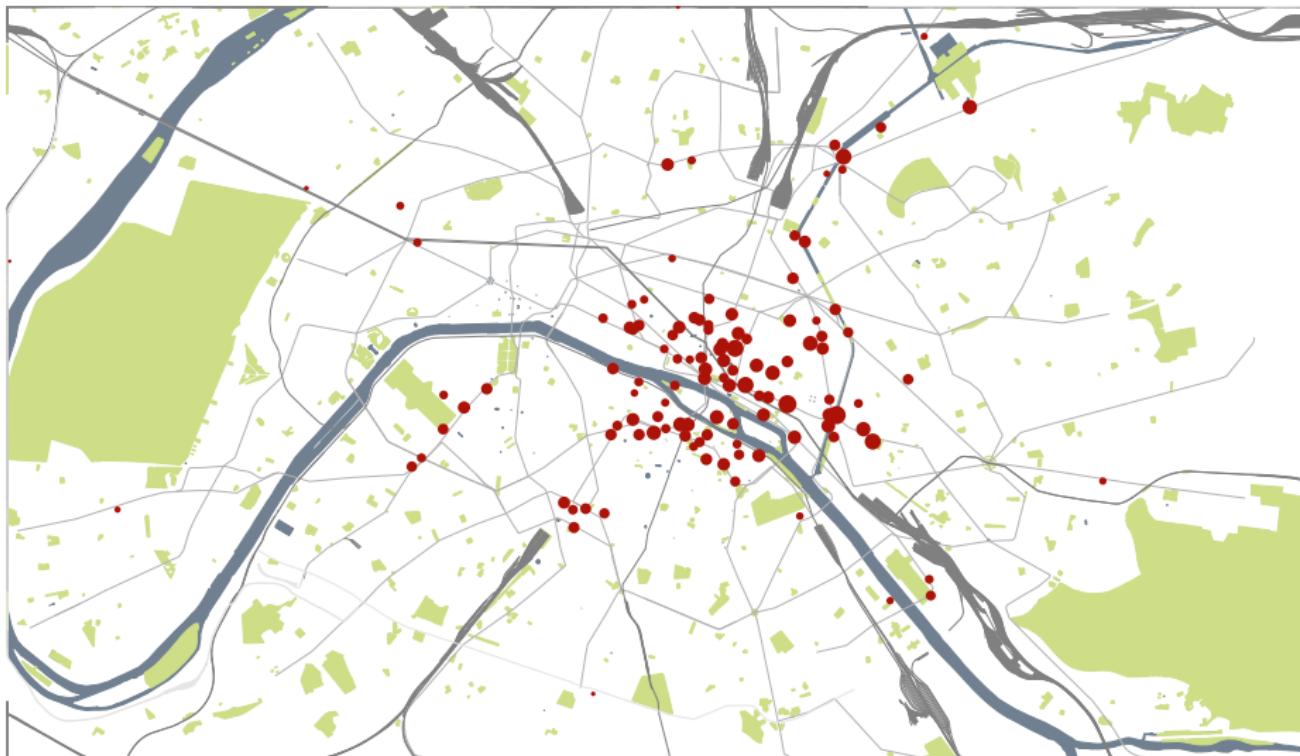
Sorties nocturnes



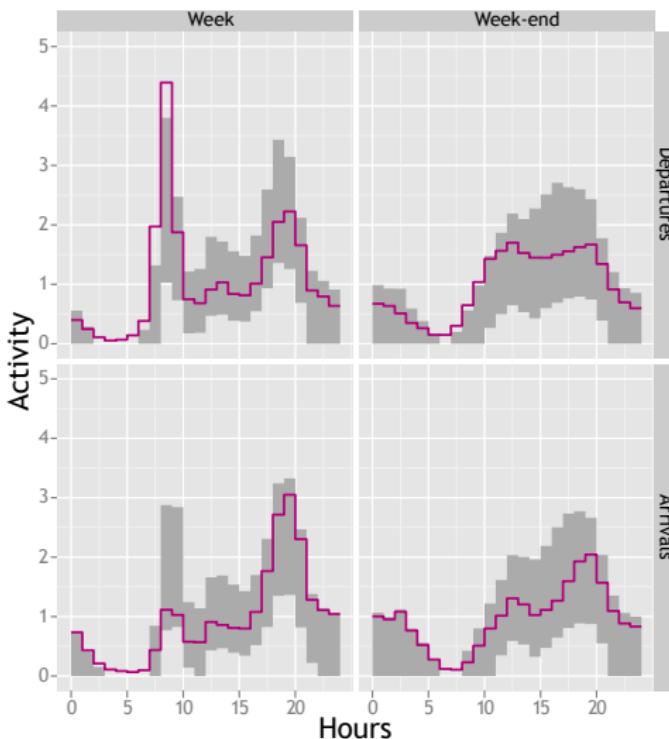
Sorties nocturnes et week-end



Sorties nocturnes et week-end

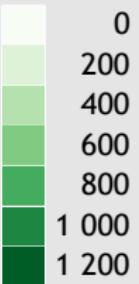


Logements

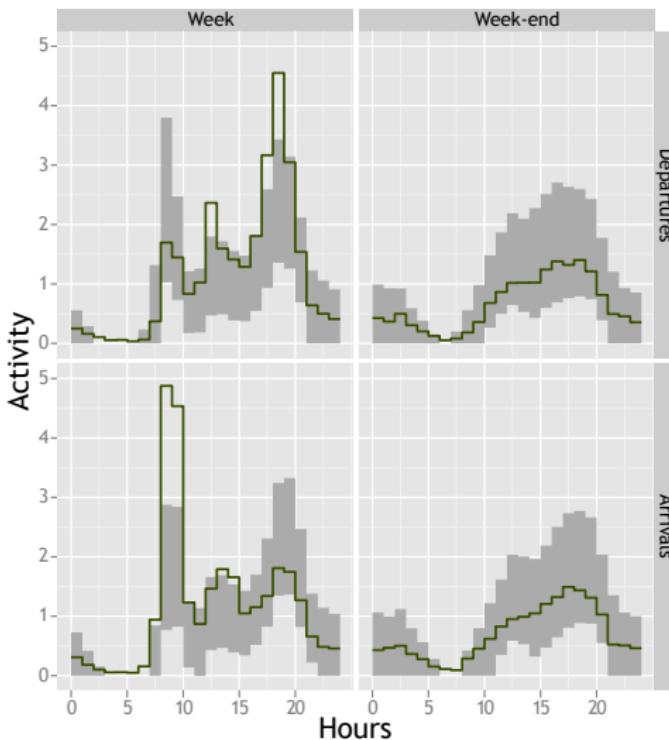


Logements

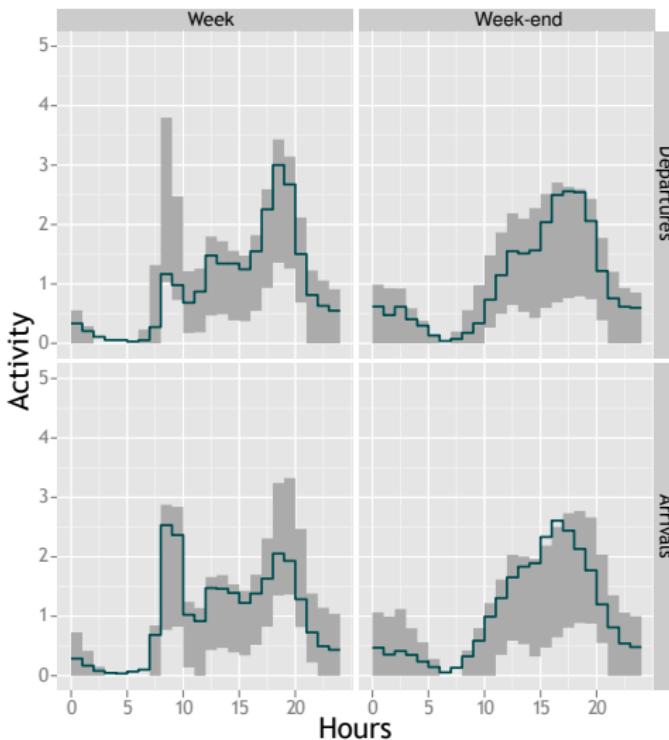
Inhabitants / ha



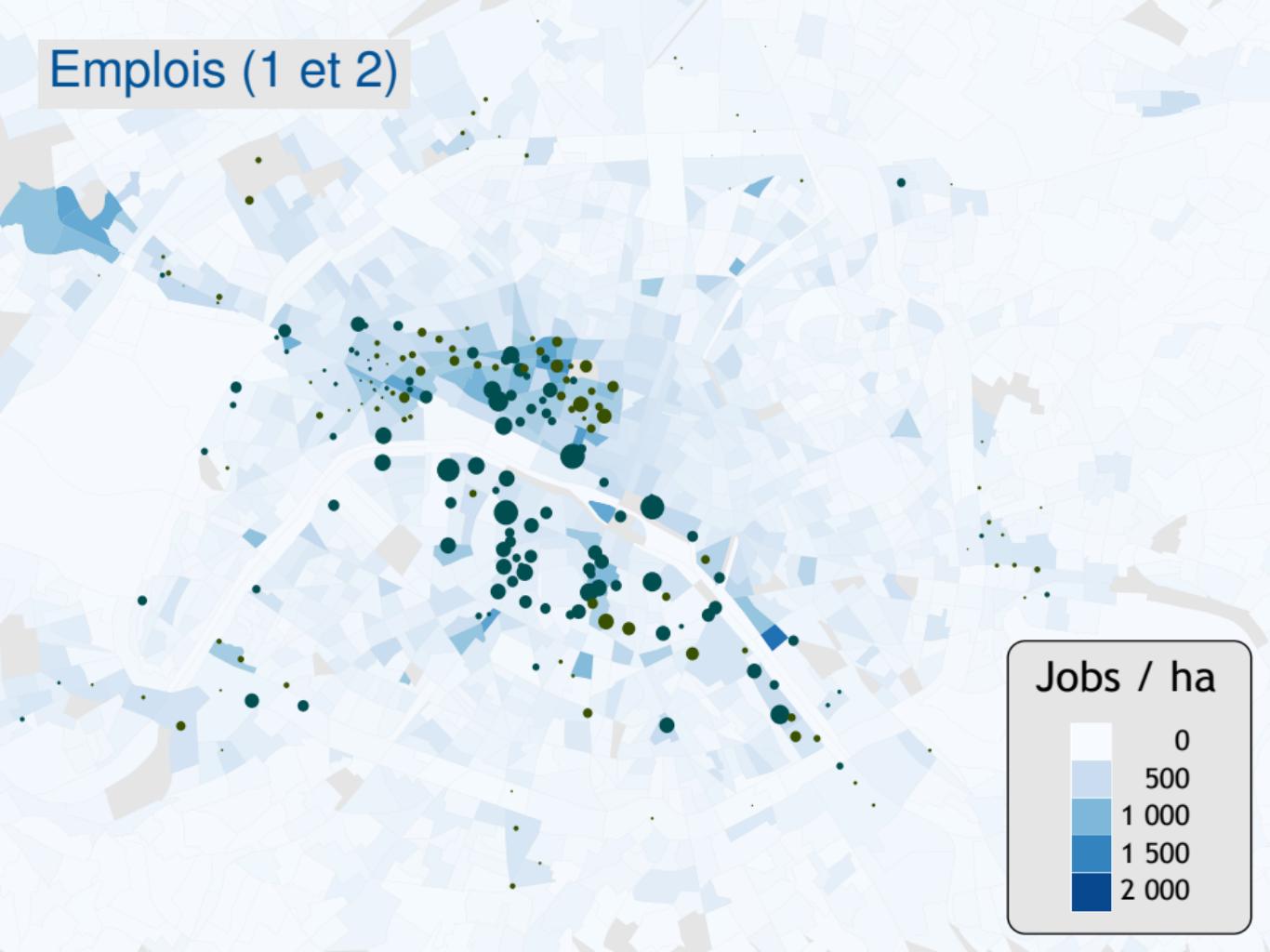
Emplois(1)



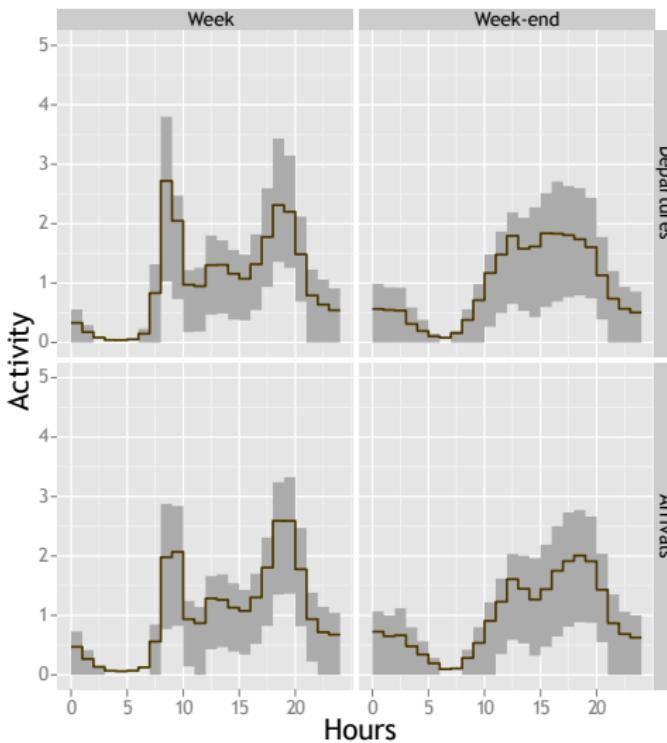
Emplois (2)



Emplois (1 et 2)



Usage mixte



Usage mixte



Croisement données population-emplois-services

	hab/ha	emp/ha	serv/ha	com/ha
	162	237	4.2	3.7
Sorties (1)	367	189	6.3	4.4
Sorties (2)	261	322	7.7	6.9
Parcs	172	90	2	1.7
Gares	209	206	2.4	1.8
Logements	375	108	3.8	2.7
Emplois(1)	138	409	4.5	2.8
Emplois(2)	157	456	5.7	5.6
Moyennes	301	163	3.8	2.8

TABLE 1: Comparaison des moyennes de densités de population, d'emplois, de services et de commerces pour les différentes groupes de stations.

Conclusions sur le clustering de stations

Résultats

- Des stations bien différenciées en termes d'usage
- Interprétation aisée des groupes de stations
- Densités de population, d'emplois et d'équipements explicatives des groupes de stations
- Profils temporels des clusters interprétables et informatifs

Clustering de matrices O-D dynamiques

Objectifs

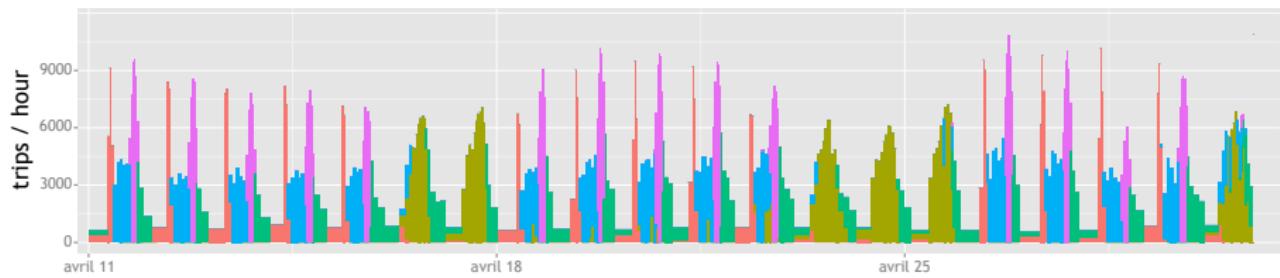
- Représentation des données : matrices OD dynamiques
⇒ Recherche de stationnarités et de points de changement dans les dynamiques de matrice OD
- Modèle utilisé “Latent Dirichlet Allocation”

Résultats

- Segmentation temporelle et cycles
- Distribution spatiale des flux (OD de référence) / segment
⇒ Permet de caractériser le déséquilibre du réseau / segment

Représentation synthétique de la dynamique du réseau

Interprétation temporelle



Remarques

- La cyclostationnarité est clairement visible
- Faible mélange entre les différentes activités latentes
- Interprétation temporelle des 5 groupes obtenus : Domicile ↔ Travail, Déjeuner, Travail ↔ Domicile, Loisirs nocturnes, Loisirs

Interprétation spatiale : déséquilibre du réseau



FIGURE 4: Activité latente "Domicile → Travail", déséquilibre du réseau pour $N_{dep} = 10\,000$

"Travail → Domicile"

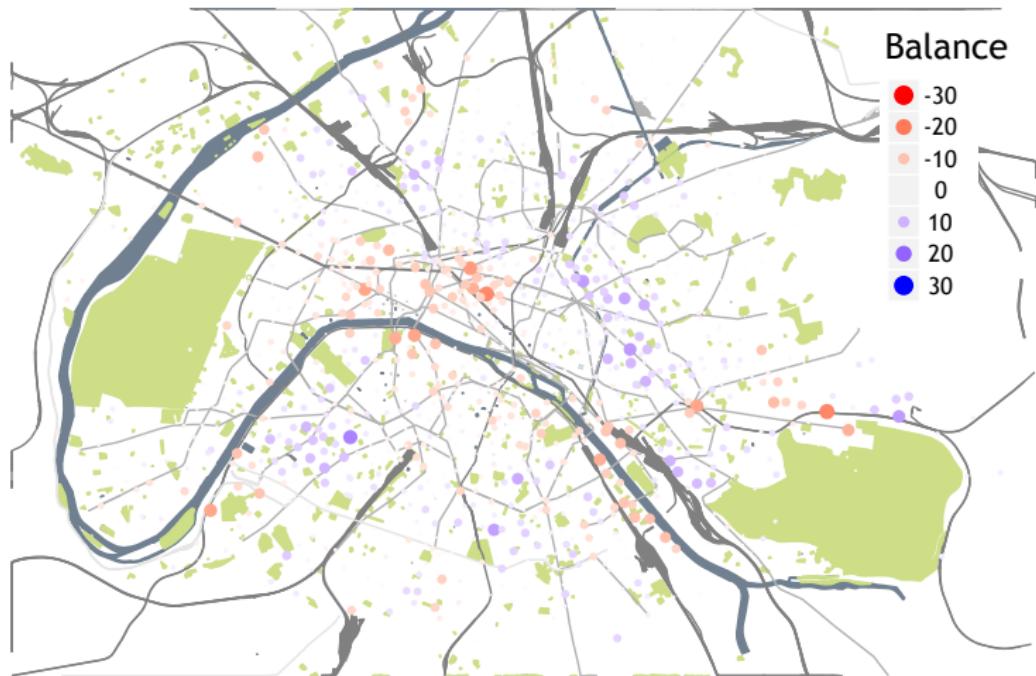


FIGURE 5: Activité latente "Travail → Domicile", déséquilibre du réseau pour $N_{dep} = 10\,000$

"Déjeuner"



FIGURE 6: Activité latente "Déjeuner", déséquilibre du réseau pour $N_{dep} = 10\,000$

"Loisirs nocturnes"

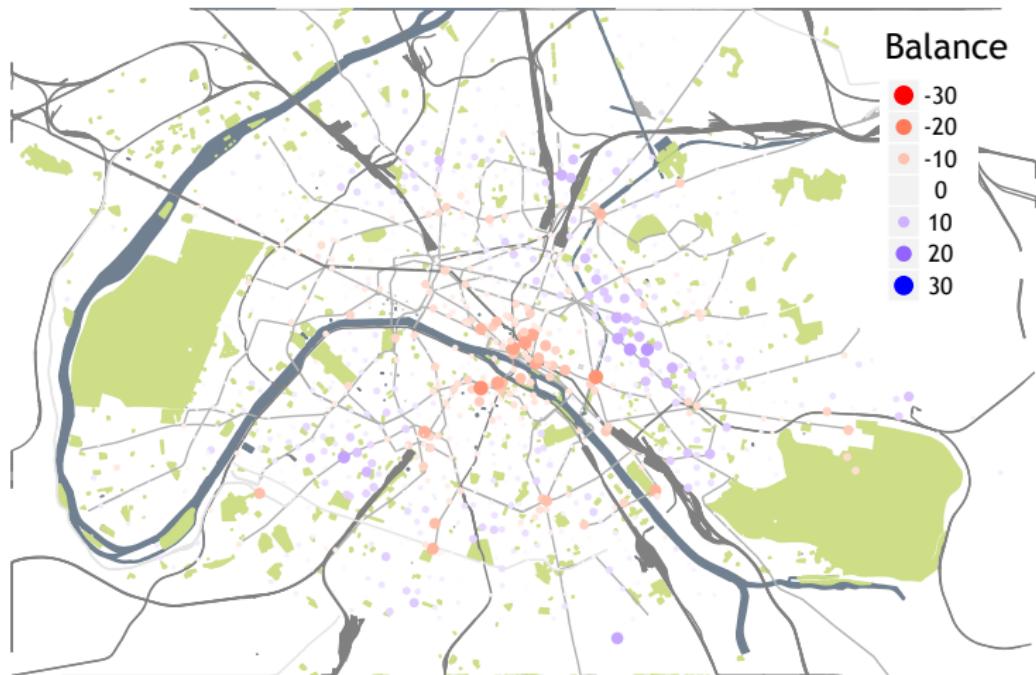


FIGURE 7: Activité latente "Loisirs nocturne", déséquilibre du réseau pour $N_{dep} = 10\,000$

"Loisirs"



FIGURE 8: Activité latente "Loisirs", déséquilibre du réseau pour
 $N_{dep} = 10\,000$

"Loisirs", stations avec un fort flux entrant



FIGURE 9: Stations incoming specificity

"Loisirs", stations avec un fort flux sortant

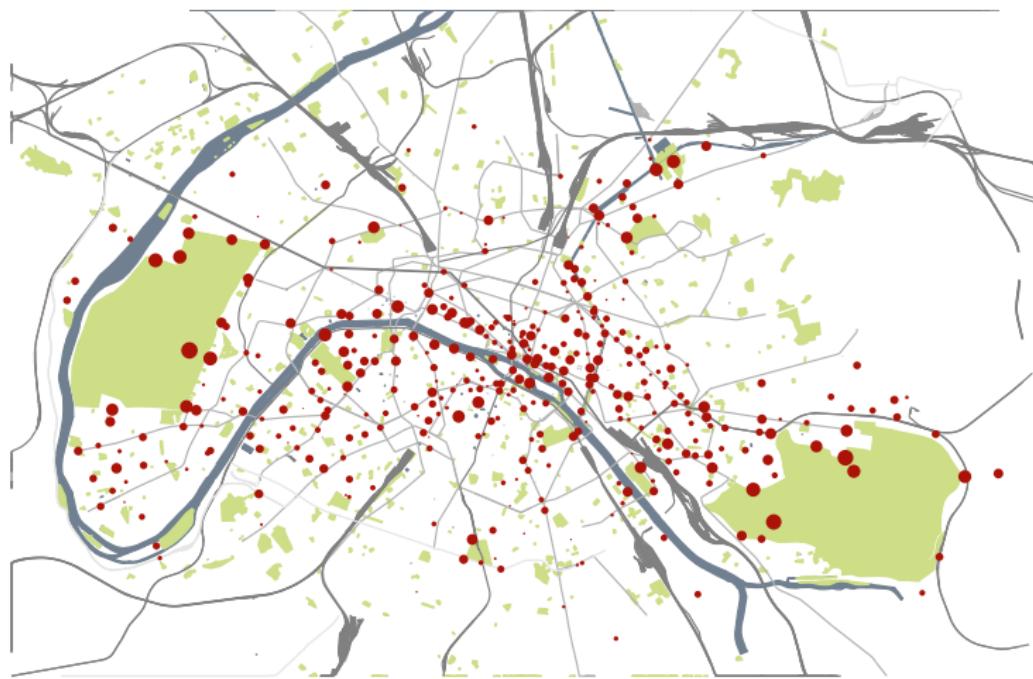


FIGURE 10: Stations outgoing specificity

Conclusion

Résultats

- Meilleure compréhension de la dynamique du réseau
- Modèle synthétique de la dynamique
- Mise en évidence des cycles

Limites

- Prise en compte uniquement de la demande satisfaite
- Pas de liens explicites avec les données socio-économiques et géographiques

Visualisation exploratoire et interactive du jeu de données

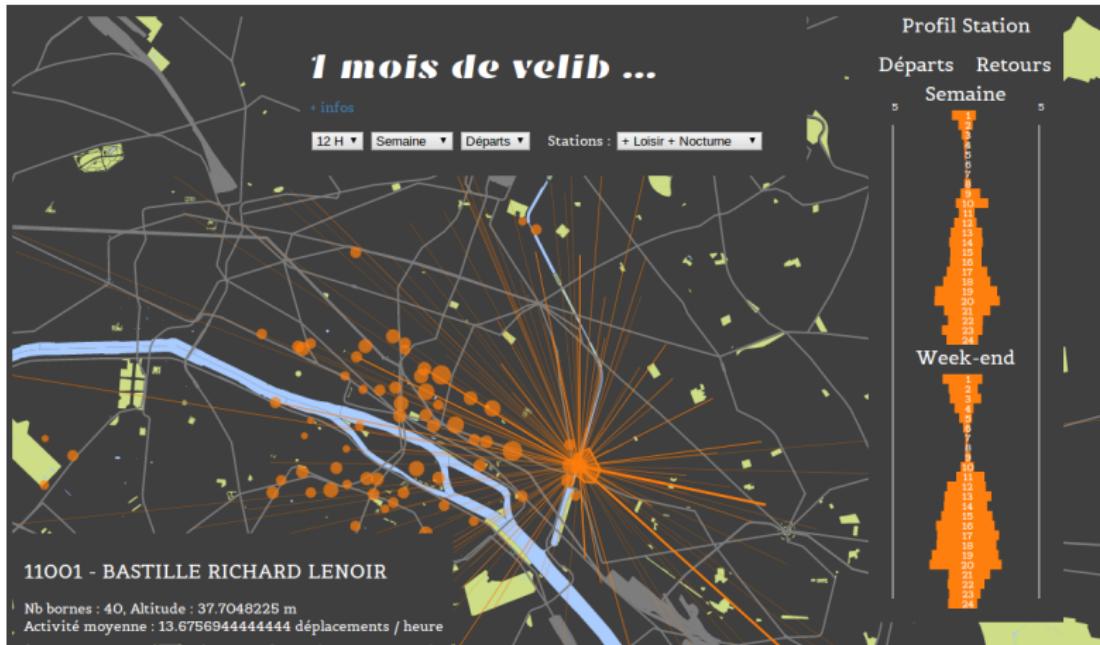


FIGURE 11: <http://www.comeetie.fr/galerie/velib/>

Recherche d'itinéraire (OpenTripPlanner)

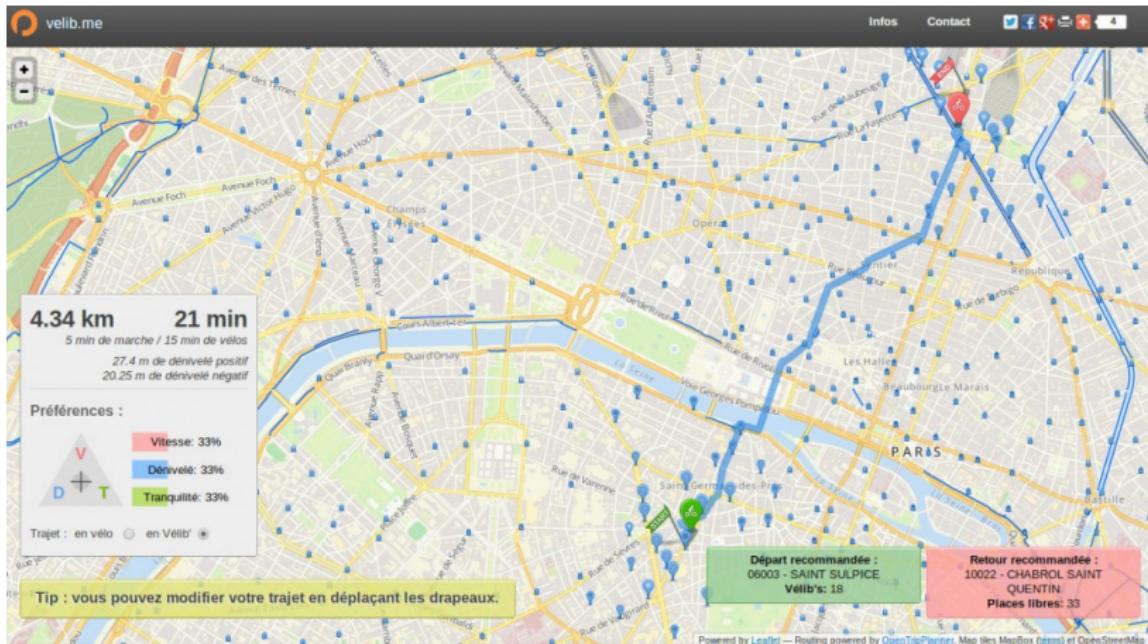


FIGURE 12: <http://velibme.com>

Projection de l'usage sur le réseau routier et cyclable



Travaux actuels & futurs

Modèle de prédiction

- Objectif : information usagers + amélioration qualité de service
- Besoins : + de données (aspect saisonnier, météo, ...)

Modèle spatio-temporel enrichi

Prise en compte explicite de données socio-économiques et géographiques

- Objectif : outils pour le dimensionnement et l'extension d'un réseau
- Ouverture : comparaisons de différentes villes (Londres, Washington D.C.)

Travaux actuels & futurs

Aspects intermodaux

Synergie entre VLS et TC

- Objectif : comprendre et mesurer l'intermodalité grâce à des données de billetterie

Projet PREDIT Mobilletic déposé (cas d'étude sur la ville de Rennes)

Estimation de la demande (satisfait + satisfaite)

- Objectif : mesure et optimisation de la qualité de service
- Besoins : données conjointes stocks et déplacements

Merci pour votre attention

latifa.oukhellou@ifsttar.fr, etienne.come@ifsttar.fr, @comeetie

Ifsttar

Centre de Marne-la-Vallée

Batiment le "Bienvenue"

14-20 Bd Newton Cité Descartes, Champs sur Marne F-77447 Marne la Vallée Cedex 2

Tél. +33 (0)1 81 66 87 19

Site : www.ifsttar.fr

