



**HAL**  
open science

## Commuter's individual travel time estimation: using AFC data of Paris transit system

Xiaoyan Xie, Fabien Leurent, Vincent Aguiléra

### ► To cite this version:

Xiaoyan Xie, Fabien Leurent, Vincent Aguiléra. Commuter's individual travel time estimation: using AFC data of Paris transit system. COST Action TU1004 Modelling Public Transport Passenger Flows in the Era of Intelligent Transport Systems, May 2015, Paris, France. hal-01262143

**HAL Id: hal-01262143**

**<https://hal.science/hal-01262143v1>**

Submitted on 15 Nov 2016

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Commuter's individual travel time estimation: using AFC data of Paris transit system

Xiaoyan Xie <sup>a,\*</sup>, Fabien Laurent <sup>a</sup>, and Vincent Aguiléra <sup>a,b</sup>,

<sup>a</sup>Université Paris Est, Laboratoire Ville Mobilité Transport, Ecole des Ponts ParisTech, Marne-la-Vallée, France

<sup>b</sup>Cerema, 110 rue de Paris, Provins, France

\*Corresponding authors: [xiaoyan.xie@enpc.fr](mailto:xiaoyan.xie@enpc.fr)

## Introduction

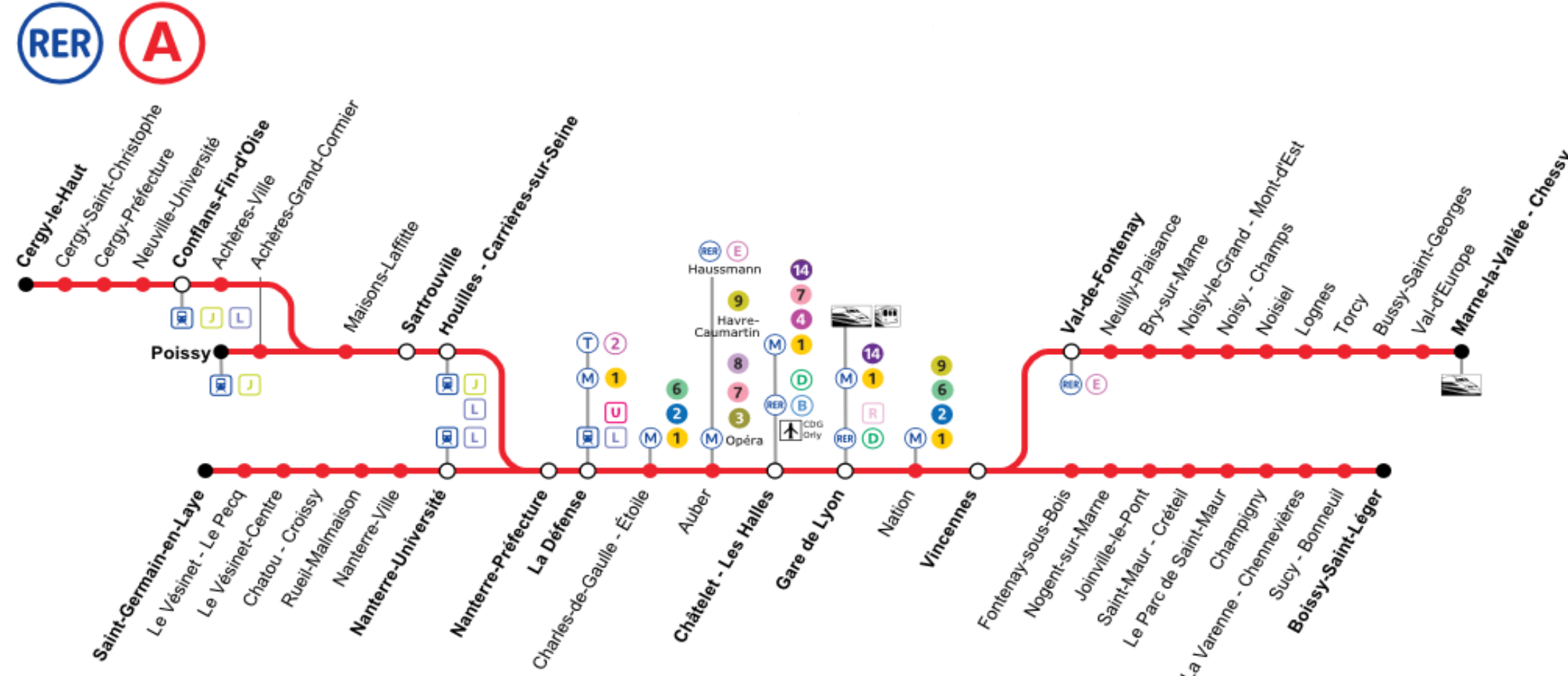
Automated Fare Collection (AFC) data of public transport details more passenger behavior information than manual survey. Based on a AFC dataset of line RER A of Paris transit system, commuter's Origin-Destination (O-D) pair has been generated by two loops. Passenger's gate-to-gate individual travel time between access station and egress station is derived directly from the generated O-D pair.

**Key words:**  
public transport,  
travel time,  
AFC data,  
O-D inference.

## AFC Data and Data Processing

### Case study and AFC data presentation

(i) RER A services's topological structure



(ii) Paris AFC system records commuter's farecard number which is anonymous, validation moments at access/egress stations' gates, the date, access/egress stations (including transfer station), and station serial number.

### Data processing

Data processing is realized by **3 principal steps** combining TCL and SQLite:

#### Step 1: Data filtering

Extraction of all records of line RER A from mixed data.

#### Step 2: O-D pair inference

**Loop 1:** through each farecard number  
**Loop 2:** through each trip

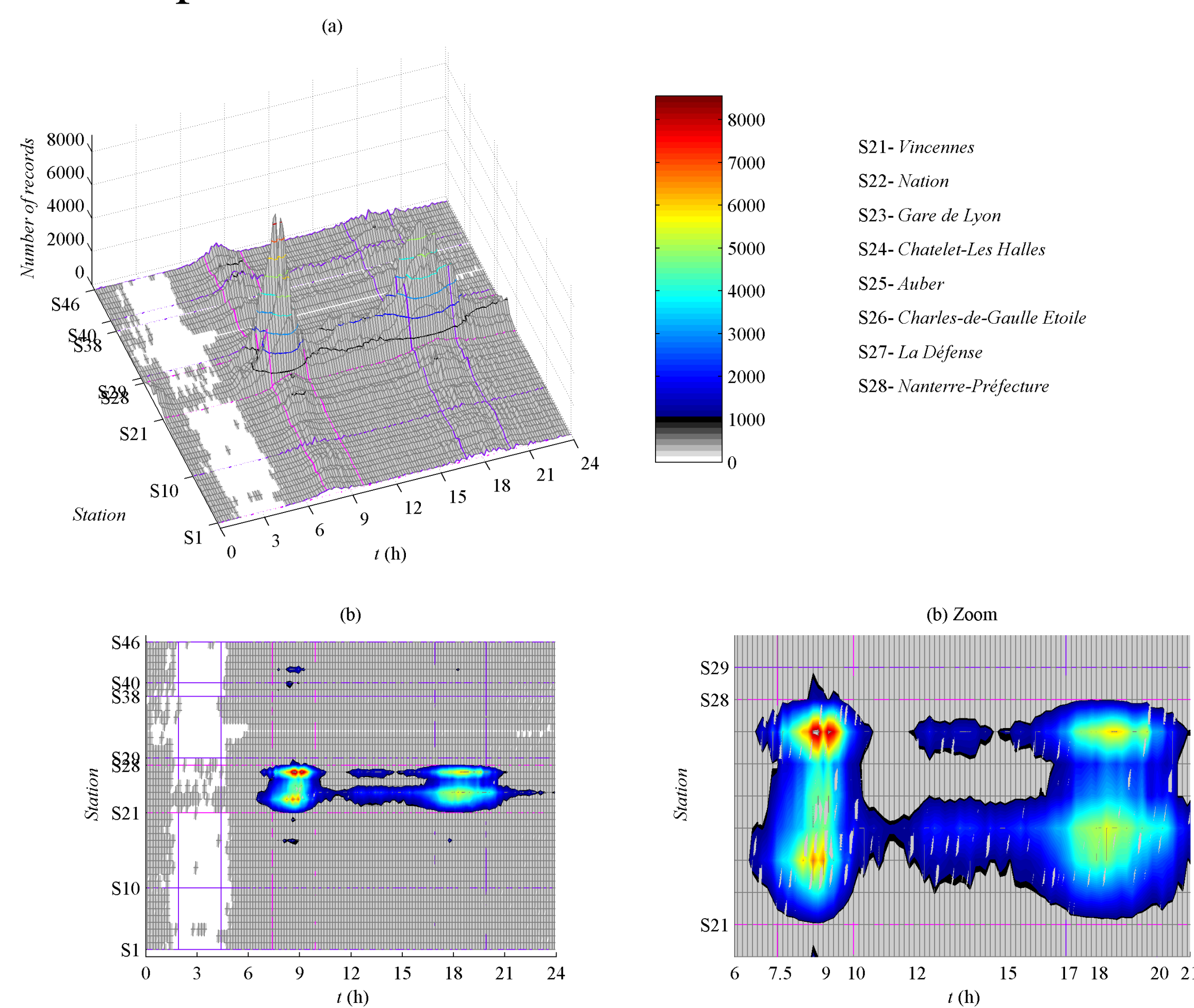
#### Step 3: Travel time calculation

$TT = t_j - t_i$   
where  $t_i$  validation time at access station gate;  
 $t_j$  validation time at egress station gate.

## Results

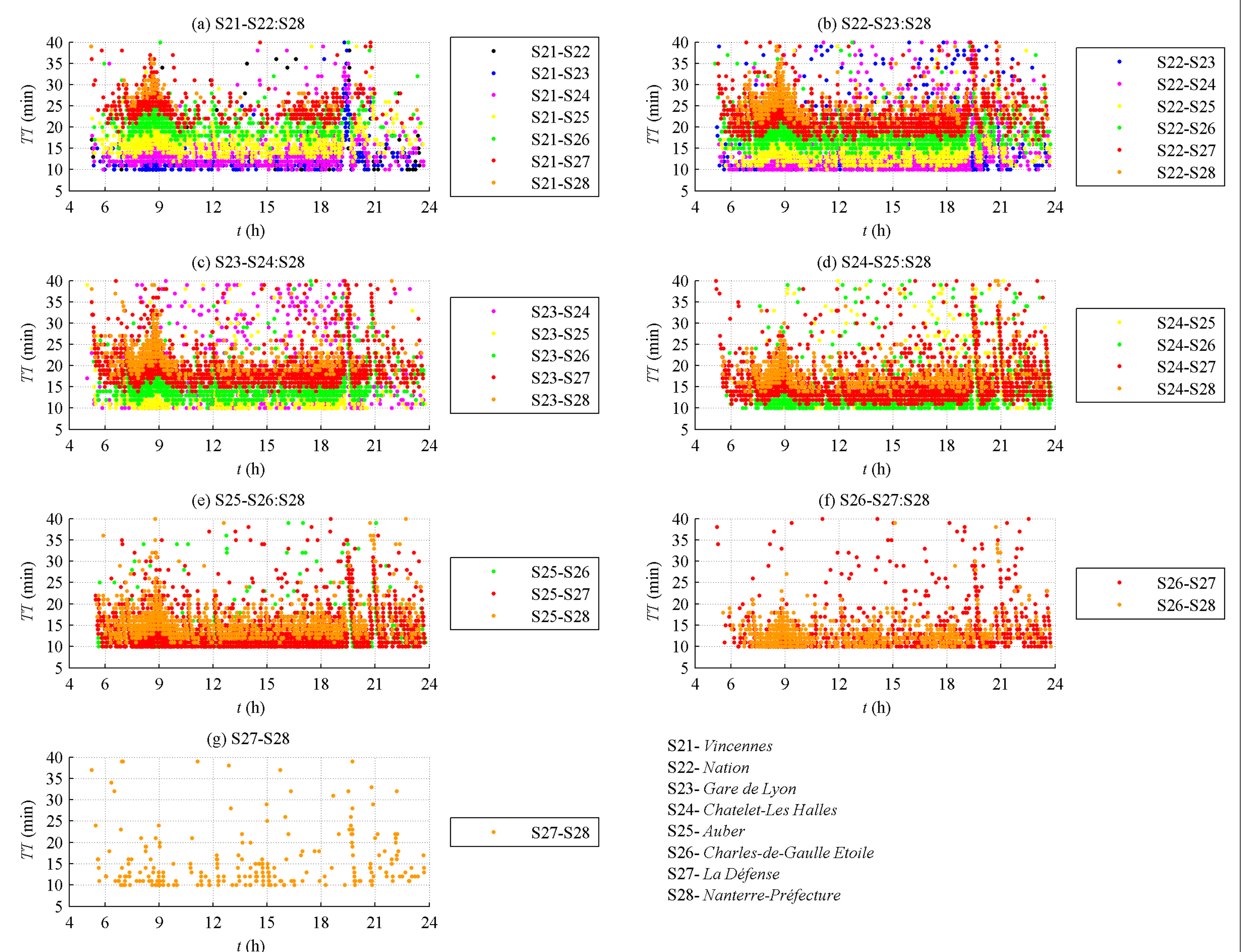
### Extracted data of RER A

About 2 millions records of about 750 thousands distinct farecard numbers on April 7<sup>th</sup> 2011.



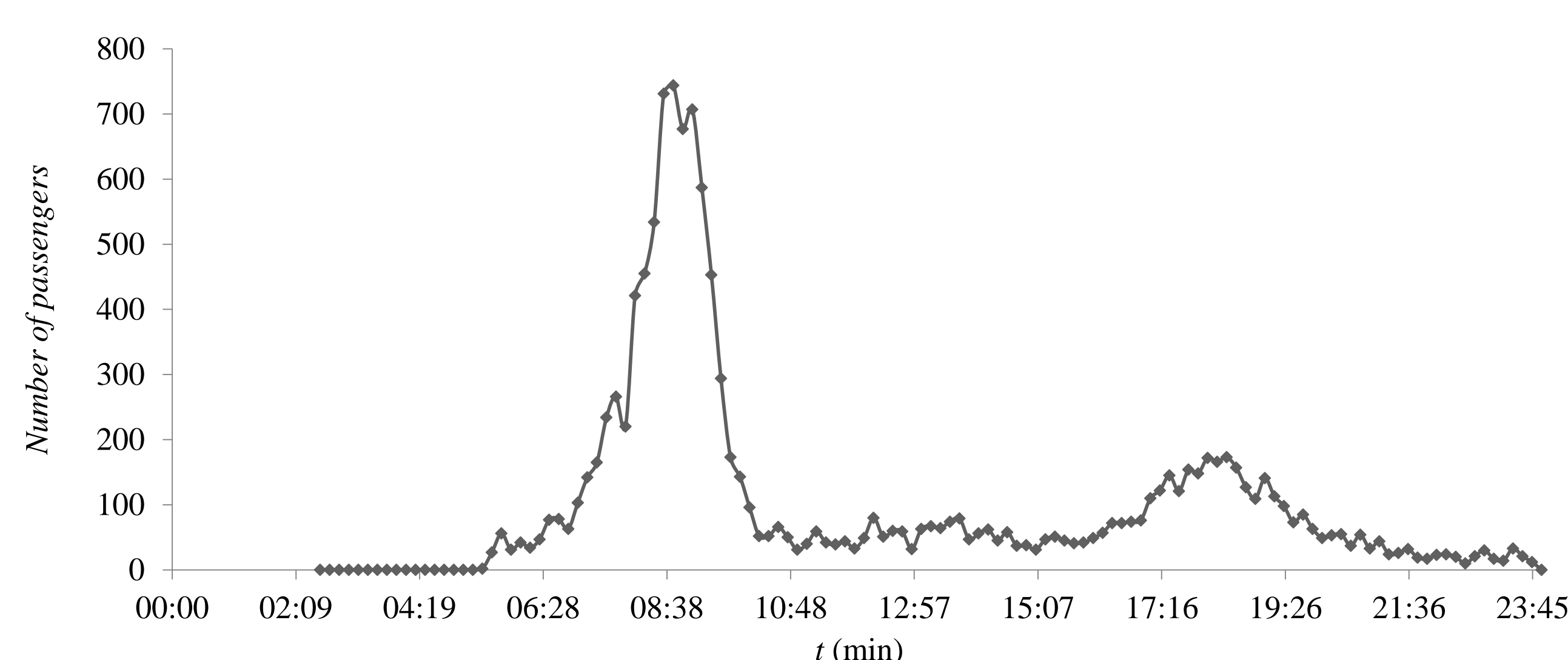
### Passenger's gate-to-gate individual travel time

Passenger's gate-to-gate individual travel time ( $TT$ ) distribution between each O-D pair of central segment of line RER A from Vincennes to Nanterre-Préfecture between 5h00 and 24h00.



### Demand distribution of one O-D pair

Passenger's demand from Auber to La Défense.



- $TT$  increases not only with travel distance, but also with the demand while the network is congested during peak hours;
- The morning  $TT$  peaks are sharper than evening ones;
- The dispersion of the  $TT$  points motivate a further study about the error sources of input AFC data.

## Conclusions and perspectives

Passenger's gate-to-gate individual travel time ( $TT$ ) between each O-D pair has been estimated for high frequency congested transit line, line RER A of Paris transit system, by using AFC data. AFC data provides passenger behavior information at microscopic level. The growing of  $TT$  of each O-D pair displays different traffic congestion conditions during morning and evening peak hours.

This work could be continue in two directions: (i) a study about the error sources of input AFC data; and (ii) an statistical analysis on output data.

### Acknowledgements:

This research was founded by the 'Chaire STIF-ENPC'.