



**HAL**  
open science

## Estimating pollutant emissions from aggregated traffic variables: the influence of data sources and sampling methods

Delphine Lejri, Nor Amimy Ibrahim, Cécile Becarie, Ludovic Leclercq

### ► To cite this version:

Delphine Lejri, Nor Amimy Ibrahim, Cécile Becarie, Ludovic Leclercq. Estimating pollutant emissions from aggregated traffic variables: the influence of data sources and sampling methods. ISTS 2014, International Symposium of Transport Simulation, Jun 2014, Ajaccio, France. 10p. hal-01922820

**HAL Id: hal-01922820**

**<https://hal.science/hal-01922820v1>**

Submitted on 14 Nov 2018

**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.

## **Estimating pollutant emissions from aggregated traffic variables: the influence of data sources and sampling methods**

**Delphine Lejri, Nor Amimy Ibrahim, Cécile Bécarie, Ludovic Leclercq**

*Laboratoire Ingénierie Circulation Transport LICIT*

*IFSTAR / ENTPE – Université de Lyon*

*Rue Maurice Audin, F-69518 Vaulx-En-Velin*

*E-mail: delphine.lejri@entpe.fr*

### **Abstract**

This paper investigates the sensibility of pollutant estimations to errors when determining the traffic states at large urban scale. Two key figures are required to estimate traffic-related pollutants at this scale: the mean speed with the considered region and the travel production. Loop detectors and probes can be used to determine these variables. In this paper, the estimation methods for these variables and the influence of the sampling rate will be questioned. The analysis revealed that (i) probes data provide accurate results even for low sampling rate and that (ii) the precision of travel production's estimate is mainly linked to the precision of the sampling ratio's assessment. The latter will be approximated combining information from probes and loops.

**Keywords:** pollutant emission, estimation method, loop detector, probe vehicle, sampling

### **1. Introduction**

Emissions due to road traffic are known to make large contribution to air pollution in urban areas and exposure to pollutants concentrations has been demonstrated to have detrimental impacts on human health (WHO, 2011 and Ren & al., 2008). In the last few years, many efforts have been made to quantify and reduce the contribution of greenhouse gas emissions and other pollutants from transportation. Indeed traffic management systems at large urban scale are not only meant to improve transportation systems, but also to reduce their environmental impact. To attempt this objective, the assessment of the pollutant emissions requires a satisfying knowledge of traffic-related data. The key issue at stake is to describe accurately the traffic dynamic and especially the congestion periods.

Classical methods for assessing traffic-related pollutants emissions are based on an aggregated description of the vehicle behavior (André & al., 2000 and EMEP-EEA, 2013). Emission model like COPERT (COPERT 4, 2007) only need mean speed and travel

production (total distance travelled by vehicle for a given period of time) to estimate the related emissions. Indeed, the total emissions are calculated as the product of the “activity” (travel production) and the unitary emission factors (elementary emission for a vehicle technology or a vehicle fleet) that are depending on the mean speed.

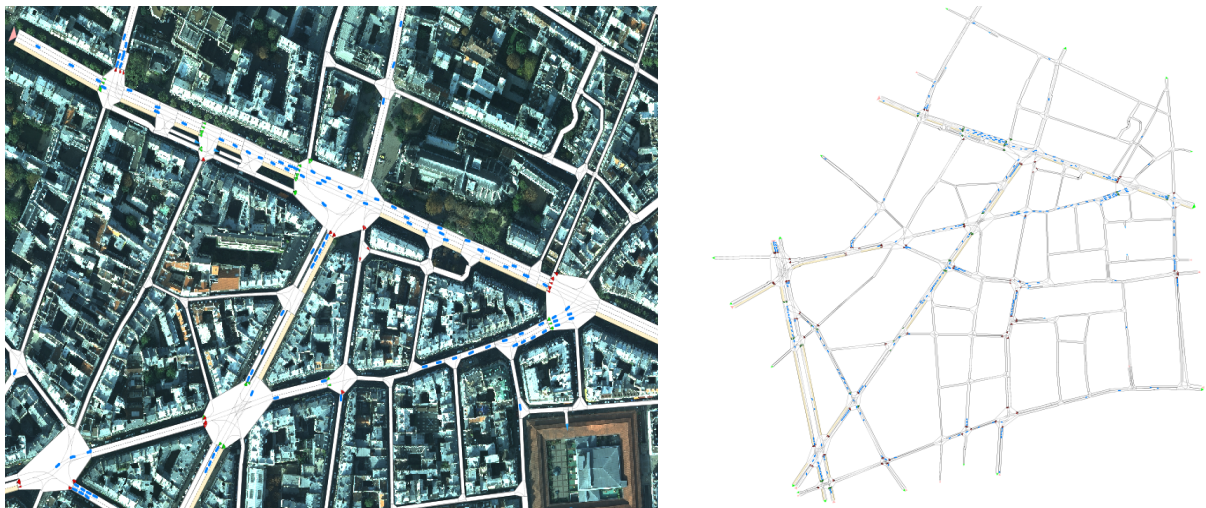
However, when the traffic description comes from a static approach, the consequences of traffic congestion are often misestimating. Congestion periods induce important speed reduction and may also modify the travel production due to dynamic re-routing. It is then important to get accurate estimates for both traffic mean speed and travel production when calculating the emissions. Dynamic traffic simulators can directly provide good estimates of these two variables but at the cost of extensive works for setting and calibrating the simulations. Traffic data from loops or probe vehicles can also be used to assess the variables. However, monitoring technics never provide a full picture of the traffic behaviour over the whole network because (i) loop detectors only give punctual information for equipped links and (ii) probe vehicles often correspond to a small subset of all vehicles.

The aim of this paper is to quantify the errors that occur when assessing mean speed and travel production from partial information and the associated errors on emission calculations. The objective is to draw some comparisons between several estimating methods of those two-targeted variables. Then we will assess what bias, due to an approximation on these traffic variables, is consequently introduced on the associated emissions.

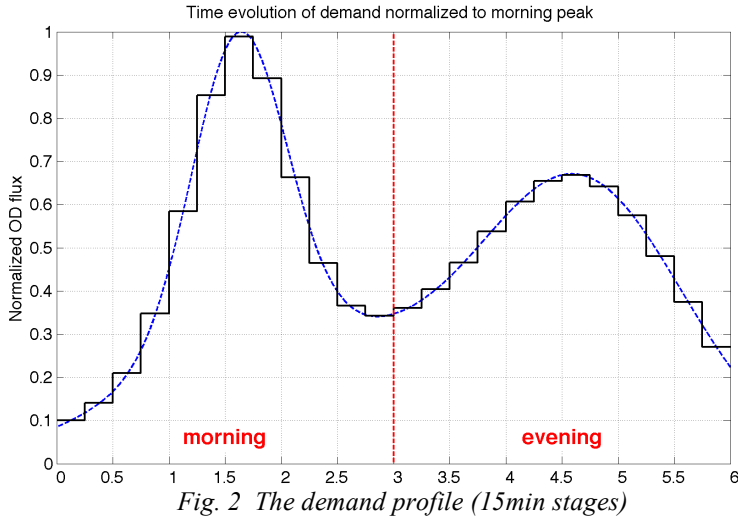
## 2. Method

### 2.1 Case study

The case study corresponds to the simulation of the 6<sup>th</sup> district of Paris (fig.1). The network in the dynamic traffic simulator, the Symuvia platform (Leclercq & al., 2007 and Laval & al., 2008), is composed of 234 sections and 93 crossroads. Passenger cars are the only vehicle class that have been modelled, even if reserved bus lanes and the restriction to the flow associated are implemented.



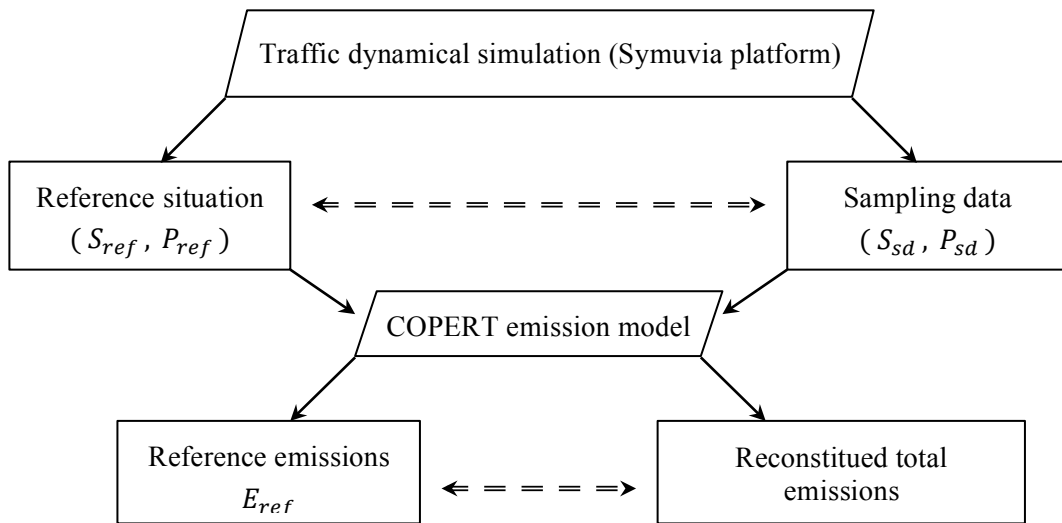
*Fig. 1 The 6th district of Paris traffic simulation.*



The demand profile synthesizes the most significant six hours of a typical day and especially the onset and the offset of congestion during the morning (evening) peak hour. As described on fig.2, the demand varies each 15 minutes and two different periods were used for route assignment.

Here, we aim to determine the total pollutant emission for the whole district every 15 min. This means that our indicators are spatially aggregated on the entire quarter; whereas they are temporally aggregated every 15min. Twenty-four periods of 15min will then describe the variables in more or less congested conditions.

### 2.2 The modelling chain



The proposed methodology is the following. Mean speed and travel production will be gained through a dynamic traffic simulator, the Symuvia platform. This simulation provides the trajectories for all vehicles. With these data, we can properly calculate the aggregated traffic variables by applying Edie's definitions (Edie, 1963). We then obtain the mean reference speed  $S_{ref}$  and the total travel production  $P_{ref}$  summing on all the vehicles  $i$  present in the simulation during the period  $T$ :

$$S_{ref}(T) = \frac{\sum di}{\sum ti} \quad (1)$$

$$P_{ref}(T) = \sum di \quad (2)$$

In order to assess exhaust emissions, the European emission model named COPERT has been used. The software provides emissions factors associated to the different technologies of vehicles, thanks to chassis dynamometer measurements.

The emission factors  $EF$  were determined for the French urban fleet of the year 2015 (IFSTTAR fleet, 2013) and various speed values (fig.4). The emissions are then calculated for each time period, as:

$$E_{\text{ref}}(T) = P_{\text{ref}}(T) \cdot EF(S_{\text{ref}}(T)) \quad (3)$$

In urban conditions the mean network speed is relatively low. On the 6<sup>th</sup> district of Paris, the 15min mean speed observed is between 9 and 21 km/h. In this speed range, the emission factors are highly variable. For instance the CO emission factor varies between 5 and 30% in this speed range if an error of 1km/h is committed on the mean speed (fig.4).

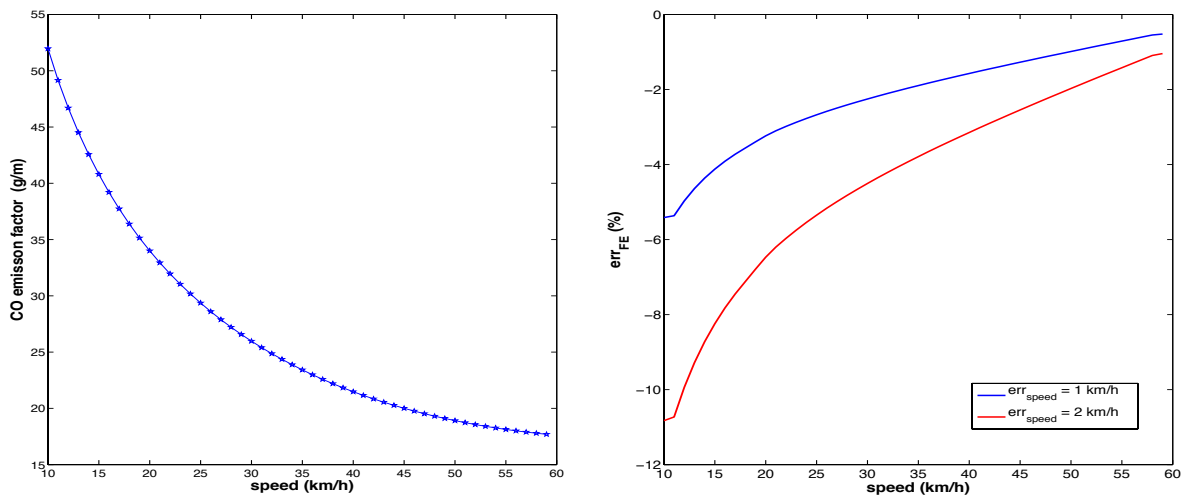


Fig. 3 (a) Evolution of the CO emission factor versus speed ; (b) Error on this EF when the error on speed is 1 or 2 km/h

### 2.3 The different sampling tests

In real world, such a fine traffic data is almost never available at large scale. And even if it were, the huge amount of data, when we consider the trajectories of all the vehicles, would conduct to a considerably long calculation time. This approach could be considered too complex when the stake is to assess air pollution.

Another approach is to use partial data: (i) sample the trajectories (representing probe data) or (ii) sample the equipped links (representing loop detectors). We will focus on the first case. A small subset of vehicles is chosen and these sampling traffic data, will be used to determine mean speed and travel production.

#### Constant sampling rate

First, we looked at the case where the sampling data are randomly selected to correspond to a fixed and exogenously known rate of all the vehicles present in the  $i^{\text{th}}$  time period. It is already known (Gayah & al., 2013 and Leclercq & al., 2014) that probes are very efficient even for low sampling rates inferior to 10% when calculating the mean network speed. However, probes are useless when calculating the travel production expect if the sampling ratio is known and steady enough.

### Variable sampling rate

In reality, the sampling rate is unknown and variable over the time periods. However, this ratio is essential to scale up the total travel production and to estimate the total emissions. Different tests have been performed to see the influence of the uncertainties on this particular ratio on the emission calculations.

We also suggest looking at different monitoring scheme based of virtual loops and/or probes to get estimates for these two variables and compare with the reference. This part aims to determine the optimal sampling rates and the optimal mix between loop and probe information to derive accurate estimations of the two targeted traffic variables. Discrepancies will be studied in terms of related errors on emission calculations. In this paper, we will only present the CO emissions' results. However, we also looked at NO<sub>x</sub>, PM<sub>2.5</sub> and PM<sub>10</sub> emissions and the trends are the same.

## 3. Results

### Constant sampling rate

In this case, the sampling rate is constant over the 15min periods and a priori known. For each period we draw randomly a percentage of the present vehicles and we estimate the mean speed of this subset of vehicles thanks to Edie's definition. The numerical experiments were reconducted 100 times. The results of all these replications are shown on fig. 4.

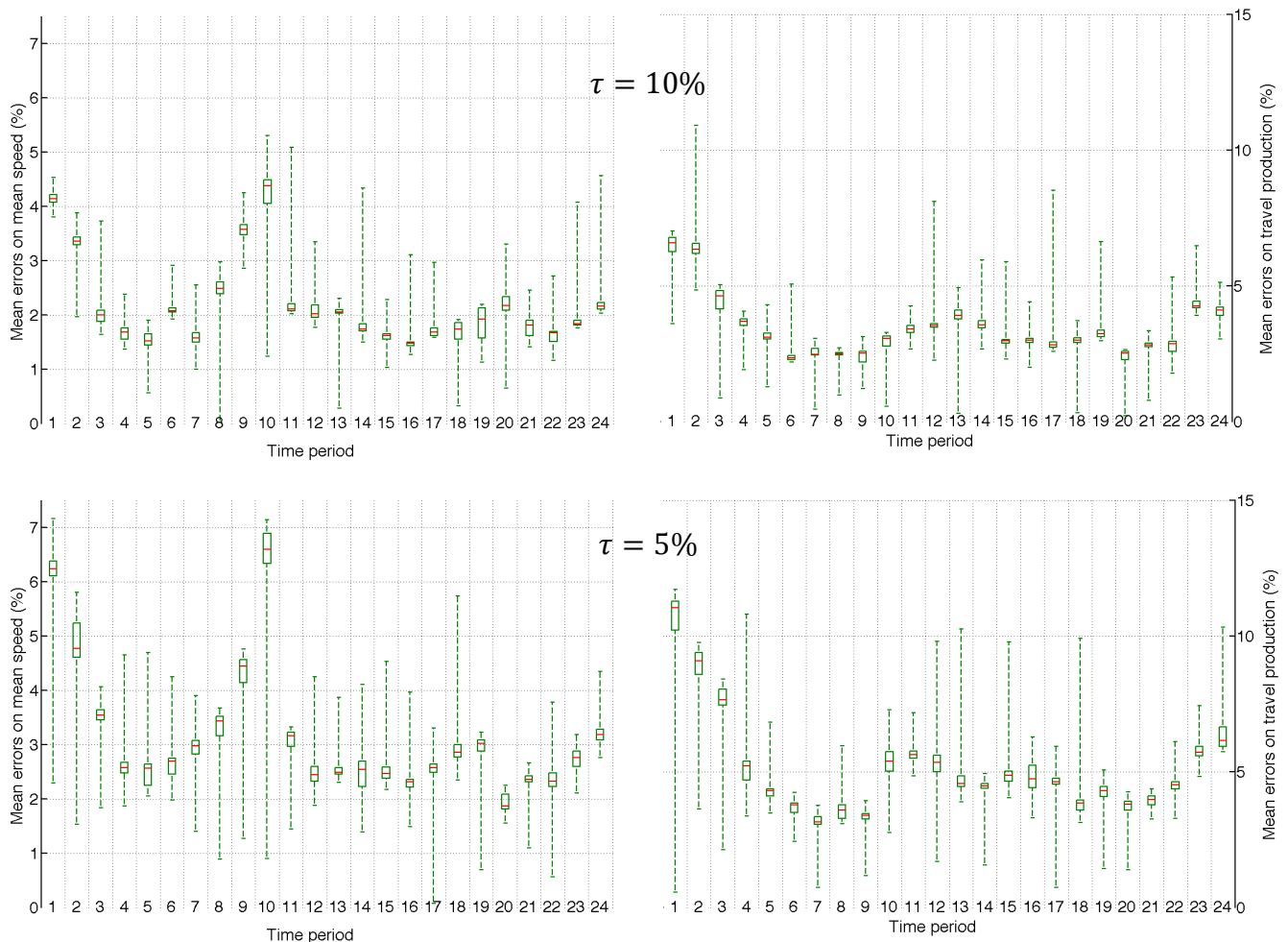


Fig. 4 Mean errors on mean speed (left) and travel production (right) with a sampling rate of 10 and 5%

When the sampling rate is constant over the periods and known, the results are very satisfying even if the ratio is lower than 10%. This is especially the case for mean speed for which the mean errors are less than 4.4% (respectively 6.6%) for a sampling rate of 10% (respectively 5%). Besides, the estimate of the travel production is more deteriorated when the sampling rate decreases and the maximum mean error on travel production reaches 11% for  $\tau = 5\%$ .

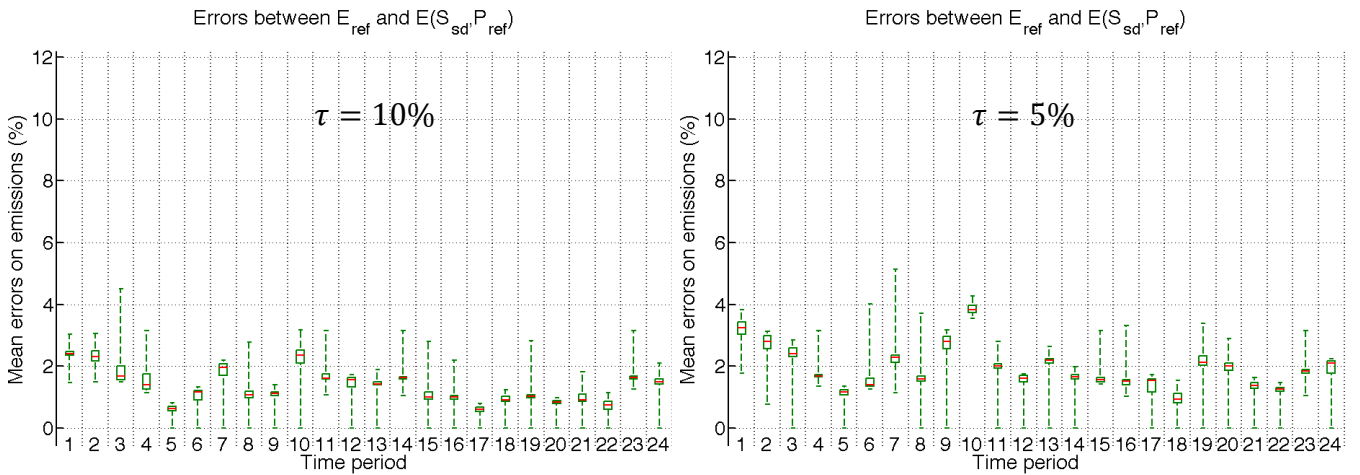


Fig. 7(a)

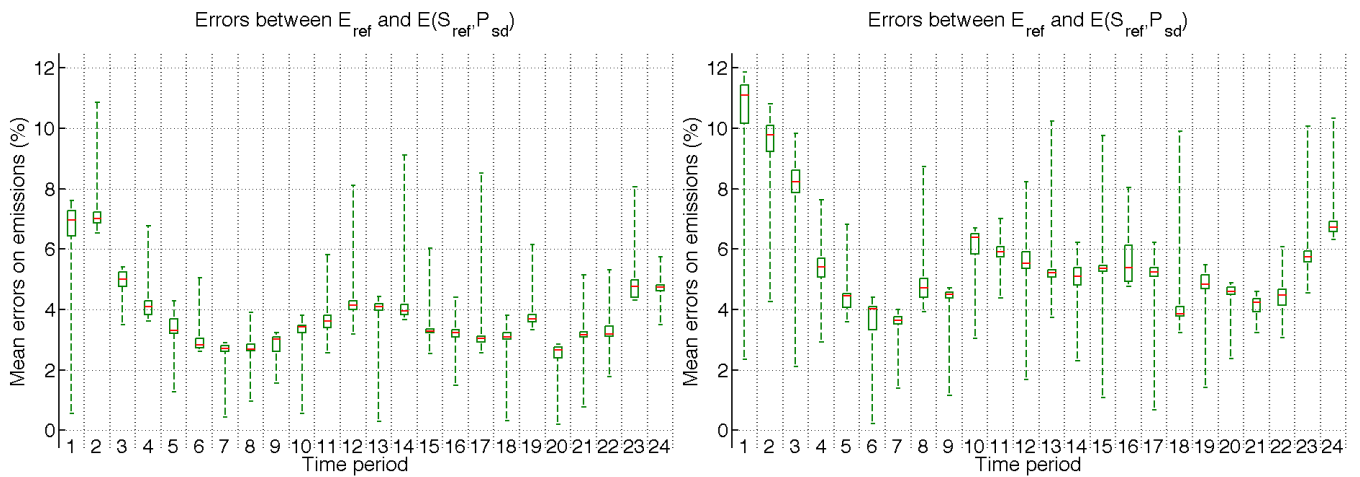


Fig. 7(b)

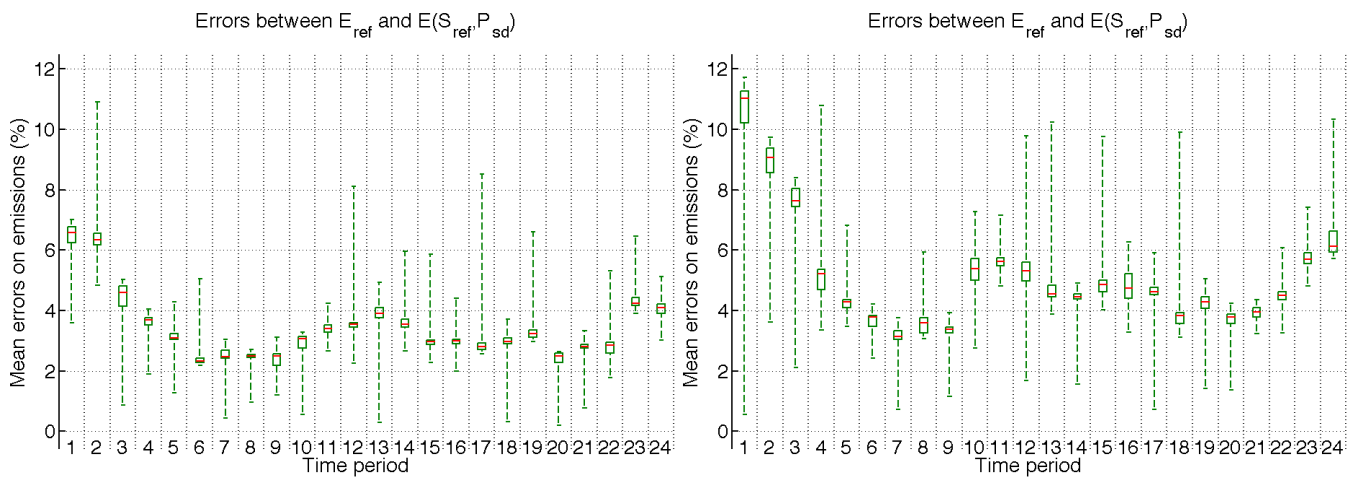


Fig. 7(c) Mean errors on emissions

When it comes to the impact of these errors on emission calculations, we can see in fig. 5 that the effect is not symmetric between the two variables. The bias due to an error on mean speed induces errors lower than 2.4% (respectively 3.8%) for a sampling rate of 10% (respectively 5%). This reduced value, in comparison to the mean errors on mean speed, is due to the fact that we used emission factors defined by 1km/h step. This means that the smallest errors have no influence on the EF values. On the contrary, the bias due to the estimation of the travel production (fig.5 (b)) represents almost the total error on emission calculations (fig. 5(c)), which is less than 6.6% (respectively 11.1%) for a sampling rate of 10% (respectively 5%).

**Variable sampling rate**

We now consider that the ratio variable over the 15min periods. For each period, we draw randomly a ratio value around 10% thanks to a normal distribution, so that the sampling rate remains between 8 and 12%.

In that case, there is no major impact on the assessment of mean speed (fig.6). When it comes to travel production, we need an estimate of the sampling rate to scale up the values associated to the probes. Figure 7 shows the results with the exact sampling rate and with a mean sampling rate, which is here 10%. With this last value, the results are deteriorated and the mean error can reach 18%.

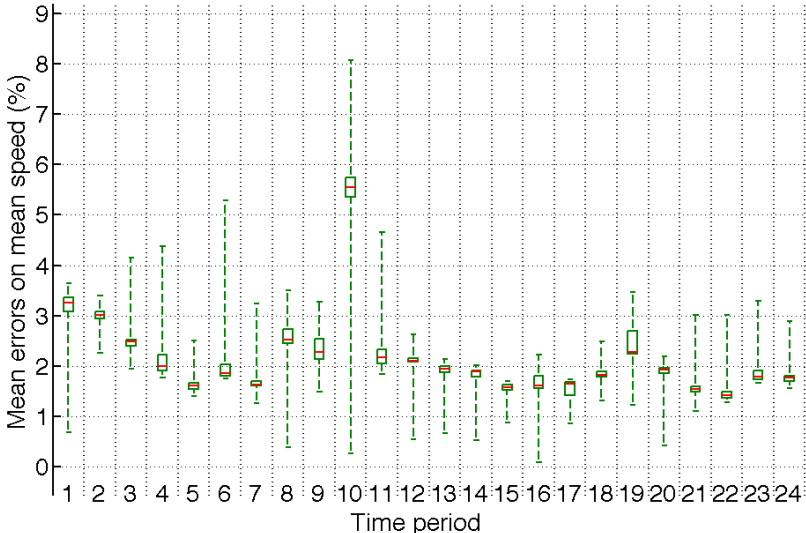


Fig. 6 Mean errors on mean speed with a variable sampling rate

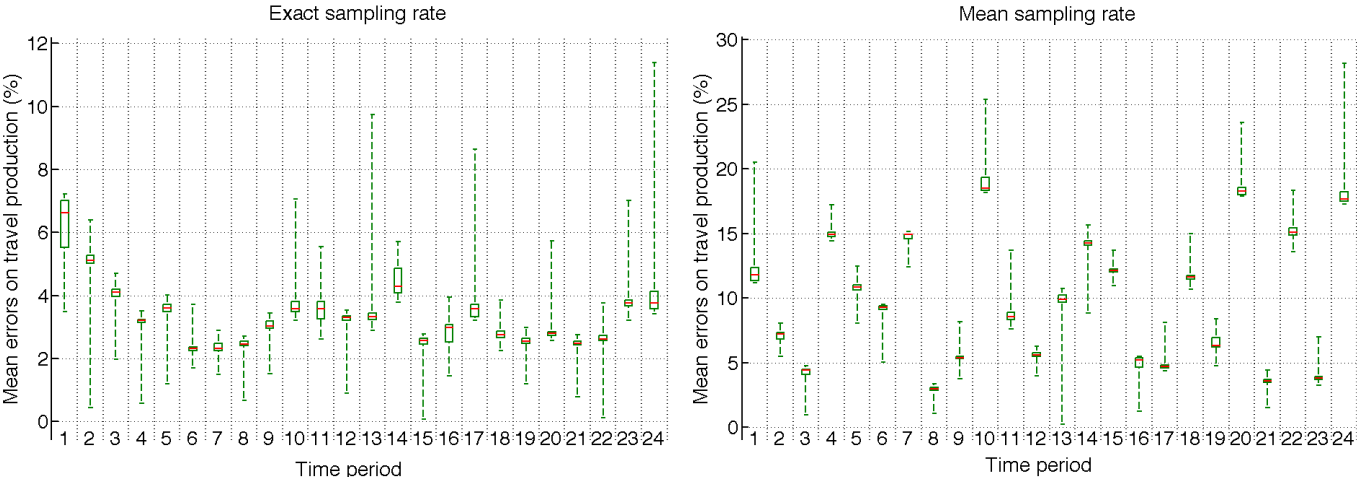


Fig. 7 Mean errors on travel production with a variable sampling rate



We can look at the impact of such bias on macroscopic traffic variables on CO emissions calculations. The conclusions are the same (fig.8); the results are deteriorated with the mean sampling rate and the mean error can reach 18%.

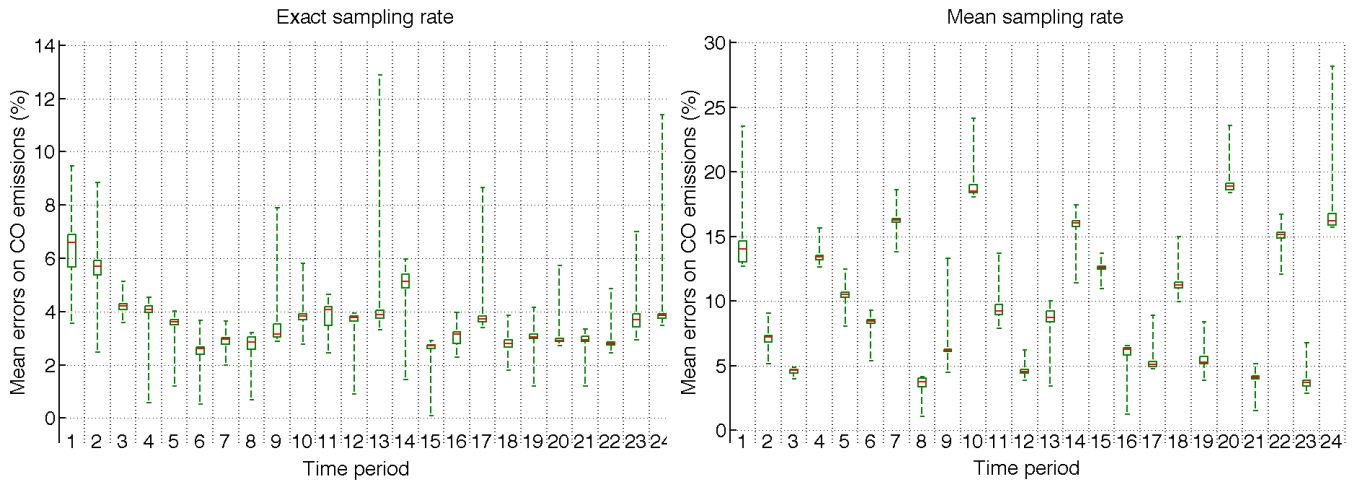


Fig. 8 Mean errors on CO emissions with a variable sampling rate

In order to ameliorate these results, we suggest combining the information available on loop detectors with our probes data. Our purpose is to apply a fishing method (Geroliminis & al., 2008) to derive from this new data source, a better estimate of the sampling rate  $\tau^*$ .

There is a loop detector on each link in the simulation. We then drew randomly a small subset of these loop detectors and looked at the number of probes seen by the  $k$  detectors ( $N_{probes}^k$ ) and the total number of vehicles seen by the same ones ( $N_{vehicles}^k$ ).

$$\tau^*(T) = \frac{\sum_k N_{probes}^k(T)}{\sum_k N_{vehicles}^k(T)} \quad (4)$$

We compared the results obtained with 100%, 5% and 2% of the loop detectors (fig.9).

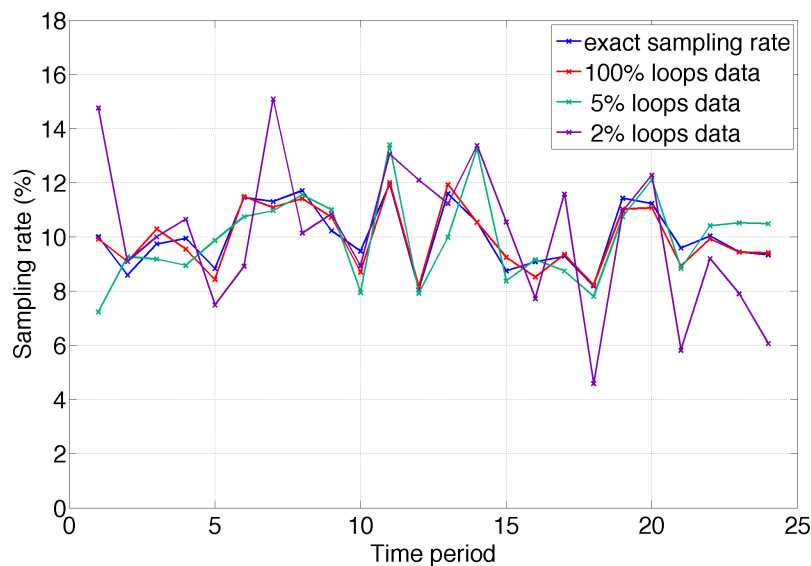


Fig. 9 Estimation of a variable sampling rate with loop detectors.

On first approach a low percentage of loop detectors chosen randomly does not totally improve the results on CO emissions. Indeed the errors are still significantly important with 2% loops data and better, but still not satisfying with 5% loops data. The results need to be confirmed by a complete analysis of the data.

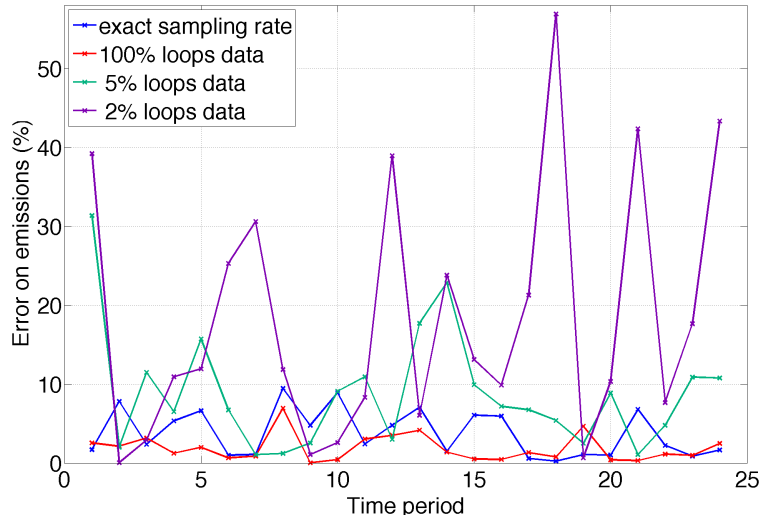


Fig. 10 Errors on CO emissions with a variable sampling rate and fishing method

#### 4. Synthesis

In this paper we aimed to use traffic data from loops or probe vehicles in order to assess traffic variables required to estimate the related emissions: mean speed and total travel production at a quarter scale (6<sup>th</sup> district of Paris) and for each 15min periods.

We saw that the mean speed can be accurately estimated with a low percentage of probe data. On the other hand the travel production can be also accurately assessed at the condition that we precisely estimate the sampling rate. For this we suggested to use a fishing method on loops detectors. The results have to be confirmed and can certainly be improved by adding some criteria for the loops detector's choice.

This work gave the opportunity to understand better what is at stake in the dynamical traffic variables estimation used for pollutants assessments. The results can be improved considering looking at the optimal mixed information between probes and loops data. Further, the robustness of the method would be tested in taking into account the bias introduced by measurements errors.

#### Acknowledgments

The authors particularly thank Daniel Villegas for his help throughout the writing of this paper.

## References

André M., Hammarström, U. (2000). Driving speeds in Europe for pollutant emission estimation. . In: *Transportation Research Part D*, Vol.5, pp. 321–335.

COPERT 4 Computer programme to calculate emissions from road transport. User manual. Dimitrios Gkatzoflias, Chariton Kouridis, Leonidas Ntziachristos and Zissis Samaras. December 2007

Geroliminis N, Daganzo C. (2008). Existence of urban-scale macroscopic fundamental diagrams: Some experimental findings. In: *Transportation Research Part B: Methodological*, Vol. 42, No. 9, pp 759-770

Edie, L.C., (1963). Discussion of traffic stream measurements and definitions. In: Almond, J. (Ed.), *Proceedings of the 2nd International Symposium on the Theory of Traffic Flow*. OECD, Paris, France, pp. 139–154.

EMEP-EEA air pollutant emission inventory guidebook 2013 "Part B1 sectoral guidance-energy"

Gayah, V.V. , Dixit, V.V. (2013), Using mobile probe data and the macroscopic fundamental diagram to estimate network. . In: *Transportation Research Record 2390*, pp. 76-96

Laval, J. A. and L. Leclercq (2008), Microscopic modeling of the relaxation phenomenon using a macroscopic lane-changing model. . In: *Transportation Research Part B: Methodological*, Vol. 42, No. 6, pp. 511 – 522.

Leclercq, L., Chiabaut, N., Trinquier, B. (2014), Macroscopic Fundamental Diagrams: A cross-comparison of estimation methods. In: *Transportation Research Part B: Methodological*, Vol. 62, pp. 1–12

Leclercq, L., J. Laval, and E. Chevallier (2007), The Lagrangian coordinate system and what it means for first order traffic flow models. . In: *Proceedings of the 17th ISTTT*, pp. 393–415.

Ren C., Tong S. (2008) Health effects of ambient air pollution - recent research development and contemporary methodological challenges. In: *Environmental Health*, Vol. 7, p. 56

WHO (2011), World Health Organization *Fact sheet* N°313. Air Quality and Health Updated September 2011 <http://www.who.int/mediacentre/factsheets/fs313/en/index.html>