

# **CADDY — Contrôle et Acquisition de Données temporelles massives, stockage et modèles DYNamiques**

**F. DE VUYST**<sup>1</sup>, **G. JOMIER**<sup>2</sup>, **G. HÉBRAIL**<sup>3</sup>,  
**O. CARLES**<sup>4</sup>, **B. HUGUENEY**<sup>1</sup>, **M. MANOUVRIER**<sup>2</sup>,  
**Y. NAIJA**<sup>2</sup>, **G. SCEMAMA**<sup>4</sup>, **L. STEFFAN**<sup>2</sup>

<sup>1</sup> MAS, École Centrale de Paris

<sup>2</sup> LAMSADE, Université Paris-Dauphine

<sup>3</sup> ENST

<sup>4</sup> INRETS

# Problématique

- analyser,
- comprendre,
- contrôler,
- améliorer

*des systèmes complexes*

au comportement décrit par des

*séries chronologiques interdépendantes.*

# Équipe et compétences

## CADDY : projet multidisciplinaire

- \* Mathématiques Appliquées :  
MAS, École Centrale de Paris
- \* Bases de Données, Optimisation :  
LAMSADE Université ParisDauphine
- \* Analyse de Données : ENST
- \* Spécialistes des Transports : INRETS

# Plan

- Présentation des données utilisées
- Résumés numériques
- Classification et résumés symboliques
- Entrepôt de données
- Perspectives

# Réseau urbain étudié

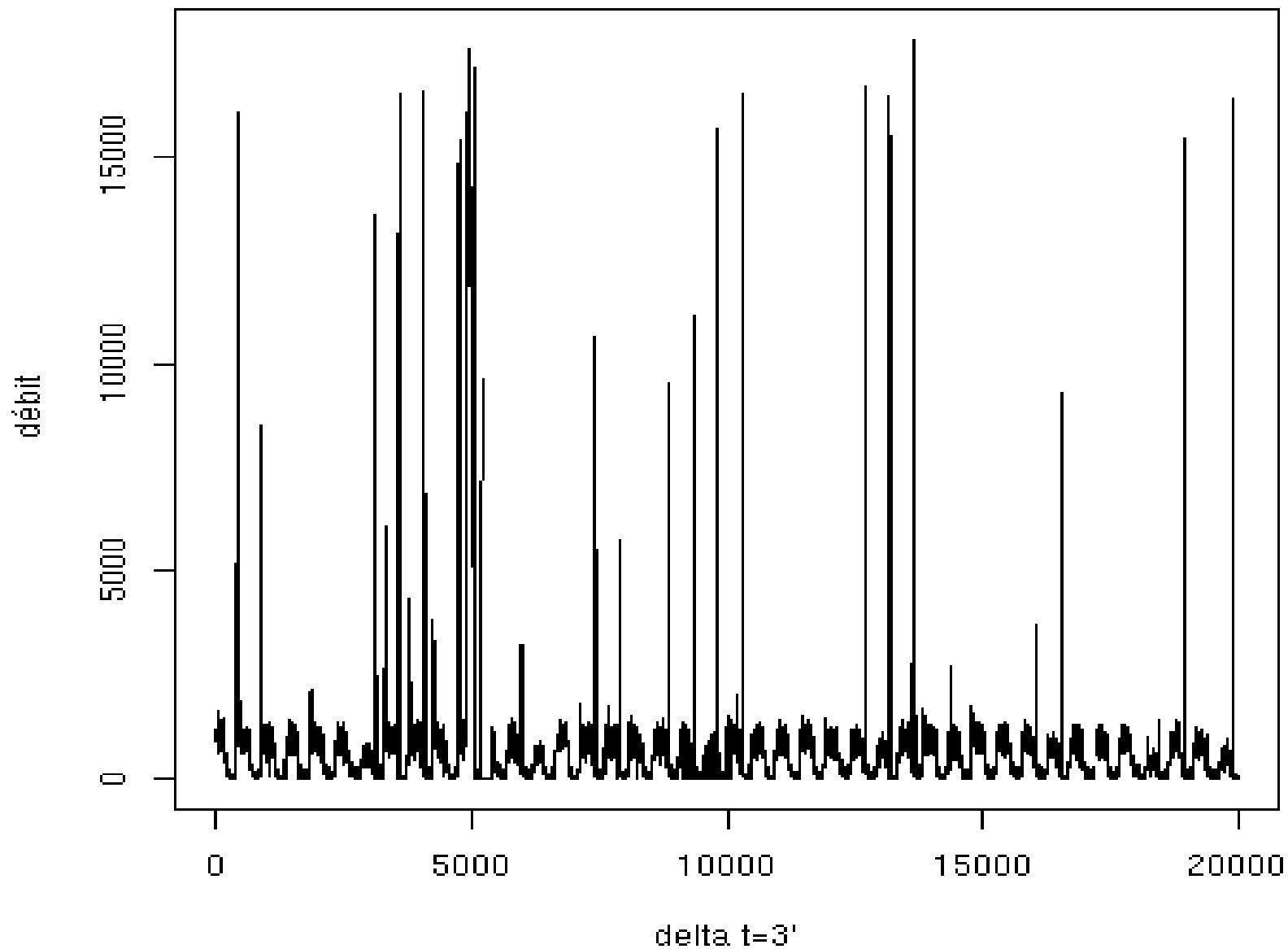
**1 000 capteurs**

**960 000 données  
élémentaires/jour**

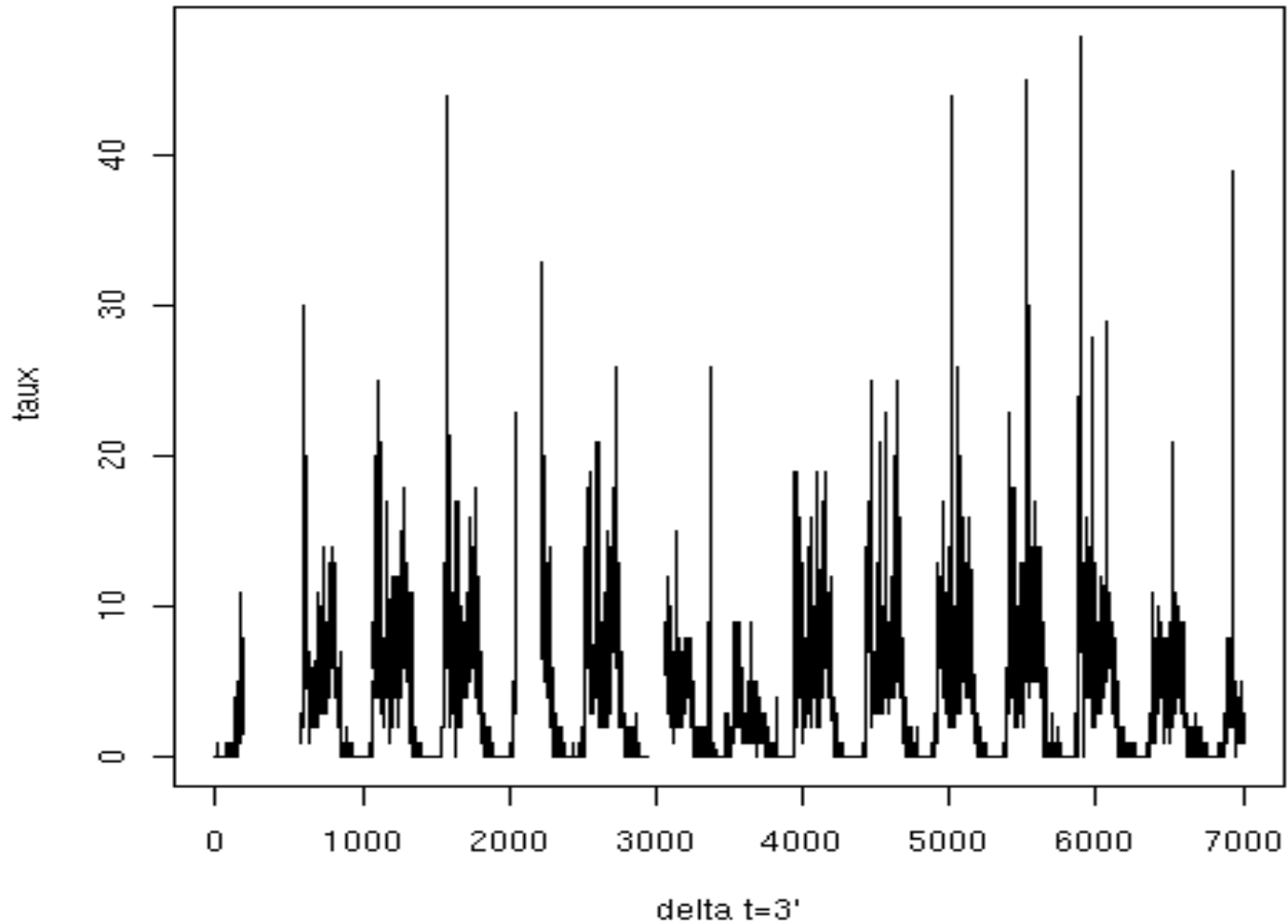
QuickTime™ et un décompresseur  
PNG sont requis pour visualiser  
cette image.

**main courante**

# Débit d'un capteur



# Taux d'occupation du sol



# Objectif : une méthodologie

- Concevoir des algorithmes de fouille de données (spatio-) temporelles
- Définir un entrepôt de données spatio-temporelles
- Elaborer des requêtes performantes pour analyser le comportement du système réel
- Imaginer une visualisation adaptée aux utilisateurs, à fin d'exploration, de contrôle, d'aide à la décision
- Simuler et confronter au réel à fin de prévision
- Adapter à l'évolution par apprentissage
- Réaliser un prototype logiciel



# Résumés numériques : PAA (1)

QuickTime™ et un décompresseur  
PNG sont requis pour visualiser  
cette image.

# Résumés numériques : PAA (2)

QuickTime™ et un décompresseur  
PNG sont requis pour visualiser  
cette image.

# Résumés numériques : PLA (1)

QuickTime™ et un décompresseur  
PNG sont requis pour visualiser  
cette image.

# Résumés numériques : PLA (2)

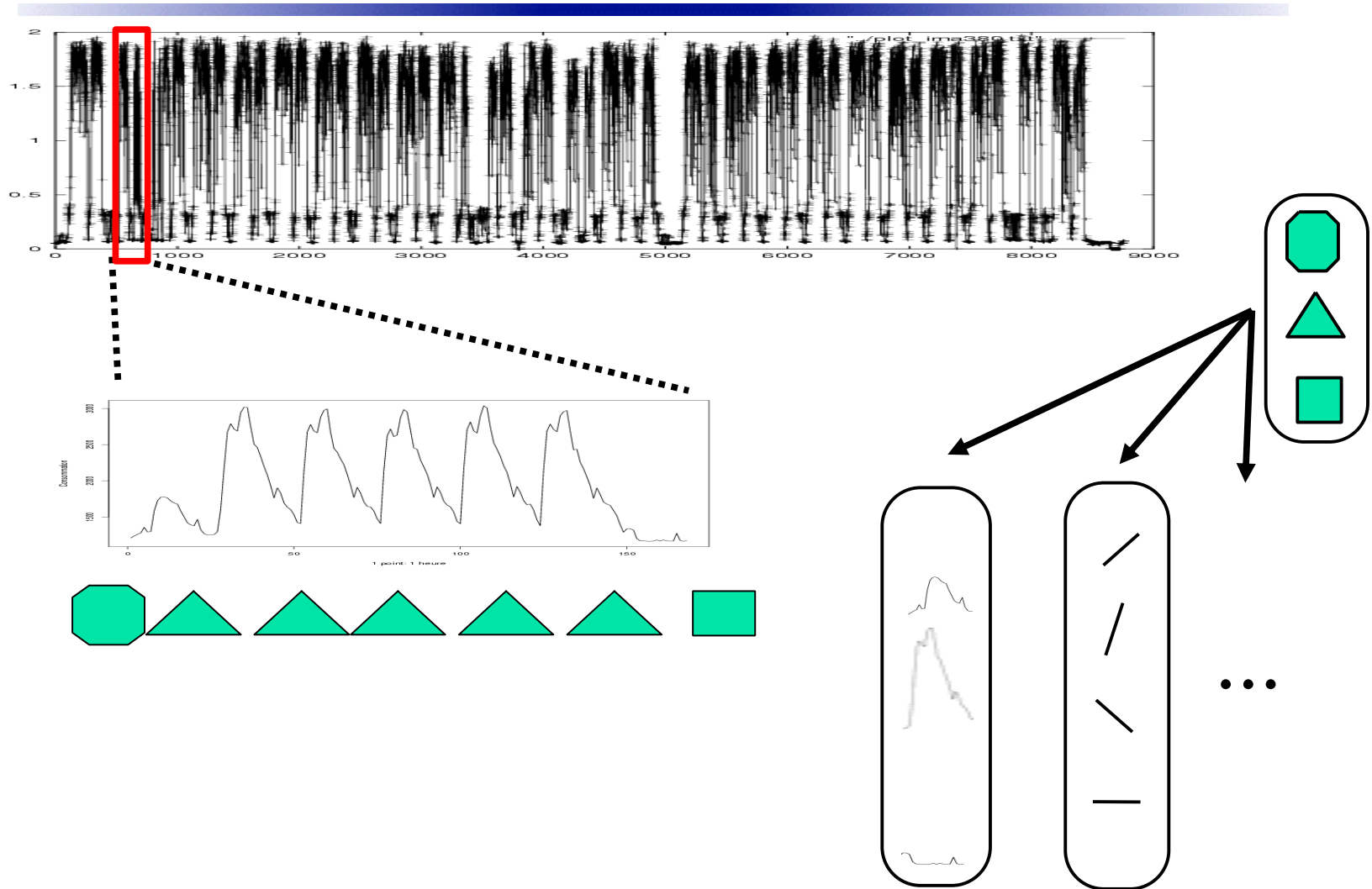
QuickTime™ et un décompresseur  
PNG sont requis pour visualiser  
cette image.

# Classification

QuickTime™ et un décompresseur  
PNG sont requis pour visualiser  
cette image.

1 capteur : débit journalier

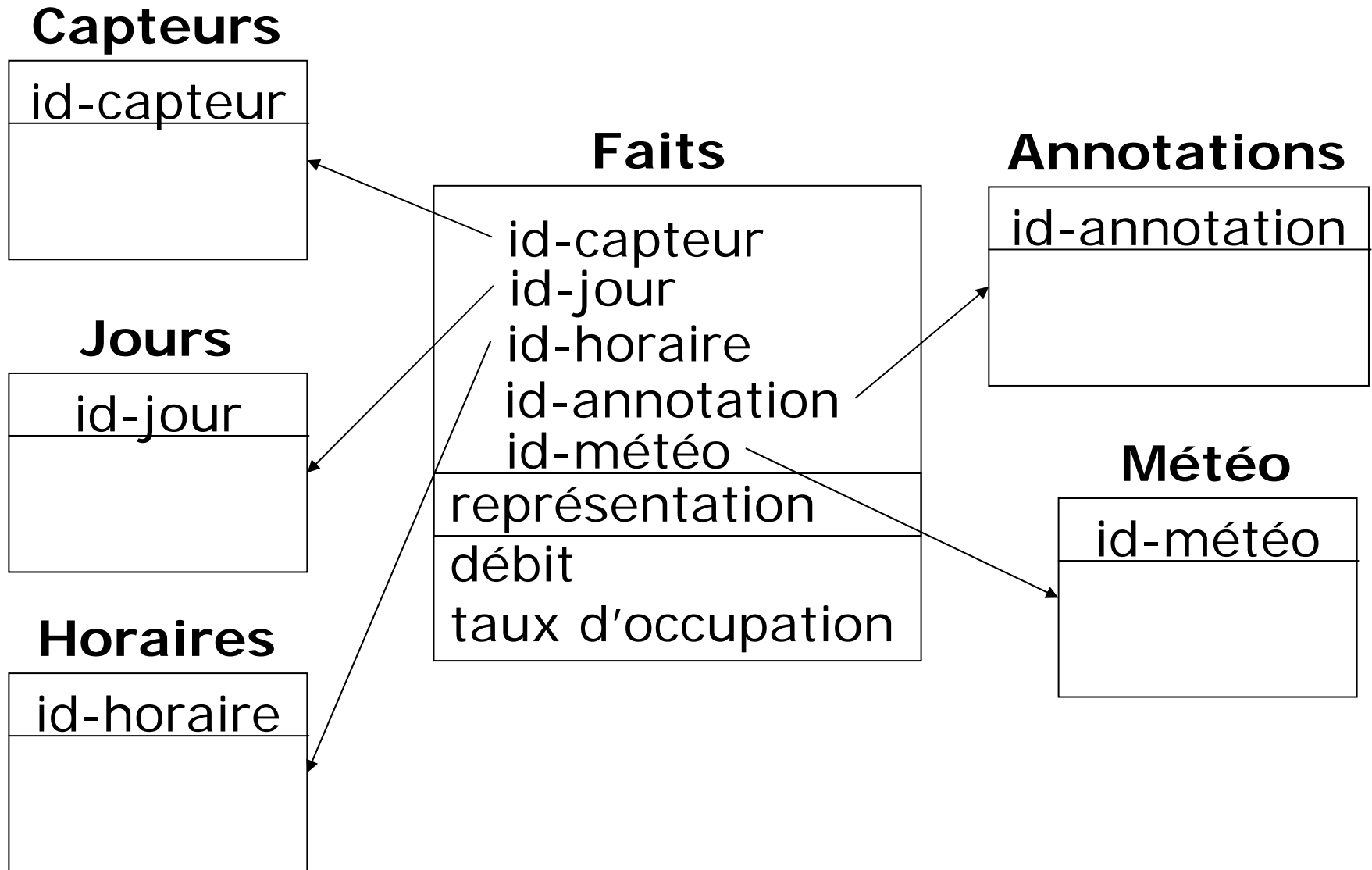
# Représentations symboliques



# Architecture

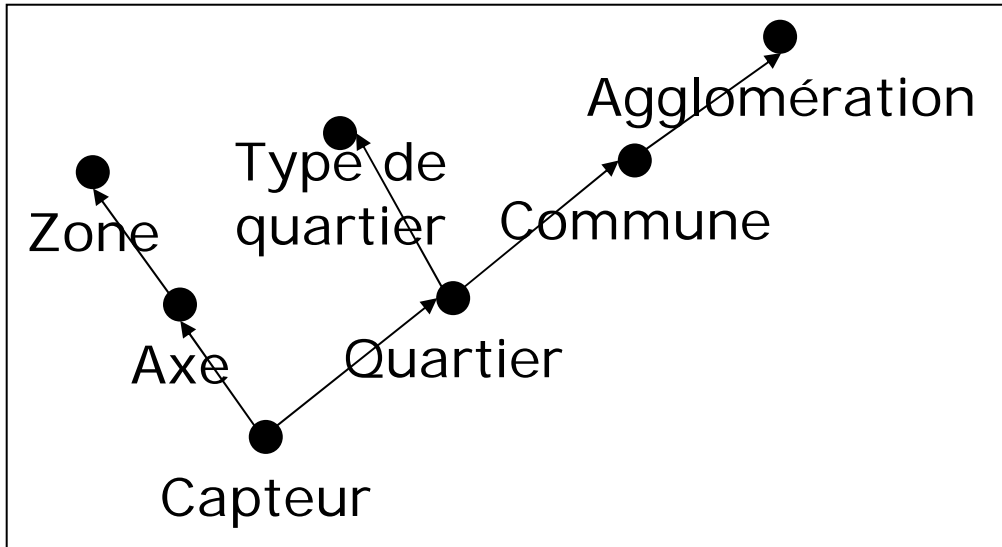
QuickTime™ et un décompresseur  
PNG sont requis pour visualiser  
cette image.

# Schéma de l'entrepôt





# Dimensions d'agrégation



Météo

...

Horaire (3 mn)

Quart d'heure

Heures

Épisodes heures

Épisodes a priori

Annotations

...

Jour

*Calendrier*

Semaine

Mois

} Année

*Activité scolaire*

Vac. Noël, février, ...

Vacances

Périodes scolaires

Temps scolaire

*Activité professionnelle*

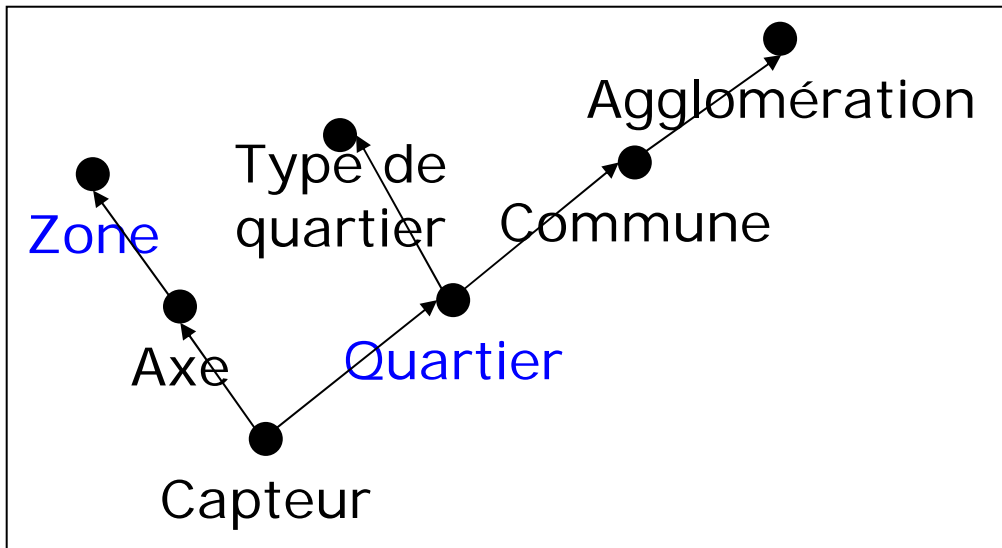
Fériés / Ouvrables

Week-end / Semaine

*Activité socio-culturelle*

...

# Dimensions d'agrégation



Météo  
...

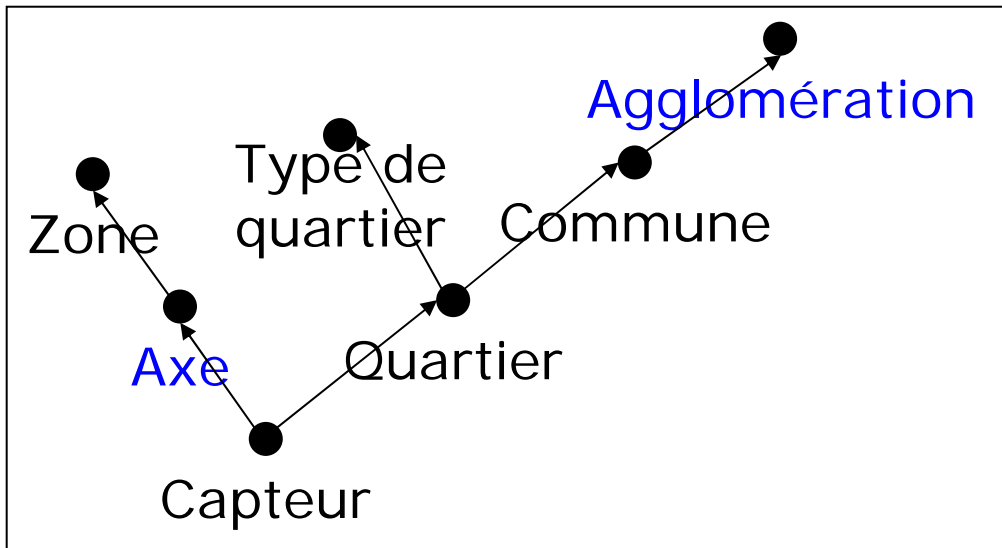
Horaire (3 mn)  
Quart d'heure  
Heures  
Épisodes heures  
Épisodes a priori

Annotations  
...

Jour  
*Calendrier*  
Semaine } Année  
Mois }  
*Activité scolaire*  
Vac. Noël, février, ...  
Vacances  
Périodes scolaires  
Temps scolaire  
*Activité professionnelle*  
Fériés / Ouvrables  
Week-end / Semaine  
*Activité socio-culturelle*  
...

Quelles sont les zones congestionnées d'un quartier en fin de soirée lors de manifestations sportives ?

# Dimensions d'agrégation



Météo  
...

Horaire (3 mn)  
Quart d'heure  
Heures  
Épisodes heures  
Épisodes a priori

Annotations  
...

Jour  
Calendrier  
Semaine } Année  
Mois }  
Activité scolaire  
Vac. Noël, février, ...  
Vacances  
Périodes scolaires  
Temps scolaire  
Activité professionnelle  
Fériés / Ouvrables  
Week-end / Semaine  
Activité socio-culturelle  
...

Quels sont les axes fluides les veilles de week-end de vacances scolaires entre 15h et 17h par beau temps ?

# Thèmes abordés

- Trafic routier en réseau urbain
- Séries temporelles multivariées
- Fouille de données spatiotemporelles
- Entreposage de données
- Gestion des flux de données
- Visualisation avancée
- Supervision et aide à la décision
- Apprentissage et modèle dynamique

# **Perspectives : finaliser le projet**

- Découverte et visualisation de motifs spatio-temporels
- Analyse en ligne multi-échelle et multi-représentation
- Restitution de l'information adaptée à l'utilisateur (décideur, exploitant, usager)
- Apprentissage

# Perspectives : après CADDY

- Acquisition et traitement temps réel des flux de données
- Simulation, prévision
- Etude d'autres systèmes complexes avec corrélations spatio-temporelles entre les flux de données
  - hydrologie au Sahel (AGRHYMET, CIRAD)
  - consommation d'électricité (EDF),
  - collaboration avec le CEA