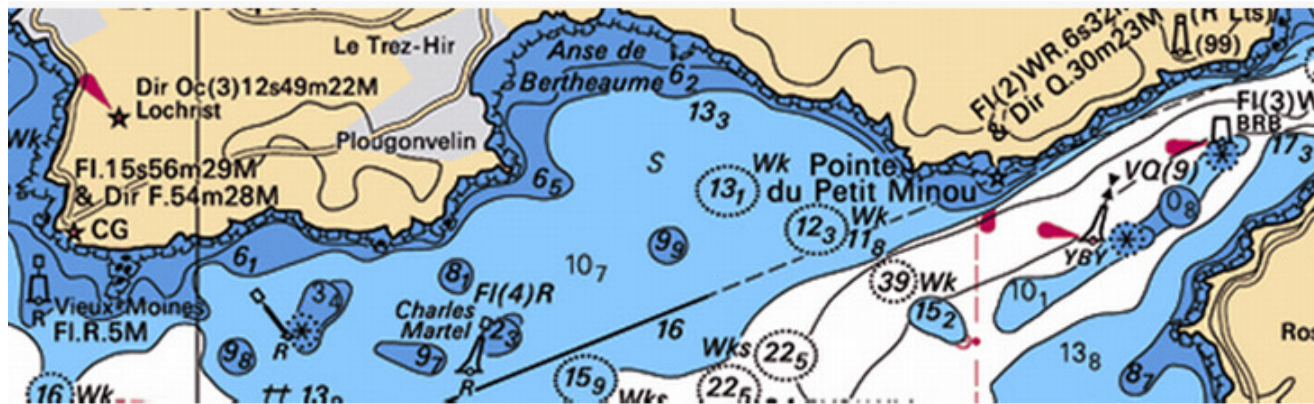




Sageo¹³_{rest}

Proposition d'une méthode de génération de matrices de flux temporelles issues de l'activité de téléphones portables

Ana-Maria OLTEANU-RAIMOND, Françoise BAHOKEN,
Thomas COURONNE et Zbigniew SMOREDA
Brest, le 25 septembre 2013

Le contexte

La génération classique de matrices de flux OD (enquêtes et recensements) pose différents problèmes d'exhaustivité, d'actualisation, de coût...

Les capteurs numériques, une source complémentaire

Collecte, stockage de masses de données géo-localisées

À un faible coût : téléphones portables, GPS , Bluetooth, IP (internet), ...

Des réponses par de nouvelles méthodes de génération de matrices de flux OD, en utilisant des traces de capteurs (GPS, mobiles)

Byeong-Seok et *al.*, 2005

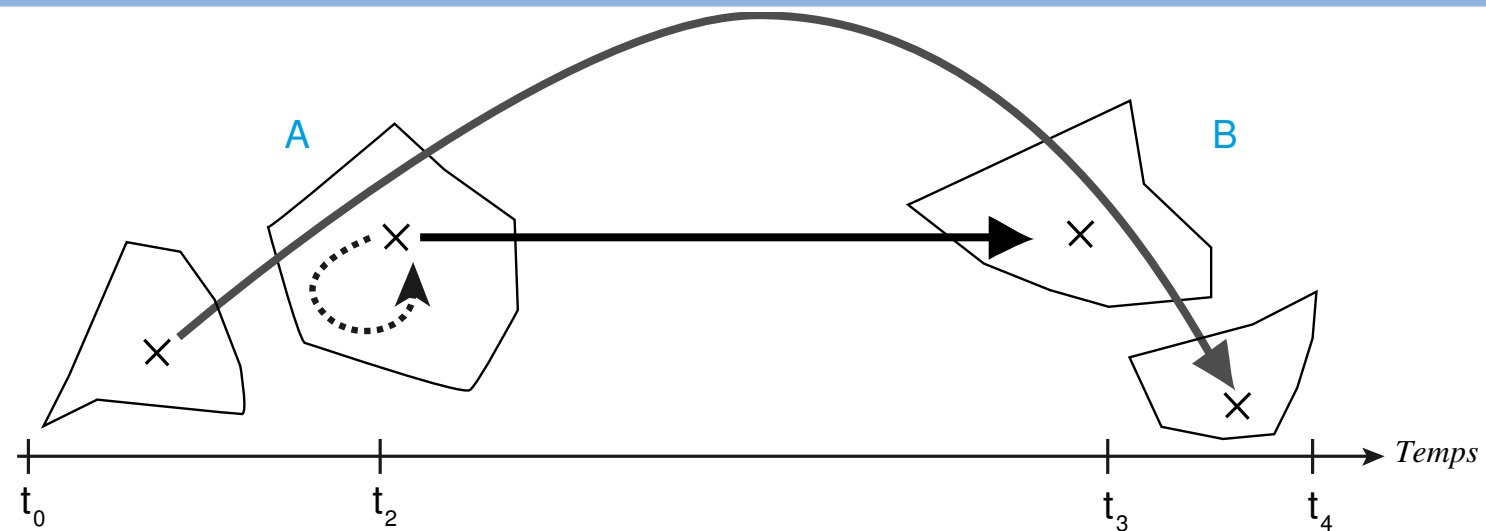
Caceres et *al.*, 2007

Zhang et *al.*, 2010

Giannotti et *al.*, 2011

Frias-Martinez et *al.*, 2012

Les objectifs



 Flux Origine-Destination (entre 2 lieux)  Trajectoire (entre plusieurs lieux)
 Flux intra zone (1 lieu)

$\begin{smallmatrix} D \\ \diagdown \\ O \end{smallmatrix}$	A	...	B
A	30	5	1
...	10	7	1
B	1	3	0

avec :

$A, \dots, B$: identifiants des lieux

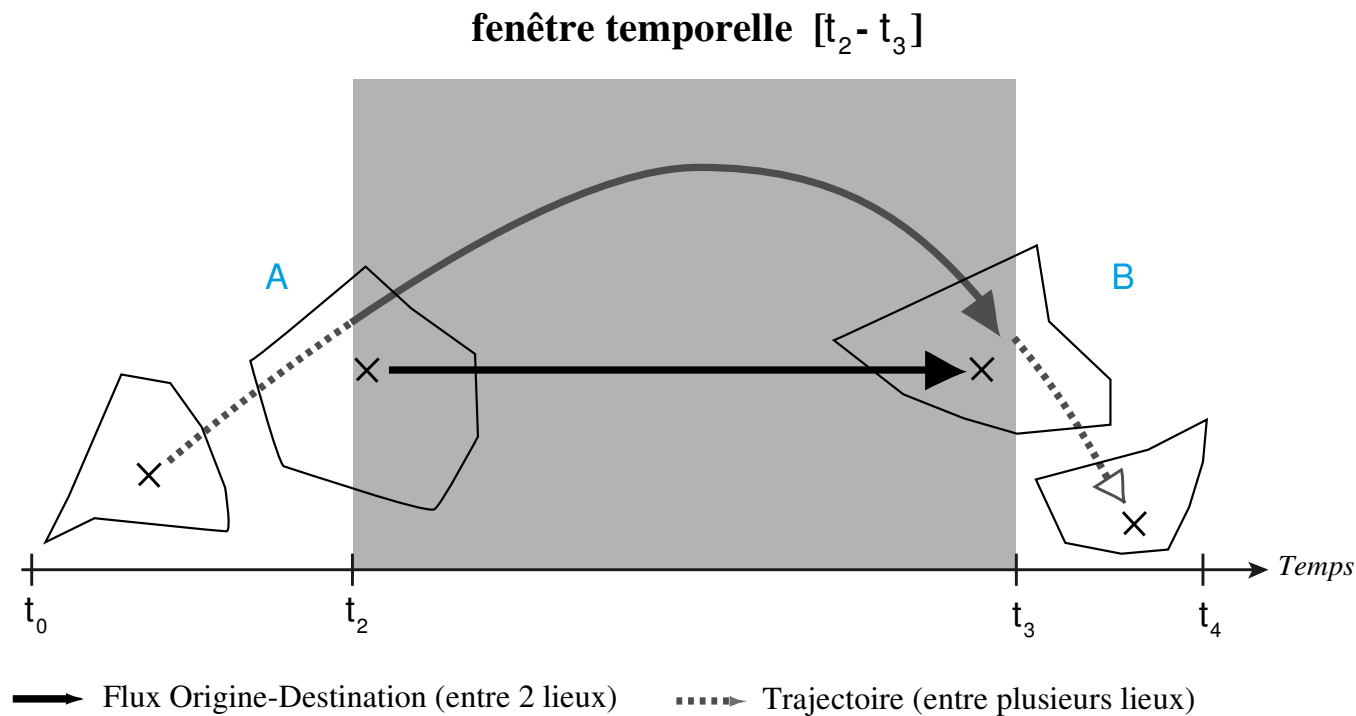
5 : valeur du flux interzonal

30 : valeur du flux intrazonal

→ Objectif 1 : Comment générer les valeurs de la matrice ?

Les objectifs

- Cas général : 1) filtrage temporel puis 2) spatial de trajectoires individuelles
 - Fenêtre temporelle : découpage temporel
 - *je sélectionne les déplacements qui se sont produits dans une période ...*
 - Fenêtre spatiale : découpage territorial en maille (qui forment l'aire d'étude)
 - *... et qui ont franchi au moins deux limites de mailles, situées dans ma zone d'étude*



→ on ne retient que les déplacements entre $[t_1 - t_2]$ entre les mailles (A et B)
→ flux interzonaux

Les objectifs

- Cas général
 - Fenêtre spatiale : souvent des découpages administratifs, imposés
 - Fenêtre temporelle : souvent définie *a priori* (ex : mobilité quotidienne)
- Problème : des découpages spatio-temporels potentiellement inadaptés aux données issues de la téléphonie mobile (hétérogénéité des données)

→ Objectif 2 : Développer une méthode de définition de la fenêtre temporelle* fondée sur le signal (l'activité du mobile)

*Le filtrage spatial et ses effets sont abordés dans (Bahoken et Olteanu-Raimond , 2013)

Plan de l'exposé

- **Présentation de traces de mobiles**
- **Proposition de génération de matrices OD temporelles**
- **Cas d'application**



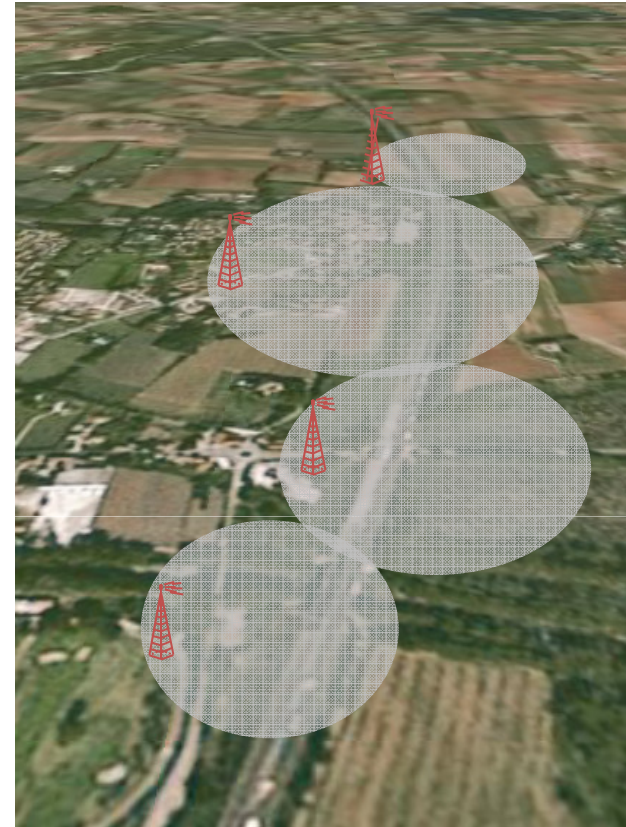
Les données

→ Opérateur Orange France©

- de 160 000 à 280 000 usagers / jour
- entre 2,3 et 5 millions d'enregistrements / jour

→ Plusieurs types de données

- Comptes rendus d'appel (CRA)
- Sondes 2G
- Sondes 3G



→ Les données décrivent la position du mobile*

- rapportée à une antenne localisée
- à un moment donné.

* On fait l'hypothèse qu'à un mobile correspond un seul utilisateur

Les données

→ Une collecte de données issue d'un enregistrement automatique

- **Appel** : reçu ou émis

CRA

Sondes 2G

Sondes 3G

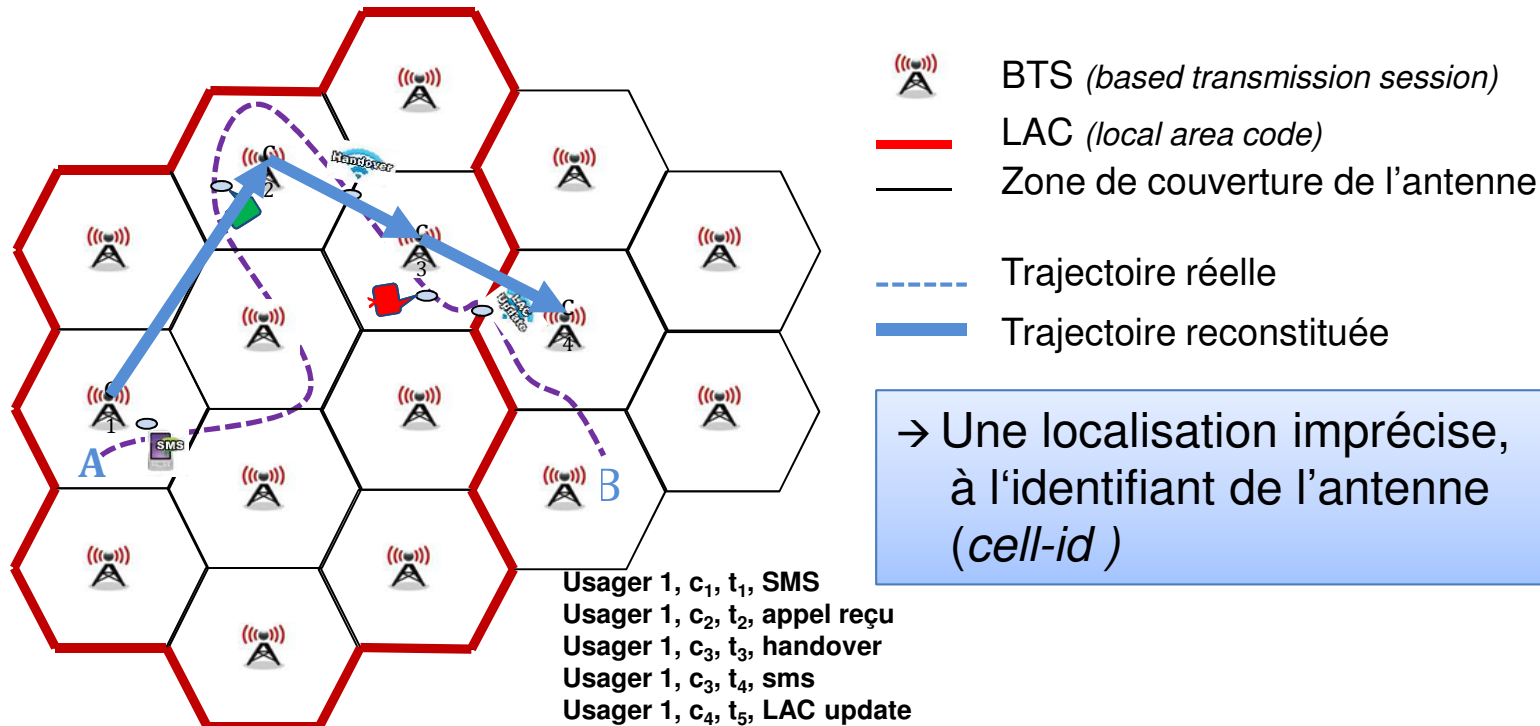
- **SMS** : reçu ou émis

- **Handover** : changement d'antenne pendant la communication

- Mise à jour de la localisation : **changement de LAC** et **inactivité** > 3h

- Allumage/ extinction de téléphone portable

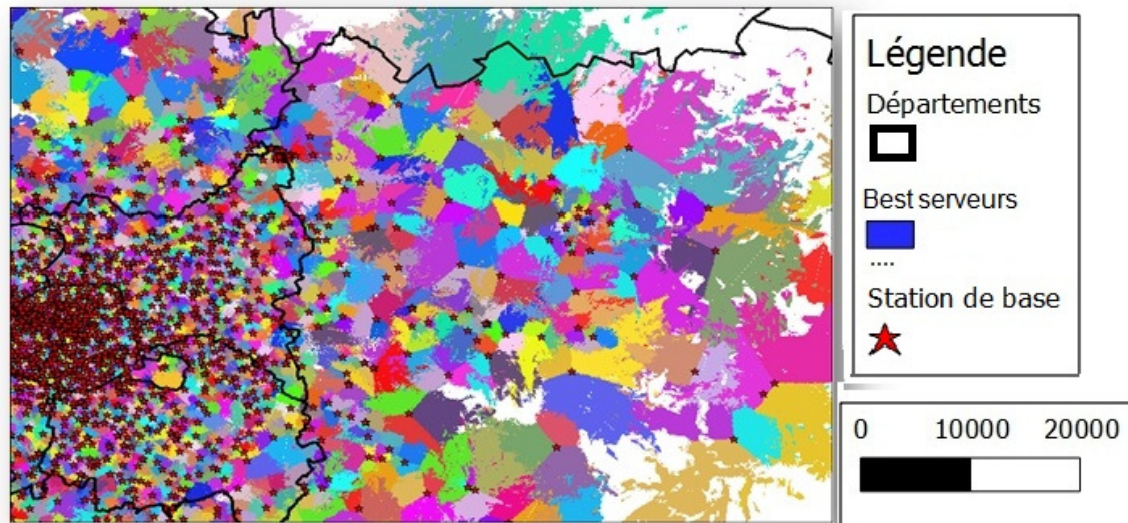
- *Connexion data*



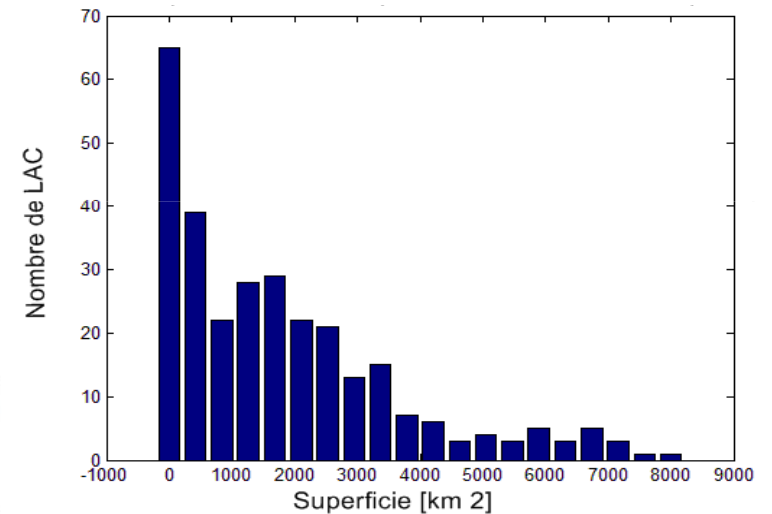
Les données

→ Une granularité spatiale variable

Distribution des stations de base et de leurs zones de réception (*best serveurs*)



Répartition des LAC en métropole



Les données

→ Une granularité temporelle variable

Id	Heure et Date	Evénement	Cellule id	LAC	Position
1	17/03/2011 12:00:00	1	1932	2	(x_1, y_1)
1	17/03/2011 22:09:00	6	793	3	(x_2, y_2)
1	17/03/2011 22:10:00	2	891	3	(x_2, y_2)



→ L'absence de sémantique :

- sur les pratiques de l'individu aux lieux ;
- sur les « stations » ;
- sur le mode de transport,
- etc.

→ De nombreux avantages :

- échantillon important ;
- enregistrement pseudo-continu ;
- collecte massive à faible coût ;
- etc.

Plan de l'exposé

- Présentation de traces de mobiles
- Proposition de génération de matrices OD temporelles
- Cas d'application



La construction de la matrice

→ L'estimation de flux OD à partir de trajectoires

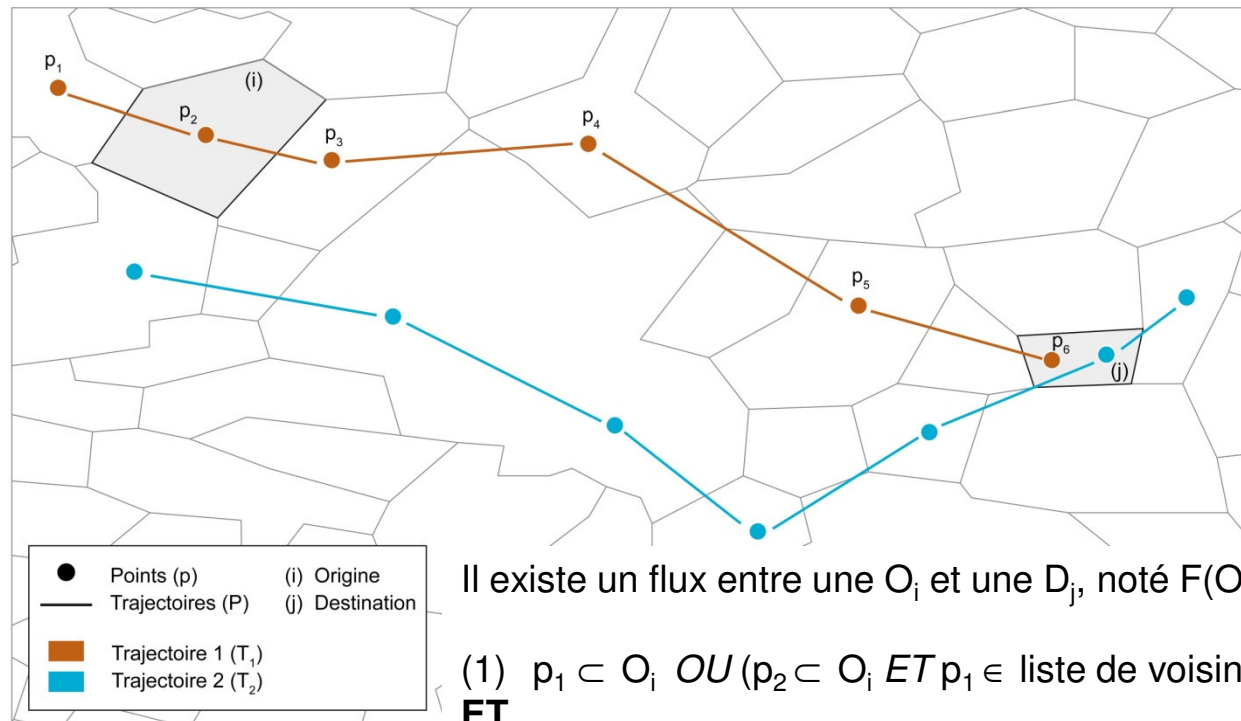
Soit une trajectoire composée d'un ensemble de n points successifs,

$$T_k = \{p_1, p_2, \dots, p_i, \dots, p_{n-1}, p_n\}, \text{ où}$$

$p_i = (x_i, y_i, t_i)$, un point de coordonnées (x, y) à un moment t

$$t_1 < t_2 < \dots < t_i < \dots < t_{n-1} < t_n$$

$i = 1..n$, n = le nombre de points composant la trajectoire T_k



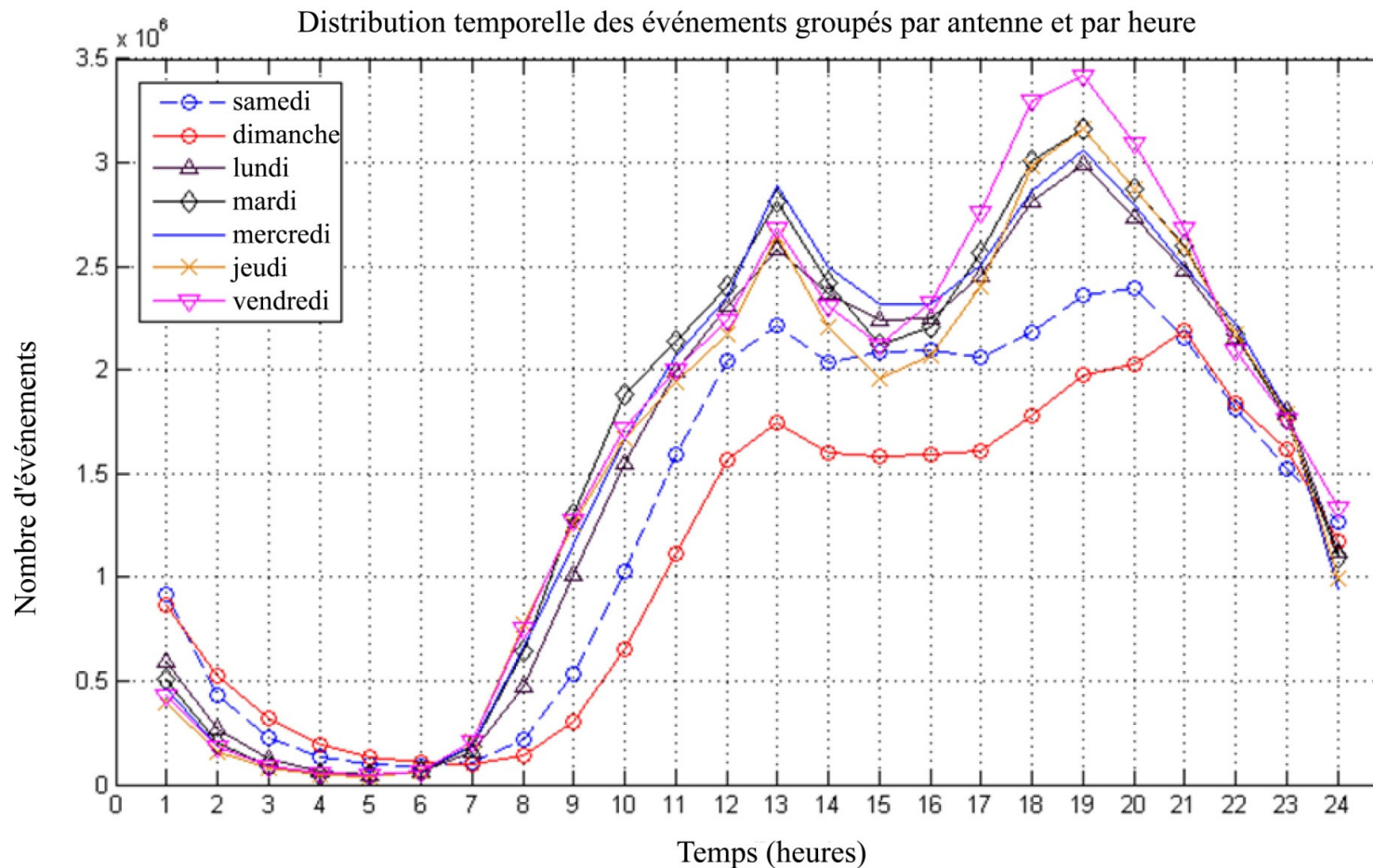
Il existe un flux entre une O_i et une D_j , noté $F(O_i/D_j)$ si et seulement si :

(1) $p_1 \subset O_i$ OU ($p_2 \subset O_i$ ET $p_1 \in$ liste de voisins de 1^{er} ordre de p_2)

ET

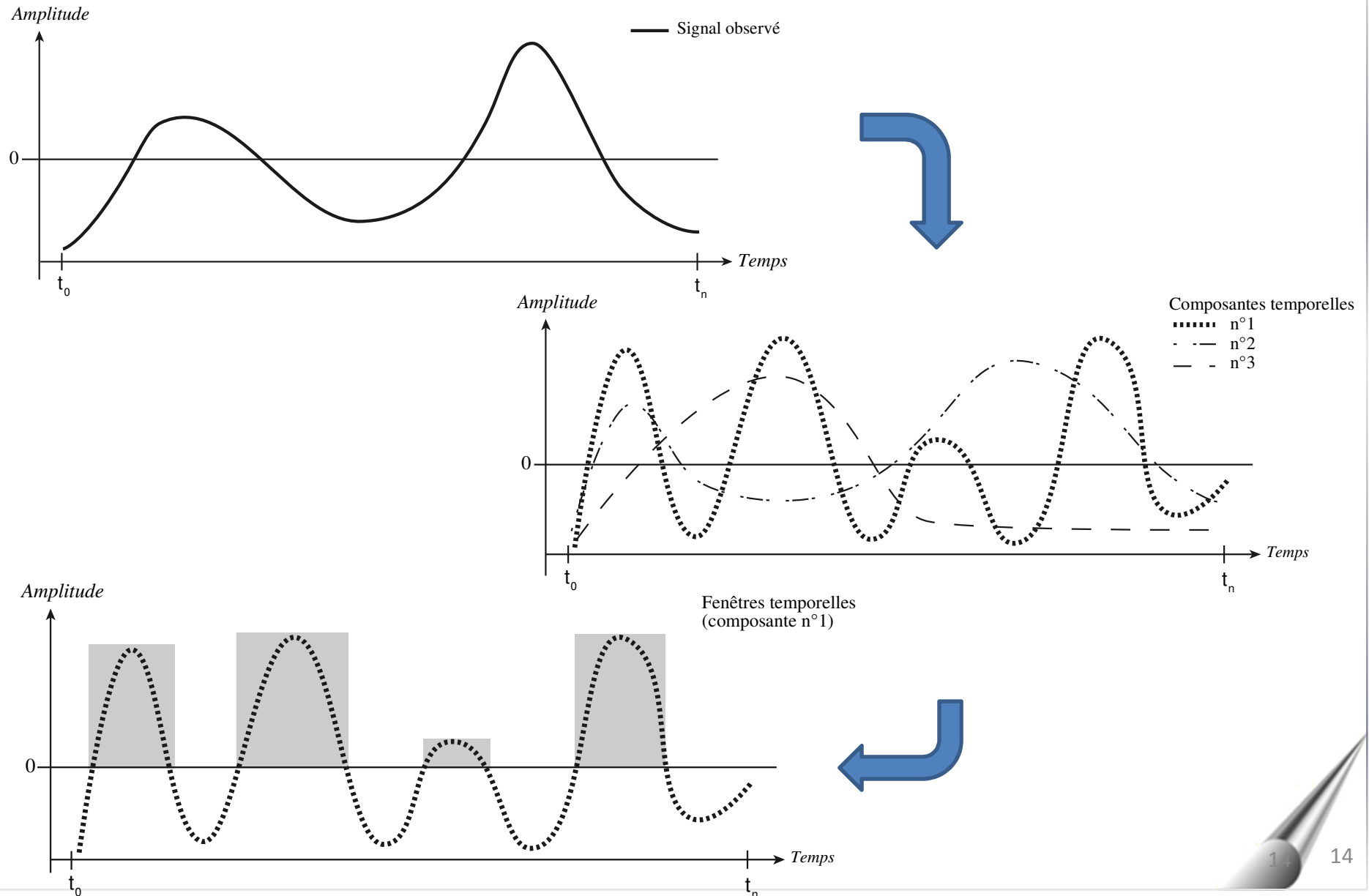
(2) $p_n \subset D_j$ OU ($p_{n-1} \subset D_j$ ET $p_n \in$ liste de voisins de 1^{er} ordre de p_{n-1})

La construction de la matrice



→ d'où l'idée de générer des fenêtres temporelles
à partir de l'activité des mobiles

La construction de la matrice → La méthode EMD (Huang et al. 1971)



Plan de l'exposé

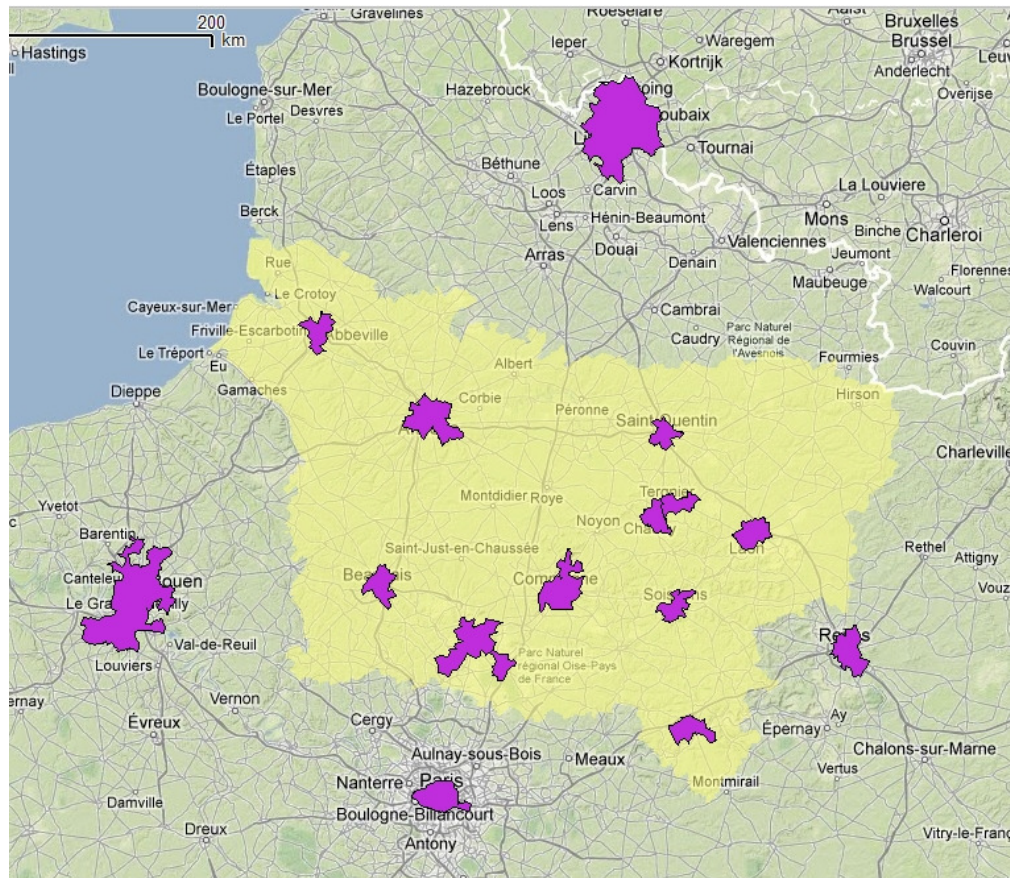
- Présentation de traces de mobiles
- Proposition de génération de matrices OD temporelles
- **Cas d'application**



Un cas d'application

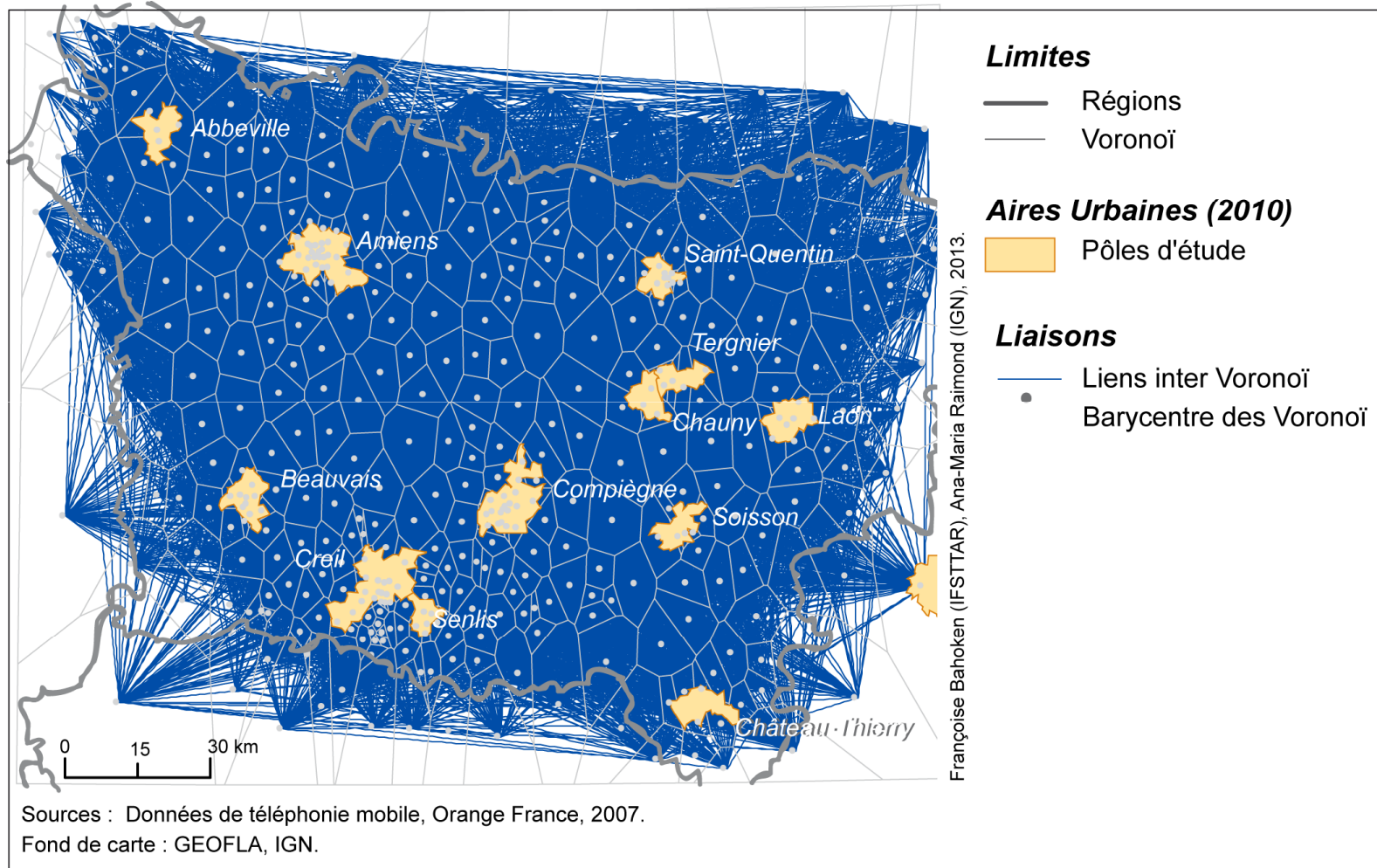
Mesurer les flux *picards, inter pôles*, quotidiens

Filtres spatiaux : région Picardie, échelle des pôles urbains (Insee)



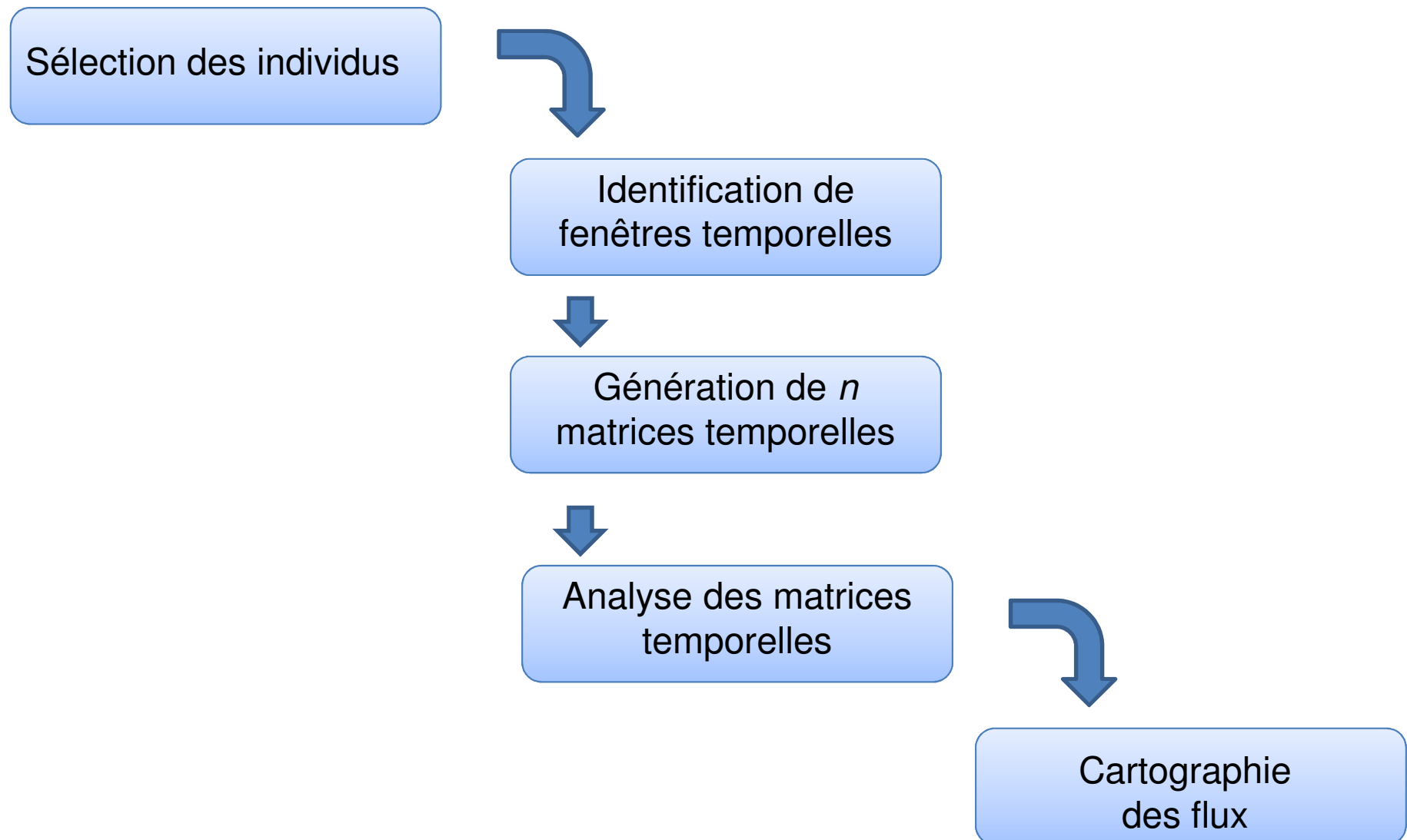
La construction de la matrice

Sans filtrage temporel et spatial



Un cas d'application

→ La procédure



Un cas d'application

→ La procédure

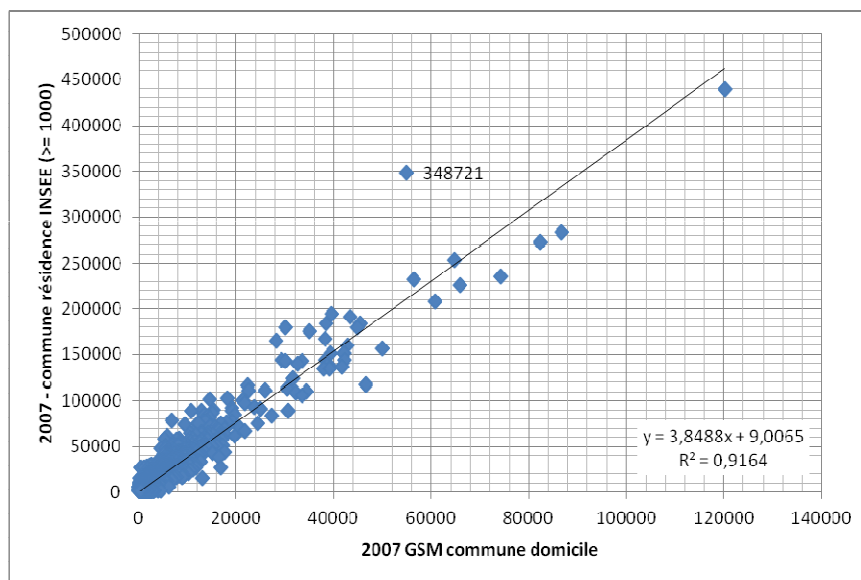
Sélection des individus

Zone d'étude : France entière

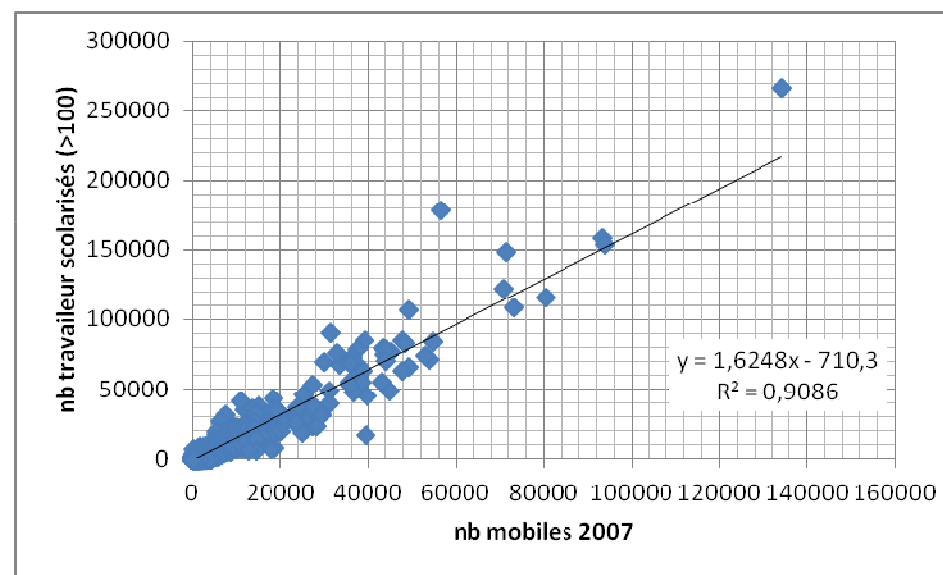
Période : de septembre à novembre 2007 (CRA)

Proposition d'une démarche déductive

- semaine : de 20h à 6h -> **lieu de résidence**
- jours ouvrés : de 6h à 20h -> **lieu de travail / étude**



Corrélation de **0,91** entre le lieu de résidence détecté et les lieux de résidence (RP, 2007, Insee)



Corrélation de **0,90** entre le lieu de travail / études détecté et les lieux de « travailleurs » et « scolarisés » (RP, 2007, Insee)

Effectif total : 372 936 individus picards qui travaillent dans l'un des pôles

(parmi 10 155 916 individus qui résident en Picardie)

La construction de la matrice

Sélection des individus

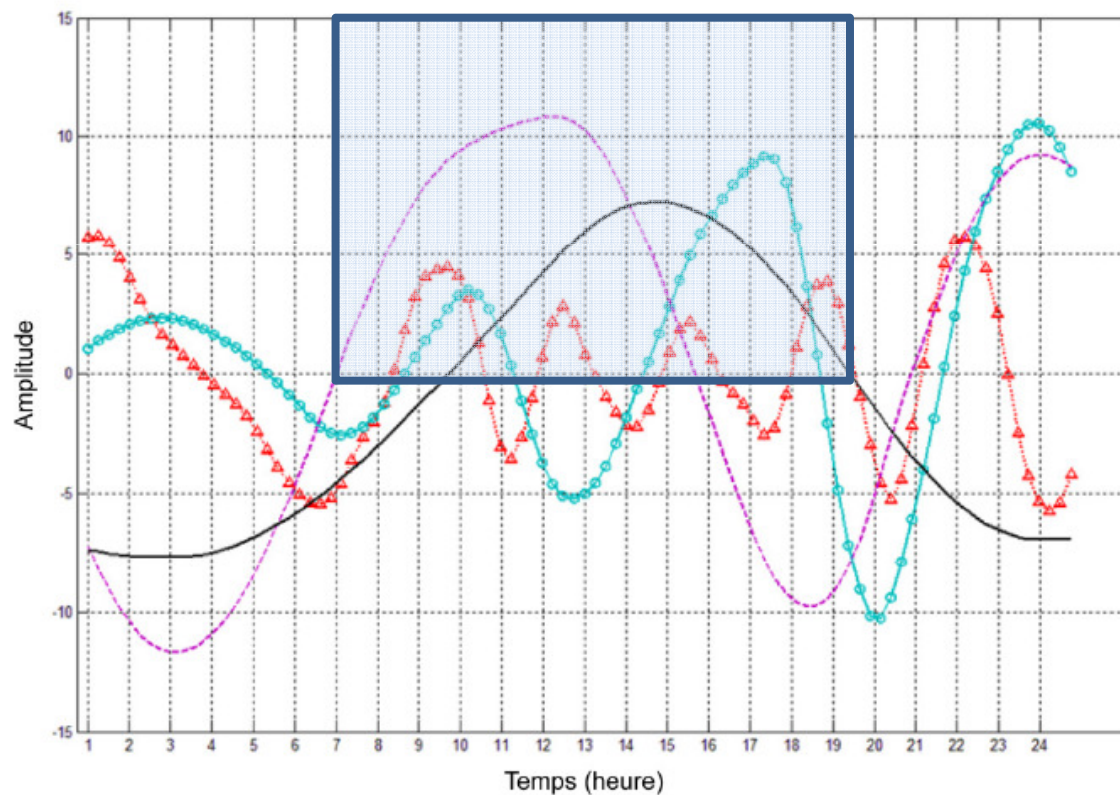


Identification de
fenêtres temporelles



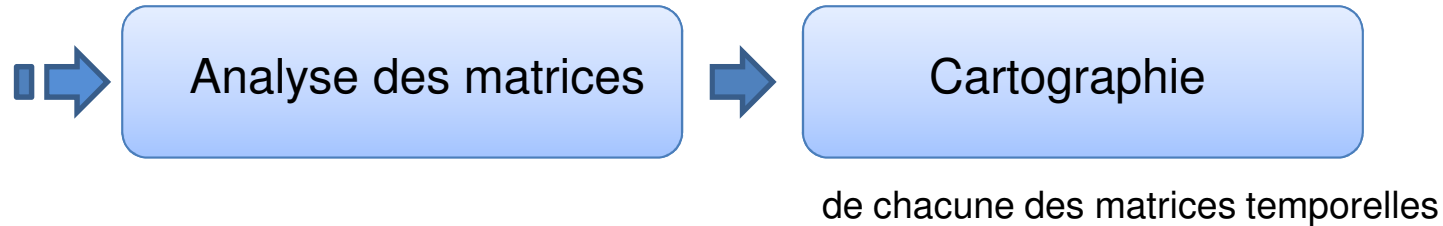
Génération de n
matrices temporelles

Les 4 composantes primitives obtenues à partir du signal observé,



Soit : 8 fenêtres temporelles pendant la journée : de 7h00 à 19h30
8 matrices de flux

La construction de la matrice



→ Quatre groupes de journées :

Groupe 1) lundi-mardi-mercredi

Groupe 2) samedi-dimanche (we)

Groupe 3) lundi-mardi-mercredi avec le jeudi

Groupe 4) le vendredi.



Journée du vendredi



Heures de pointe du matin



Heures de pointe matin et soir

Conclusion

→ Une méthode de construction de matrices OD temporelles

- prise en compte de l'hétérogénéité spatio-temporelle des données
- génération de fenêtres temporelles à partir du signal produit par l'activité téléphonique

→ Application sur les flux des picards à l'échelle des pôles urbains

- détection des lieux de résidence et de travail

→ Méthode généralisable

→ Perspectives

- analyser davantage les différentes matrices temporelles obtenues
- analyser la robustesse de la méthode EMD aux types de données et d'aggrégation

→ une nouvelle demande du CR Picardie : estimer les moyens de transports des flux générés

Merci de votre attention.

Françoise BAHOKEN
Université Paris-Est / IFSTTAR
Boulevard Newton,
77420 Champs-sur-Marne
Tél. +33 (0)1 81 66 87 82
www.ifsttar.fr
francoise.bahoken@ifsttar.fr

Ana-Maria OLTEANU-RAIMOND
IGN - COGIT
ORANGELab'
Tél. +33 (0)1 43 98 80 00 +7550
www.ign.fr
ana-maria.raimond@ign.fr