Validation of a Per-Lane Traffic State Estimation Scheme for Highways with Connected Vehicles

Sofia Papadopoulou, Claudio Roncoli, Nikolaos Bekiaris-Liberis,
Ioannis Papamichail and Markos Papageorgiou, *Fellow, IEEE*

***Abstract*— This study presents a thorough microscopic simulation investigation of a recently developed model-based approach for per-lane density estimation, as well as on-ramp and off-ramp flow estimation, for highways in the presence of connected vehicles. The estimation methodology is mainly based on the assumption that a certain percentage of vehicles is equipped with Vehicle Automation and Communication Systems (VACS), which provide the necessary measurements used by the estimator, namely vehicle speed and position measurements. In addition, a minimum number of conventional flow detectors is needed. In the investigation, a calibrated and validated, with real data, microscopic multi-lane model is employed, which concerns a stretch of motorway A20 from Rotterdam to Gouda in the Netherlands. It is demonstrated that the proposed methodology provides satisfactory estimation performance even for low penetration rates of connected vehicles.**

I. Introduction

Most cities around the world experience ever-growing traffic congestion in urban areas and motorway networks. Congestion may be mitigated by optimizing the performance of the traffic infrastructure through traffic management and operational strategies. Real-time traffic information is a prerequisite for traffic operations, such as freeway ramp metering control, dynamic route guidance, incident detection, and variable message sign operations. In recent years, VACS are all the more receiving considerable attention since they may create new principles in traffic management, as they are capable of communicating real-time information and execute novel control tasks [1]. Considering that density distribution may be highly heterogeneous among the different lanes of a highway, real-time lane assignment strategies may have significant advantages in traffic management. Lane policies and lane advice may be achieved if real-time traffic state information per lane is available [2], [3].

Recently, research on exploiting the innovative characteristics of VACS as a source of traffic data in traffic state estimation has drawn some attention, primarily due to the low cost, wide coverage and high accuracy of the extracted data. Connected vehicles use localization technologies that can provide these data such as Dedicated Short-Range Communications (DSRC), as well as Global Positioning Systems (GPS), cellular and Bluetooth. Data stemming from connected vehicles may contain a variety of essential dynamic transportation information, while the most commonly used are vehicle position (longitude, latitude, and altitude) and vehicle speed. Global Positioning System (GPS) receivers are stated as the most popular communication system because they are low-cost, efficient and are already commonplace in many vehicles, in use for navigation. GPS systems have stated accuracy ranging from 5 to 15 meters in geographical positioning [4], [5], [6]. But most modern methods adopt a hybrid positioning system, combining differential GPS (DGPS) with map-matching and dead-reckoning, which improved vehicle position data up to 1 to 5 meters accuracy [7], sufficient for lane-based applications. Speed measurements are mostly reported to be quite accurate with a precision error lower than 1 km/h [8], [4], while some studies claim a tendency of underestimation in speed measurements and a reported error around 5 km/h [9].

Traffic state estimation utilizing floating car data has been investigated in numerous studies, such as, for example, [10]–[28]. However, the existing studies that deal with lane-based traffic state estimation are rare in the traffic literature, while they mainly assume data obtained from conventional detectors [29]–[32] with the exception of [33].

Previous work by [23], [27], [28] employed connected-vehicle data to estimate the aggregate-lane traffic density on highway segments. The distinguishing feature of that work was the lack of any empirical modelling approach, such as fundamental diagrams, that would call for tedious calibration procedures. A similar approach, appropriately extended to enable per-lane density estimation, is considered in this paper and is thoroughly evaluated via microscopic simulation. The estimation scheme uses a data-driven macroscopic model for per-lane traffic density and employs real-time measurements obtained from connected vehicles as well as a minimum number of spot flow measurements. The flow measurements from detectors are needed to guarantee the observability property of the underlying model, see [34] for details. Density estimation is performed via the employment of a Kalman Filter. The performance of the estimation scheme is examined under various penetration rates of connected vehicles, using data retrieved from a microscopic multi-lane model, which was calibrated and validated in [35] with real data from a stretch of motorway A20 from Rotterdam to Gouda in the

S. Papadopoulou, N. Bekiaris-Liberis, I. Papamichail and M. Papageorgiou are with the Dynamic Systems and Simulation Laboratory, School of Production Engineering and Management, Technical University of Crete, Chania, 73100, Greece. Email addresses: spapadopoulou@isc.tuc.gr, nikos.bekiaris@dssl.tuc.gr, ipapa@dssl.tuc.gr and markos@dssl.tuc.gr.

C. Roncoli is with the Department of Built Environment, School of Engineering, Aalto University, 02150 Espoo, Finland. Email address: claudio.roncoli@aalto.fi.

This research was supported by the European Research Council under the European Union's Seventh Framework Programme (FP/2007-2013)/ERC Advanced Grant Agreement n. 321132, project TRAMAN21.

Netherlands. The case-study highway stretch includes several on-ramps, off-ramps, and a lane-drop, while the employed simulation scenario is characterized by both congested and free-flow traffic conditions. It is worth mentioning that, in the investigation, simple algorithms are employed in case of inconsistencies in the probe vehicle data (such as, for instance, in the case there are temporarily no measurements available from connected vehicles). The performance of the tested estimation scheme is shown to be satisfactory even for low penetration rates.

The remainder of the paper is organized as follows. The model for the per-lane density dynamics and the proposed estimation scheme are presented in Section II. The description of the microscopic simulation configuration as well as the highway network under study and the traffic conditions are presented in Section III, which includes also the details of the computation of the data employed by the estimator. Subsequently, the results of the estimation are presented in Section IV. Finally, in Section V the main findings of this study are summarized.

II. Per-Lane Traffic State Estimation Using a Data-Driven Model

A. General Set-Up

We consider a highway stretch consisting of M lanes, indexed by $j = 1, \dots, M$, subdivided into N segments, indexed by $i = 1, \dots, N$. We define a cell (i, j) to be the highway part that corresponds to lane j of segment i. The length of each segment is denoted by Δ_i, $i = 1, \dots, N$.

The following variables are extensively used in the paper:

- Average speed $\left[\frac{\text{km}}{\text{h}}\right]$ of vehicles in cell (i, j), denoted by $v_{i,j}$, for $i = 1, \dots, N$ and $j = 1, \dots, M$.
- Total traffic density $\left[\frac{\text{veh}}{\text{km}}\right]$ at cell (i, j), denoted by $\rho_{i,j}$, for $i = 1, \dots, N$ and $j = 1, \dots, M$.
- Total longitudinal inflow $\left[\frac{\text{veh}}{\text{h}}\right]$ of cell $(i+1, j)$, denoted by $q_{i,j}$, for $i = 0, \dots, N-1$ and $j = 1, \dots, M$.
- Total on-ramp flow $\left[\frac{\text{veh}}{\text{h}}\right]$ entering at cell (i, j), denoted by $r_{i,j}$, for $i = 1, \dots, N$ and $j = 1, \dots, M$.
- Total off-ramp flow $\left[\frac{\text{veh}}{\text{h}}\right]$ exiting from cell (i, j), denoted by $s_{i,j}$, for $i = 1, \dots, N$ and $j = 1, \dots, M$.
- Total lateral flow $\left[\frac{\text{veh}}{\text{h}}\right]$ at segment i that enters lane j_2 from lane j_1, denoted by $L_{i,j_1 \to j_2}$, for $i = 1, \dots, N$, $j_1 = 1, \dots, M$, and $j_2 = j_1 \pm 1$.

Note that, the attribute "total" refers to the total population of both connected and conventional vehicles.

B. Available Information from Connected Vehicle Reports

The data-driven model presented in the next subsection, requires the availability of the following measurements:

- Average speed of connected vehicles at cell (i, j), denoted by $v^{\text{c}}_{i,j}$, for $i = 1, \dots, N$ and $j = 1, \dots, M$.
- Density of connected vehicles at cell (i, j), denoted by $\rho^{\text{c}}_{i,j}$, for $i = 1, \dots, N$ and $j = 1, \dots, M$.
- Lateral flow of connected vehicles at segment i that enters lane j_2 from lane j_1, denoted by $L^{\text{c}}_{i,j_1 \to j_2}$, for $i = 1, \dots, N$, $j_1 = 1, \dots, M$, and $j_2 = j_1 \pm 1$.

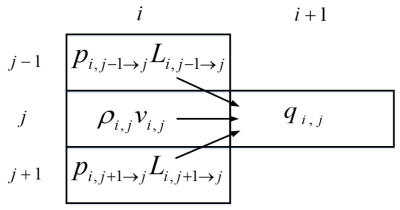

Fig. 1. Model of the exiting longitudinal flow from cell (i, j) as described in (2).

Average speeds, densities, and lateral flows of connected vehicles may be readily obtained from position and speed reports.

C. Model Description for the Density and Ramp Flow Dynamics

The conservation equation yields the following model for the density dynamics in each cell (i, j)

$$\begin{aligned}\rho_{i,j}(k+1) = \rho_{i,j}(k) + \frac{T}{\Delta_i} \big(& q_{i-1,j}(k) - q_{i,j}(k) \\ & + L_{i,j-1\to j}(k) + L_{i,j+1\to j}(k) - L_{i,j\to j-1}(k) \\ & - L_{i,j\to j+1}(k) + r_{i,j}(k) - s_{i,j}(k)\big), \end{aligned} \tag{1}$$

where T is the time discretization step. For convenience, we assume $r_{i,j} \equiv s_{i,j} \equiv 0$, $\forall i$ and $1 \le j \le M-1$, where M denotes the right-most lane (assuming right-hand traffic); we have $L_{i,j_1\to j_2} \equiv 0$ if either j_1 or j_2 equals zero or $M+1$. Moreover, in case there is a lane-drop at cell (i, j), then we define $q_{i,j}(k) \equiv 0$. We note that the per-lane inflows at the highway entry, namely, $q_{0,j}$, $j = 1, \dots, M$, are treated as measured inputs to system (1).

The following relation is employed for total flows (Fig. 1)

$$\begin{aligned} q_{i,j}(k) \;=\; & v_{i,j}(k)\rho_{i,j}(k) + p_{i,j-1\to j} L_{i,j-1\to j}(k) \\ & + p_{i,j+1\to j} L_{i,j+1\to j}(k) + \bar{p}_{i,j} r_{i,j}(k), \end{aligned} \tag{2}$$

for $i = 1, \dots, N$, $j = 1, \dots, M$, where $p_{i,j_1\to j_2}, \bar{p}_{i,j} \in [0, 1]$, $\forall (i, j)$, $j_1 = 1, \dots, M$, and $j_2 = j_1 \pm 1$, indicate the percentages of "diagonal" lateral movements, including lateral flows from an on-ramp acceleration lane, for each specific cell. While the first term in (2) is well-known (see, e.g., [36]), the motivation for the rest of the terms may be less obvious. Their choice is guided from the fact that, at locations featuring strong lateral flows (e.g., at cells where an on-ramp is located or at segments that feature lane-drops), a significant amount of the lateral flow may appear close to the cell end (e.g., in the former case, at the acceleration lane end). As a result, the flow modeling may be more accurately described considering that a percentage of lateral or on-ramp flows actually acts as additional exiting longitudinal flow. This formulation is also employed in other works, e.g., [37].

For the lateral flows, we employ the following relation

$$L_{i,j_1\to j_2}(k) = \frac{L^{\text{c}}_{i,j_1\to j_2}(k)}{\rho^{\text{c}}_{i,j_1}(k)} \rho_{i,j_1}(k), \tag{3}$$

for $i = 1, \ldots, N$, $j_1 = 1, \ldots, M$, and $j_2 = j_1 \pm 1$. Equation (3) is based on the reasonable assumption that the behavior of the population of connected vehicles in a given cell, with respect to lateral movements, is representative for the total vehicle population in that cell. This allows one to quantify the total lateral movements from a cell using (3), namely, by scaling the lateral movements of connected vehicles with the *inverse* of the percentage of connected vehicles in that cell.

Plugging (2) and (3) into (1), we get for all (i, j)

$$\begin{aligned}\rho_{i,j}(k+1)=&\left(1-\frac{T}{\Delta_i}v_{i,j}(k)-\frac{T}{\Delta_i}\frac{L^{\mathrm{c}}_{i,j\to j-1}(k)}{\rho^{\mathrm{c}}_{i,j}(k)}-\frac{T}{\Delta_i}\right.\\&\left.\times\frac{L^{\mathrm{c}}_{i,j\to j+1}(k)}{\rho^{\mathrm{c}}_{i,j}(k)}\right)\rho_{i,j}(k)+\frac{T}{\Delta_i}v_{i-1,j}(k)\\&\times\rho_{i-1,j}(k)+\frac{T}{\Delta_i}\left((1-p_{i,j-1\to j})\right.\\&\times\frac{L^{\mathrm{c}}_{i,j-1\to j}(k)}{\rho^{\mathrm{c}}_{i,j-1}(k)}\rho_{i,j-1}(k)+(1-p_{i,j+1\to j})\\&\left.\times\frac{L^{\mathrm{c}}_{i,j+1\to j}(k)}{\rho^{\mathrm{c}}_{i,j+1}(k)}\right)\rho_{i,j+1}(k)+\frac{T}{\Delta_i}\left(p_{i-1,j-1\to j}\right.\\&\times\frac{L^{\mathrm{c}}_{i-1,j-1\to j}(k)}{\rho^{\mathrm{c}}_{i-1,j-1}(k)}\rho_{i-1,j-1}(k)+p_{i-1,j+1\to j}\\&\left.\times\frac{L^{\mathrm{c}}_{i-1,j+1\to j}(k)}{\rho^{\mathrm{c}}_{i-1,j+1}(k)}\rho_{i-1,j+1}(k)\right)+\frac{T}{\Delta_i}\left((1-\bar{p}_{i,j})\right.\\&\left.\times r_{i,j}(k)+\bar{p}_{i-1,j}r_{i-1,j}(k)-s_{i,j}(k)\right).\end{aligned}\tag{4}$$

We adopt, as usual in absence of a descriptive dynamic model (see [38]) a random walk to describe the dynamics of on-ramp and off-ramp flows. The deterministic parts of such models read

$$r_{i,M}(k+1) = r_{i,M}(k) \tag{5}$$
$$s_{i,M}(k+1) = s_{i,M}(k). \tag{6}$$

We write next compactly the overall system (4)–(6). For this, we define first the state vector x as follows

$$\begin{aligned}x = &(\rho_{1,1}, \cdots, \rho_{N,1}, \cdots, \rho_{1,M} \cdots, \rho_{N,M}, r_{1,M}, \cdots,\\&r_{N,M}, s_{1,M}, \cdots, s_{N,M})^{\mathrm{T}}.\end{aligned}\tag{7}$$

The average speed of connected vehicles is representative of the average cell speed, as motivated in [28] and justified with real data and in microscopic simulation in [27] and [23], respectively, even for connected-vehicle penetrations as low as 2%. Thus, the unmeasured cell speeds $v_{i,j}$ may be replaced by the corresponding measured speeds $v^{\mathrm{c}}_{i,j}$; and, using (7), we re-write (4)–(6) in a compact form as

$$x(k+1) = A\left(v^{\mathrm{c}}(k), L^{\mathrm{c}}(k), \rho^{\mathrm{c}}(k)\right)x(k) + Bu(k), \tag{8}$$

where v^{c}, L^{c}, and ρ^{c} denote vectors that incorporate all average cell speeds of connected vehicles $v^{\mathrm{c}}_{i,j}$, lateral flows of connected vehicles $L^{\mathrm{c}}_{i,j_1\to j_2}$, and densities of connected vehicles $\rho^{\mathrm{c}}_{i,j}$, respectively; while u denotes the vector of inflows at the highway entrance, namely $u = (q_{0,1}, \cdots, q_{0,M})^{\mathrm{T}}$, $A \in \mathbb{R}^{(N\times M+2N)\times(N\times M+2N)}$, and $B \in \mathbb{R}^{(N\times M+2N)\times M}$.

Together with (8), we associate an output vector y, which holds all mainstream total flows that are measured by corresponding mainstream fixed detectors and, as follows from (2) and (3), is given by

$$y(k) = C\left(v^{\mathrm{c}}(k), L^{\mathrm{c}}(k), \rho^{\mathrm{c}}(k)\right)x(k), \tag{9}$$

where $C \in \mathbb{R}^{(M+l_r+l_s-1)\times(N\times M+2N)}$, with l_r and l_s being the number of on-ramps and off-ramps, respectively. The minimum number of rows of C equals $M + l_r + l_s - 1$ in order for system (8), (9) to be observable (see [28], [34] for details). Note that we assume $2N$ ramp flows. In the case where there are less on-ramp or off-ramp flows, the dimensions of the matrices A, B, and C are reduced accordingly.

D. Per-Lane Total Density and Ramp Flow Estimation Utilizing a Kalman Filter

We employ a standard Kalman filter utilizing model (8), (9) for per lane total density estimation. Defining the vector $\hat{x}$ as the system state to be estimated,

$$\begin{aligned}\hat{x} = &(\hat{\rho}_{1,1}, \cdots, \hat{\rho}_{N,1}, \cdots, \hat{\rho}_{1,M} \cdots, \hat{\rho}_{N,M}, \hat{r}_{1,M}, \cdots,\\&\hat{r}_{N,M}, \hat{s}_{1,M}, \cdots, \hat{s}_{N,M})^{\mathrm{T}},\end{aligned}\tag{10}$$

the filter equations are

$$\begin{aligned}\hat{x}(k+1) =&A\left(v^{\mathrm{c}}(k), L^{\mathrm{c}}(k), \rho^{\mathrm{c}}(k)\right)\hat{x}(k) + Bu(k)\\&+ A\left(v^{\mathrm{c}}(k), L^{\mathrm{c}}(k), \rho^{\mathrm{c}}(k)\right)K(k)(z(k)\\&- C\left(v^{\mathrm{c}}(k), L^{\mathrm{c}}(k), \rho^{\mathrm{c}}(k)\right)\hat{x}(k))\end{aligned}\tag{11}$$

$$\begin{aligned}K(k) =&P(k)C\left(v^{\mathrm{c}}(k), L^{\mathrm{c}}(k), \rho^{\mathrm{c}}(k)\right)^{\mathrm{T}}\\&\times\left(C(v^{\mathrm{c}}(k), L^{\mathrm{c}}(k), \rho^{\mathrm{c}}(k))P(k)\right.\\&\left.\times C\left(v^{\mathrm{c}}(k), L^{\mathrm{c}}(k), \rho^{\mathrm{c}}(k)\right)^{\mathrm{T}} + R\right)^{-1}\end{aligned}\tag{12}$$

$$\begin{aligned}P(k+1) =&A\left(v^{\mathrm{c}}(k), L^{\mathrm{c}}(k), \rho^{\mathrm{c}}(k)\right)(I - K(k)\\&\times C\left(v^{\mathrm{c}}(k), L^{\mathrm{c}}(k), \rho^{\mathrm{c}}(k)\right))P(k)\\&\times A\left(v^{\mathrm{c}}(k), L^{\mathrm{c}}(k), \rho^{\mathrm{c}}(k)\right)^{\mathrm{T}} + Q,\end{aligned}\tag{13}$$

where z is a noisy version of the measurement y defined in (9), whereas $Q = Q^T \succ 0$ and $R = R^T \succ 0$ are tuning parameters. The Kalman Filter is initialized as

$$\hat{x}(k_0) = \mu, \quad P(k_0) = H, \tag{14}$$

where μ and $H = H^T \succ 0$.

III. Microscopic Simulation Set-Up

The behavior of the proposed estimation scheme is examined and evaluated through microscopic simulation using AIMSUN [39]. Traffic measurements are extracted from a specific subset of the whole population of vehicles in the network that are considered to be connected. Vehicles entering the network are all of the same vehicle type, featuring a probabilistic distribution of movement behaviors, and are randomly marked as connected according to an assumed penetration rate, based on a uniform distribution.

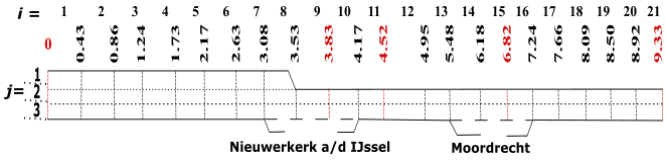

Fig. 2. Schematic representation of the case study network. Field detector positions are indicated as the distance (in km) from the network entrance. The detectors used by the estimator for obtaining flow measurements within the case study are colored in red.

A. Network and Experimental Configuration

The case-study network (see Fig. 2) is a stretch of motorway A20 from Rotterdam to Gouda, in the Netherlands. The multi-lane microscopic model employed in AIMSUN was designed and calibrated in [35] with real, lane-specific traffic data obtained from detectors [40], thus providing a realistic ground truth scenario.

The considered highway stretch, shown in Fig. 2, constitutes a challenging test-bed for the proposed estimation scheme, as it incorporates a non-trivial combination of a lane-drop, on-ramps and off-ramps, which trigger a variety of corresponding lane-changing behaviours. The considered stretch is about 9.33 km in length, comprises 3 homo-directional lanes until 3.53 km, where there is a lane-drop. For the purpose of estimation, the stretch is space-discretised in $N = 21$ segments. Two on-ramps and two off-ramps are located at 3.08 km, 5.48 km and 4.17 km, 7.24 km, respectively.

The ground truth in our experiments, used to evaluate the performance of the developed estimation scheme, is represented by the total density in each cell and the total ramp flows. The cell densities $\rho_{i,j}$ are computed by counting the number of all vehicles that are present within cell (i, j) at a time instant kT, divided by the segment length Δ_i; whereas all ramp flows are computed by counting the number of vehicles that cross the corresponding location within the time interval $(kT, (k+1)T]$. However, since lane-based densities and ramp flows are very noisy, a moving average of the 6 latest available measurements is considered as ground truth. Average segment speeds $v_{i,j}$ that represent ground truth are computed by averaging arithmetically at time step kT the instant speeds of all vehicles present in a segment.

B. Employed Scenario

The employed scenario utilizes available real demand measurements and replicates traffic conditions, whereby a strong congestion is created at 4 km at around 6:30 AM because of the increased flow entering from on-ramp "Nieuwerkerk a/d IJssel". The congestion spills back, strengthens at the lane-drop at 3.6 km, and covers the stretch up to 1.2 km. From 7:00 AM until 7:30 AM, the congestion persists downstream of the lane-drop, while congestion upstream thereof dissolves. Fig. 3, illustrates the speeds along each lane from 5:00 AM to 9:00 AM. This congestion pattern allows to test and evaluate the proposed estimator under varying traffic conditions, which include the formation and dissipation of a stretch-internal congestion, which is not visible at the stretch boundaries.

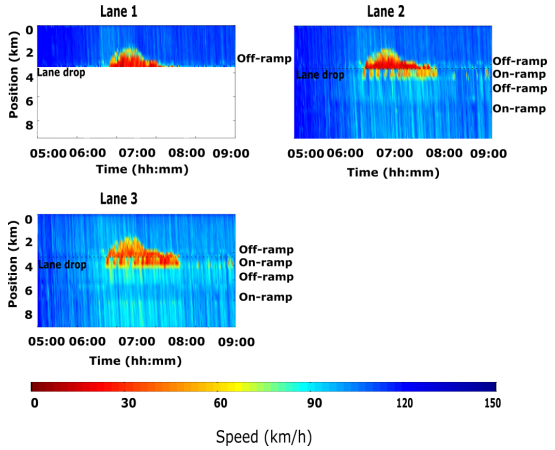

Fig. 3. Contour plot of the per lane simulated speed, Wednesday 26-05-2010, as calibrated in [35].

C. Computation of Data Employed by the Estimator

Prior to the performance evaluation of the proposed estimation scheme, we detail the information provided to the estimator. The estimation performance depends critically on the quality of this information, thus we employ simple algorithms to ensure that this information is reliable and as representative as possible.

Especially at low penetrations rates, only few or even no connected vehicles may be present in a cell. Therefore, a moving average, utilizing available speed measurements from previous times steps, is considered to compensate for potential large variations or nonexistence of speed measurements utilized by the estimator. Specifically, we feed the Kalman Filter with a moving average of the last n available speed measurements as

$$v^{\mathrm{c}}_{i,j}(k) = \sum_{l=0}^{n-1} \frac{\nu^{\mathrm{c}}_{i,j}(k-l)}{n}, \tag{15}$$

where $\nu^{\mathrm{c}}_{i,j}(k)$ is the average speed of connected vehicles at cell (i, j), computed from speed reports that are collected every 2 s and averaged arithmetically every time step kT. In cases where there are no connected vehicle reports available at cell (i, j) during a time interval $((k-1)T, kT]$, we replace this $\nu^{\mathrm{c}}_{i,j}(k)$ with the speed reported at the previous time step as, i.e., we set $\nu^{\mathrm{c}}_{i,j}(k) = \nu^{\mathrm{c}}_{i,j}(k-1)$. Within our experiments, at low penetration rates or light traffic, a cell may feature complete absence of connected vehicle's reports, for more than one minute. For this reason, a moving average (15) of the last $n = 12$ speed measurements available is chosen.

Lateral flows of connected vehicles $L^{\mathrm{c}}_{i,j_1 \to j_2}$ are computed based on position reports of connected vehicles. Specifically, we count first the number of connected vehicles moved from lane j_1 to lane j_2 within a time interval of 2 s; then, every time instant kT, these lateral flow measurements are accumulated for the time interval $((k-1)T, kT]$ to produce an intermediate lane-changing measurement $\bar{L}^{\mathrm{c}}_{i,j_1 \to j_2}(k)$.

This lateral flow vector of connected vehicles $\bar{L}^{\mathrm{c}}$, may exhibit some spiky behaviour due to the rare appearance of connected vehicle lateral movements; and therefore, this value may not be representative for the occurring lane-changing flow of the total vehicle population. To account for this fact and obtain representative, though averaged lateral flows, we feed the estimation scheme with an exponentially smoothed version of the lateral flows of connected vehicles, rather than the original measured lateral flows. Thus, for each lateral flow measurement, we have for all $i = 1, \ldots, N$, $j_1 = 1, \ldots, M$, and $j_2 = j_1 \pm 1$

$$L^{\mathrm{c}}_{i,j_1\to j_2}(k+1) = (1-a)L^{\mathrm{c}}_{i,j_1\to j_2}(k) + a\bar{L}^{\mathrm{c}}_{i,j_1\to j_2}(k), \quad (16)$$

where the smoothing factor $a \in [0, 1]$ is chosen based on statistical analysis (see [41]) and $L^{\mathrm{c}}_{i,j_1\to j_2}(0) = 0$.

Also for the density measurements $\rho^{\mathrm{c}}_{i,j}(k)$, that are fed to the Kalman Filter every time step k, we employ, for similar reasons as above, a moving average of the $m = 6$ (time window of 1 min) latest available measurements as

$$\rho^{\mathrm{c}}_{i,j}(k) = \sum_{l=0}^{m-1} \frac{\tilde{\rho}^{\mathrm{c}}_{i,j}(k-l)}{m}, \quad (17)$$

where $\tilde{\rho}^{\mathrm{c}}_{i,j}(k)$ is the instant density at cell (i, j) for every time step k. In cases where there are no connected vehicle reports available at cell (i, j) at a time instant kT, we replace the corresponding $\tilde{\rho}^{\mathrm{c}}_{i,j}(k)$ with the density reported at the previous time step, i.e., we set $\tilde{\rho}^{\mathrm{c}}_{i,j}(k) = \tilde{\rho}^{\mathrm{c}}_{i,j}(k-1)$.

Assuming that the (smoothed) lateral flow of connected vehicles exiting from the lane-drop cell, namely, $L^{\mathrm{c}}_{8,1\to2}$, is always non-zero, it is shown in our companion paper [34] that the utilized flow measurement configuration, shown in Fig. 2, guarantees observability of the underlying model (8), (9) (see also [28] for the aggregated-lane case). Specifically, sets of detectors (one in each lane) located at the highway entrance (0 km) and exit (9.33 km) are used to obtain the input and output of system (8), (9), respectively. All ramp flows are assumed unmeasured; therefore, to establish observability, an additional flow measurement on the right-most lane of a (arbitrarily chosen) segment between every pair of consecutive unmeasured ramps is needed. In the present evaluation, we exceed these minimum measurement requirements for observability by assuming the presence of fixed flow detectors at all lanes (cross-section), rather than only the right-most lane, of the aforementioned segments; this is deemed reasonable, since detectors are usually installed for a cross-section. In conclusion, we employ a set of additional flow detectors (one per lane) located at 3.83 km (segment $i = 9$), 4.52 km (segment $i = 11$), and 6.82 km (segment $i = 15$).

IV. Performance Evaluation for Varying Penetration Rates of Connected Vehicles

Based on the microscopic environment configuration described in Section III, we simulate traffic conditions featuring various penetration rates of connected vehicles. To account for a variety of possible current and future traffic scenarios, the performance of the estimation scheme is evaluated for a wide range of penetration rates of connected vehicles, more specifically, for $2\%, 5\%, 10\%, 20\%$, and 50%.

We employ the estimation scheme (11)–(13) for per-lane total density and ramp flow estimation. We choose $Q = 10$ and $R = 10.000$, which, according to our experiments achieve a satisfactory estimation performance. In general, the estimation scheme is expected to not be very sensitive to the choice of Q and R, see [23] for details.

It should be noted that, since we expect strong diagonal flows mainly at the lane-drop cell (in comparison with the rest of the cells), we set all percentages $p_{i,j_1\to j_2}, \bar{p}_{i,j}$ equal to zero except for the percentage

$$p \equiv p_{8,1\to2}, \quad (18)$$

which corresponds to the diagonal lateral flow of cell $(8, 2)$.

A. Quantitative Performance Measure

To assess the overall performance of the suggested estimation scheme, a performance index formulated as the Coefficient of Variation (CV) of the root mean square error of the 60-second moving averages of estimated densities $\hat{\rho}_{i,j}$ and ramp flows $\hat{\theta}_i$ with respect to the corresponding ground truth, is adopted as

$$CV_\rho = \frac{\sqrt{\frac{1}{MNK}\sum_{i=1}^{N}\sum_{j=1}^{M}\sum_{k=1}^{K}\left[\hat{\rho}_{i,j}(k) - \rho_{i,j}(k)\right]^2}}{\frac{1}{MNK}\sum_{i=1}^{N}\sum_{j=1}^{M}\sum_{k=1}^{K}\rho_{i,j}(k)} \quad (19)$$

$$CV_{r,s} = \frac{\sqrt{\frac{1}{K(l_r+l_s)}\sum_{k=1}^{K}\sum_{i=1}^{l_r+l_s}\left[\hat{\theta}_i(k) - \theta_i(k)\right]^2}}{\frac{1}{K(l_r+l_s)}\sum_{k=1}^{K}\sum_{i=1}^{l_r+l_s}\theta_i(k)}, \quad (20)$$

where $\theta_1 = r_{10}, \theta_2 = r_{16}, \theta_3 = s_{18}, \theta_4 = s_{14}$, and $M = 3, N = 21$, $l_r = l_s = 2$, $K = 1440$.

B. Performance Evaluation for a Baseline Case

For the sake of brevity, only specific results of the cell densities and ramp flow estimation for 20% penetration rate of connected vehicles are presented, as illustrated in Figs. 4 and 5, respectively. The results are obtained with $p = 0.3$ and $a = 0.05$, chosen after a sensitivity analysis presented in [41].

It is evident from the plots that the proposed scheme successfully estimates and tracks the dynamics of both segment densities and ramp flows under various traffic conditions, including congestion and free-flow, as well as for (short-lived) time intervals where no information from connected vehicle reports is available. Density estimation is characterized by a performance index $CV_\rho = 34.4\%$, whereas ramp flow estimation is characterized by a performance index $CV_{r,s} = 78.0\%$.

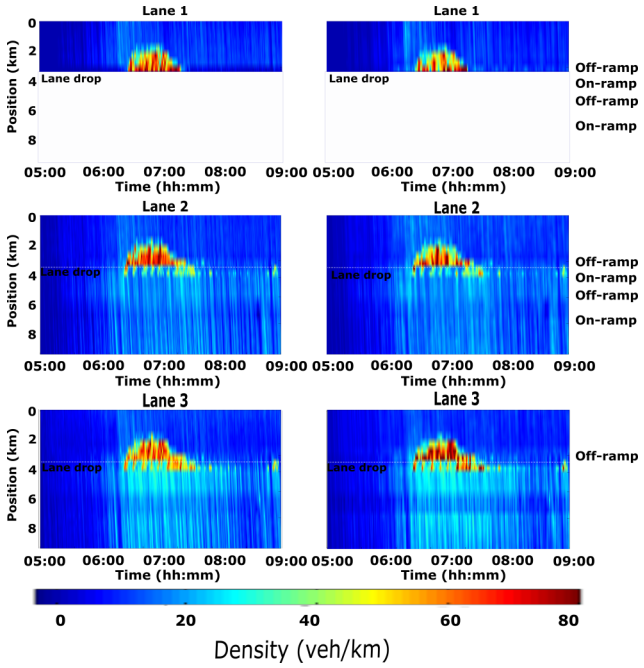

Fig. 4. Contour plot for the ground truth (left column) and estimated (right column) densities.

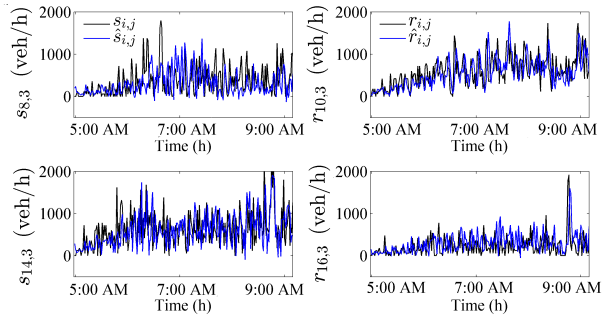

Fig. 5. Comparison between ground truth (black line) and estimated (blue line) ramp flows for all on-ramps and off-ramps in the network.

C. Performance Evaluation for Various Penetration Rates

Fig. 6 illustrates the performance for density and ramp flow estimation with regard to different penetration rates of connected vehicles. A moderate sensitivity is observed in the performance of the estimation scheme, which is seen to deteriorate with decreasing penetration rates of connected vehicles present at the highway. This is mainly due to the accordingly reduced adequacy of the available traffic information feeding the estimation scheme. In fact, the proportion of time intervals without any connected vehicle in a cell increases with decreasing penetration rate, reaching up to 70% for a penetration rate of 2% (see also [41]).

V. Conclusions

The validity of a per-lane traffic density as well as ramp flow estimation scheme, which is based on an appropriate multi-lane traffic flow model and a standard Kalman filter, has been thoroughly tested using the AIMSUN microscopic traffic simulator. The proposed scheme is mainly based on speed and position information obtained from connected vehicle reports. The effectiveness of the proposed methodology was examined in carefully designed experiments for a real highway stretch and real demand scenarios. The obtained results demonstrate that the estimation scheme captures the onset of congestion with accurate timing and more generally, reproduces reliably the challenging traffic conditions in space and time. Density estimation is satisfactory even for penetration rates as low as 2%.

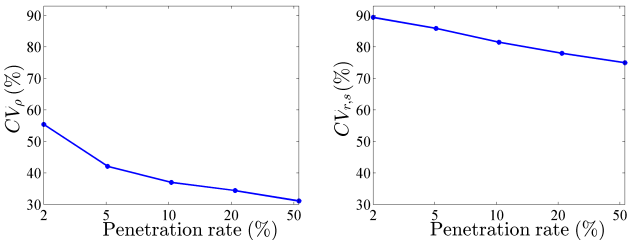

Fig. 6. Performance comparison of density (left) and ramp flow (right) estimations for various penetration rates of connected vehicles.

The presented approach has several advantages for possible future real-world applications, including:

- the use of a macroscopic model that has now calibration requirements (e.g. no fundamental diagram)
- the extensive use of low-cost connected vehicle data that are already available and are expected to increase in the near future;
- the use of only a limited amount of fixed flow sensors.

It should be emphasized that the availability of real-time traffic density per lane estimates is a prerequisite for the application of lane-based traffic control algorithms.

Acknowledgment

The authors would like to thank Prof. Bart van Arem and his group for their support in providing information related to the network used in the case study.

References

[1] M. Papageorgiou, C. Diakaki, I. Nikolos, I. Ntousakis, I. Papamichail, and C. Roncoli, "Freeway traffic management in presence of vehicle automation and communication systems (VACS)," in *Road Vehicle Automation 2*. Springer, 2015, pp. 205–214.

[2] C. Roncoli, N. Bekiaris-Liberis, and M. Papageorgiou, "Optimal lane-changing control at motorway bottlenecks," in *IEEE Conference on Intelligent Transportation Systems*, 2016, pp. 1785–1791.

[3] ——, "Lane-changing feedback control for efficient lane assignment at motorway bottlenecks," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2625, pp. 20–31, 2017.

[4] R. Zito, G. D'este, and M. Taylor, "Global positioning systems in the time domain: How useful a tool for intelligent vehicle-highway systems?" *Transportation Research Part C: Emerging Technologies*, vol. 3, no. 4, pp. 193–209, 1995.

[5] S. Turksma, "The various uses of floating car data," 2000.

[6] K. Liu, T. Yamamoto, and T. Morikawa, "An analysis of the cost efficiency of probe vehicle data at different transmission frequencies," *International Journal of ITS Research*, vol. 4, no. 1, pp. 21–28, 2006.

[7] B. Waterson and S. Box, "Quantifying the impact of probe vehicle localisation data errors on signalised junction control," *IET Intelligent Transport Systems*, vol. 6, no. 2, pp. 197–203, 2012.

[8] T. J. Chalko, "High accuracy speed measurement using gps (global positioning system)," *NU Journal of Discovery*, vol. 4, pp. 1–9, 2007.

[9] X. Zhao, K. Carling, and J. Håkansson, *Reliability of GPS based traffic data: an experimental evaluation*. Högskolan Dalarna, 2014.

[10] D. B. Work, O.-P. Tossavainen, S. Blandin, A. M. Bayen, T. Iwuchukwu, and K. Tracton, "An ensemble kalman filtering approach to highway traffic estimation using gps enabled mobile devices," in *IEEE Conference on Decision and Control*, 2008, pp. 5062–5068.

[11] C. De Fabritiis, R. Ragona, and G. Valenti, "Traffic estimation and prediction based on real time floating car data," in *IEEE Conference on Intelligent Transportation Systems*, 2008, pp. 197–203.

[12] J. C. Herrera, D. B. Work, R. Herring, X. J. Ban, Q. Jacobson, and A. M. Bayen, "Evaluation of traffic data obtained via gps-enabled mobile phones: The mobile century field experiment," *Transportation Research Part C: Emerging Technologies*, vol. 18, no. 4, pp. 568–583, 2010.

[13] M. Rahmani, H. N. Koutsopoulos, and A. Ranganathan, "Requirements and potential of gps-based floating car data for traffic management: Stockholm case study," in *IEEE Conference on Intelligent Transportation Systems*, 2010, pp. 730–735.

[14] T. Qiu, X.-Y. Lu, A. Chow, and S. Shladover, "Estimation of freeway traffic density with loop detector and probe vehicle data," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2178, pp. 21–29, 2010.

[15] T. Schreiter, H. van Lint, M. Treiber, and S. Hoogendoorn, "Two fast implementations of the adaptive smoothing method used in highway traffic state estimation," in *IEEE Conference on Intelligent Transportation Systems*, 2010, pp. 1202–1208.

[16] M. Treiber, A. Kesting, and R. E. Wilson, "Reconstructing the traffic state by fusion of heterogeneous data," *Computer-Aided Civil and Infrastructure Engineering*, vol. 26, no. 6, pp. 408–419, 2011.

[17] Y. Yuan, J. Van Lint, R. E. Wilson, F. van Wageningen-Kessels, and S. P. Hoogendoorn, "Real-time lagrangian traffic state estimator for freeways," *IEEE Transactions on Intelligent Transportation Systems*, vol. 13, no. 1, pp. 59–70, 2012.

[18] C. P. van Hinsbergen, T. Schreiter, F. S. Zuurbier, J. Van Lint, and H. J. van Zuylen, "Localized extended kalman filter for scalable real-time traffic state estimation," *IEEE Transactions on Intelligent Transportation Systems*, vol. 13, no. 1, pp. 385–394, 2012.

[19] W. Deng, H. Lei, and X. Zhou, "Traffic state estimation and uncertainty quantification based on heterogeneous data sources: A three detector approach," *Transportation Research Part B: Methodological*, vol. 57, pp. 132–157, 2013.

[20] A. Anand, G. Ramadurai, and L. Vanajakshi, "Data fusion-based traffic density estimation and prediction," *Journal of Intelligent Transportation Systems*, vol. 18, no. 4, pp. 367–378, 2014.

[21] B. Piccoli, K. Han, T. L. Friesz, T. Yao, and J. Tang, "Second-order models and traffic data from mobile sensors," *Transportation Research Part C: Emerging Technologies*, vol. 52, pp. 32–56, 2015.

[22] T. Seo and T. Kusakabe, "Probe vehicle-based traffic flow estimation method without fundamental diagram," *Transportation Research Procedia*, vol. 9, pp. 149–163, 2015.

[23] M. Fountoulakis, N. Bekiaris-Liberis, C. Roncoli, I. Papamichail, and M. Papageorgiou, "Highway traffic state estimation with mixed connected and conventional vehicles: Microscopic simulation-based testing," *Transportation Research Part C: Emerging Technologies*, vol. 78, pp. 13–33, 2017.

[24] F. Rempe, P. Franeck, U. Fastenrath, and K. Bogenberger, "Online freeway traffic estimation with real floating car data," in *IEEE Conference on Intelligent Transportation Systems*, 2016, pp. 1838–1843.

[25] R. Wang, D. B. Work, and R. Sowers, "Multiple model particle filter for traffic estimation and incident detection," *IEEE Transactions on Intelligent Transportation Systems*, vol. 17, no. 12, pp. 3461–3470, 2016.

[26] M. Wright and R. Horowitz, "Fusing loop and gps probe measurements to estimate freeway density," *IEEE Transactions on Intelligent Transportation Systems*, vol. 17, no. 12, pp. 3577–3590, 2016.

[27] C. Roncoli, N. Bekiaris-Liberis, and M. Papageorgiou, "Highway traffic state estimation using speed measurements: case studies on ngsim data and highway A20 in the netherlands," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2559, pp. 90–100, 2016.

[28] N. Bekiaris-Liberis, C. Roncoli, and M. Papageorgiou, "Highway traffic state estimation with mixed connected and conventional vehicles," *IEEE Transactions on Intelligent Transportation Systems*, vol. 17, no. 12, pp. 3484–3497, 2017.

[29] M.-F. Chang and D. C. Gazis, "Traffic density estimation with consideration of lane-changing," *Transportation Science*, vol. 9, no. 4, pp. 308–320, 1975.

[30] B. Coifman, "Estimating density and lane inflow on a freeway segment," *Transportation Research Part A: Policy and Practice*, vol. 37, no. 8, pp. 689–701, 2003.

[31] K. Singh and B. Li, "Discrete choice modelling for traffic densities with lane-change behaviour," *Procedia-Social and Behavioral Sciences*, vol. 43, pp. 367–374, 2012.

[32] M. Yılan, "Multilane traffic density estimation with KDE and nonlinear LS and tracking with scalar kalman filtering," Master's thesis, İstanbul Şehir University, 2016.

[33] Z. Zhou and P. Mirchandani, "A multi-sensor data fusion framework for real-time multi-lane traffic state estimation," in *Transportation Research Board 94th Annual Meeting*, no. 15-0186, 2015.

[34] N. Bekiaris-Liberis, C. Roncoli, and M. Papageorgiou, "Structural observability of multi-lane traffic with connected vehicles," in *IEEE Conference on Intelligent Transportation Systems, Yokohama, JAPAN*, 2017.

[35] G. Perraki, C. Roncoli, I. Papamichail, and M. Papageorgiou, "Evaluation of an MPC strategy for motorway traffic comprising connected and automated vehicles," *In IEEE Conference on Intelligent Transportation Systems, Yokohama, JAPAN*, 2017.

[36] M. Papageorgiou and A. Messmer, "Metanet: A macroscopic simulation program for motorway networks," *Traffic Engineering & Control*, vol. 31, no. 8-9, pp. 466–470, 1999.

[37] J. A. Laval and C. F. Daganzo, "Lane-changing in traffic streams," *Transportation Research Part B: Methodological*, vol. 40, no. 3, pp. 251–264, 2006.

[38] Y. Wang and M. Papageorgiou, "Real-time freeway traffic state estimation based on extended kalman filter: a general approach," *Transportation Research Part B: Methodological*, vol. 39, no. 2, pp. 141–167, 2005.

[39] *Aimsun 8 Dynamic Simulators Users' Manual*, Transport Simulation Systems, 2014.

[40] W. J. Schakel and B. Van Arem, "Improving traffic flow efficiency by in-car advice on lane, speed, and headway," *IEEE Transactions on Intelligent Transportation Systems*, vol. 15, no. 4, pp. 1597–1606, 2014.

[41] S. Papadopoulou, "Microscopic simulation-based validation of a per lane traffic state estimation scheme on highways with connected vehicles," Master's thesis, School of Production Engineering and Management, Technical University of Crete, 2017.