Dynamic density and flow reconstruction in large-scale urban networks using heterogeneous data sources

Martin Rodriguez-Vega, Carlos Canudas de Wit, Hassen Fourati

Dynamic density and flow reconstruction in large-scale urban networks using heterogeneous data sources

Martin Rodriguez-Vega[a,*], Carlos Canudas-de-Wit[a], Hassen Fourati[a]

[a] *Univ. Grenoble Alpes, CNRS, Inria, Grenoble INP, GIPSA-Lab, 38000 Grenoble, France*

Abstract

This work proposes an estimator for the vehicle density of every road section of a large urban traffic network using a data-based approach. We assume that a limited number of flow and turning ratio sensors can be installed and that aggregate Floating Car Data are available. To validate the estimator, real data obtained in the city of Grenoble in France is used. The results show that the trajectories of the real and estimated flows and density are very close, showing the good performance of the proposed methods.

Keywords: Traffic density estimation, Sensor location, Large networks

1. Introduction

The development of transportation networks is one of the main issues in the evolution of modern cities. As urban areas become larger and more populated, traffic congestions occur more frequently with higher levels of cost in today's life. Intelligent Transportation Systems (ITS) refers to the use of advanced technologies and applications to offer innovative solutions to the problems of traffic congestion. Some techniques under study by ITS consist of traffic control (ramp metering, pricing, information provision, and traffic light cycle control), and strategic transportation planning. Traffic State Estimation (TSE) is one of the main ingredients required to provide vital information to the efficient and accurate application of traffic control strategies. TSE refers to the use of partially observed and noisy traffic data to infer the value of traffic variables, such as flow, density, velocity, traveling time, and others (Seo et al. (2017)).

First proposed works with classical TSE were applied to highways (Ferrara et al. (2018)). These works are usually based on the well-known LWR model proposed independently by Lighthill and Whitham (1955) and Richards (1956), and on its discrete counterpart Cell Transmission Model (CTM) proposed in Daganzo (1994). These models are based on a conservation law applied to the traffic density with the addition of the Fundamental Diagram (FD), originally proposed in Greenshields et al. (1934) and is an empirical relation between the equilibrium density and flow of a road section. Tampère and Immers (2007) proposes the use of an Extended Kalman Filter (EKF) coupled with the CTM to estimate the density of sections of highways, using data from stationary sensors located at certain sections of the road. This approach is based on linearizing the CTM around a current state, and then compare the measurements with the predictions of the linearized model. Later, Herrera and Bayen (2010) made use of the increase of availability of GPS data to complement the stationary sensor data. In this formulation, a Lagrangian model was used as input to the EKF. More recently, Sun and Work (2014) and Seo et al. (2015) show how information from connected vehicles can be provided to a KF approach to improve highway state estimation. In Bekiaris-Liberis et al. (2016), the authors propose the direct use of speed measurements from connected vehicles to approximate the mean speed of each highway section, and model the traffic dynamics as a linear time-varying system, without requiring the use of the FD. A more complete review of TSE techniques can be found in Seo et al. (2017).

However, the development of TSE approaches to the case of urban networks has received less attention (Seo et al. (2017)). One of the reasons for this is the additional complexity encountered in the modeling of flow exchanges at intersections. One of the most commonly used junction models is the flow-maximizing model, which was proposed in Daganzo (1995) as an extension of the CTM to networks. This model assumes that drivers cooperate such that the total flow that can be delivered from an intersection is

*Corresponding author. E-mail address: Martin.Rodriguez-Vega@gipsa-lab.fr

maximized subject to the downstream road supply constraints. The flow-maximizing model is very commonly used in the analysis of traffic networks. However, it has been found that this model generally overestimates the flows that would be produced in real scenarios, as it is unable to describe vehicle decelerations when approaching intersections (Liou et al. (2017)). Other models have been proposed that incorporate other phenomena such as traffic light cycles (Jabari (2016)). Nevertheless, realistic junction models remain an open problem for the treatment of traffic networks.

Despite the model limitations for the case of urban networks, some works have proposed TSE approaches. He (2013) provides a method to estimate the equilibrium flow for each road from sensor data using graph-based algorithms. In Lovisari et al. (2016), a method to estimate the dynamical evolution of the density of each road is proposed by using both stationary sensor measurements and Floating Car Data (FCD). This is done by first giving a flow estimate for each road as if the network is in steady-state, and then use the speed-density FD to provide density pseudo-measurements. These estimates are then the input to a Luenberger-like observer. In Nantes et al. (2016), an approach based on the CTM is coupled with a KF using measurements from flow sensors, BT detectors, and GPS traces, to estimate the density in a section of an urban arterial road. In Ladino et al. (2018), FCD is used to determine if each road is in free-flow or congested regime, and then uses the flow-density FD and stationary sensor data to provide density estimates to each road in the traffic network. More recently, Rostami Shahrbabaki et al. (2020) proposed a data-based density estimation method that does not require the use of an FD. Instead, data from connected vehicles is used to estimate the outflow of each road and then uses the density conservation law to estimate the traffic state.

Our contribution in this paper is the proposal of a data-based estimator for the vehicle density of each road section of a large urban traffic network. We assume that aggregate FCD is available, such that the space-mean speed of each road can be estimated, and that a limited number of flow and Turning Ratio (TR) sensors are available. We propose a method to locate TR sensors, which takes as input previous low-quality estimates of the turn rates, and then assigns each intersection a weight according to the effect on the total density estimation error caused by perturbations between the a priori and actual TR values. We present the study case of the downtown area of the city of Grenoble in France to validate the estimator and sensor location scheme. The selected area can provide real traffic data from sensors located using the proposed methodology.

This paper is organized as follows. Section 2 presents the underlying traffic model used to develop the TSE approach. Section 3 the heterogeneous data sources used to feed the density estimator. Additionally, this Section proposes a sensor location scheme for the measurement of necessary parameters. Section 4 describes the practical implementation of the estimator. Section 5 presents the results of the estimator validation using both simulated and real data. Section 6 discusses the limitations of the proposed estimator during the experimental analysis, and compares the proposed approach to similar works in the literature. Lastly, Section 7 concludes the paper.

2. Model

We define a traffic network as a directed graph $\mathcal{G} = \{\mathcal{N}, \mathcal{E}\}$ where the nodes $\mathcal{N} \subset \mathbb{N}$ correspond to intersections, and the edges $\mathcal{E} \subset \mathcal{N} \times \mathcal{N}$ correspond to road sections. Additionally, for every road section i, there are associated parameters such as the road length ℓ_i, number of lanes Γ_i, and maximum velocity $v_i^{\max}$.

The traffic state refers to the collection of road densities, inflows and outflows for all roads, which we denote with vectors $\boldsymbol{\rho}(t), \boldsymbol{\varphi}^{\text{in}}(t), \boldsymbol{\varphi}^{\text{out}}(t) \in \mathbb{R}^{|\mathcal{E}|}$. These variables and their interactions at intersections are visualized in Fig. 1. Intersections are modeled as 0-dimensional points that do not store vehicles. To model the exchange of inflows and outflows of the different roads at the intersections we use the TR parameters.

Definition 1. Turning Ratios: Let $\mathcal{I}(n)$ be the set of incoming roads to some intersection $n \in \mathcal{N}$ and $\mathcal{O}(n)$ be the set of outgoing roads from n. A TR $r_{i,j}$ for $i \in \mathcal{I}(n)$ and $j \in \mathcal{O}(n)$ defines the proportion of vehicles exiting i that enters j.

As intersections do not store vehicles, flow conservation implies that

$$\sum_{j \in \mathcal{O}(n)} r_{i,j} = 1 \ , \quad \forall n \in \mathcal{N} \quad \forall i \in \mathcal{E} \setminus \mathcal{E}^{\text{out}} \tag{1}$$

where $\mathcal{E}^{\text{out}}$ are outgoing roads at the boundary of the network, hence have no downstream roads. From the definition, the inflow for each road can be expressed as the sum of the outflows of the upstream roads, weighted by the corresponding TRs,

$$\varphi_i^{\text{in}}(t) = \sum_{j \neq i} r_{j,i} \varphi_j^{\text{out}}(t). \tag{2}$$

Let $R \in \mathbb{R}^{|\mathcal{E}| \times |\mathcal{E}|}$ be the TR matrix with elements $r_{i,j}$. If there is no connection between roads i and j, $R_{i,j} = 0$.

Traffic dynamics are governed by the conservation law

$$\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{\rho}(t) = L^{-1}(\boldsymbol{\varphi}^{\text{in}}(t) - \boldsymbol{\varphi}^{\text{out}}(t)) \tag{3}$$

where L is a diagonal matrix containing the road lengths, i.e., $L = \text{diag}(\boldsymbol{\ell})$. Applying (2) to all roads, we obtain the system of linear equations

$$\boldsymbol{\varphi}^{\text{in}}(t) = R^{\top} \boldsymbol{\varphi}^{\text{out}}(t) + B\mathbf{u}(t) \tag{4}$$

where B is a selection matrix that indicates the location of boundary inflows $\mathcal{E}^{\text{in}}$, and $\mathbf{u}(t)$ is a vector containing the external input demands at the boundaries. Combining (3) and (4), we have

$$\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{\rho}(t) = L^{-1}(R^{\top} - \mathbb{I})\boldsymbol{\varphi}^{\text{out}}(t) + L^{-1}\boldsymbol{\varphi}^{\text{ext}}(t) \tag{5}$$

Define $v_i(t)$ as the space-mean speed of road i, that is, the average of the speeds of vehicles inside of road i at time t. Denote $\mathbf{v}(t)$ as the speed vector with elements $v_i(t)$. Using the hydrodynamic relation, we can approximate the outflows as

$$\boldsymbol{\varphi}^{\text{out}}(t) \approx V(t)\boldsymbol{\rho}(t) \tag{6}$$

where $V(t) = \text{diag}(\mathbf{v}(t))$. This relation applies accurately when considering very short distances, or when the spatial variations in vehicle speed and density are negligible. We make the following assumption,

Assumption 1. *The density and speed in each road section are distributed uniformly during each time interval.*

Assumption 1 does not ignore the inter-cell spatial distribution of speed and density. Nevertheless, each cell is only represented by a unique density and speed value, and the assumption is made to justify the hydrodynamic equation, which is applicable only to point locations. In practice, however, the validity of the hydrodynamic relation for vehicles only makes sense when referring to section lengths large enough that macroscopic variables can be defined, but small enough that these variables are approximately constant (see Treiber and Kesting (2013), Chapter 7). Road sections in dense urban networks are usually short enough that the assumption does not introduce significant error. Nevertheless, for longer sections, there can be a large deviation between the speed spatial distribution and the section mean value. In such cases, it is advised to divide long sections into smaller ones, at the cost of augmenting the dimension of the system. Assumption 1 is commonly made in many TSE approaches, both in data-based

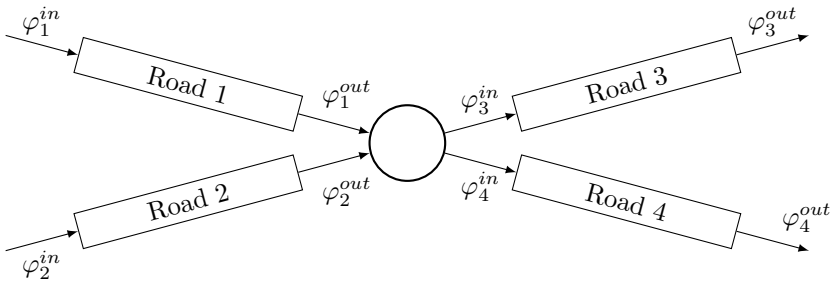

Figure 1: Flow exchange at an intersection.

methods as in Rostami Shahrbabaki et al. (2020), and in model-based methods such as the CTM, where vehicle speeds in each cell are determined by the FD.

Substitution of (6) into (5) yields

$$\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{\rho}(t) = L^{-1}(R^\top - \mathbb{I})V(t)\boldsymbol{\rho}(t) + L^{-1}B\mathbf{u}(t) \tag{7}$$

In this paper, we use an open-loop observer following the same dynamics of (7). To achieve this goal, we require knowledge about the TRs, road speeds, and external flows. The measurement of these variables will be examined in Section 3. For now, denote $\hat{R}$ as the estimate for the TR matrix, $\hat{V}(t)$ as the estimate for the speed matrix, and $\hat{\mathbf{u}}(t)$ as the flow measurements at the boundary inputs. Thus, the estimator for the system's density is

$$\frac{\mathrm{d}}{\mathrm{d}t}\hat{\boldsymbol{\rho}}(t) = L^{-1}(\hat{R}^\top - \mathbb{I})\hat{V}(t)\hat{\boldsymbol{\rho}}(t) + L^{-1}B\hat{\mathbf{u}}(t) \tag{8}$$

where $\hat{\boldsymbol{\rho}}(t)$ is the estimated density. The error associated to this estimate is defined as

$$\mathbf{e}(t) = \boldsymbol{\rho}(t) - \hat{\boldsymbol{\rho}}(t). \tag{9}$$

with dynamics,

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{e}(t) \quad &= \quad L^{-1}(R^\top - \mathbb{I})V(t)\mathbf{e}(t) - L^{-1}(R^\top - \mathbb{I})\Omega(t)\hat{\boldsymbol{\rho}}(t) \\ &\quad -L^{-1}\Xi^\top V(t)\hat{\boldsymbol{\rho}}(t) - L^{-1}\Xi^\top\Omega(t)\hat{\boldsymbol{\rho}}(t) + B(\mathbf{u}(t) - \hat{\mathbf{u}}(t))\end{aligned} \tag{10}$$

where $\Xi = \hat{R} - R$ and $\Omega(t) = \hat{V}(t) - V(t)$. The performance of the proposed estimator depends on the stability of (10) and how it is affected by the error terms. Consider first the case where all information is known and measurements are ideal, i.e., $\Xi = \mathbf{0}$, $\Omega(t) = \mathbf{0}$, and $\mathbf{u}(t) = \hat{\mathbf{u}}(t)$. Theorem 1 shows that under some assumptions on the velocity $V(t)$, the estimator is asymptotically stable for any initial condition.

Theorem 1. *Consider the dynamic system*

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{e}(t) = L^{-1}(R^\top - \mathbb{I})V(t)\mathbf{e}(t) \tag{11}$$

If $V = diag(\mathbf{v}(t))$ such that $\mathbf{v}(t)$ is continuous, bounded by

$$\mathbf{0} < v_{min}\mathbf{1} \leq \mathbf{v}(t) \leq v_{max}\mathbf{1}, \tag{12}$$

and is continuously differentiable with derivative bounded by

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{v}(t) \leq \frac{v_{min}^2(1-\epsilon)}{\ell_{max}||(\mathbb{I} - R^\top)^{-1}||_\infty}\mathbf{1} \tag{13}$$

where $0 < \epsilon < 1$ is arbitrary, then (11) is asymptotically stable, i.e., $\mathbf{e}(t) \to \mathbf{0}$ as $t \to \infty$.

The proof of Theorem 1 is shown in Appendix A. The restrictions imposed on the velocity are very strict, but they are sufficient conditions that may be conservative, and the system may still be stable when the conditions are not satisfied. This stability result shows that if the measured TRs, velocities, and boundary inflows are close to the real values, the estimator can give a close estimation of the network's states. In Section 3, we discuss how these variables are measured, and how to locate sensors such the resulting error is minimized.

The proposed estimation approach is data-based, as the outflows of each road section are not given by a function of density as in the LWR model, and instead are calculated via the hydrodynamic relation (6) using road speed data. This has the advantage that no flow-density relation is used, which usually assumes that flow and speed are functions of only the density, which is not necessarily valid in urban scenarios, where pedestrian crossings, bicycles, and other factors have an important effect on vehicle speeds.

Although the effect of traffic lights is not explicitly taken into account due to lack of precise cycle information, their effect can be perceived by changes in the speed profiles. However, as the quality of speed data can be limited by technical considerations and traffic light cycles can affect other variables such as the TRs for short durations, the flow and density estimates may produce meaningful results for aggregated time windows as discussed in Section 5.

3. Input data

In the following subsections, we describe the available data sources, and the sensor location methods used to decide their locations.

3.1. Stationary counting sensors

Stationary sensors are placed in a fixed position in a road section and collect information of the vehicles passing through that point, such that they provide the flow at the specific location. We assume that sensors are located at the boundaries of the network such that the input demands $\mathbf{u}(t)$ can be measured directly. This information is required as an input to the estimation algorithm. Furthermore, several sensor technologies can measure the flow in multiple lanes and both directions, so we also assume that the flow of vehicles leaving the network at the boundaries is at least partially available. Data from these locations is denoted by $\mathbf{y}(t)$, i.e., the boundary outflows.

There are several technologies of stationary flow sensors that are used in practical applications, such as inductive loops, microwave radars, infrared detectors, and cameras. In practice, the flow measurements are not continuous. Instead, the sensors report the number of vehicles counted during a discrete interval. Let $\{t_0^\varphi\ ,\ t_1^\varphi\ ,\ t_2^\varphi\ ,\ \dots\}$ denote the times at which flow measurements are available. Let $\mathbf{u}[m]$ be the measured input demands during the interval $[t_{m-1}^\varphi, t_m^\varphi]$. Then

$$\mathbf{u}[m] = \frac{1}{(t_m^\varphi - t_{m-1}^\varphi)} \int_{t_{m-1}^\varphi}^{t_m^\varphi} \mathbf{u}(\tau)\mathrm{d}\tau. \tag{14}$$

The discrete-time outflow measurements $\mathbf{y}[m]$ are defined analogously.

3.2. Floating car data

Floating Car Data (FCD) are obtained from trajectories of individual vehicles collected via devices such as GPS navigators. Due to privacy policies, data from multiple users is aggregated.

Consider that a fraction of the vehicles in the network are equipped with devices (e.g. a GPS navigator) that periodically report to a centralized server the vehicles' position and velocity at a given time. Define by $\mathcal{V}_i(t)$ the set of vehicles indexes providing FCD that are inside of road i at time t. The total number of vehicles in i cannot be estimated with this information as the penetration rate of vehicles that provides FCD is unknown. However, as the velocity of a vehicle is affected by the velocities of surrounding vehicles, this can be used to estimate the space-mean speed of the section. Let ν_α be the speed of a vehicle indexed by α. We define the aggregated speed for road section i from FCD data by

$$v_i^{\mathrm{FCD}}(t) = \frac{1}{|\mathcal{V}_i(t)|} \sum_{\alpha \in \mathcal{V}_i(t)} \nu_\alpha(t) \tag{15}$$

where $|\mathcal{V}_i(t)|$ is the cardinality of $\mathcal{V}_i(t)$, corresponding to the number of vehicles in road i at time t. We assume that FCD speeds are direct measurements of the space-mean road speed, and thus

$$\Omega(t) = \mathbf{0} \tag{16}$$

However, due to technical constraints, this information is not provided continuously, but at discrete times denoted by $\{t_0^v\ ,\ t_1^v\ ,\ t_2^v\ ,\ \dots\}$. Thus let $\mathbf{v}^{\mathrm{FCD}}[i]$ be the average of the reported vehicle velocities for all road sections between times t_i^v and t_{i-1}^v. Denote $\hat{V}[i] = \mathrm{diag}(\mathbf{v}^{\mathrm{FCD}}[i])$. Furthermore, data is not provided for each road. Instead, data is collected from vehicles contained in collections of roads forming straight paths, with an average length of 200m. Roads contained in the same collection are given the same speed value.

We assume in what follows that the aggregated speeds are available for all roads. In the case where the information is not available for some roads due to a lack of reporting vehicles, there are two possible cases: *(a)* the road is empty or uncongested; or *(b)* the road is congested. Assuming that reporting vehicles are distributed randomly and independently of route choice, congested roads would have a higher probability of containing reporting vehicles, making case *(b)* less likely than case *(a)*. Therefore, for this case we use the value of the maximum speed (or free-flow speed) $v^{\max}$, based on the idea that when there are few vehicles on a road, each vehicle's behavior is not affected by the others, and it can move at its desired speed. Nevertheless, the possible occurrences of case *(b)* contributes to the error of the estimator as unobservable noise.

3.3. Turning ratio parameters

The estimator requires knowledge of the TR parameters for all intersections in the traffic network. This is however unfeasible in practice as it requires providing every intersection with sensors, which is impossible for networks containing more than a few intersections. In the literature, several methods have been proposed to estimate the TRs based on different data sources. For instance, GPS traces have been utilized to identify route choices, from which TRs can be reconstructed. However, this approach requires extremely detailed FCD, which is not common: as previously mentioned, privacy regulations require the aggregation of individual vehicle data, and usually section-level average speed is available.

Another possible way to recover route choice without the use of vehicle traces is the use of Traffic Assignment (TA) techniques. These methods take as input the Origin/Destination (O/D) matrices to determine the optimal routes for each O/D pair, such that some user cost is minimized (e.g. travel time), given the structure and characteristics of the road network. Although these methods can be very powerful and provide accurate values for the TRs, they require input data that may be hard to obtain, such as the O/D matrices, and require detailed modeling of user behavior and road cost functions. Furthermore, route-based techniques may require a very high computational power for large networks, as the number of O/D pairs and the number of routes between each pair increases.

3.3.1. TR estimation using network parameters

To circumvent these restrictions, some authors have proposed methods to provide estimates for the TRs using only network parameters which are more readily obtainable (Furth (1990)). Due to the simplicity of this approach, there can be a high uncertainty in the produced results. This can be improved by additional information, such as flow or TR measurements that can be used to tune the provided values. For instance Schaefer (1988) proposed two ways to provide initial estimates for the TR values: *a)* the use of equal ratios for all outbound roads; and *b)* the use of arbitrary standard ratios for each direction (right, straight, left) (e.g. the 25/50/25 rule). Later, Furth (1990) proposed the use of additional network parameters such as the number of roads in an area, the angle between inbound and outbound roads, and the presence of "shortcut" roads. Both works used the initial values in conjunction with flow measurements to determine the most likely TR values.

More recently, Krumm (2010) estimated the TRs based on the assumption that drivers are less likely to turn to destination-only roads (dead-ends), in comparison to roads that offer more destination options. This method was then tested with collected data for 40 different intersections with a median error of 0.142. This work supports the underlying intuition that the design and function of a road in a traffic network have an impact on the route choice of its users.

The main advantage of network-based approaches to estimate TR values is their simplicity of use, which makes them suitable for very large networks, as the estimates only depend on local variables of each intersection and its adjacent roads. In the following paragraphs, we will describe two heuristic methods to estimate the TRs for each intersection. To address the uncertainty introduced by these methods, Section 3.3.2 will describe a TR sensor location method that finds the optimal location of a limited number of sensors to reduce the flow and density estimation error.

Turning ratio estimation using road capacities. As a first working value, we distribute the TR values proportionally to the capacity (maximum flow) in each outbound road from an intersection, calculated using the maximum speed v^{max} and number or lanes Γ,

$$r_{i,j}^{\text{Cap}} = \frac{v_j^{max}\Gamma_j}{\sum_{k\in\mathcal{O}(n_i)} v_k^{max}\Gamma_k} \tag{17}$$

where n_i is the intersection to which the inbound road i is adjacent, and $\mathcal{O}(n_i)$ is the set of outbound roads.

This heuristic is based on the assumption that roads with higher capacity are more likely to be chosen by vehicles, as they can be seen as greedy choices to reduce the traveling time for each driver in the absence of congestion. In previous tests (see Appendix B of Rodriguez-Vega (2021)), we found that this model provided accurate results for the TRs in highways, with median errors between 0.05 and 0.15. However, as will be discussed in Section 5.3, this choice provided inaccurate results for the case of urban networks.

Table 1: Description of the road classes provided by TomTom.

Class	Short description	Long description
1	Major roads of high importance	Roads of high importance that are used for international and national traffic.
2	Other major roads	Roads used to travel between neighboring country regions.
3	Secondary roads	Roads used to travel between parts of the same region.
4	Local connecting roads	Roads making settlements accessible or making parts of a settlement accessible.
5	Local roads of high importance	Local roads that are the main connections in a settlement.
6	Local roads	Roads used to travel within a part of a settlement.
7	Local roads of minor importance	Roads that only have a destination function.

Turning ratio estimation using road classification. Generally, roads are designed for a specific function within a network. According to this function, the road will have different design considerations such as maximum speed, capacity, and land use (U.S Federal Highway Administration (2013)). This results in a hierarchical relation of roads, such as collector roads that then lead to arterial ones. The categorization of roads according to their roles and functions is called Functional Road Classification (FRC) (D'Andrea et al. (2014)). The considered classes vary according to the desired granularity and application. A simple categorization consists only of arterial, collector, and local roads (U.S Federal Highway Administration (2013)) for the case of highways, whereas a more detailed classification including urban areas is shown in Table 1, used by TomTom[1] which is used in this work. This classification differentiates between major roads that experience heavy traffic from a variety of O/D pairs and minor roads which are inside a residential area and experience light traffic only.

We consider, thus, to calculate the TRs in each intersection as a function of the FRC classes of each of the outbound roads. This heuristic is based on the idea that driver route-choice is influenced by the network design. As identified by Krumm (2010), the role of a road in the network is a key factor on the probability of drivers turning to it: dead-ends are destination only, and therefore less used than highways. In comparison to the first described model, TRs based on capacity, there are cases of roads with the same maximum speed and number of lanes, but where one leads to an arterial road and the other leads to a closed residential neighborhood. Conversely, we consider that FRC implicitly includes factors such as v^{max} and Γ.

For each FRC class in the set $\{1, 2, \ldots, 7\}$, we define a weight $\theta \in (0, 1]$. Let $\boldsymbol{\theta} \in (0, 1]^7$ be the vector of class weights. Suppose that the TRs at each intersection are distributed proportionally to the class weights of each of its outgoing roads. Thus, TRs are computed as

$$r_{i,j}^{\text{FRC}} = \frac{\theta_{\text{FRC}(j)}}{\sum_{k \in \mathcal{O}(n_i)} \theta_{\text{FRC}(k)}} \tag{18}$$

where FRC(i) is the class of road i.

To compute the value of $\boldsymbol{\theta}$, we consider the following optimization problem that uses flow measurements at the network boundaries,

$$\begin{aligned} \min_{\boldsymbol{\theta}} \quad & ||\bar{\mathbf{y}} - \hat{\mathbf{y}}|| \\ \text{subject to} \quad & \hat{\mathbf{y}} = C(\mathbb{I} - \hat{R}^{\top}(\boldsymbol{\theta}))^{-1} B \bar{\mathbf{u}}, \\ & \hat{R}_{i,j}(\boldsymbol{\theta}) = r_{i,j}^{\text{FRC}} \\ & \boldsymbol{\theta} \in (0, 1]^7, \quad \theta_1 = 1, \end{aligned} \tag{19}$$

[1] https://developer.tomtom.com/traffic-stats/support/faq/what-are-functional-road-classes-frc

where C is a selection matrix which identifies the outgoing boundary roads $\mathcal{E}^{\text{out}}$, $r_{i,j}^{\text{BT}}$ is computed using (42), $r_{i,j}^{\text{FRC}}$ is computed using (18), and

$$\bar{\mathbf{u}} = \frac{1}{T}\int_0^T \mathbf{u}(t)\mathrm{d}t \quad , \quad \bar{\mathbf{y}} = \frac{1}{T}\int_0^T \mathbf{y}(t)\mathrm{d}t. \tag{20}$$

are the daily ($T = 1$ day) average flows from the input and output sets, respectively.

In this problem postulation, $\hat{\mathbf{y}}$ corresponds to estimated steady-state boundary outflows, computed using the mean boundary inflows and the TR estimates. The problem tries to match the observed flows with the estimates, by changing the relative FRC weights. Intersections for which BT data were collected are given fixed TR values. The condition θ_1 is set arbitrarily without loss of generality, as only the relative differences between the weights are important. If knowledge of real values of TR for a subset of intersections is available, the corresponding values of $\hat{R}$ can be fixed, to take into account the additional data.

3.3.2. Location of turning ratio sensors

Suppose that a limited number of sensors is available. The selection of intersections to locate sensors is done to minimize the error in the density estimation caused by deviations between the real and estimated TRs. For this, we propose to identify which intersections generate the highest error given some perturbations in its TRs.

Let $\hat{r}_{i,j}$ be the a priori value proposed for some TR. Let the deviation between the a priori and the actual values be given by $\xi_{i,j}$ such that,

$$\hat{r}_{i,j} = r_{i,j} + \xi_{i,j} \tag{21}$$

The deviations $\xi_{i,j}$ will cause an error in the state estimation. Define $\hat{R}$ and Ξ as matrices with elements $\hat{r}_{i,j}$ and $\xi_{i,j}$, respectively. Note that $\hat{R} = R + \Xi$. To quantify the sensitivity of the TRs in the error, we neglect the effects of time variations of the external flows and velocities, hence $V(t) = V$ and $B\mathbf{u}(t) = B\mathbf{u}$.

Estimation error. We want to express the effect of the TR perturbations Ξ in the estimation error $\mathbf{e}(t)$. To simplify the notation, let $M = (\mathbb{I} - R^\top)V$ and $\hat{M} = (\mathbb{I} - \hat{R}^\top)V$, such that $\hat{M} = M - \Xi^\top V$. Under the assumption that section speeds are constant, the dynamics (7) and (8) become time-invariant, for which a closed-form solution is known,

$$\begin{aligned}
\boldsymbol{\rho}(t) &= \mathrm{e}^{-L^{-1}Mt}\boldsymbol{\rho}(0) + \left(\int_0^t \mathrm{e}^{-L^{-1}M\tau}\mathrm{d}\right) L^{-1}B\mathbf{u} \\
\hat{\boldsymbol{\rho}}(t) &= \mathrm{e}^{-L^{-1}\hat{M}t}\hat{\boldsymbol{\rho}}(0) + \left(\int_0^t \mathrm{e}^{-L^{-1}\hat{M}\tau}\mathrm{d}\right) L^{-1}B\mathbf{u}
\end{aligned} \tag{22}$$

Therefore, the error can be written as

$$\mathbf{e}(t) = e^{-L^{-1}Mt}\boldsymbol{\rho}(0) - e^{-L^{-1}\hat{M}t}\hat{\boldsymbol{\rho}}(0) + \int_0^t \left(e^{-L^{-1}M\tau} - e^{-L^{-1}\hat{M}\tau}\right)\mathrm{d}\tau L^{-1}B\mathbf{u} \tag{23}$$

In Rodriguez-Vega et al. (2019), we showed that $\mathbb{I} - R^\top$ is an invertible M-matrix, with all eigenvalues having positive real parts. Therefore, it can be shown that $-L^{-1}M$ has eigenvalues with negative-real parts, so it is a stable matrix. Thus, the asymptotic error is

$$\mathbf{e} = \lim_{t\to\infty} \mathbf{e}(t) = (M^{-1} - \hat{M}^{-1})B\mathbf{u} \tag{24}$$

where we have used the facts that the exponential function of a stable matrix goes to zero as time goes to infinity, and that

$$\int_0^\infty e^{At}\mathrm{d}t = -A^{-1} \tag{25}$$

for any stable invertible matrix A.

Sensitivity of turning ratio deviations. Suppose that the deviations $\xi_{i,j}$ are small, and that we want to calculate the error resulting from small nudges in a single value,

$$\frac{\partial \mathbf{e}}{\partial \xi_{i,j}} = \frac{\partial}{\partial \xi_{i,j}} \left((M^{-1} - \hat{M}^{-1}) B\mathbf{u} \right) = - \left(\frac{\partial}{\partial \xi_{i,j}} \hat{M}^{-1} \right) B\mathbf{u} \tag{26}$$

Consider the following theorem.

Theorem 2. *(Magnus and Neudecker (2019)) Let $A(t)$ be an invertible matrix that depends on a scalar parameter t. Then,*

$$\frac{\mathrm{d}A^{-1}}{\mathrm{d}t} = -A^{-1} \frac{\mathrm{d}A}{\mathrm{d}t} A^{-1} \tag{27}$$

Using Theorem 2, (26) can be written as

$$\frac{\partial \mathbf{e}}{\partial \xi_{i,j}} = \hat{M}^{-1} \frac{\partial \Xi^\top}{\partial \xi_{i,j}} V \hat{M}^{-1} B\mathbf{u} \tag{28}$$

Assume that the TR deviations are independent from each other. Thus,

$$\frac{\partial \Xi^\top}{\partial \xi_{i,j}} = \boldsymbol{\mu}_j \boldsymbol{\mu}_i^\top \tag{29}$$

where $\boldsymbol{\mu}_i$ is the i-th column of the identity matrix of suitable dimensions. Hence,

$$\frac{\partial \mathbf{e}}{\partial \xi_{i,j}} = \hat{M}^{-1} \boldsymbol{\mu}_j \boldsymbol{\mu}_i^\top V \hat{M}^{-1} B\mathbf{u} \tag{30}$$

Selection of intersections. Consider an intersection $n \in \mathcal{N}$ with incoming roads $\mathcal{I}(n)$ and outgoing roads $\mathcal{O}(n)$. Let $\boldsymbol{\xi}_i = \{\xi_{i,j}\}_{j \in \mathcal{O}(n)}$ for each $i \in \mathcal{I}(n)$. To quantify the effect of the perturbations $\xi_{i,j}$ on the error $\mathbf{e}$ we propose the following procedure:

1. Calculate the Jacobian matrix
$$J_i = \frac{\partial \mathbf{e}}{\partial \boldsymbol{\xi}_i} \tag{31}$$
for each incoming road $i \in \mathcal{I}(n)$.
2. Calculate the error energy due to perturbations in the TRs of i using the Frobenius norm of the Jacobian,
$$||J_i||_F^2 = \sum_{k \in \mathcal{E}} \sum_{j \in \mathcal{O}(n)} \left(\frac{\partial e_k}{\partial \xi_{i,j}} \right)^2. \tag{32}$$
3. Calculate the intersection weight as the total error energy due to its incoming roads,
$$w_n = \sum_{i \in \mathcal{I}(n)} ||J_i||_F^2 \tag{33}$$
4. Locate the available sensors in intersections with the highest values of w_n.

Note that the intersection weight can be simplified to

$$w_n = \sum_{i \in \mathcal{I}(n)} \sum_{j \in \mathcal{O}(n)} \sum_{k \in \mathcal{E}} \left(\sum_{p \in \mathcal{E}} \hat{M}_{k,j}^{-1} \hat{M}_{i,p}^{-1} v_i u_p \right)^2 \tag{34}$$

This heuristic procedure is a greedy approach as it assumes that deviations in the TR values are independent of each other.

4. Estimator implementation

4.1. Discrete-time estimator with variable input sampling times

To be able to use the estimator in practical applications, it is required to provide a discrete-time version of the dynamic equation. Furthermore, as the sampling times for the different data sources might vary, we propose an algorithm able to incorporate these measurements to estimate the traffic state.

Define the discrete time density $\boldsymbol{\rho}[k]$ such that it a sampled version of the density at times $t = k\Delta t$. Assuming that Δt is small enough, (8) can be discretized as

$$\hat{\boldsymbol{\rho}}[k+1] = \hat{\boldsymbol{\rho}}[k] + \Delta t A[k]\hat{\boldsymbol{\rho}}[k] + L^{-1}B\mathbf{u}[k] \tag{35}$$

where

$$A[k] = L^{-1}(\hat{R}^{\top}[k] - \mathbb{I})\hat{V}[k] \tag{36}$$

The discretization step Δt must be chosen such that the estimator is numerically stable. The Courant-Friedrichs-Lewy criterion establishes that a necessary condition for the discretization step is $\Delta t < \min_{\mathrm{i}}(\ell_i/v_i^{max})$, that is, the discretization time is less than the minimum traversal time for all roads in the network. Algorithm 1 allows using the discrete-time estimator with input data provided at arbitrary discrete times.

Algorithm 1. Density estimation

Inputs:

- Initial density condition $\boldsymbol{\rho}[0]$.
- Discretization time Δt.
- Graph $\mathcal{G}$ and road lengths L.
- Sensor data:
  - External inflows: Update times t_m^{φ} for $m \in \{1, 2, \ldots, m^{max}\}$ and data $\mathbf{u}[m]$.
  - Aggregate speeds: Update times t_i^v for $i \in \{1, 2, \ldots, i^{max}\}$ and data $\hat{V}[i]$.
  - Turning ratio data $\hat{R}$.

Outputs:

- Discrete density vector $\boldsymbol{\rho}[k]$ for $k \in \{0, 1, 2, \ldots, k^{max}\}$, where

$$k^{max} = \lceil \min(t_{i^{max}}^v, t_{m^{max}}^{\varphi})/\Delta t \rceil.$$

Steps:

1. Initialize $(i, k, m, T) \leftarrow (1, 0, 1, 0)$.
2. While $(i \leq i^{max}) \ \wedge \ (m \leq m^{max})$
   2.1. $T \leftarrow \min(t_i^v, t_m^{\varphi})$
   2.2. $n \leftarrow \left\lceil \dfrac{T}{\Delta t} \right\rceil$
   2.3. While $k < n$
      2.3.1. $V \leftarrow \hat{V}[i]$
      2.3.2. $A \leftarrow L^{-1}(\hat{R}^{\top} - \mathbb{I})V$
      2.3.3. $\boldsymbol{\rho}[k+1] \leftarrow (\mathbb{I} + \Delta t A)\boldsymbol{\rho}[k] + \Delta t L^{-1}B\mathbf{u}[m]$
      2.3.4. $k \leftarrow k + 1$
   2.4. If $T \geq t_i^v$ then $i \leftarrow i + 1$
   2.5. If $T \geq t_m^{\varphi}$ then $m \leftarrow m + 1$
3. Return $\boldsymbol{\rho}[k]$ $\forall k$.

To deal with the variable input-data times, Algorithm 1 uses the sample-and-hold scheme to update the density estimates.

4.2. Applicability to large-scale networks

The proposed models and algorithm in this paper have the objective of being applicable for large-scale networks. For this, two main requirements should be considered: availability of input data, and computational complexity.

4.2.1. Data availability

The proposed estimator requires the physical properties and layout of the traffic network (graph), measurement of the boundary flows, measurement of the road speeds, and knowledge of the TRs for all intersections. The graph and boundary flows are a widespread need among all TSE approaches. Graph information is generally readily available with some manual adjustments, and base topological data can be found in open sources. Flow sensor data is generally costly, and economical constraints can limit the size of networks that can be studied by the proposed methods in this work.

The availability of road mean speeds from FCD has greatly expanded during recent years and can be obtained from GPS providers in many countries. This data source has received great attention both in practical applications and in the traffic research literature. Although this data is not available for all links of a network, usually data is available for the most important roads that have a greater contribution to the overall dynamics.

The requirement of TR data is a very strict condition and is a very important limitation in many TSE approaches. To reduce these limitations, Section 3.3.1 discusses methods to estimate TR values using network properties that are easier to obtain in practice, such as the number of lanes, maximum road speed, and FRC. To complement these estimates, TR sensors can be located using the previously described method, though there is a trade-off between cost and overall estimation accuracy. As discussed below, the computational complexity of these methods is not high, and therefore they can be implemented for arbitrary network sizes.

4.2.2. Computational complexity

The main computation in Algorithm 1 is done in step 2.3.3. where there is the multiplication of a square matrix of dimension $|\mathcal{E}| \times |\mathcal{E}|$ with a vector of dimension $|\mathcal{E}|$, that needs to be repeated with a period less than Δt for a real-time application. Fortunately, each road in a traffic network has a limited number of downstream neighbors (usually between 1 and 4), so the corresponding graph $\mathcal{G}$ and TR matrix R are sparse. Define by η the average outdegree of the network's intersections. Then, the online computational complexity per unit time of the proposed algorithm in Big-O notation is

$$O\left(\frac{1}{\Delta t}\eta|\mathcal{E}|^2\right) \tag{37}$$

which represents an improvement of the naive non-sparse implementation of complexity $O(|\mathcal{E}|^3/\Delta t)$.

Regarding the offline complexity, two important stages are the calculation of the TR values using the methods of Section 3.3.1, and the TR sensor location method of Section 3.3.2. For instance, using Eqs. (18) and (19) to estimate TRs using road FRCs, requires the solution of an optimization problem, whose cost function involves the inversion of a $|\mathcal{E}| \times |\mathcal{E}|$ matrix, and has thus a computational complexity of at least $O(|\mathcal{E}|^3)$. Nevertheless, this operation is only required to be done once offline, and due to the limited number of parameters (number of considered FRC classes), this problem can be solved with common optimization solvers in a reasonable time.

For the calculation of the TR sensor locations, the proposed method requires the inversion of a $|\mathcal{E}|\times|\mathcal{E}|$ matrix, and then a series of iterated sums for the sensitivity weight in (34). Overall, this method has a computational complexity of

$$O(\eta^2|\mathcal{N}||\mathcal{E}| + |\mathcal{E}|^3). \tag{38}$$

This process is done offline.

5. Results

5.1. Setup description

The Grenoble Traffic Lab for urban networks (GTL-Ville)[2] is an experimental platform for real-time collection of traffic data coming from a network of sensors installed in the city of Grenoble. This platform also provides real-time traffic indicators and analysis oriented towards the users of the city, traffic operators, and researchers, which are available for download.

For this work, a section of the city center was selected to validate the proposed estimation approaches, which spans an area of 1.4 Km by 1 Km, and is composed of 463 intersections, 804 road sections, and 1234 allowed turns. This area is a representative section of the city, as it contains a variety of important and residential areas, traffic lights, shopping centers, etc. allowing the inclusion of many different traffic phenomena. This region is equipped with 40 stationary flow sensors comprising two different sensing technologies:

Induction loops. This technology consists of a coil buried under the pavement that perceives a change in the magnetic induction when a vehicle passes over it. In the section of interest, 17 single-loop sensors are available. These sensors are administered by the city's traffic administration (Metromobilité), which allows the use of this data. Information is received in real-time, with an aggregation time of 6 minutes. For each time period, received data is the number of detected vehicles and occupancy.

Microwave radars. These sensors periodically emit a pulse of radiation and then measure the reflections of these pulses. By measuring the properties of the reflected radiation, this technology can accurately measure the velocity and length of vehicles passing through the specified location, for every lane and direction. These sensors are leased and administered by a subcontractor. Data from these sensors are available in real-time, with an aggregation time of 1 minute. For each passing vehicle, the sensors save the time of passage, its velocity, length, and lane. 23 radar locations are available.

The layout of the zone of study and the flow sensor locations are shown in Figure 2. Furthermore, flow data is divided into three categories: inflow sensors count the vehicles entering the network $\mathbf{u}(t)$, outflow sensors count the vehicles exiting the network $\mathbf{y}(t)$, and validation sensors are located in the interior of the network and will be used to evaluate the performance of the estimator.

5.2. Validation using simulated vehicle trajectories

To evaluate the performance of the estimator under ideal conditions, we used simulated traces of individual vehicles using the well-known traffic microsimulator Aimsun (Barceló and Casas (2005)). This type of software models the position, speed, and acceleration of each vehicle according to the interaction with the other vehicles in each section and intersection. A model of the real network layout was constructed, which contains information about the number of lanes, speed limit, direction, possible turns at intersections, and length of each road.

The simulation inputs are the input demands and the TR values. As input demands, an arbitrary but realistic time-varying profile was applied to each boundary inflow. The TR parameters were initialized using the values described in Section 3.3.1. However, to simulate the uncertainty in the TRs, a uniform disturbance between -10% and 10% was added to the initial values. These random values were unknown by the estimator.

As simulation outputs, the position and speed of each vehicle were collected at each time step. From this raw data, the average road speeds and boundary flows are calculated by (15) and (14). For validation, the ground-truth density for road i is calculated as

$$\rho_i[k] = \frac{1}{\ell_i \tau} \int_{(k-1)\tau}^{k\tau} |\mathcal{V}_i(t)| \mathrm{d}t \tag{39}$$

where $\mathcal{V}_i(t)$ is the set of vehicle indexes contained in road i at time t, and τ is the duration of a time window that represents the measurement aggregation time.

For illustration purposes, Figure 3 shows the results for an arbitrary road, assuming that TR measurements are available for all intersections. The trajectories of the real and estimated density and flow

[2] http://gtlville.inrialpes.fr/

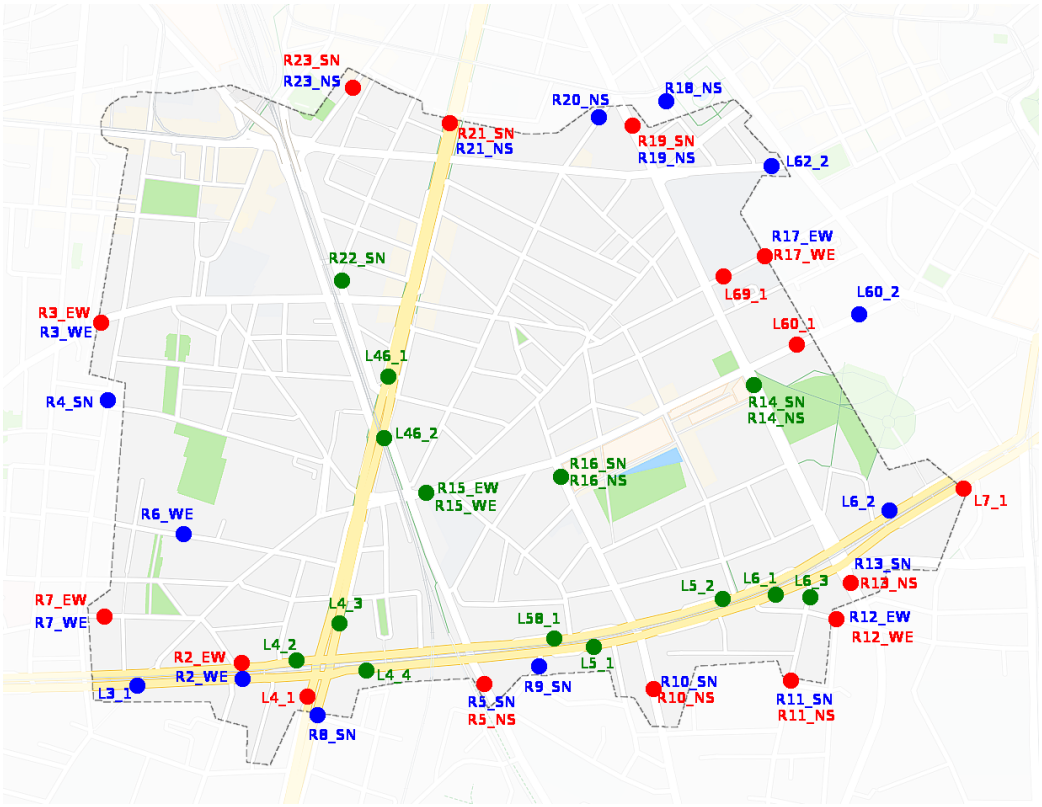

Figure 2: Location of flow sensors. Text indicators starting with "L" correspond to Induction Loops, whereas those starting with "R" indicate Radars. Blue sensors provide Inflow data, red sensors provide Outflow data, and green sensors provide Validation data.

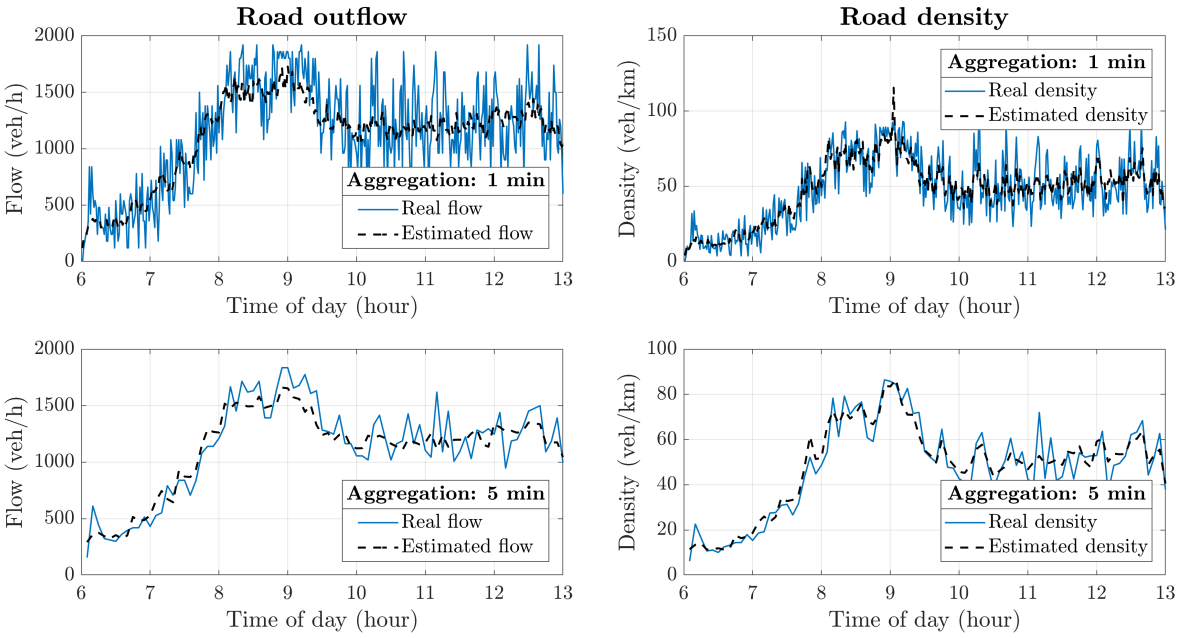

Figure 3: Time series for the trajectories of the real and estimated density and flow for one road using different aggregation times. **Top row** uses a window of 1 minute, whereas the **bottom row** uses one of 5 minutes.

are shown for two aggregation times, 1 and 5 minutes. Note that for both cases the estimates are close to the real values. Nevertheless, the real values present a higher dispersion, which increases as the aggregation time decreases. This dispersion is due to the stochastic behavior of individual vehicles that are not captured by the macroscopic model. As the aggregation time increases, the individual vehicle decisions are smoothed out, so the estimation error decreases. This shows that the proposed method behaves as a low-pass filter of the real flow and density profiles, which will provide better approximations for larger aggregation times.

To validate the proposed TR sensor location method, we considered multiple scenarios by varying the number of intersections to be equipped with TR sensors. The method presented in Section 3.3.2 to calculate the sensitivity of each intersection was used, taking as inputs the average boundary inflows, the road speeds, and the a priori TR estimates. The resulting sensitivity weights are shown in Fig. 4. Note that the values decrease very quickly, so it is expected that diminishing returns will be observed as more sensors are located. Thus, a few intersections could provide a value trade-off.

To quantify the error we consider as error metrics the Relative Mean Error (RME) and the Relative Absolute Error (RAE) which are defined as,

$$\mathrm{RME}_i = \frac{\left| \int_0^T \varphi_i^{\mathrm{out}}(t) - \hat{\varphi}_i^{\mathrm{out}}(t) \mathrm{d}t \right|}{\int_0^T \varphi_i^{\mathrm{out}}(t) \mathrm{d}t} \quad , \quad \mathrm{RAE}_i = \frac{\int_0^T \left| \varphi_i^{\mathrm{out}}(t) - \hat{\varphi}_i^{\mathrm{out}}(t) \right| \mathrm{d}t}{\int_0^T \varphi_i^{\mathrm{out}}(t) \mathrm{d}t} \ . \tag{40}$$

To take into account all roads, we use the Cumulative Distribution Function (CDF) for each metric: $y = \mathrm{CDF}(x)$ is the fraction of roads y which have an error value less than or equal to x.

Figure 5 shows the CDF for the density RME and RAE for 11 sensor configurations, from 0 intersections equipped with sensors, to 240 intersections equipped with sensors (100% of nodes)[3], with regular increments of 24 intersections (10%).

As is to be expected, the estimation accuracy improves by increasing the number of TR measurements, as the corresponding curves lie above-left of curves with fewer TR sensors. Note, however, that for most

[3]Only nodes with more than 1 outgoing road are considered.

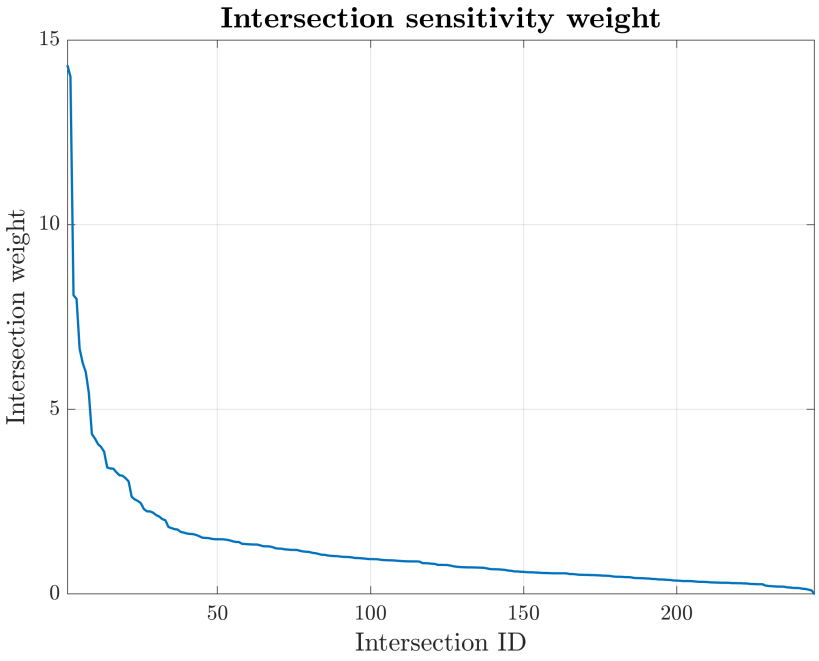

Figure 4: Sensitivity weight of each intersection: effect in the estimation error by small perturbations in the TRs. Intersection identifiers were sorted according to the weight.

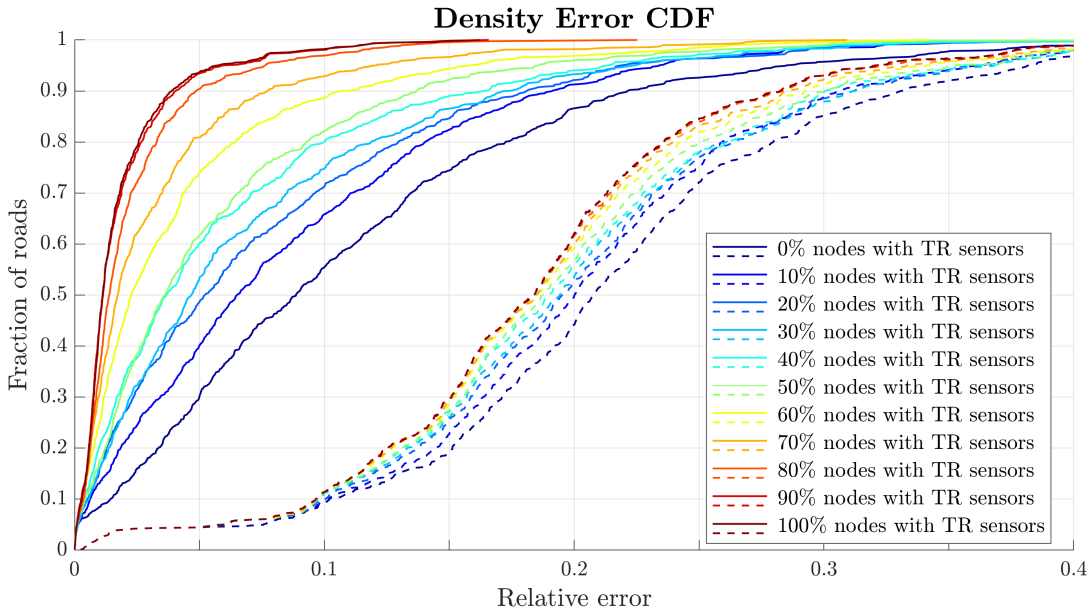

Figure 5: Relative error CDF for the density estimation using varying numbers of intersections (nodes) equipped with TR sensors, with increments of 24 sensors. Solid lines correspond to the RME, and dashed lines correspond to the RAE.

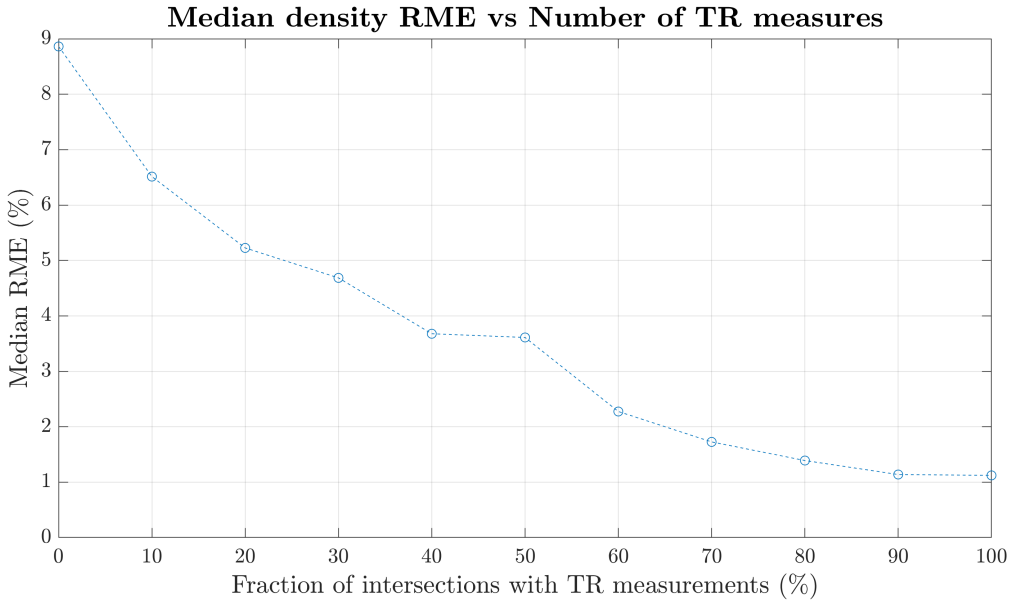

Figure 6: Median RME for density estimation using varying number of intersection equipped with TR sensors.

cases, the spacing between curves decreases as the number of sensors increases. For example, the spacing between the curves corresponding to 0% and 10% nodes with sensors is less than the spacing between the curves corresponding to 60% and 70%. For a more detailed analysis, Fig. 6 shows the median error obtained for the varying number of TR sensors. The slope of the dashed line is steeper at the beginning of the plot (small number of TR sensors) and then tends to flatten. This supports the claim that there are diminishing returns in the addition of more TR sensors, once the most important intersections are measured. This was previously predicted from Fig. 4. Additionally, it can be seen that for all scenarios, the estimator achieves good accuracy, as they all have a median RME below 9% and a median RAE below 22%. However, this last claim depends on the magnitude of the TR deviations of the unmonitored intersections.

For further evaluation of the method, Fig. 7 shows the CDF for the RME and RAE of the density estimation using only 12 TR sensors. The equipped intersections are chosen randomly, and by using the proposed optimal approach. The figure shows that our sensor location method obtains a better result compared to a random choice, as the median RME decreases from 9% to 7% (22% of relative improvement).

5.3. Validation using real data

In this section, we discuss the implementation of the real flow sensors described in Section 5.1 and the addition of new TR sensors, to evaluate the performance of the TR estimation methods of Section 3.3.1, and the flow-density estimation algorithms.

5.3.1. Turning ratio measurements from real data

Due to budget constraints, a set of 12 intersections were selected to locate TR sensors using the proposed method in Section 3.3.2, using as inputs the flow measurements from the inflow sensors, and road speeds from FCD. The TR sensor locations are shown in Fig. 8, corresponding to the intersections with the highest sensitivity weights. In Section 3.3.1, we proposed two methods to provide estimated values for the TRs using network properties:

- **Method "Cap":** TRs are proportional to outbound road capacities, using (17).
- **Method "FRC":** TRs are proportional to the FRC class weights of outbound roads, using (18).

Remark 1. Although, we claim that Method FRC provides better accuracy for the flow-density estimation (see Section 5.3.2), the TR estimates of Method Cap were used to compute sensor locations. This

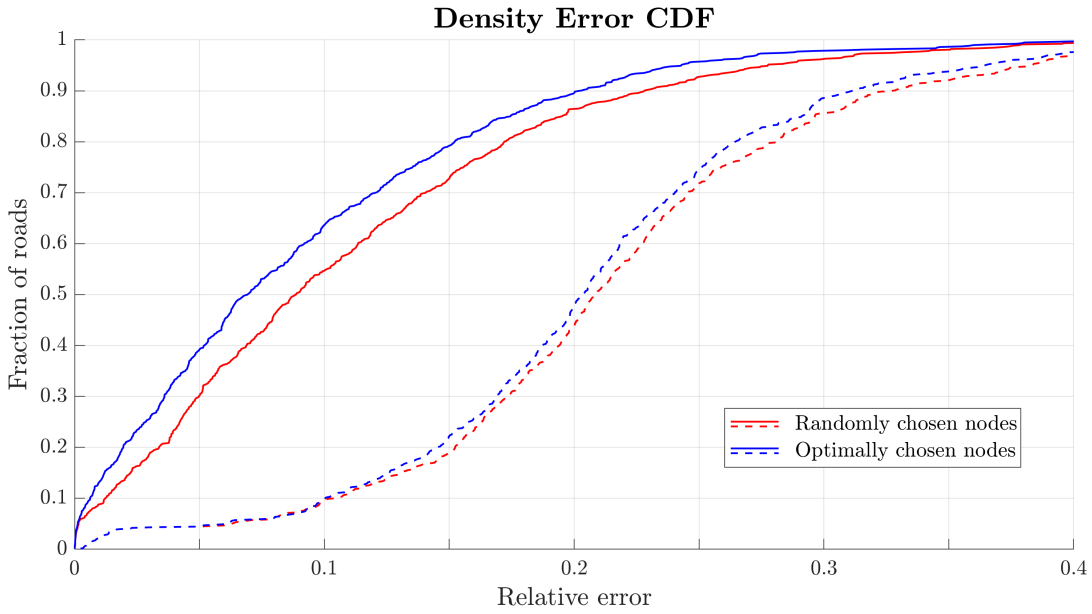

Figure 7: Relative error CDF for the density estimation by locating 12 TR sensors according to *a)* a random selection, and *b)* the proposed approach. Solid lines correspond to the RME, and dashed lines correspond to the RAE.

is because at the time of deciding the locations, Method FRC had not yet been proposed, and its need was identified after the fact, during testing. However, both methods provide similar sensor locations, as intersections with higher connectivity are heavily weighted by the sensor location method, and many intersections consist of roads with the same FRC and capacity, for which both methods predict the same TRs.

To obtain direct measurements of TRs, Bluetooth (BT) reader devices were located at the adjacent roads for a selection of intersections. These sensors work by detecting the IDs of BT devices in a vehicle and then match this identity with those of other locations, such that information about the number of counts and passing times are available. This technology has received recent attention as in Bhaskar and Chung (2013) and Abbott-jard et al. (2013), where the use of BT to provide complementary traffic is explored. Although these sensors are used mostly to measure traveling time, their functionality also allows the measurement of O/D matrices and TRs.

Due to technical and economical constraints, this data is not available in real-time. Instead, campaigns of one-week duration were organized for each intersection, for a total of 4 campaigns. The timetable for these campaigns is shown in Table 2.

To illustrate the data collection process, Fig. 9 shows the location of BT devices around one of the chosen intersections. For each sensor pair, the detected vehicle IDs are compared with their corresponding timestamps. Thus, it is possible to assign the origin and destination roads for each of the detected vehicles. This data is aggregated during a time period of 1 hour, so the available information is the

Table 2: Measurement campaigns for the collection of BT data, in 2020.

Campaign	Intersections	Begin date	End date
1	5, 6, 7	Sep 21 - 12h00	Sep 28 - 13h00
2	1, 4, 10, 11	Sep 30 - 12h00	Oct 08 - 8h00
3	3, 9, 12	Oct 12 - 13h00	Oct 21 - 8h00
4	2, 8	Oct 21 - 14h00	Oct 28 - 14h00

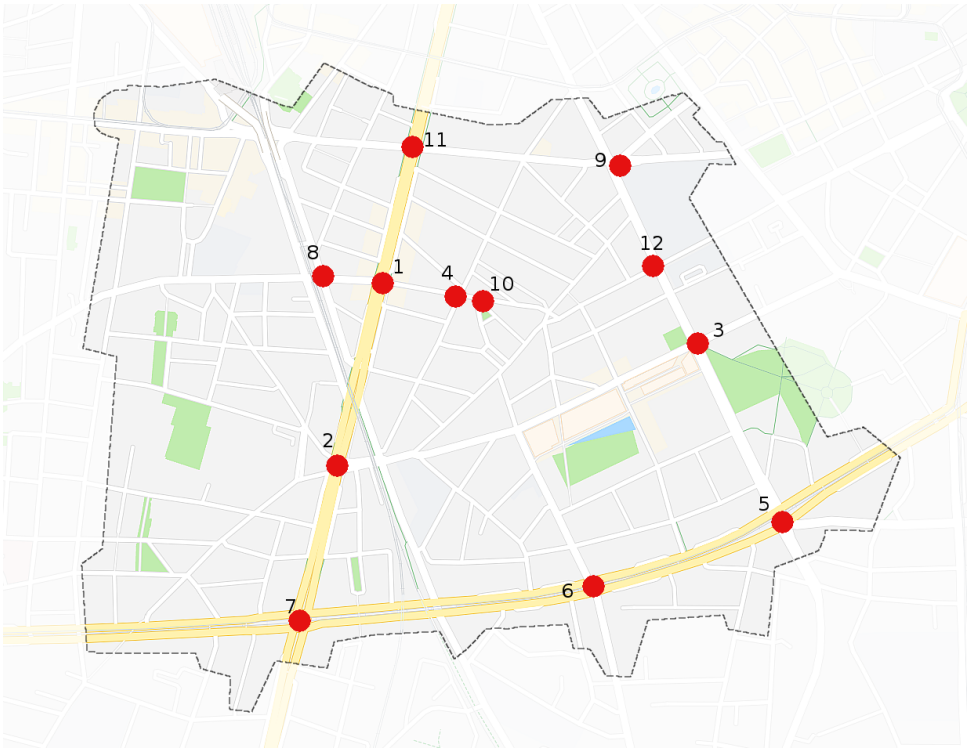

Figure 8: Intersections used to collect TR data using BT identifiers in red, with their location ID.

number of counts for each of the possible turns in the intersection. From this, TRs are computed as

$$r_{i,j}^{\mathrm{BT}}(t) = \frac{\mathrm{Counts}(i,j,t)}{\sum_k \mathrm{Counts}(i,k,t)} \tag{41}$$

where $\mathrm{Counts}(i,j,t)$ is the number of counts from road i to road j via the BT during the time interval $[t-\Delta t, t]$, with $\Delta t = 1$ hour.

To evaluate the quality of the data, Fig. 10 shows the time series of the TR values for one road. A zoom of this figure for only one day (Oct 22, 2020) is shown in Figure 11. Note that during the day hours (6h00 to 20h00), the TR values are fairly constant, lying in a range of ± 0.07 around the mean. The peaks that can be seen between 23h00 and 6h00 are due to the low vehicular flow during these times, which results in a very low number of measurements and the estimation may not be good.

The mean TR values are computed using the total number of vehicle detections during the duration of the campaign,

$$r_{i,j}^{\mathrm{BT}} = \frac{\mathrm{TotalCounts}(i,j)}{\sum_k \mathrm{TotalCounts}(i,k)}. \tag{42}$$

Nevertheless, these sensors are only able to detect vehicles equipped with BT devices, for which the penetration rate is unknown. In this paper, we assume that vehicle route-choice is independent of BT availability. Thus, BT measurements provide an unbiased estimate of the TRs. Some authors have reported BT penetration rates of 12% (Cvetek et al. (2019)) and up to 30% in Europe (Steinmaßl et al. (2021)). Taking into account the daytime hours (6h00 to 20h00) for intersection 2, the turn with the minimum number of counts reported an average of 8 detections per hour, and the turn with the maximum number of counts reported an average of 92 detections per hour. Over all turns, the average is 26 detections per hour. During the duration of the campaigns, this results in a total average count of 2500 per turn. Therefore, the estimated TRs have a very high probability of having a close value to the real values.

Figure 9: Location of BT devices around intersection 2, with their corresponding ID in database.

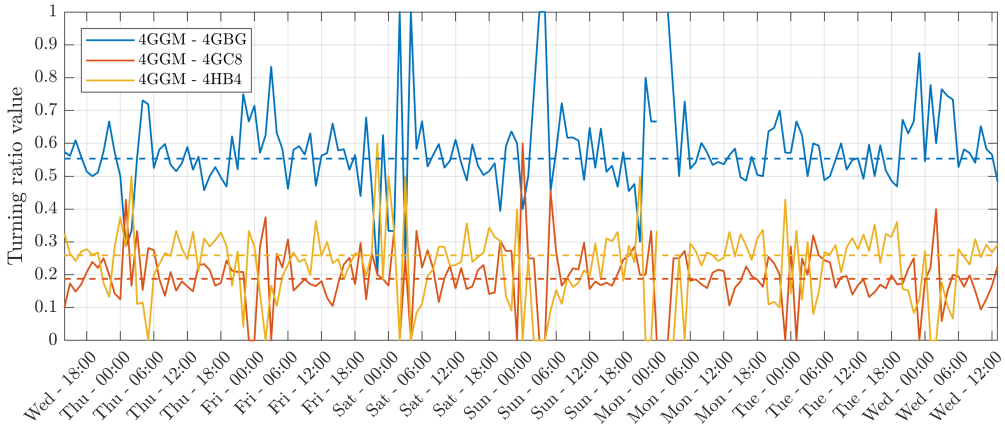

Figure 10: Time series for the estimated TRs for one road in intersection 2, from Oct 21 to Oct 28, 2020. Dashed lines are the mean values for each turn.

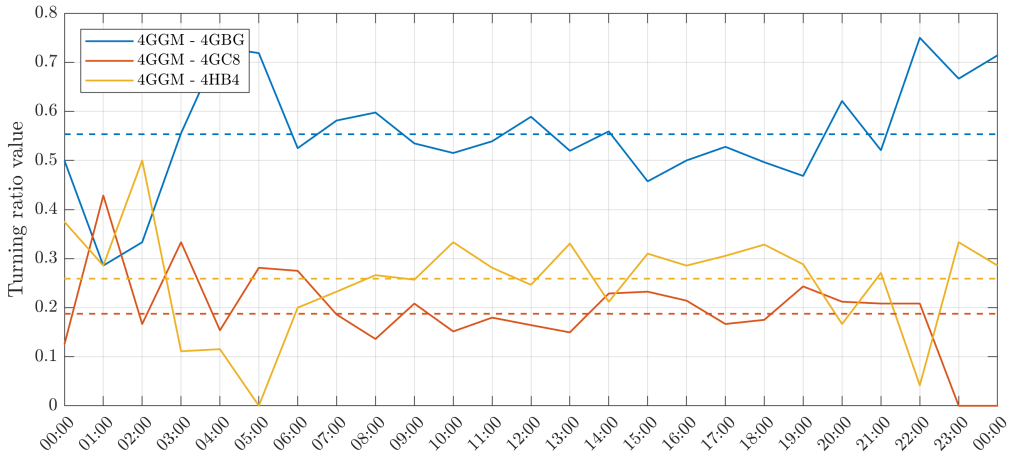

Figure 11: Time series for the estimated TRs for one road in intersection 2, for Thursday Oct 22, 2020. Dashed lines are the mean values for each turn.

Table 3: Value of FRC weights for the estimation of TR parameters.

Class index	1	2	3	4	5	6	7
Class weight θ	1.00	N/A	0.99	0.50	0.23	0.13	0.03

Comparison of TR estimates with measurements. The TR values obtained during the measuring campaigns using the BT sensors were used to evaluate the accuracy of the estimates provided by Methods Cap and FRC. For the former, the input data are the number of lanes and max. speed of each road, which are readily known, and no further calibration is required.

For Method FRC, Figure 12 shows the classification of each road in the zone of interest. For this area, there are no roads with FRC 2. The weight vector $\boldsymbol{\theta}$ needs to be calibrated using the optimization problem in (19). The resulting values are shown in Table 3. Figure 13 shows a visualization of the weights as a function of the FRC. Note that as the importance of the road decreases, so does the corresponding class weight as is to be expected. The missing values for FRC 2 are due to the lack of roads with this class in the area of interest.

For each Method, the error is calculated by the difference between the predicted estimates and the measured values for each TR value, for a total of 142 error samples. We found that both error signals followed very similar Gaussian distributions: Method Cap presented a mean of 0.00 and a standard deviation of 0.18; Method FRC presented a mean of 0.00 and a standard deviation of 0.22. Further comparison of both Methods is done in the following section.

5.3.2. Flow and density estimation

For evaluation purposes, we considered the data collected for January 8, 2021. Figure 14 shows the time series for the real and estimated flows for a selection of validation sensors. The flow estimates were obtained using Algorithm 1 using the flow and speed input data, and the TR values obtained from the BT measuring campaigns for the selected intersections, and using Methods Cap and FRC for the non-measured ones. The figure only shows the estimates using Method FRC for illustration purposes. The real and estimated trajectories are shown with a time aggregation of 10 minutes, to smooth out the dispersion due to traffic lights. Note that for most cases, the estimated and real values have a very similar trajectory. The mismatches obtained for some of the sensors can be attributed to several factors. The main source of error is due to deviations between the real and estimated turning ratios. As these parameters were computed using a simplifying hypothesis using the FRC, there are intersections for which the obtained values present error. However, this method provides a simple to use manner to compute these parameters for large networks with easily obtainable information and provides good initial results for a large number of locations which can be improved with time. Another possible error source is the presence of internal sources and sinks of traffic flow which are not taken into account.

Figure 12: FRC for the Grenoble network, provided by TomTom.

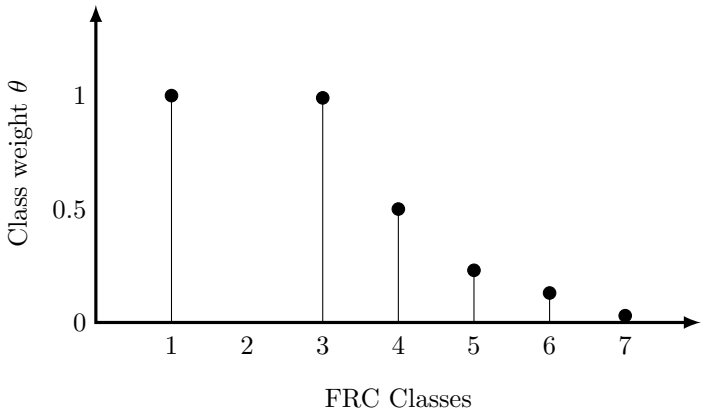

Figure 13: Visualization of the class weights for each FRC class.

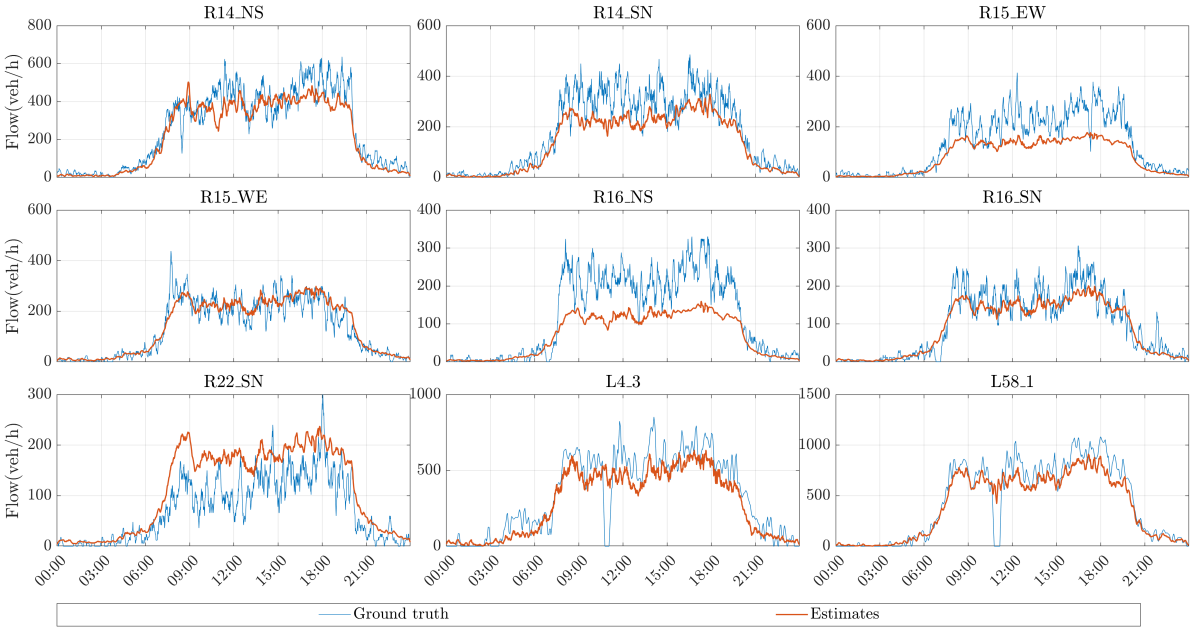

Figure 14: Values for the real and estimated flows.

Figure 15 shows the obtained estimation RME and RAE using the available validation sensors. The metrics are shown for 4 scenarios:

1. Method FRC used to calculate TR values, without including the BT measurements.
2. Method FRC used to calculate TR values, including the BT measurements for 12 intersections.
3. Method Cap used to calculate TR values, without including the BT measurements.
4. Method Cap used to calculate TR values, including the BT measurements for 12 intersections.

Method FRC outperforms the results of Method Cap, with and without the inclusion of the measured TR values. Furthermore, the inclusion of BT measures greatly improves the accuracy of Method FRC, but only slightly increases the accuracy of Method Cap for low error locations.

For the best-case scenario, the RME shows that the proposed estimator provides close estimates to the real values, with a median of 0.16, and a maximum of 0.44. When considering the RAE, the error increases as this metric considers not only the differences between the mean trajectories but also takes into account the dispersion of the real data. Nevertheless, the median RAE is 0.29, and maximum RAE is 0.46, showing a good agreement of the estimation with the real data. These results are similar to those obtained for the simulated scenario in Figs. 5 and 7. The error obtained with the real data is higher, but this is due to the added uncertainty in this case, and that only a small number of validation locations are available.

5.4. Real time demonstrator

This work uses data obtained from the experimental platform GTL-Ville. This platform also has implemented the flow and density estimation method presented in this work, to provided real-time indicators to the users of the network, city traffic management operators, and researchers. Currently, a real-time demonstrator of the proposed methods is implemented. The density estimates are visualized in three different ways, as described below.

Roads. This visualization shows the direct results of the estimator. Each road is given a density value according to the implementation of Algorithm 1. This visualization is shown in Fig. 16. Note from the figure that only the considered zone of the city center is selected, according to the sensor instrumentation. In the top right, the menu shows the different indicators, visualizations, and data sources that can be selected. For the scope of this paper, the entries in the menu "Traffic occupancy" are of interest.

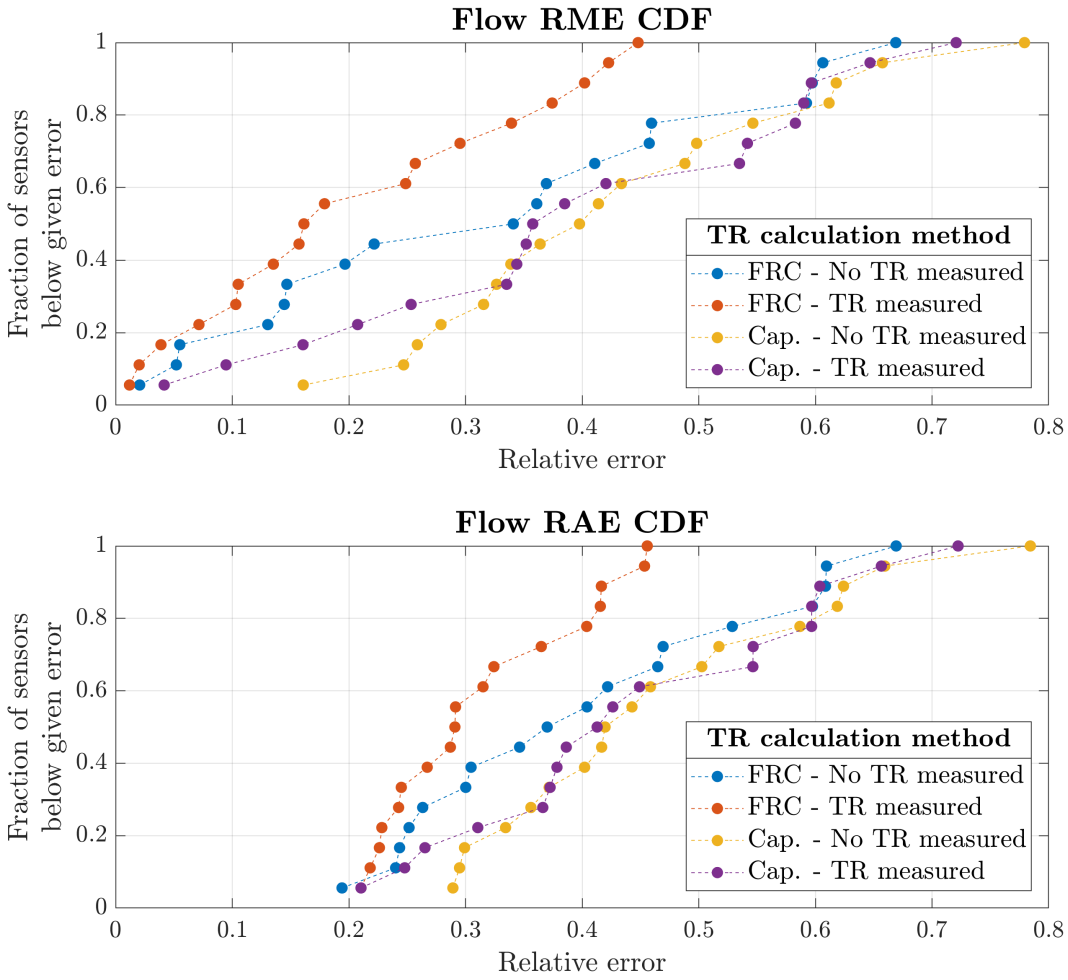

Figure 15: Relative estimation error CDF for the validation sensors using different TR values.

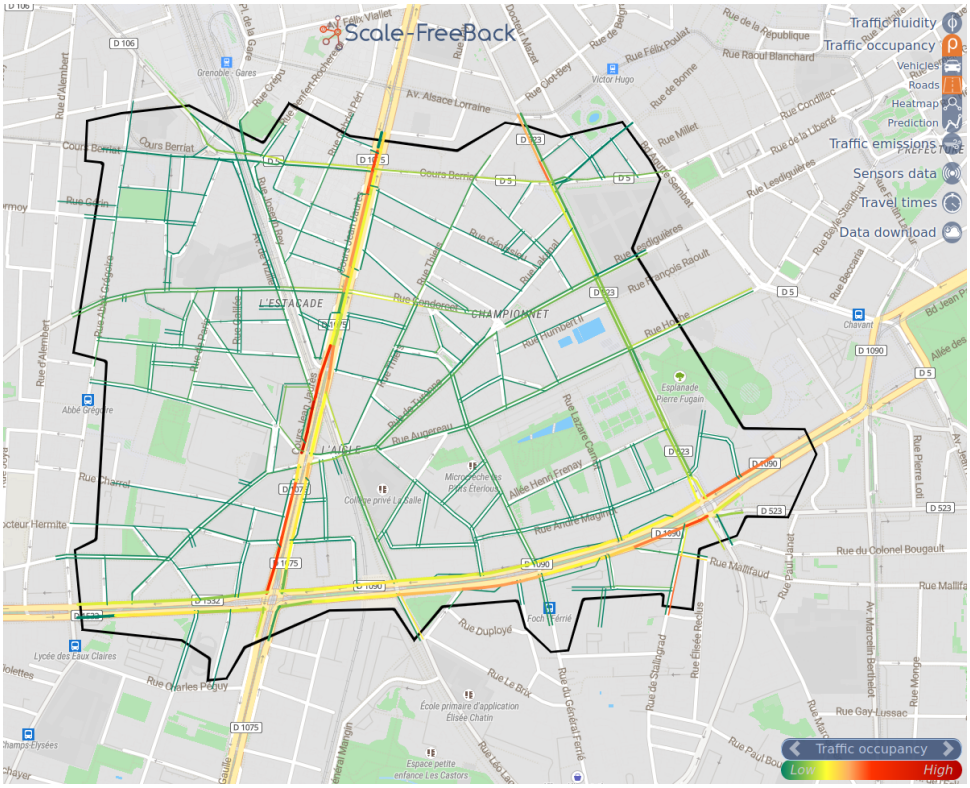

Figure 16: Demonstration of density reconstruction in real time. Taken from GTL-Ville website.

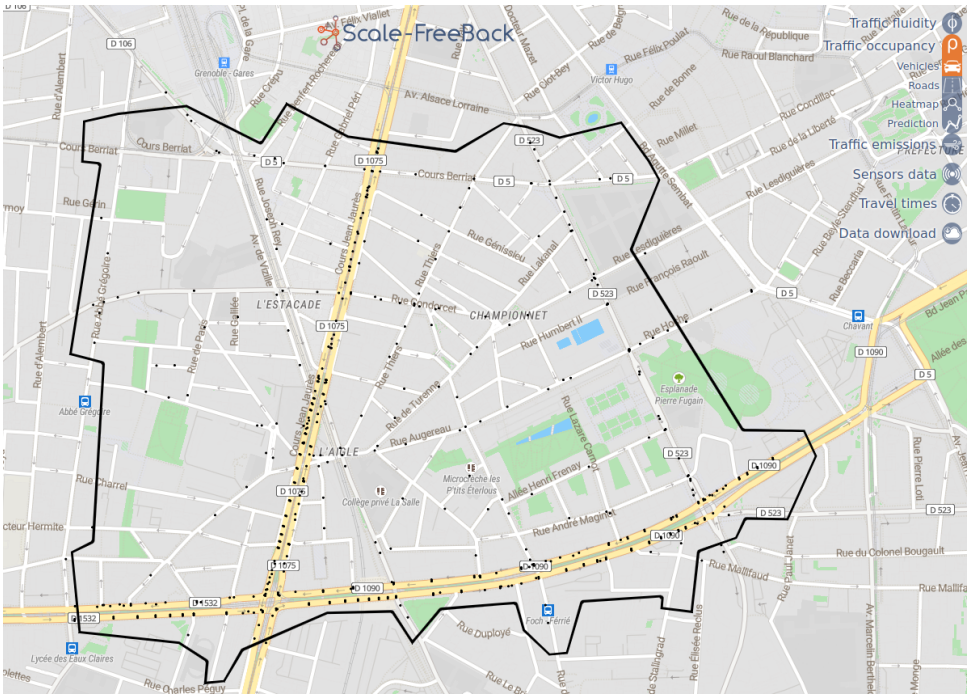

Figure 17: Estimated position of vehicles in roads based on density estimation. Taken from GTL-Ville website.

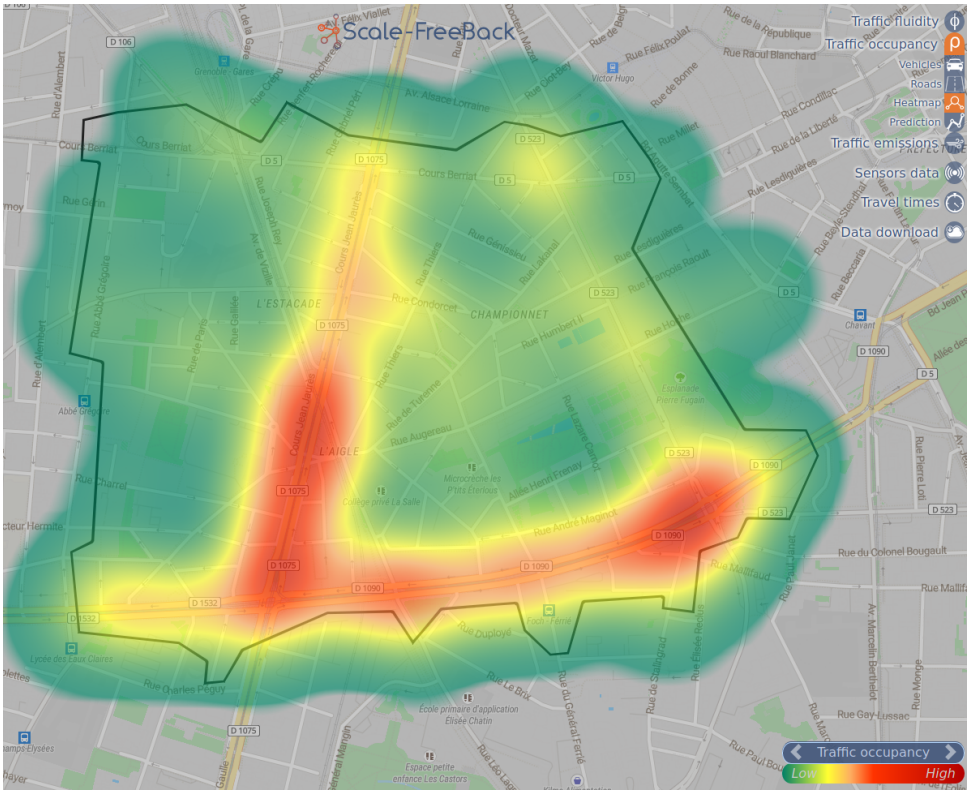

Figure 18: Two-dimensional density heat map. Taken from GTL-Ville website.

Vehicles. As shown in Fig. 17, this visualization shows the likely position of vehicles inside of each road represented by the black dots. From density data on each road, it is possible to calculate the estimated vehicle mass using the linear conversion

$$\mathbf{m}(t) = L\boldsymbol{\rho}(t) \tag{43}$$

where $\mathbf{m}(t)$ is the number of vehicles contained in each road at time t. With this information, the specified number of dots corresponding to vehicles are placed along each road according to an arbitrary distribution. Although the precise vehicle positions are unknown, this visualization is a ballpark estimate which can be more intuitive to users.

Heatmap. The previous visualizations correspond to the 1D density distribution inside each road section. However, these representations are hard to visualize easily. To make the application easier to use for the general public and users of the traffic network, we propose a 2D visualization of the traffic density, as shown in Fig. 18. In this visualization, the 1D density is transformed into a new variable that refers to the number of vehicles per unit area. This visualization allows to easily identify regions of the zone that exhibit high congestion, and evidence the more traversed regions of the city center. However, this representation is no longer constrained to the original graph of the network. This 2D visualization is obtained using the Density Kernel Estimation method presented in Mollier et al. (2019), which is a convolution of the vehicle positions with a two-dimensional gaussian. This representation can be useful to infer areas of concentrated vehicle emissions and noise pollution.

6. Discussion

As discussed in the simulated and real-data results, the proposed estimator can produce close estimates to the ground-truth trajectories for most of the roads of the considered traffic network. However, some limitations of the method are also seen. As constant TRs are used for all intersections, and as FCD speeds are collected during a significant time interval in practice, the estimator exhibits a low-pass filter behavior and is unable to accurately reconstruct high-frequency components. This is more visible in the

validation with real data, where traffic light signals present in the real system introduce dispersion in the ground truth data. Therefore, the proposed approach provides meaningful results when considering aggregated variables during a time window, in the order of 5 to 10 minutes to smooth out the dispersion due to traffic lights. Furthermore, it is seen that uncertainty in the TR parameters can generate large errors in some locations. However, this uncertainty can be decreased by increasing the number of intersections where TRs are measured.

Another important source of uncertainty is the quality of the FCD to produce speed information, especially in the case of very congested situations due to the queue spillback of roads at maximum density. When a road becomes full, the vehicles of the upstream roads would not be able to flow out, reducing their speed to 0. During a correct functioning of the system with accurate speed data, this is captured by the estimation algorithm, expanding the congestion upstream. However, it was seen in some situations that a delay in the update of the speed values, such that the maximum density of the congested road is surpassed, until the correct speed of the upstream road is updated. To avoid this unrealistic behavior, an extra step can be added to the algorithm to stop inflows from coming to a fully congested road.

To contextualize our proposed method in the TSE literature, we compare our approach to other related works. For instance, Bekiaris-Liberis et al. (2016) proposed a very similar model, where mean speed information from connected vehicles is used directly in combination with the density estimates to compute the outflow of each section. Although similar to our proposed method, their results are restricted to the case of highways only, and their mathematical results on the observability and convergence of the estimator cannot be extended to the general case of urban networks. The authors tested their approach using simulations based on the macroscopic second-order model METANET (Messmer and Papageorgiou (1990)), showing good results with a reported error ranging between 6.75 and 7.1 %. These results are similar to our microscopic simulation results in Section 5.2 with a mean error between 7 and 9% using 12 TR sensors. Although the authors used real data to calibrate the speed measurement noise, no validation of density or flow reconstruction using real data was provided.

Although TSE approaches applied to the case of urban networks are limited in comparison to the case of highways, some works deal with a similar problem as the one discussed in this paper, but using different approaches. For instance, Nantes et al. (2016) considers the estimation of the flow and density for each road section in an urban corridor. Their approach is based on the CTM model, so they require FD. Furthermore, they require information about the traffic light cycles. Regarding input data, this work also considered road speeds from a mixture of GPS and BT sources, and flow data from stationary sensors. Data and model are integrated using a Kalman Filter. Using simulations, the authors reported the method's accuracy using the Similarity score metric, with a best-case relative-mean-similarity of 77.29% (22.71% error) between ground truth and estimates.

In contrast to this method, our proposed approach does not depend on the FD. Instead, the direct use of the speed data allows us to consider the effects of different causes on the speed values, other than only density. Also, as we do not use the KF approach, the iterative cycles in the implementation of the algorithm have a smaller computational cost, so it is easier to implement for larger networks. Although the accuracy of both methods is hard to compare as the reported metrics and underlying study cases differ, the simulated results are not dissimilar, as we reported a median RME error of 7% and a median RAE of 21% in the simulated scenario in Fig. 7. The work of Nantes et al. (2016) has no real data validation.

A very similar data-based TSE approach is presented in Rostami Shahrbabaki et al. (2018) and expanded in Rostami Shahrbabaki et al. (2020). In these works, the authors also propose the use of the speed data, from connected vehicles, to calculate the outflows of each road, without the need for the FD. However, these works use much more detailed input data, as they assume the availability of position and speed information about individual vehicles (traces). The authors can subdivide each road section into two cells: a congested cell (queue) and an uncongested cell. Although very interesting, such detailed information may be difficult to obtain in many practical applications, as privacy regulations in some cities do not allow the processing of this type of data. The authors validated their approach using simulated data with good accuracy, although no validation with real data was presented.

7. Conclusions

In this work, we proposed a method to estimate the dynamical evolution of vehicular flow and density in large urban traffic networks by using heterogeneous data sources: counting flow sensors, Bluetooth

vehicle identifiers, and Floating Car Data. We discussed the implementation of these types of data in real case scenarios such as varying sampling rates and time intervals, and how to take them into account in the estimator implementation.

The proposed estimator is data-based, as it is only based on the basic flow conservation law and road outflows are computed using road speed data directly from FCD. As such, this method does not use the Fundamental Diagram. However, the knowledge of the TRs for every intersection is required. To simplify this requirement, we introduce a TR sensor location scheme that takes as input a priori values for these parameters to identify the intersections for which small perturbations in the TR values would generate the highest error.

We tested the proposed method using both, microsimulations using the well-known software Aimsun, and the experimental platform GTL-Ville that provides real traffic data for the city of Grenoble. The results show that the estimator performs well, as it can estimate the overall trajectory of the traffic states, with many roads presenting little errors. Although the problem of TSE in large networks is challenging, the obtained results are encouraging as the estimated flow for individual roads is very close to the ground truth data provided by sensors.

Acknowledgments

This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement 694209), http://scale-freeback.eu.

Appendix A. Proof of Theorem 1

Consider the linear transformation $\mathbf{x}(t) = L\mathbf{e}(t)$, subject to

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{x}(t) = (R^\top - \mathbb{I})L^{-1}V(t)\mathbf{x}(t) \tag{A.1}$$

note that L is positive definite, and therefore (11) is asymptotically stable if and only if (A.1) is asymptotically stable. Consider the Lyapunov candidate function

$$f(\mathbf{x}, t) = \mathbf{x}^\top P^{-1}(t)\mathbf{x} \tag{A.2}$$

where

$$P(t) = \mathrm{diag}(\mathbf{p}(t)) \tag{A.3}$$

and

$$\mathbf{p}(t) = V^{-1}(t)L(\mathbb{I} - R^\top)^{-1}\mathbf{1} \tag{A.4}$$

As $\mathbb{I} - R^\top$ is an invertible M-matrix, its inverse elements are non-negative. Moreover, as $V(t)$ has positive diagonal entries and is bounded, $P(t)$ and $P^{-1}(t)$ are bounded diagonal matrices with positive diagonal entries, so f is positive definite. The derivative of f is

$$\frac{\mathrm{d}}{\mathrm{d}t}f(\mathbf{x}, t) = \mathbf{x}^\top \left(V(t)L^{-1}(R - \mathbb{I})P^{-1}(t) + P^{-1}(t)(R^\top - \mathbb{I})L^{-1}V(t) + \frac{\mathrm{d}}{\mathrm{d}t}P^{-1}(t) \right)\mathbf{x}. \tag{A.5}$$

Consider

$$Q(t) = P(t)V(t)L^{-1}(R - \mathbb{I}) + (R^\top - \mathbb{I})L^{-1}V(t)P(t) - \frac{\mathrm{d}}{\mathrm{d}t}P(t) \tag{A.6}$$

Note that

$$\frac{\mathrm{d}}{\mathrm{d}t}f(t) = \mathbf{x}^\top(t)P^{-1}(t)Q(t)P^{-1}(t)\mathbf{x}(t) \tag{A.7}$$

where we have use the fact that

$$\frac{\mathrm{d}}{\mathrm{d}t}P^{-1}(t) = -P^{-1}(t)\left(\frac{\mathrm{d}}{\mathrm{d}t}P(t)\right)P^{-1}(t) \tag{A.8}$$

Furthermore, as $P(t)$ is positive definite, $\frac{\mathrm{d}}{\mathrm{d}t}f(\mathbf{x},t)$ is negative-definite if and only if $Q(t)$ is negative-definite.

Consider the first two terms of $Q(t)$,

$$Q_1(t) = P(t)V(t)L^{-1}(R-\mathbb{I}) + (R^\top - \mathbb{I})L^{-1}V(t)P(t) \tag{A.9}$$

Let $q_{i,j}(t)$ be the element in the (i,j) position of $Q_1(t)$. According to the Gershgorin circle theorem, the i-th eigenvalue of $Q_1(t)$, $\lambda_i(t)$, is contained in a disk with center $q_{i,i}(t)$ and radius

$$\sum_{j\neq i} |q_{i,j}(t)|. \tag{A.10}$$

In this specific case, we have

$$q_{i,i} = -2\frac{p_i(t)v_i(t)}{\ell_i} \tag{A.11}$$

and

$$\begin{aligned} \sum_{j\neq i} |q_{i,j(t)}| &= \sum_j \left(\frac{p_i(t)v_i(t)}{\ell_i} r_{i,j}\right) + \sum_j \left(\frac{p_j(t)v_j(t)}{\ell_j} r_{j,i}\right) \\ &\leq \frac{p_i(t)v_i(t)}{\ell_i} + \sum_j \left(\frac{p_j(t)v_j(t)}{\ell_j} r_{j,i}\right) \end{aligned} \tag{A.12}$$

where we have used the property

$$\sum_j r_{i,j} = \begin{cases} 1 & \text{if} \quad i \notin \mathcal{E}^{\text{out}} \\ 0 & \text{if} \quad i \in \mathcal{E}^{\text{out}} \end{cases} \tag{A.13}$$

and $r_{i,i} = 0$ for all i.

As the eigenvalues of $Q_1(t)$ are real, they can be upper-bounded by

$$\lambda_i(t) \leq z_i(t) \tag{A.14}$$

where

$$z_i(t) = -\frac{p_i(t)v_i(t)}{\ell_i} + \sum_{j\neq i} \left(\frac{p_j(t)v_j(t)}{\ell_j} r_{j,i}\right) \tag{A.15}$$

is the sum of the center and radius of the Gershgorin disk. This results in a system of equations of the form

$$\mathbf{z}(t) = -(\mathbb{I} - R^\top)L^{-1}V(t)\mathbf{p}(t) \tag{A.16}$$

Substitution of (A.4) into (A.16) gives $\mathbf{z}(t) = -\mathbf{1}$. This implies that all eigenvalues of $Q_1(t)$ are bounded above by -1, and thus,

$$Q_1(t) \leq -\mathbb{I} \tag{A.17}$$

Now consider $\frac{\mathrm{d}}{\mathrm{d}t}P(t)$, whose i-th diagonal entry is

$$\frac{\mathrm{d}}{\mathrm{d}t}p_i(t) = -\frac{\ell_i}{v_i^2(t)}\frac{\mathrm{d}}{\mathrm{d}t}v_i(t)\sum_j m_{i,j} \tag{A.18}$$

where $m_{i,j}$ is the (i,j) entry of $(\mathbb{I} - R^\top)^{-1}$. Using (13),

$$\frac{\mathrm{d}}{\mathrm{d}t}p_i(t) \geq -\frac{v_{min}^2}{v_i^2(t)}\frac{\ell_i}{\ell_{max}}\frac{\sum_j m_{i,j}}{||(\mathbb{I}-R^\top)^{-1}||_\infty}(1-\epsilon) \geq -(1-\epsilon). \tag{A.19}$$

The last inequality follows from the fact that each of the fractions in the equation are less than or equal to 1. This implies

$$\frac{\mathrm{d}}{\mathrm{d}t}P(t) \geq -(1-\epsilon)\mathbb{I}. \tag{A.20}$$

Substitution of (A.17) and (A.20) into

$$Q(t) = Q_1(t) - \frac{\mathrm{d}}{\mathrm{d}t}P(t) \tag{A.21}$$

gives

$$Q(t) \leq -\epsilon\mathbb{I} \tag{A.22}$$

Thus, the Lyapunov function $f(\mathbf{x}, t)$ is positive-definite with negative-definite derivative, which implies that (A.1) (and by extension (11)) is asymptotically stable.

References

Abbott-jard, M., Shah, H., Bhaskar, A., 2013. Empirical evaluation of Bluetooth and Wifi scanning for road transport, in: Australasian Transport Research Forum, Brisbane, Australia. pp. 1–14.

Barceló, J., Casas, J., 2005. Dynamic Network Simulation with AIMSUN, in: Simulation Approaches in Transportation Analysis. Springer-Verlag, New York, pp. 57–98. doi:10.1007/0-387-24109-4_3.

Bekiaris-Liberis, N., Roncoli, C., Papageorgiou, M., 2016. Highway Traffic State Estimation With Mixed Connected and Conventional Vehicles. IEEE Transactions on Intelligent Transportation Systems 17, 3484–3497. doi:10.1109/TITS.2016.2552639.

Bhaskar, A., Chung, E., 2013. Fundamental understanding on the use of Bluetooth scanner as a complementary transport data. Transportation Research Part C: Emerging Technologies 37, 42–72. doi:10.1016/j.trc.2013.09.013.

Cvetek, D., Bojić, V., Jelušić, N., Muštra, M., 2019. Initial Bluetooth probe vehicle penetration rate analysis: a case study in the city of Zagreb, in: International Scientific Conference "Science and Traffic Development" (ZIRP 2019), Opatija, Croatia. pp. 113–123.

Daganzo, C.F., 1994. The cell transmission model: A dynamic representation of highway traffic consistent with the hydrodynamic theory. Transportation Research Part B 28, 269–287. doi:10.1016/0191-2615(94)90002-7.

Daganzo, C.F., 1995. The cell transmission model, part II: Network traffic. Transportation Research Part B 29, 79–93. doi:10.1016/0191-2615(94)00022-R.

D'Andrea, A., Cappadona, C., La Rosa, G., Pellegrino, O., 2014. A functional road classification with data mining techniques. TRANSPORT 29, 419–430. doi:10.3846/16484142.2014.984329.

Ferrara, A., Sacone, S., Siri, S., 2018. Freeway Traffic Modelling and Control. Advances in Industrial Control. 1st editio ed., Springer International Publishing.

Furth, P.G., 1990. Model of Turning Movement Propensity. Transportation research record , 195–204.

Greenshields, B.D., Bibbins, J.R., Channing, W.S., Miller, H.H., 1934. A study of traffic capacity, in: Highway Research Board, Washington, USA. pp. 448–477.

He, S.X., 2013. A graphical approach to identify sensor locations for link flow inference. Transportation Research Part B: Methodological 51, 65–76. doi:10.1016/j.trb.2013.02.006.

Herrera, J.C., Bayen, A.M., 2010. Incorporation of Lagrangian measurements in freeway traffic state estimation. Transportation Research Part B: Methodological 44, 460–481. doi:10.1016/j.trb.2009.10.005.

Jabari, S.E., 2016. Node modeling for congested urban road networks. Transportation Research Part B: Methodological 91, 229–249. doi:10.1016/j.trb.2016.06.001.

Krumm, J., 2010. Where will they turn: Predicting turn proportions at intersections. Personal and Ubiquitous Computing 14, 591–599. doi:10.1007/s00779-009-0248-1.

Ladino, A., Canudas-de Wit, C., Kibangou, A., Fourati, H., Rodriguez, M., 2018. Density and flow reconstruction in urban traffic networks using heterogeneous data sources, in: IEEE European Control Conference (ECC), Limassol, Ciprus. pp. 1679–1684. doi:10.23919/ECC.2018.8550267.

Lighthill, M.J., Whitham, G.B., 1955. On kinematic waves II. A theory of traffic flow on long crowded roads. Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences 229, 317–345. doi:10.1098/rspa.1955.0089.

Liou, H.T., Hu, S.R., Peeta, S., 2017. Estimation of Time-Dependent Intersection Turning Proportions for Adaptive Traffic Signal Control under Limited Link Traffic Counts from Heterogeneous Sensors. Technical Report. NEXTRANS Center, Purdue University.

Lovisari, E., Canudas-de Wit, C., Kibangou, A.Y., 2016. Density/Flow reconstruction via heterogeneous sources and Optimal Sensor Placement in road networks. Transportation Research Part C: Emerging Technologies 69, 451–476. doi:10.1016/j.trc.2016.06.019.

Magnus, J.R., Neudecker, H., 2019. Matrix Differential Calculus with Applications in Statistics and Econometrics. 3rd ed.

Messmer, A., Papageorgiou, M., 1990. METANET: A macroscopic simulation program for motorway networks. Traffic Engineering and Control 31, 466–470.

Mollier, S., Delle Monache, M.L., Canudas-de Wit, C., Seibold, B., 2019. Two-dimensional macroscopic model for large scale traffic networks. Transportation Research Part B: Methodological 122, 309–326. doi:10.1016/j.trb.2019.02.016.

Nantes, A., Ngoduy, D., Bhaskar, A., Miska, M., Chung, E., 2016. Real-time traffic state estimation in urban corridors from heterogeneous data. Transportation Research Part C: Emerging Technologies 66, 99–118. doi:10.1016/j.trc.2015.07.005.

Richards, P.I., 1956. Shock Waves on the Highway. Operations Research 4, 42–51. doi:10.1287/opre.4.1.42.

Rodriguez-Vega, M., 2021. Optimal sensor placement and density estimation in large-scale traffic networks. Phd thesis. Université Grenoble Alpes.

Rodriguez-Vega, M., Canudas-de Wit, C., Fourati, H., 2019. Location of turning ratio and flow sensors for flow reconstruction in large traffic networks. Transportation Research Part B: Methodological 121, 21–40. doi:10.1016/j.trb.2018.12.005.

Rostami Shahrbabaki, M., Safavi, A.A., Papageorgiou, M., Papamichail, I., 2018. A data fusion approach for real-time traffic state estimation in urban signalized links. Transportation Research Part C: Emerging Technologies 92, 525–548. doi:10.1016/j.trc.2018.05.020.

Rostami Shahrbabaki, M., Safavi, A.A., Papageorgiou, M., Setoodeh, P., Papamichail, I., 2020. State estimation in urban traffic networks: A two-layer approach. Transportation Research Part C: Emerging Technologies 115, 102616. doi:10.1016/j.trc.2020.102616.

Schaefer, M.C., 1988. Estimation of Intersection Turning Movements from Approach Counts. ITE Journal 58, 41–46.

Seo, T., Bayen, A.M., Kusakabe, T., Asakura, Y., 2017. Traffic state estimation on highway: A comprehensive survey. Annual Reviews in Control 43, 128–151. doi:10.1016/j.arcontrol.2017.03.005.

Seo, T., Kusakabe, T., Asakura, Y., 2015. Traffic State Estimation with the Advanced Probe Vehicles Using Data Assimilation, in: 2015 IEEE 18th International Conference on Intelligent Transportation Systems, IEEE. pp. 824–830. doi:10.1109/ITSC.2015.139.

Steinmaßl, M., Kranzinger, S., Rehrl, K., 2021. Analyzing Travel Time Reliability from Sparse Probe Vehicle Data: A Case Study on the Effects of Spatial and Temporal Aggregation. Transportation Research Record: Journal of the Transportation Research Board doi:10.1177/03611981211031538.

Sun, Y., Work, D.B., 2014. A distributed local Kalman consensus filter for traffic estimation, in: 53rd IEEE Conference on Decision and Control, IEEE. pp. 6484–6491. doi:10.1109/CDC.2014.7040406.

Tampère, C.M., Immers, L.H., 2007. An extended Kalman filter application for traffic state estimation using CTM with implicit mode switching and dynamic parameters, in: IEEE Intelligent Transportation Systems Conference, Seattle, USA. pp. 209–216. doi:10.1109/ITSC.2007.4357755.

Treiber, M., Kesting, A., 2013. Traffic flow dynamics: Data, models and simulation. Springer Berlin Heidelberg, Berlin, Heidelberg. doi:10.1007/978-3-642-32460-4.

U.S Federal Highway Administration, 2013. Highway function classification concepts, criteria and procedures. Technical Report.