

# Data fusion algorithms for Density Reconstruction in Road Transportation Networks

Enrico Lovisari, Carlos Canudas de Wit, Alain Y. Kibangou

► **To cite this version:**

Enrico Lovisari, Carlos Canudas de Wit, Alain Y. Kibangou. Data fusion algorithms for Density Reconstruction in Road Transportation Networks. 54th IEEE Conference on Decision and Control (CDC 2015), Dec 2015, Osaka, Japan. <hal-01185523>

**HAL Id: hal-01185523**

**<https://hal.archives-ouvertes.fr/hal-01185523>**

Submitted on 18 Sep 2015

**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.

# Data fusion algorithms for Density Reconstruction in Road Transportation Networks

Enrico Lovisari, Carlos Canudas de Wit, and Alain Y. Kibangou

**Abstract**—This paper addresses the problem of density reconstruction in traffic networks with heterogeneous information sources. The network is partitioned in cells in which vehicles flow from their origin to their destination. The state of the network is represented by the densities of vehicles in each cell. Density estimation is of crucial importance in future Intelligent Transportation Systems for monitoring, control, and navigation purposes. However, deploying fixed sensors for this purpose can be very expensive. Therefore, most of fixed sensors networks are rather sparse. On the contrary, recent technologies have enormously increased the availability of relatively inexpensive Floating Car Data. A data fusion algorithm is then proposed to incorporate the two sources of information into a single observer of density of vehicles. The efficiency of the proposed algorithm is shown in a real scenario using data from the Grenoble Traffic Lab fixed sensor network and INRIX Floating Car Data on the Rocade Sud in Grenoble.

**Index Terms**—Road Transportation systems; Dynamical flow network; Density reconstruction; Floating Car Data.

## I. INTRODUCTION

The increase of the number of vehicles observed in the last decades has steered crucial roads towards a state of near saturation, exhibiting on a daily basis periods of congested traffic [1], in turn causing severe economical and environmental losses. Built upon recent technological and theoretical advancements, Intelligent Transportation Systems (ITSs) are expected to provide robust techniques for real-time monitoring, prediction and actuation of traffic networks.

This paper studies the problem of estimating the density of vehicles in traffic networks. Such an information allows to forecast travel time and traffic evolution, therefore providing navigation suggestion to drivers and allowing traffic managers to control the network through traffic lights, ramp metering and speed limits [2], [3], [4], [5].

The main tool to estimate the state of the network are fixed sensors, such as induction loops or magnetometers. These devices are able to 1) count the number, 2) estimate the density, and 3) measure the average speed of the vehicles that cross a certain section of road in a period of time.

In addition, the spread of wireless devices allows new sensing and communication capabilities. In particular for the traffic application, any vehicle equipped with GPS devices can act as a probe in the traffic and provide Floating Car Data (FCD). If a non negligible fraction of vehicles acts as probe, the collected data provides an estimate of the evolution of

speed in the network. Due to privacy reasons, single vehicles traces are usually not directly used, but rather aggregated as average speed of vehicles in segments of road. Advanced methodologies, such as the one employed by INRIX, ensure a very fine spatial partition of the network, with segments as short as 250 meters ([6]). Compared to fixed sensors, this technology is less precise, but since it exploits existing communication systems it is relatively less expensive and already covers all major traffic networks.

In this paper we propose an algorithm that aims to reconstruct the traffic density by fusing fixed sensors measurements and Floating Car Data. We employ a macroscopic model, partitioning the network in cells and assigning to each cell a density of vehicles. The latter evolves dynamically according to a first order mass-conservation law.

The most celebrated macroscopic traffic model is the PDE based Lighthill-Whitham and Richards (LWR) model [7], which is able to reproduce crucial phenomena such as traffic shock waves, and a well known discretization of the LWR-PDE is the Cell Transmission Model (CTM) [8]. Huge efforts have been put in the last 15 years to calibrate the CTM [9] and to unveil its system-theoretical properties [10]. Fusion of flow, density and speed measurements, mostly considering single vehicles traces, has also been addressed. Approaches range from signal processing techniques such as the generalized Treiber-Helbing filter [11], to nonlinear versions of the Kalman filter [12], to stochastic versions of the three-detector model [13].

Our approach inherits from the CTM the cell-based topology, but we do not directly employ the resulting dynamical model. Instead, inflows and outflows are estimated on the basis of the available flow measurements only, and speed measurements are employed to compute a pseudo-measurement of the density. These quantities are the inputs for the density observer. In addition, we propose a gradient descent method to calibrate the Fundamental Diagram, and we implement the proposed solution using real fixed sensor measurements from the Grenoble Traffic Lab [14], a sensing network deployed along the Rocade Sud in Grenoble, France, and speed FCD measurements provided by INRIX, one of the most well known traffic solutions companies.

The remainder of the paper is organized as follows: Section II formulates the problem and details the available measurements. Section III describes the proposed solution, while Section IV presents our numerical experiments. Finally, Section V draws the conclusions and presents several future research directions.

The authors are with Univ. Grenoble Alpes, Gipsa-Lab, with CNRS, Gipsa-Lab, and with Inria Grenoble Rhône-Alpes, F-38000 Grenoble, France [enrico.lovisari](mailto:enrico.lovisari), [carlos.canudas-de-wit](mailto:carlos.canudas-de-wit), [alain.kibangou@gipsa-lab.fr](mailto:alain.kibangou@gipsa-lab.fr). This work has been partially supported by the EU project SPEEDD FP7 n. 619435.

### A. Notation

The symbols  $\mathbb{R}^n$ ,  $\mathbb{R}_+^n$  and  $\mathbb{R}^{n \times m}$  denote the sets of real and positive real valued vectors of dimension  $n$ , and of real valued matrices of dimension  $n \times m$ , respectively. For a finite set  $\mathcal{A}$ , the symbol  $\mathbb{R}^{\mathcal{A}}$  denotes the set of vectors indexed by elements of  $\mathcal{A}$ , and  $|\mathcal{A}|$  its cardinality. The transpose of  $A \in \mathbb{R}^{n \times m}$  is  $A^T$ . For a vector  $x \in \mathbb{R}^n$ ,  $x \geq 0$  is meant component-wise. The symbol  $I$  denotes the identity matrix of suitable dimension.

A graph  $\mathcal{G}$  is a pair  $(\mathcal{V}, \mathcal{E})$  where  $\mathcal{V}$  is called the set of nodes and  $\mathcal{E}$  the set of edges. Edges are equipped with two functions  $t : \mathcal{E} \rightarrow \mathcal{V}$  and  $h : \mathcal{E} \rightarrow \mathcal{V}$ , the head and tail functions, respectively, so that  $e$  is the directed edge between  $t(e)$  and  $h(e)$ . We allow for parallel edges, namely, edges having the same head and tail, but not for loops, namely, edges whose head and tail coincide. A path of length  $n \geq 2$  is a sequence of edges  $e_1, \dots, e_n$  that are consecutive, namely, such that  $h(e_i) = t(e_{i+1})$  for  $i = 1, \dots, n-1$ . A path of length 1 is a path made of a single link.

## II. PROBLEM FORMULATION

A Transportation Network is a graph  $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ , where junctions  $v \in \mathcal{V}$  are interfaces between cells  $e \in \mathcal{E}$ , and multi-lane roads are partitioned in sequences of parallel cells. Cells that carry vehicles in and out from the network are called onramps and offramps, respectively, and each cell  $e$  lies on the path from an onramp to an offramp. Time is discrete and slotted in intervals of length  $T > 0$ . On each cell  $e \in \mathcal{E}$  we denote by  $\rho_e(t)$  the density of vehicles, in vehicles per km (veh/km), during the  $t$ -th time slot. The state of the network  $\rho(t) = [\rho_e(t)]_{e \in \mathcal{E}}$  changes dynamically in time according to the first-order model

$$\rho_e(t+1) = \rho_e(t) + \frac{1}{\ell_e} (f_e^{\text{in}}(t) - f_e^{\text{out}}(t)) \quad (1)$$

where  $\ell_e$  is the length of cell  $e$ , and  $f_e^{\text{in}}(t)$  and  $f_e^{\text{out}}(t)$  are the inflow and the outflow at cell  $e$  during the  $t$ -th time slot.

The fraction of vehicles on cell  $e$  that turns into the cell  $j$ , the splitting ratio from  $e$  to  $j$ , is denoted  $R_{ej} \geq 0$ . If  $j$  does not follow  $e$ , then  $R_{je} = 0$ , and  $\sum_{j \in \mathcal{E}} R_{ej} \leq 1$ , with strict inequality at offramps. By definition of splitting ratios, one has  $f_e^{\text{in}} = \sum_{j \in \mathcal{E}} R_{je} f_j^{\text{out}}$ , which, stacking inflows and outflows into vectors  $f^{\text{in}}$  and  $f^{\text{out}}$ , is written in matrix form

$$f^{\text{in}} = R^T f^{\text{out}} \quad (2)$$

where the matrix  $R = [R_{ej}]$  is the *matrix of splitting ratios*. In this paper we assume that  $R$  is fixed and predetermined.

As a Godunov scheme for the LWR-PDE model, the CTM and derived models postulate that the outflow  $f_e^{\text{out}}$  on cell  $e$  depends on the density of vehicles on  $e$  and on cells immediately following  $e$ . We adopt a different point of view and renounce to explicitly determine the flow-density relation. The sole assumption that we make is the standard

$$f_e^{\text{out}} = \rho_e v_e(\rho), \forall e \in \mathcal{E} \quad (3)$$

namely that the outflow from a cell is the product of density and average speed of vehicles on the same cell. While in

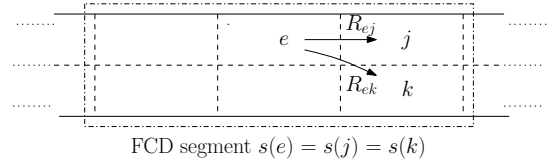


Fig. 1. A stretch of road partitioned in cells and FCD segments. Splitting ratios are shown from a cell  $e$  to following cells,  $j$  and  $k$ . A FCD segment including, among others, cells  $e$ ,  $j$  and  $k$ , is also shown.

the CTM the speed is a function of the density of vehicles around cell  $e$ , we will leave it unmodelled. The reason is that, as it will be clarified in the following, we assume having measurements of speed in each cell, so there is no need to model its dependence on the state of the network. We shall consider from now on the dynamics of the real system to be dictated by Eqs. (1)-(3), where  $v_e(\rho)$ ,  $e \in \mathcal{E}$ , is an unmodelled function of the local state of the network.

For data fusion and estimation purposes, we make a further simplifying assumption, writing

$$\varphi_e = \varphi_e(\rho_e), \forall e \in \mathcal{E} \quad (4)$$

where  $\varphi_e$  is the flow of vehicles *at the sensor locations*. The graph of  $\varphi_e(\cdot)$  is called the Fundamental Diagram on cell  $e$ , and it is usually supposed to be a concave function with  $\varphi(0) = \varphi(\rho^{\text{jam}}) = 0$ , where  $\rho^{\text{jam}}$  is the jam density (standard values for  $\rho^{\text{jam}}$  varying from 200 to 300 veh/km).

### A. Available measurements

As already mentioned, in this paper we consider a scenario in which heterogeneous measurements are available.

1) *Flow and density measurements*: Measurements of flows and of density are obtained via fixed sensors, which count the number (flow) and estimate the density of vehicles that, during a time slot, cross the section of road where they are placed. Without loss of generality and for sake of simplicity, we assume that a new set of flow and density measurements is available at every time slot, so we can write

$$\begin{aligned} \varphi_e^m(t) &= f_e^{\text{out}}(t) + \omega_e^\varphi(t), & e \in \mathcal{E}^m \\ \rho_e^m(t) &= \rho_e(t) + \omega_e^\rho(t), & e \in \mathcal{E}^m \end{aligned} \quad (5)$$

where  $\mathcal{E}^m \subseteq \mathcal{E}$  is the set of cells equipped with sensors,  $\varphi_e^m(t)$  and  $\rho_e^m(t)$  are flow and density measures at time  $t$ , and  $\omega_e^\varphi(t)$  and  $\omega_e^\rho(t)$  are measurement errors whose stochastic properties depend on the performance of the sensor as well as on road and weather conditions, with sources of noise ranging from temporary inability to detect changes of the magnetic field, too fast or too slow vehicles, etc.

2) *Speed measurements*: As mentioned, vehicles equipped with GPS devices can act as probes and communicate to ITSs Floating Car Data including their position and speed (traces). Since for privacy reasons traces of private vehicles cannot be used, roads are divided into segments, let  $\mathcal{S}$  denote the set of all segments, and data are aggregated as average speed per segment  $s \in \mathcal{S}$ . Notice that the resulting measurement is an average of the speed of a *fraction* of vehicles in the road.

Floating Car Data are obtained through existing communication architecture, making them less expensive than

fixed sensors and available on great part of the major roads. However, their precision is related to their penetration rate and therefore is not as high as fixed sensors, and moreover they cannot distinguish vehicles on different lanes. Finally, FCD are usually averaged over a relatively long period of time. For example, while the GTL sensor network provides measurement every 15 seconds, FCDs are aggregated by INRIX every minute, standard practice being around 5/10 minutes. We will assume that new speed aggregate data is available every  $N$  time instants, i.e., at times  $N, 2N, 3N, \dots$ , corresponding to the average speed in the periods  $[0, N - 1], [N, 2N - 1], \dots$ , respectively. As such, speed measurements can be formally written as<sup>1</sup>

$$v_e^{\text{FCD}}(t) = \begin{cases} v_e^{\text{ff}}, & t \in [0, N - 1] \\ v_{s(e)}^{\text{FCD}}(k), & t \in [kN, (k + 1)N - 1] \end{cases} \quad (6)$$

where

- $v_e^{\text{ff}} > 0$  is the freeflow speed on cell  $e$ , namely, the speed of vehicles in low density regime;
- $v_{s(e)}^{\text{FCD}}(k)$  is given by

$$v_{s(e)}^{\text{FCD}}(k) = \frac{1}{N|s(e)|} \sum_{j \in s(e), \tau \in \mathcal{I}_t} v_j(\tau) + \omega_{s(e)}^{\text{FCD}}(k)$$

where  $s(e)$  denotes the segment of which  $e$  is one of the cells (see Fig. 1),  $\omega_{s(e)}^{\text{FCD}}(k)$  is a measurement error whose stochastic properties depend on the performance of the sensor as well as on road and weather conditions, and  $\mathcal{I}_t = \{\tau : \lfloor \frac{t}{N} \rfloor - 1 \leq \frac{\tau}{N} < \lfloor \frac{t}{N} \rfloor\}$ .

### III. A NONLINEAR OBSERVER FOR TRAFFIC NETWORKS

Eqs. (1)-(2) cannot be used as they are to observe the system, as errors in flow measurement, unavoidable in real systems, are integrated in the dynamics, therefore possibly producing unbounded estimation errors. To overcome this difficulty, we consider the following Luenberger-like observer, in which  $\varphi^m(t) = [\varphi_e^m(t)]_{e \in \mathcal{E}_m}$  and  $v^{\text{FCD}}(t) = [v_e^{\text{FCD}}(t)]_{e \in \mathcal{E}}$  are the stacked versions of the measurements,

$$\begin{cases} \hat{\rho}_e(t + 1) = \hat{\rho}_e(t) + \frac{1}{\ell_e} (\hat{f}_e^{\text{in}}(t) - f_e^{\text{out}}(t)) \\ \quad + \kappa (\tilde{\rho}_e(t) - \hat{\rho}_e(t)) \\ \hat{f}_e^{\text{in}}(t) = \hat{f}_e^{\text{in}}(\varphi^m(t)) \\ \hat{f}_e^{\text{out}}(t) = \hat{f}_e^{\text{out}}(\varphi^m(t)) \\ \tilde{\rho}_e(t) = \tilde{\rho}_e(\varphi^m(t), v^m(t)) \end{cases} \quad \forall e \in \mathcal{E} \quad (7)$$

where

- $\hat{\rho}_e(t)$  is the estimate of the density on cell  $e$  at time  $t$ ;
- $\hat{f}_e^{\text{in}}(t), \hat{f}_e^{\text{out}}(t)$  are estimates, based on the flow measurements, of inflow and outflow in cell  $e$  at time  $t$ ;
- $\tilde{\rho}_e(t)$  is an instantaneous estimate, based on flow and speed measurements, of the density on cell  $e$  at time  $t$ ;
- $\kappa$  is a tunable gain trading off between flow and instantaneous density estimates. Notice that it should not be too large, as to avoid meaningless negative densities.

<sup>1</sup>The choice to set the speed measurement to freeflow speed for the first  $N$  samples, when no information is available, is done uniquely for the sake of simplicity, and can be replaced by any other educated guess.

Our aim is to design the maps

$$\begin{aligned} \hat{f}_+^{\text{in}} &= \{\hat{f}_e^{\text{in}}\}_{e \in \mathcal{E}} : \mathbb{R}_+^{\mathcal{E}^m} \rightarrow \mathbb{R}^{\mathcal{E}} \\ \hat{f}_+^{\text{out}} &= \{\hat{f}_e^{\text{out}}\}_{e \in \mathcal{E}} : \mathbb{R}_+^{\mathcal{E}^m} \rightarrow \mathbb{R}^{\mathcal{E}} \\ \tilde{\rho} &= \{\tilde{\rho}_e\}_{e \in \mathcal{E}} : \mathbb{R}_+^{\mathcal{E}^m} \times \mathbb{R}_+^{\mathcal{E}} \rightarrow \mathbb{R}^{\mathcal{E}} \end{aligned}$$

in such a way that the observer provides a good estimate of the real density of the system.

#### A. Proposed solution

The procedure consists of two steps, namely offline calibration and online update.

1) *Offline calibration*: In this section we propose a simple solution for calibrating the Fundamental Diagram, namely, for estimating the function  $\varphi_e(\cdot)$  on the cells  $e \in \mathcal{E}^m$ . We adopt a linear-quadratic Fundamental Diagram

$$\varphi_e(\rho) = \begin{cases} v_e^{\text{ff}} \rho, & \rho \leq \rho_e^c \\ a_e \rho^2 + b_e \rho + c_e, & \rho > \rho_e^c \end{cases}$$

where

- $\rho_e^c$  is the critic density, which separates the freeflow low-density region  $[0, \rho_e^c]$ , in which vehicles do not influence one each other, from the high-density congested region  $(\rho_e^c, \rho_e^{\text{jam}}]$ , in which speed decreases with density;
- $\rho_e^{\text{jam}}$  is the jam density, at which vehicles stop;
- $v_e^{\text{ff}} > 0$  is the freeflow speed on cell  $e$ ;
- the flow in congested region is a convex quadratic function of the density, where for consistency

$$\begin{cases} a_e (\rho_e^c)^2 + b_e \rho_e^c + c_e = v_e^{\text{ff}} \rho_e^c \\ a_e (\rho_e^{\text{jam}})^2 + b_e \rho_e^{\text{jam}} + c_e = 0 \\ a_e > 0 \end{cases}$$

This choice is driven by the empirical observation that the standard triangular Diagram tends to overestimate the flow in congestion. Other solutions include the inverted- $\lambda$  Fundamental Diagram [15], but the resulting model involves hysteresis and becomes more complex.

Let  $e \in \mathcal{E}^m$  be a cell equipped with sensors, and denote by  $\{(\rho_k, \varphi_k)\}_{k \in \mathcal{K}}, \mathcal{K} = 1, \dots, K$ , the set of  $K$  density and flow measurements used as learning set. Calibration of the Fundamental Diagram involves two steps (we write  $\rho^c$  and  $C$  instead of  $\rho_e^c$  and  $C_e$  for sake of notation)

- Gradient descent algorithm for estimation of  $\rho^c$  and  $C = v^{\text{ff}} \rho^c$ : in the first step, we estimate the critical density  $\rho^c$  and the capacity  $C$ , namely, the *nominal* maximum flow, by solving the non-linear and non-convex minimization problem

$$\begin{aligned} \min_{(\rho^c, C)} \quad & V_{(\rho^c, C)} = \sum_k (\varphi_k - \varphi_{(\rho^c, C)}(\rho_k))^2 \\ \text{s.t.} \quad & 0 < \rho^c < \rho^{\text{jam}} \\ & C > 0 \end{aligned}$$

$$\varphi_{(\rho^c, C)}(x) = \begin{cases} \frac{C}{\rho^c} x, & x \leq \rho^c \\ \frac{C(\rho^{\text{jam}} - x)}{\rho^{\text{jam}} - \rho^c}, & x > \rho^c \end{cases} \quad (8)$$

To solve (8) we propose the following gradient descent with diminishing stepsize algorithm

- Basic step: initialize  $\rho_0^c, C_0$ . A choice is  $\rho_0^c = 20$  veh/km (vehicles start influencing each other when their relative distance is less than 50 meters), and  $C_0 = v_e^{\text{limit}} \rho_0^c$ , where  $v_e^{\text{limit}}$  is the speed limit on cell  $e$  normalized by the sampling time  $T$ ;
- $n$ -th step: let  $(\rho_n^c, C_n)$  descend along the gradient of the cost, namely

$$\begin{aligned}\rho_{n+1}^c &= \rho_n^c + \frac{\delta}{n} \nabla_{\rho^c} V_{(\rho^c, C)} \\ C_{n+1} &= C_n + \frac{\delta}{n} \nabla_C V_{(\rho^c, C)}\end{aligned}$$

where the gradients  $\nabla_{\rho^c} V_{(\rho^c, C)}$  and  $\nabla_C V_{(\rho^c, C)}$  are computed at  $(\rho^c, C) = (\rho_{n-1}^c, C_{n-1})$ , and  $\delta > 0$  is a fixed initial step size;

- Stopping criterion: stop if  $\left\| \begin{bmatrix} \rho_n^c \\ C_n \end{bmatrix} - \begin{bmatrix} \rho_{n-1}^c \\ C_{n-1} \end{bmatrix} \right\| < \varepsilon$  for some small threshold  $\varepsilon > 0$ .
- Calibration of the congested part: given  $(\rho^c, C)$ , calibration of the quadratic function for the congested region can be cast as the following quadratic problem

$$\begin{aligned}\min_{(a,b,c)} & \sum_{k \in \mathcal{I}^{\text{FF}}(\rho^c)} (\varphi_k - (a\rho_k^2 + b\rho_k + c))^2 \\ \text{s.t.} & a(\rho^c)^2 + b\rho^c + c = C \\ & a(\rho^{\text{jam}})^2 + b\rho^{\text{jam}} + c = 0 \\ & a \geq 0\end{aligned}\quad (9)$$

which can be readily solved using off-the-shelf tools.

The Fundamental Diagrams can be extended on cells  $e \in \mathcal{E} \setminus \mathcal{E}_m$  by convex interpolation of the parameters  $\rho^c, C, a, b$  and  $c$ , using those calibrated in the closest cells  $e \in \mathcal{E}_m$ , with weights proportional to the relative distance.

2) *Online density reconstruction algorithm*: The online algorithm proceeds as follows

- at the beginning of the  $t$ -th time slot
  - receive the measurements  $\{\varphi_e^m(t)\}_{e \in \mathcal{E}^m}$ . The vector of estimate of the outflows  $\hat{f}^{\text{out}}(t)$  is computed as solution of the following minimization problem

$$\begin{aligned}\min_{\hat{f}^{\text{out}}} & \|(I - R^T)\hat{f}^{\text{out}}\|^2 \\ & + \gamma \sum_{e \in \mathcal{E}^m} (\hat{f}_e^{\text{out}} - \varphi_e^m(t))^2 \\ \text{s.t.} & \hat{f}_e^{\text{out}} \geq 0\end{aligned}\quad (10)$$

which aims both to match outflows and measurements and to balance the outflows according to the splitting ratios, as if the network were at steady state. The parameter  $\gamma$  trades off between the two: the higher, the more flows on cells in  $\mathcal{E}^m$  will be forced to match the measurements, the lower, the more the solution will be as if the system were at steady state. Finally,  $\hat{f}^{\text{in}}(t) = R^T \hat{f}^{\text{out}}(t)$ .

- receive the measurements  $\{v_e^m(t)\}_{e \in \mathcal{E}}$  when available, or hold the previous measurements;
- For each cell  $e$ , assume that the outflow can be approximated by the local flow  $\varphi_e = \varphi_e(\rho_e)$  determined by the Fundamental Diagram. Compute the two possible densities  $\rho_e^1$  (freeflow) and  $\rho_e^2$  (congested) corresponding to flow  $\hat{f}_e^{\text{out}}(t)$ ;

- For each cell  $e$ , compute the two corresponding velocities  $v_e(\rho_e^1) = \frac{\hat{f}_e^{\text{out}}}{\rho_e^1}$  and  $v_e(\rho_e^2) = \frac{\hat{f}_e^{\text{out}}}{\rho_e^2}$ ;
- Set  $\tilde{\rho}_e(t) = \arg \min_{i=1,2} \{|v_e(\rho_e^i) - v_e^m(t)|\}$  as a pseudo-measurement of the density, Since it is solely based on latest measurements of flows and speed, it is very noisy, especially in congestion regime. Therefore, we don't use it directly;
- For each cell  $e \in \mathcal{E}$ , let the density estimate evolve according to the observer equation (7).

#### IV. DENSITY RECONSTRUCTION - EXPERIMENTAL RESULTS

Our experimental setting is the Grenoble Traffic Lab (GTL) [14], a network of sensors deployed for monitoring and research purposes along the ‘‘Rocade Sud’’, a 12 km long freeway encircling the town of Grenoble, France (Fig. 2, upper panel). The network consists of 135 magnetometers buried in the asphalt on both lanes of the main line and on each onramp and offramp, totalling 68 *sensing locations*. Fig. 2, upper panel, shows the position of the 22 sections of the main line equipped with sensors. With few exceptions, each sensing location consists of two magnetometers deployed in pairs at a known distance. On each sensing location and every  $T = 15$  seconds, the system provides flow, occupancy and speed measurements. Occupancy  $o_e$  is defined as the percentage of the last period of  $T = 15$  seconds a vehicle was sitting over the sensor, and is approximatively related to the density as  $\rho_e \approx \frac{o_e}{100\ell_{\text{ave}}}$ , where  $\ell_{\text{ave}}$  is the average length of a vehicle (in km). Therefore, we assume that density is measured directly. Fig. 2, lower panel, shows a stylized representation of the Rocade Sud, the positions of the 68 sensing locations, and the distance between consecutive measurement sections along the main line. For a detailed report on the GTL, we refer to [14].

We partition the network in cells in such a way that each cell comprises the space between two consecutive sensing locations. As such, the numbered circles in Fig. 2 also represent cells. We use Floating Car Data provided by INRIX, one of the main navigation and traffic monitoring companies. The Rocade has been partitioned into *FCD segments*, and average speed is computed in each FCD segment every 1 minute. FCD segments cover the whole main line of the Rocade and most onramps and offramps, but lanes are not distinguished along the main line. FCD segments are represented in Fig. 2 as rectangles encircling several sensing locations/cells.

For our experiments, and to prove that the method performs well even with sparse equipment, we consider a very limited subset of the available sensors. In particular, we only use the sensors on the 9 sections shown in light blue in Fig. 2, of which 8 correspond to positions in which loops inductors have been deployed by the Government Agency *Centre national d'information routière (CNIR)* [16] for monitoring purposes. No information from ramps is used.

We calibrate the Fundamental Diagrams employing data from April 10th, 2014, a working day (a Thursday) exhibiting standard traffic pattern:

- very limited night time traffic;

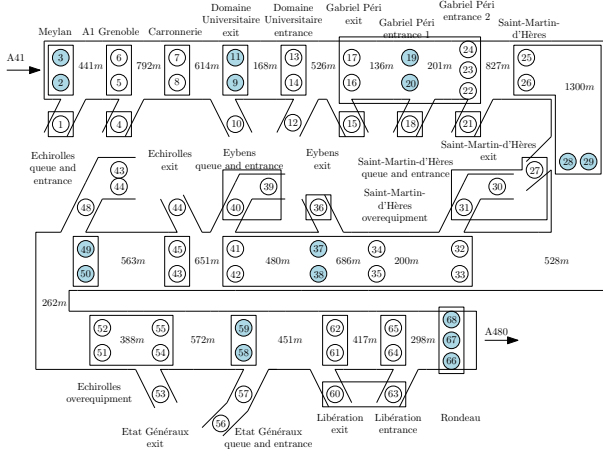
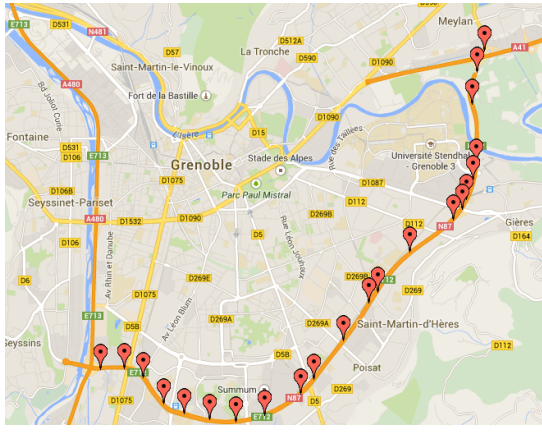


Fig. 2. The experimental setting: the town of Grenoble and the Rocade Sud (upper panel), and a stylized version of the freeway (lower panel). The positions along the main line of the 22 sections of the main line in which sensors have been placed is shown as red pins. The positions of the 68 fixed sensing locations are shown in the stylized map. Each sensing location also corresponds to a cell. Light blue circles denote fixed sensors that are used in our implementation. Each rectangle represents one FCD segment, often providing average speed measurements over more than one cell.

- a peak of congestion in the morning (8:00 - 10:00), triggered by vehicles exiting the Rocade at the offramp of Eybens (cells 37/38) and spilling back until Meylan;
- a second congestion triggered by vehicles entering in A480 at Rondeau (unable to do so due to the high traffic on A480), and spilling back until around Libération;
- a third, smaller, congestion triggered at Eybens around 14:00-15:00;
- medium/heavy but fluid traffic from 10:00 to 16:00
- a second peak of congestion in the afternoon, again triggered by congestion at Rondeau at around 16:00, spilling back on the whole freeway in around 60 minutes, and lasting around two hours.

The matrix of splitting ratios is set according to the following rule of thumb:

- let  $e$  be a fast lane cell, and  $j$  and  $k$  be the following fast and slow lane cells. Then  $R_{ej} = R_{ek} = 0.5$ ;
- let  $e$  be a slow lane cell, and  $j$  and  $k$  be the following fast and slow lane cells. If among the cells that follow  $e$  there is not an offramp, then  $R_{ej} = R_{ek} = 0.5$ .

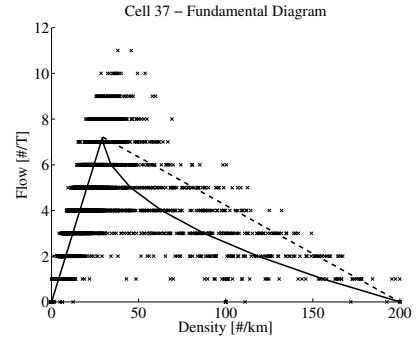


Fig. 3. Calibration of the Fundamental Diagram on the cell Eybens exit - slow lane. The linear-convex Fundamental Diagram, calibrated using data from April 10th, 2014, is shown in thick line. Dashed thick line represents the corresponding linear Fundamental Diagram in congested regime. Each cross is a (flow, density) pair measured on April 24th, 2014, one for each time slot of  $T = 15$  seconds during the whole day.

Otherwise,  $R_{ej} = R_{ek} = 0.4$  and  $R_{er} = 0.2$ , where  $r$  is the offramp cell that follows  $e$ ;

- if  $e$  is an onramp cell and  $j$  is the following slow ramp cell, then  $R_{ej} = 1$ .

#### A. Validation

Our method was validated over another Thursday working day, April 24th 2014. Calibration of the Fundamental Diagram at sensing locations is illustrated in Fig. 3, which shows in thick black the linear-quadratic Fundamental Diagram, in dashed thick black the standard linear Fundamental Diagram in congestion regime, and as crosses the pairs (density, flow) measured on April 24th, 2014. As standard, data in freeflow are in good accordance with the linear part, while data in congested regime are much more scattered. As it can be observed, a standard triangular Fundamental Diagram would overestimate the flows in congested regime (dashed thick line), while the convex quadratic curve better captures the average flow-density relation.

The algorithm was implemented in Matlab on a commercial laptop with 2.1 GHz i7-4600U CPU and 8 GB RAM. The optimization problems were solved using Matlab functions and the optimization system CVX [17], [18].

The results are reported in Fig. 4. For validation purposes only, density measurements from all GTL fixed sensors are considered ground truth. As such, the upper panel shows the evolution of the “true” measured density in all the cells on the main line, over the whole day. On the  $x$ -axis, the 46 sensing locations along the main line (numbers correspond to the labels in Fig. 2), on the  $y$ -axis, the 5760 time slots over the whole validation day (one slot every  $T = 15$  seconds). Colors vary from green to yellow to red as density increases, with a minimum of 0 vehicles per km (green) to a maximum of  $\rho^{jam} = 200$  vehicles per km (red). In the lower panel, we show, in the same format, the results of the density reconstruction. The estimation algorithm captures the four congestions described above in a reasonably good way, given the limited amount of information employed; in particular, the two small congestions at Rondeau during

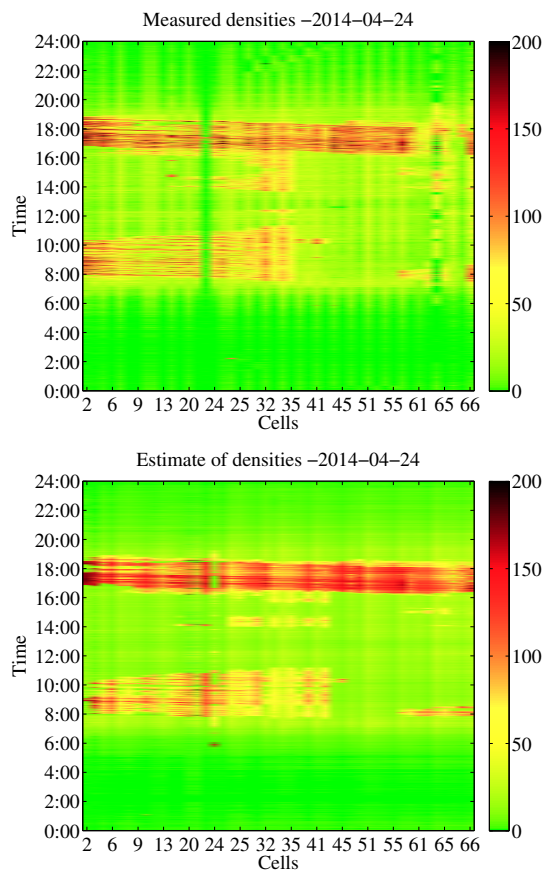


Fig. 4. Numerical results of density estimation on all cells: measured densities (upper panel) and estimated densities (lower panel).

the morning and at Eybens during early afternoon are both detected. The biggest discrepancy between estimates and measured data is around Eybens (cells 37-38) and can be explained as follows: The offramp of Eybens is a critical point at which many vehicles exit the Rocade, thus causing congestion before it. However, that point belongs to a long FCD segment running from Saint-Martin-d'Hères (cells 32-33) to Eybens entrance (cells 41-42), which thus provides a unique, rather low speed measurement, in turn yielding a high density estimate over the whole segment. A second obvious discrepancy is the smoothness of the reconstructed densities in congested regime compared with more scattered density measurements. While the former is related to the smoothing effect of the first-order observer, the latter is due to the high measurement rate of sensors, which well captures stop-and-go phenomena on the Rocade.

## V. CONCLUSIONS

This formalizes the problem of data fusion of heterogeneous sources of information for density reconstruction, and proposes an easily implementable solution that employs sparse fixed sensor measurements and ubiquitous Floating Car Data. Calibration algorithms for the Fundamental Diagram are discussed. Future research directions include and are not limited to estimation of statistical properties of

measurement noises, development of theoretical guarantees on maximum mean-square reconstruction error, calibration of the matrix of splitting ratios, and extension of the considered numerical scenario to part of the town of Grenoble.

## VI. ACKNOWLEDGEMENTS

The authors wish to gratefully thank INRIX Europe, whose collaboration and permission to use data from the Grenoble area led to the results presented in this work.

## REFERENCES

- [1] M. Papageorgiou, M. Ben-Akiva, J. Bottom, P.H.L. Bovy, S. P. Hoogendoorn, N. B. Hounsell, A. Kotsialos, and M. McDonald. ITS and traffic management. In C. Barnhart and G. Laporte, editors, *Handbooks in Operations Research and Management Science: Transportation*, chapter 11, pages 715–774. Elsevier, 2007.
- [2] M. Papageorgiou, C. Diakaki, V. Dinopoulou, A. Kotsialos, and Wang Y. Review of road traffic control strategies. *Proc. of the IEEE*, 91(12):2043–2067, 2003.
- [3] M. Papageorgiou, H. Hadj-Salem, and J.-M. Blosseville. ALINEA: A local feedback control law for on-ramp metering. *Transportation Research Record*, (1320):58–64, 1991.
- [4] D. Piasarski and C. Canudas de Wit. Optimal balancing of road traffic density distributions for the cell transmission model. In *Proc. of the 51th IEEE Conference on Decision and Control (CDC'12)*, page 6969–6974, Maui, Hawaii, USA, 2012.
- [5] G. Como, E. Lovisari, and K. Savla. Throughput optimal distributed routing in dynamical flow networks. In *Proc. of the 52st IEEE Conference on Decision and Control (CDC'13)*, Florence, Italy, 2013.
- [6] INRIX official website. Website: <http://www.inrix.com/xd-traffic/>.
- [7] M. J. Lighthill and G. B. Whitham. On kinematic waves. ii. a theory of traffic flow on long crowded roads. *Proc. of the Royal Society of London, Series A, Mathematical and Physical Sciences*, 229(1178):317345, 1955.
- [8] C. F. Daganzo. The cell transmission model: A dynamic representation of highway traffic consistent with the hydrodynamic theory. *Transport. Res. B-Meth.*, 28B(4):269287, 1994.
- [9] L. Muñoz, X. Sun, R. Horowitz, and L. Alvarez. A piecewise-linearized cell transmission model and parameter calibration methodology. In *Proc. of the 85th Annual Meeting of the Transportation Research Board (TRB)*, Washington D.C., USA, 2006.
- [10] F. Morbidi, L. L. Ojeda, C. Canudas de Wit, and I. Bellicot. A new robust approach for highway traffic density estimation. In *Proc. of the 13th European Control Conference (ECC'14)*, pages 2576–2580, Strasbourg, France, 2014.
- [11] J. W. C. van Lint and S. P. Hoogendoorn. A robust and efficient method for fusing heterogeneous data from traffic sensors on freeways. *Computer-Aided Civil and Infrastructure Engineering*, 25:596–612, 2010.
- [12] D.B. Work, O.-P. Tossavainen, Q. Jacobson, and A.M. Bayen. Lagrangian sensing: traffic estimation with mobile devices. In *Proc. of the American Control Conference (ACC'09)*, pages 1536–1543, June 2009.
- [13] W. Deng, H. Lei, and X. Zhou. Traffic state estimation and uncertainty quantification based on heterogeneous data sources: A three detector approach. *Transport. Res. M-Meth.*, 57:132157, Nov. 2013.
- [14] C. Canudas de Wit, F. Morbidi, L. L. Ojeda, A. Y. Kibangou, and I. Bellicot. Grenoble traffic lab: An experimental platform for advanced traffic monitoring and forecasting. *IEEE Control Syst. Mag.*, 2015.
- [15] F. L. Hall, B. L. Allen, and M. A. Gunter. Empirical analysis of freeway flow-density relationships. *Transport. Res. A-Gen.*, 20(3):197–210, 1986.
- [16] Centre national d'information routière (CNIR). Website: <http://www.bison-fute.gouv.fr/>.
- [17] M. Grant and S. Boyd. Cvx: Matlab software for disciplined convex programming, 2013. version 2.0 beta, <http://cvxr.com/cvx>.
- [18] M. Grant and S. Boyd. Graph implementations for nonsmooth convex programs. In S. Blondel, S. S. Boyd, and H. Kimura, editors, *Recent Advances in Learning and Control (a tribute to M. Vidyasagar)*, Lecture Notes in Control and Information Sciences. Springer, 2008.