TECHNICAL UNIVERSITY OF CRETE

SCHOOL OF PRODUCTION

ENGINEERING AND MANAGEMENT

Microscopic Simulation-Based Validation of a Per Lane Traffic State Estimation Scheme on Highways with Connected Vehicles

Sofia Papadopoulou

Supervisor: Prof. Markos Papageorgiou

Thesis submitted in partial fulfillment of the requirements for the degree of

Master of Science

Chania, Greece, 2017

Acknowledgments

Finalizing this thesis, I would like to thank my supervisor, Prof. Markos Papageorgiou, as well as Prof. Ioannis Papamichail, Dr. Nikolaos Bekiaris-Liberis and Prof. Claudio Roncoli for all their constructive remarks and guidance throughout the development of this work.

Most of all, I would like to thank my family for their support and encouragement all these years, as well as my friends and my boyfriend for being there for me.

This master thesis has been conducted in the frame of the project TRAMAN21, which has received funding from the European Research Council under the European Union's Seventh Framework Programme (FP/2007-2013)/ERC Advanced Grant Agreement n. 321132.

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. This investigation employs a calibrated and validated, with real data, microscopic multi-lane model and it concerns a stretch of the motorway A20 from Rotterdam to Gouda in the Netherlands. The estimation methodology is based on the assumption that a certain percentage of vehicles are equipped with Vehicle Automation and Communication Systems (VACS), which provide the necessary measurements employed by the estimator, namely, speed and position measurements. A minimum number of total flow measurements from spot sensors, which guarantee observability of the system, are also needed. The proposed methodology can provide satisfactory estimation performance, even for low penetration rates of connected vehicles, while it is shown that is little sensitive to the model parameters (two in total).

Keywords: Traffic state estimation, Connected vehicles, Kalman Filter, Microscopic simulation testing

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Motivation

Major cities around the world experience ever-growing recurrent traffic congestion in urban areas and motorway networks. Congestion adversely affects mobility, safety, and air quality, leading to direct and indirect economic losses due to delays, accidents and environmental impact. While the number of vehicles has been increasing steadily during the past decades (Dargay et al., 2007), in most cases, the capacity of the existing roadway systems cannot be increased by expanding road networks due to space, resource, or environmental constraints. A way to address congestion problem is to utilize the existing traffic systems at its best performance through traffic management and operation strategies. In order to achieve efficient traffic management, a road administrator needs to understand the traffic situation. Real-time traffic information is essential in applications, such as freeway ramp metering control, dynamic route guidance, incident detection and variable message signs operations. For this reason, traffic authorities and automobile industries are currently focusing on the development of innovative methods for traffic monitoring (Bishop,

2005).

Acquisition of road traffic data is one of the most essential roles of intelligent transport systems (ITS).A completely continuous surveillance of the network provided by information received from installed infrastructure, such as inductive loops, is practically inaccessible due to the high cost of purchase, installation and maintenance. An innovative approach, compensating these deficiencies, is utilizing vehicles that may serve both the means of information collection and transmission themselves as a source of real-time traffic data. With the advent of highly equipped vehicles and vehicle automation, Vehicle Automation and Communication Systems (VACS) aim at improving driving safety and convenience while they introduce a new principle of traffic management and control. With the evolution of VACS of various kinds, an increasing number of "connected" vehicles is introduced that is capable of sending real-time information to a local or central monitoring and control unit (MCU), offering a wider, as well as efficient, data collection range, at low-cost and thus can be exploited for the development of novel traffic estimation and control methodologies. Although the penetration rate of connected vehicles is expected to substantially increase in the future (Diakaki et al., 2015), data retrieved from probe vehicles now covers only a limited part, around 5% (Qiu et al., 2010), of the whole traffic. Thus, a combination of floating car data with conventional detectors is adopted. Connected vehicles are considered as moving sensors along the network, whereas conventional detectors are installed only at a limited number of locations, sufficient to guarantee observability of the system even in low penetration rate of connected vehicles.

1.2 Literature review

Availability of timely and efficient estimates of traffic conditions is at the core of intelligent transportation systems for traffic management and control. As such, a vast amount of studies in the literature has been focused on traffic state estimation algorithms. Various techniques to collect traffic data are considered and numerous methods to estimate traffic density are proposed in relative literature. Recently, research on exploiting the innovative characteristics of VACS, as source of traffic data, in traffic state estimation, has drawn some attention, primarily due to the wide coverage and high accuracy of the extracted data. Traffic state estimation utilizing floating car data has been investigated in numerous studies, such as, e.g. Nanthawichit et al. (2003), Work et al. (2008), De Fabritiis et al. (2008), Herrera et al. (2010), Rahmani et al. (2010), Qiu et al. (2010), Schreiter et al. (2010), Treiber et al. (2011), Yuan et al. (2012), van Hinsbergen et al. (2012), Deng et al. (2013), Anand et al. (2014), Piccoli et al. (2015), Seo and Kusakabe (2015), Fountoulakis et al. (2016), Rempe et al. (2016), Wang et al. (2016), Wright and Horowitz (2016), Roncoli et al. (2016), Bekiaris-Liberis et al. (2016), to name only a few.

Bekiaris-Liberis et al. (2017) considers a macroscopic model-based approach for estimation of vehicles density and flow, exploiting speed measurements reported from connected vehicles and a minimum number of flow measurements from fixed sensors along the highway. This approach describes the density dynamics as a linear time-varying system and employs a Kalman filter for density estimation. It is developed under the assumption that the average speed of conventional vehicles on an arbitrary segment is roughly equal to the average speed of connected vehicles reported. The proposed model is examined through simulations using the METANET (Messner

and Papageorgiou, 1990) second-order traffic flow model as ground truth for the traffic state, including the case in which the speed of connected vehicles is reported to the central monitoring and control unit (MCU) with a communication delay.

A method for dealing probe vehicle data along with conventional detector data in order to estimate the traffic state variables of traffic volume, space mean speed and density is proposed in (Nanthawichit et al., 2003). The method used a macroscopic model along with a Kalman filter. However, the findings are validated only for a small single-lane road section, using hypothetical data and under the assumption that traffic state is homogeneous within each segment.

Seo and Kusakabe (2015) proposed a traffic state estimation algorithm that relies only on the spacing information retrieved from probe vehicles. In order to estimate the flow and density, an aggregation method based on the traffic conservation law is employed. This method assumes that the number of vehicles between two specific probe vehicles is constant along a section where flow discontinuity points (e.g., a node in a road network) do not exist. Flow is assumed to be single-lane traffic that satisfy the first-in first-out (FIFO). According to the validation results presented in this work, the proposed method indicated a good performance, although the conservation law equation was not exactly satisfied because of the on/off-ramps and lane-change effect.

The problem of lane-based traffic state estimation is less documented in the traffic literature in general, with even less of these studies investigating the impact of lane change on traffic state distribution among the lanes. In addition, the existing studies mainly assume data obtained from conventional detectors.

Chang and Gazis (1975) applied a discrete Kalman filtering technique to estimate traffic state, namely vehicle counts and travel time in a multi-lane highway utilizing

speed and flow measurements from fixed sensors placed at the entrance and the exit of the section. This approach considers a lane-change model integrated in the state space model. However, their method has restricted applications as it relies on aerial data that demand technology of high cost and is more prone to measurement errors.

Sheu (1999) utilize an extended Kalman filter, truncation and normalization procedures for estimating lane changes and density values per lane. Lane change is described as a discrete-time, nonlinear stochastic model that assumes accurate measurements of lane traffic count and occupancy retrieved from conventional loop detectors operations, which in reality are proved to involve substantial measurement errors. Finally, as stated by Sheu the proposed methodology should be regarded as preliminary because of the limitations of the field data used for model tests and should be extended to cover anomalous but important traffic conditions such as lane-blocking incidents and short-term queue-overflow occurrences.

Coifman (2003), developed a method for estimating traffic density in a freeway lane between detector stations by explicitly modeling lane inflows. No dynamic traffic model was involved in the algorithm. This method, as in Sheu (1999), heavily relies on the accuracy of loop detector information that are utilized in a sparse vehicle re-identification algorithm, one that may match the measurements from as few as 5% of the vehicles that pass both detector stations.

Singh and Li (2012b) proposed a recursive traffic density estimation scheme employing Kalman filter with data retrieved from loop detectors is proposed. Lane-changing behavior is described as a Markov chain process that enhances the state space model. This approach was developed under the assumption that traffic flow is stable so that the lane-changing probabilities of the Markov chain remain approximately constant over the time. A suggestion mentioned in this work is to consider

shorter time sub-periods within period of interest during which traffic flow should be stable, which is impracticable in reality as well. Another limitation in this model is that it considers required information from dense loop detector stations deployed (every 500-1500 meters) and the performance of the model is highly sensitive to the length of the segments investigated.

This work was extended later in (Singh and Li, 2012a) to include explicit consideration of the time-varying nature of lane-change behavior. Driver's lane-change decisions are modeled using discrete choice theory exploiting vehicle speed, time headway and traffic density. The research in this paper has several limitations as well. Lane-change probabilities may not be able to capture the potential non-linearity/heterogeneity in lane-change behavior. Moreover, it is considered that lane change of each vehicle may happen only once every time interval. Finally this approach needs to be evaluated using different roadway layouts and real traffic conditions.

Yılan (2016) proposes a method for estimating density on a multi-lane highway by estimating the mean values, bandwidths, and kernel weights of clusters, which represent a group of moving vehicles. The model proposed comprises peak detection algorithm for determining speed centers, linear search method and Newton-Raphson Method for variances or kernel bandwidths estimation, Kolmogorov-Smirnov Test for empirical cumulative distribution function detection, Kernel Density Estimation is utilized to model first probability densities and finally Scalar Kalman filter is employed to estimate traffic density. However, this model is tested a real-time data that has only one lane in the road and it evaluates only loaded data.

1.3 Thesis contribution

The contribution of this thesis is to test in microscopic simulation, a per lane traffic estimation scheme, which is largely based on the presence of connected vehicles. The scheme is developed in Bekiaris-Liberis et al. (2017) and it is based on a simple data-driven macroscopic model for per lane traffic density and employs real-time traffic measurements obtained from connected vehicles and a minimum number of spot flow measurements. The model is derived from the conservation law equation, which implies that there is no requirement for a fundamental diagram, or other empirical relationships.

The performance of the estimation scheme is examined under various penetration rates of connected vehicles, using data retrieved from a microscopic multi-lane model, calibrated and validated using real data from a stretch of the motorway A20 from Rotterdam to Gouda in the 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, e.g. in the case where a limited number of measurements (or no measurements at all) are available from connected vehicles). The performance of the estimation scheme examined is shown to be satisfactory, even for low penetration rates. Finally, it is demonstrated that estimation performance is little sensitive to the choice of the model parameters (two in total).

1.4 Thesis outline

The remainder of this thesis is structured as follows. Chapter 2 derives a linear time-varying model of the density dynamics. Then, a Kalman filter is employed for the estimation of the total density of vehicles. A microscopic simulation configuration as well as the traffic network and scenario employed are analytically described in Chapter 3. Chapter 4 presents the methodologies employed for the computation of the measurements utilized by the estimation scheme. Then the evaluation of performance of the estimation scheme employed is illustrated. Under the qualitative category, graphical techniques are presented to help a visual examination of the differences between the simulation and the ground trough. Under the quantitative category employing an appropriate metric to compare the performance of the proposed methodology under different scenarios. Finally, chapter 5 summarizes the conclusions highlighting the main findings and also suggests recommendations and potential improvements for future research.

Chapter 2

Exploitation of VACS for traffic state estimation

2.1 Innovative features of VACS

Automated and connected vehicle technologies are becoming one of the most significant interdisciplinary issues that relate to the future direction of the automotive technologies. Connected vehicles incorporate cooperative systems that can be classified to Vehicle to Vehicle (V2V), Vehicle to Infrastructure (V2I) and Vehicle to both vehicle and infrastructure (V2X) communication systems (Diakaki et al., 2015). The communication flows are mainly based primarily on wireless communications technologies, typically via a GPRS/GSM network (Bishop, 2005) for short-range V2I communications and a WiFi 802.11 (Waterson and Box, 2012) network for longer range communications that serve V2V communication. 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 (Zito et al. (1995), Turksma (2000), Liu et al. (2006)). 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 (Waterson and Box, 2012), sufficient for lane-based applications. Speed measurements are mostly reported to be quite accurate with a precision error lower than 1 km/h (Chalko (2007), Zito et al. (1995)), while some studies claim a tendency of underestimation in speed measurements and a reported error around 5km/h (Zhao et al., 2014). Cellular technologies that utilize GPS-enabled mobile phones are considered as a source of traffic information but they are prone to significant inaccuracy in their reported position (Herrera et al. (2010), Cayford and Johnson (2003), Tao et al. (2012), Work et al. (2008)).

Transmission information may be distinguished as event-based (i.e. reports if traffic accident), space-based (reports as they pass observation points) and time-based (reports every pre-defined for each vehicle time instant). In this study, the time-based approach was adopted, as it offers a better surveillance of the entire network. Transmission frequencies of connected vehicles vary among different experiments and commercial systems, most frequently with simulated transmission intervals ranging from 5 seconds to a few minutes (Liu et al. (2006), Stenneth and Giurgiu (2016)).

2.2 Traffic state estimation using connected vehicles' measurements

2.2.1 Lane-based traffic density dynamics

Macroscopic description of traffic flow implies the definition of adequate flow variables expressing the average behavior of the vehicles at a specific location and time instant. In the existing study, a discrete-time multi-lane extension of the traffic conservation law for a highway segment is presented that describe the density dynamics of highway segments per lane, which is further adjusted to incorporate also the case of unmeasured ramps and lane-changing phenomena within this highway segment. Density $\rho_{i,j}(k)$ (in veh/km) of highway segment i and lane j at a time step kT, is defined as the number of vehicles per unit length of the highway segment.

$$\rho_{i,j}(k+1) = \rho_{i,j}(k) + \frac{T}{\Delta}(q_{i-1,j}(k) - q_{i,j}(k) + L_{i,j-1\to j}(k) + L_{i,j+1\to j}(k)) \quad (2.1)$$

$$- L_{i,j\to j-1}(k) - L_{i,j1\to j+1}(k) + r_{i,j}(k) - s_{i,j}(k)), \quad (2.2)$$

where $i = 1, \ldots, N$ and $j = 1, \ldots, M$, are the indexes of the specific highway segment and lane, respectively, of the network, $k = 0, 1, ..$ is the discrete time index, Δ_i (in km) is the length of segment i and T (in h) is the sample time interval; $q_{i,j}$ (in veh/h) denotes the total longitudinal inflow at segment $i+1$ and lane j during the time period $[kT, (k+1)T]$, $r_{i,j}$ and $s_{i,j}$ (both in veh/h) denote the total flow of on-ramps and off-ramps at segment i, respectively and $L_{i,j_1\to j_2}$, (in units of veh/h) for all $j_1 = 1, \ldots, M$, and $j_2 = j_1 \pm 1$ denotes the total lateral flow at segment i that enters to lane j_2 from lane j_1 during the time period $[kT, (k+1)T]$. Obviously $L_{i,j_1\to j_2} \equiv 0$, if either j_1 or j_2 equals zero or $M+1$ and $r_{i,j} \equiv s_{i,j} \equiv 0$ for all i and

$1 \leq j \leq M-1$, where lane M denotes the right most lane. Note that it is assumed that the considered network is subdivided into sections, each containing at most one on-ramp or off-ramp.

The time-discretization step T is wisely chosen, for numerical stability reasons, with respect to the the Courant-Friedrichs-Lewy (CFL) condition:

$$\max_{i,j,k} \frac{T}{\Delta_i} v_{i,j}(k) < 1, \tag{2.3}$$

for all $i = 1, \ldots, N$ and $j = 1, \ldots, M$, where $\max_{i,j,k} v_{i,j}(k)$ denotes the free-flow speed.

2.2.1.1 Model for lateral flows

The lane-change proportion between two lanes depends on the highways' time-varying traffic condition. To capture the lane changing behavior, the total lateral flow $L_{i,j_1 \to j_2}(k)$ at segment i from lane j_1 to the neighbor lane j_2 during the time interval $[(k-1)T, kT]$ are computed based on Eq.2.4. Specifically, the total lateral flow is modeled as a linear function of the inverse of the percentage of connected vehicles and the lateral flow computed on the basis of connected vehicles regularly reported positions within this segment.

$$L_{i,j_1 \to j_2}(k) = \frac{L^c_{i,j_1 \to j_2}(k)}{\rho^c_{i,j_1}(k)} \rho_{i,j_1}(k) \tag{2.4}$$

where ρ^c_{i,j_1}, for all $i = 1, \ldots, N$ and $j = 1, \ldots, M$, denotes the density measurements of connected vehicles at segment i and lane j, and $L^c_{i,j_1 \to j_2}$, for all $i = 1, \ldots, N, j_1 = 1, \ldots, M$, and $j_2 = j_1 \pm 1$, denotes the lateral flow of connected vehicles at segment i that enters to lane j_2 from lane j_1. The validity of the model is tested in Section 4.2 and in the estimation results of Section 4.3.2.

2.2.1.2 Model for total longitudinal flows

As a direct consequence of the fundamental traffic variables' definition density $\rho_{i,j}$, flow $q_{i,j}$ and speed $v_{i,j}$ (km/h), we get the relationship

$$q_{i,j}(k) = v_{i,j}(k)\rho_{i,j}(k), \tag{2.5}$$

where $v_{i,j}$, for all $i = 1, \ldots, N$ and $j = 1, \ldots, M$,, denotes the average vehicle speed at segment i and lane j at a time instant k. The main hypothesis of the model is that traffic is in an equilibrium state such that traffic flow and speed only depend on the density of traffic at the same position. Note that $q_{0,j}$, for all $j = 1, \ldots, M$, the inflows at the entry of the considered highway are assumed to be measured (via fixed sensor), and are treated as inputs to the system (2.1).

Within our experiments, we confronted an inefficiency in the longitudinal flow modeling (Eq. 2.5) under certain cases, such as, for example at cells that feature an on-ramp or a lane-drop and a significant amount of the lateral flow may appear close to the end of those cells. For instance, in case there is a lane-drop located at $(i, j-1)$, the real longitudinal flow of the cell (i, j), which is adjacent to the lane-drop cell, is higher than $\rho_{i,j} \times v_{i,j}$, see Fig. 2.1. This is caused due to the "late" lane-changing from lane $j-1$ to lane j of segment i in case of congestion. During that congested period, there is a standing queue created to the lane-drop cell. This queue produces lateral outflow towards the end of the segment ("late" lane-changing), which mainly has impact on the density dynamics of the downstream segment $i+1$ than the one that are produced.

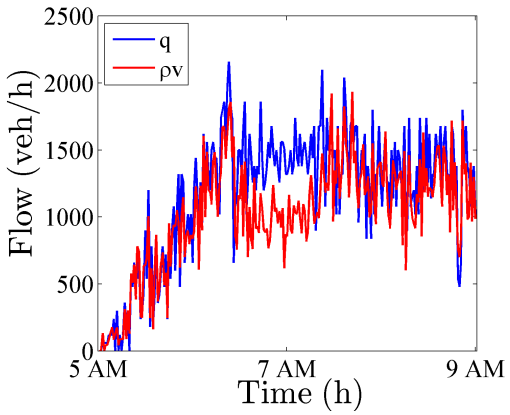


Figure 2.1: Inefficiency in longitudinal flow modeling (Eq. 2.5)

As a result, we address this inefficiency in the outflow modeling considering that a certain percentage of lateral or on-ramp flows acts as an additional exiting flow. Specifically, we have an appropriate correction of the outflow values introducing a constant parameter $p_{i,j}$ indicating the percentage of lateral inflow that splits into two different cells. Within the hypothesis that that cell $(i, j-1)$ is a cell where this "late" lane-changing phenomenon is observed, lateral flow will be considered as inflow to cells (i, j) and $(i+1, j)$ with a percentage $(1-p_{i,j})$ and $p_{i,j}$, respectively, as shown in Fig. 2.2.

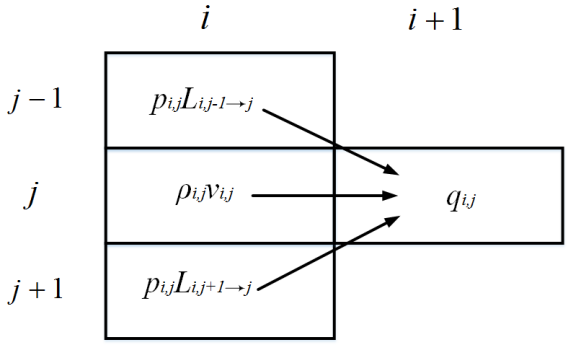


Figure 2.2: Diagonal lateral flows

The established model for the longitudinal outflow, employed in the estimation

scheme is

$$q_{i,j}(k) = v_{i,j}(k)\rho_{i,j}(k) + p_{i,j}\frac{T}{\Delta}\frac{L^c_{i,j-1\to j}(k)}{\rho^c_{i,j-1}(k)}\rho_{i,j-1}(k) + p_{i,j}\frac{T}{\Delta}\frac{L^c_{i,j+1\to j}(k)}{\rho^c_{i,j+1}(k)}\rho_{i,j+1}(k), \quad (2.6)$$

where the factor for diagonal lateral flows $p_{i,j} \in [0, 1]$, for $i = 1, \ldots, N$ and $j = 1, \ldots, M$, will be decided based on statistical analysis (see Fig. 4.8).

The percentages $p_{i,j}$ are considered as tuning parameters and may be different to each other. Within our experiments using data retrieved from microscopic simulation, it has been observed that the cell which features a significant sensitivity in the estimation performance to the variations of the parameter $p_{i,j}$ is the one adjacent to the lane-drop. However, the case where one could adopt this updated outflow model for all the highway segments, as well as in the case of the flows produced by on-ramps, has been examined. In this case terms that estimate the total lateral flows, i.e. terms L^c and ρ^c, have to be incorporated in the measurement equations (see Section 2.2.1.3), which implies additional inaccuracy (compared to the case where it depends only on the segment speeds). In addition, such modification of the measurement equation implies that observability analysis of the system may become more involved. Finally, an excessive tuning of the parameters $p_{i,j}$ would be needed in order to obtain an overall significant improvement. Note that since only one percentage value needs to be non-zero, we adopt for the rest of this paper the notation $p_{i,j} = p$.

2.2.1.3 LPV model for the density dynamics

The dynamics of the total traffic density, as described by the conservation law (2.1), are recast incorporating also (2.6), (2.4) as a linear time-varying system with known parameters that depend on the real-time measurements retrieved from connected

vehicles.

$$\begin{aligned}\rho_{i,j}(k) =& \left(1 - \frac{T}{\Delta}\upsilon_{i,j}(k) - \frac{T}{\Delta}\frac{L^c_{i,j\to j-1}(k)}{\rho^c_{i,j}(k)} - \frac{T}{\Delta}\frac{L^c_{i,j\to j+1}(k)}{\rho^c_{i,j}(k)}\right)\rho_{i,j}(k) + \frac{T}{\Delta}\upsilon_{i-1,j}(k)\rho_{i-1,j}(k) \\ &+ (1-p)\frac{T}{\Delta}\frac{L^c_{i,j-1\to j}(k)}{\rho^c_{i,j-1}(k)}\rho_{i,j-1}(k) + (1-p)\frac{T}{\Delta}\frac{L^c_{i,j+1\to j}(k)}{\rho^c_{i,j+1}(k)}\rho_{i,j+1}(k) \\ &+ p\frac{T}{\Delta}\frac{L^c_{i-1,j-1\to j}(k)}{\rho^c_{i-1,j-1}(k)}\rho_{i-1,j-1}(k) + p\frac{T}{\Delta}\frac{L^c_{i-1,j+1\to j}(k)}{\rho^c_{i-1,j+1}(k)})\rho_{i-1,j+1}(k) \\ &+ \frac{T}{\Delta}(r_{i,j}(k) - s_{i,j}(k)),\end{aligned} \tag{2.7}$$

where the terms $\upsilon_{i-1,j}(k)\rho_{i-1,j}(k)$ for all $i = 1$ and all $j = 1, \ldots, M$ are considered measured inputs of the system. Terms $\frac{L^c_{i,j\to j-1}(k)}{\rho^c_{i,j}(k)}\rho^c_{i,j}(k)$ and $\frac{L^c_{i,j-1\to j}(k)}{\rho^c_{i,j-1}(k)}\rho^c_{i,j-1}(k)$, for all i and $j = 1$ should be set to zero, whereas for all i and $j = M$ the terms $\frac{L^c_{i,j\to j+1}(k)}{\rho^c_{i,j}(k)}\rho^c_{i,j}(k)$ and $\frac{L^c_{i,j+1\to j}(k)}{\rho^c_{i,j+1}(k)}\rho^c_{i,j+1}(k)$ should be set to zero. Note that the The above formulation may be modified to incorporate different values of the parameter p.

Assuming constant flows for the unmeasured on-ramps and off-ramps, we introduce an supplementary random-walk equation for $r_{i,j}$ and $s_{i,j}$ as follows:

$$\theta_{i,M}(k+1) = \theta_{i,M}(k) + \xi^\theta_i(k), \tag{2.8}$$

where ξ^θ_i is zero-mean white Gaussian noise and

$$\theta_i = \begin{cases} \frac{T}{\Delta_i} r_{n_i,M}, & \text{if } n_i \in L_r \\ \frac{T}{\Delta_i} s_{n_i,M}, & \text{if } n_i \in L_s \end{cases} \tag{2.9}$$

where l_r and l_s are the number of on-ramps and off-ramps, respectively, whose flows are not directly measured and $L_r = \{n_1, \ldots, n_{l_r}\}$ and $L_s = \{n_{lr+1}, \ldots, n_{l_r+l_s}\}$, for all $i = 1, \ldots, l_r + l_s$, being the sets of segments, denoted by n_i, which include an unmeasured on-ramp or off-ramp, respectively.

Associating equations (2.7–2.9), the macroscopic traffic flow model of a highway

stretch can be expressed in a compact state-space form as a Linear Parameter-Varying (LPV) system. We first define vector x to describe compactly system (2.7)

$$x = (\rho_{1,1}, \ldots, \rho_{N,1}, \ldots, \rho_{1,M}, \ldots, \rho_{N,M}, \theta_{1,M}, \ldots, \theta_{l_r+l_s,M})^T. \tag{2.10}$$

Note that if some segments do not include any on-ramp or off-ramp, then vector x is reduced accordingly. The deterministic part of the dynamics of the total density given in (2.1), θ_i given in (2.9) and the dynamics of the smoothed lateral flows given in (4.5) can be written as:

$$x(k+1) = A(v^c(k), L^c(k), \rho^c(k))x(k) + Bu(k), \tag{2.11}$$

where v^c, L^c and ρ^c denote vectors that are viewed as time-varying parameters of (2.11) and that incorporate all average segment speeds $v^c_{i,j}$ computed from connected vehicles' reports, lateral flows of connected vehicles $L^c_{i,j_1 \to j_2}$, and densities of connected vehicles $\rho^c_{i,j}$, respectively, u denotes the vector of inflows at the highway, namely

$$u(k) = [q_{0,1}, ..., q_{0,M}]^T, \tag{2.12}$$

$$A(k) = \begin{cases} a_{lm} = \frac{T}{\Delta_i} \upsilon^c_{i,j}, & \text{if } l - m = 1 \text{ and } 2 \leq l \leq N \times M \\ & \text{with } i = l - \lfloor \frac{l}{N} \rfloor N - 1 \text{ and } j = \lceil \frac{l}{N} \rceil \\ a_{lm} = 1 - \frac{T}{\Delta_i}(\upsilon^c_{i,j} - \frac{L^c_{i,j\to j-1}}{\rho^c_{i,j}(k)} - \frac{L^c_{i,j\to j+1}(k)}{\rho^c_{i,j}(k)}), & \text{if } l = m \text{ and } 1 \leq l \leq N \times M \\ & \text{with } i = l - \lfloor \frac{l}{N} \rfloor N \text{ and } j = \lceil \frac{l}{N} \rceil \\ a_{lm} = p\frac{T}{\Delta_i}\frac{L^c_{i,j+1\to j}}{\rho^c_{i,j+1}(k)}, & \text{if } l = m - N \text{ and } 1 \leq l \leq N \times (M-1) \\ & \text{with } i = l - \lfloor \frac{l}{N} \rfloor N \text{ and } j = \lceil \frac{l}{N} \rceil \\ a_{lm} = p\frac{T}{\Delta_i}\frac{L^c_{i,j-1\to j}}{\rho^c_{i,j-1}(k)}, & \text{if } l = m + N \text{ and } N + 1 \leq l \leq N \times M \\ & \text{with } i = l - \lfloor \frac{l}{N} \rfloor N \text{ and } j = \lceil \frac{l}{N} \rceil \\ a_{lm} = -1, & \text{if } l = m - 2N \text{ and } N \times M - 1 \leq l \leq N \times M \\ a_{lm} = \quad 1, & \text{if } l = m - N \text{ and } N \times M - 1 \leq l \leq N \times M \\ a_{lm} = \quad 1, & \text{if } l = m \text{ and } M * N < l \leq N \times M + 2N \\ a_{lm} = \quad 0, & \text{otherwise} \end{cases} \tag{2.13}$$

$$B(k) = \begin{cases} b_{lm} = \frac{T}{\Delta_l}, & \text{if } l = m = 1 \\ b_{lm} = \frac{T}{\Delta_{n_l}}, & \text{if } l = (m-1) \times N + 1 \text{ and } m = 1, \ldots, M \\ b_{lm} = \quad 0, & \text{otherwise} \end{cases} \tag{2.14}$$

with $A \in \mathbb{R}^{(N\times M+2N)\times(N\times M+2N)}$, $B \in \mathbb{R}^{(N\times M+2N)\times M}$ and $\lfloor\zeta\rfloor, \lceil\zeta\rceil$ denote the largest previous and smallest following integer of number ζ, respectively.

Together with (2.11) we associate an output vector y, which incorporates any measured mainstream total flow, given by

$$y(k) = C(\upsilon^c(k))x(k), \tag{2.15}$$

where $C \in \mathbb{R}^{(M+l_r+l_s-1)\times(N\times M+2N)}$, is defined as

$$C(v^c(k)) = \begin{cases} c_{lm} = v^c_{m-N\times(M-1),M}, & \text{if } M+1 \le l \le M+l_r+l_s-1 \\ & \text{and some } n^*_{l-M} \le m - N\times(M-1) \le n^*_{l+1-M}-1 \\ c_{lm} = v^c_{N,l}, & \text{if } 1 \le l \le M \text{ and } m = N \times l \\ c_{lm} = 0, & \text{otherwise} \end{cases} \tag{2.16}$$

where $\bar{L}^* = \{n^*_1, n^*_2, \ldots, n^*_{l_r+l_s}\}$ is the set of $\bar{L}$ ordered by $<$, with $\bar{L} = L_r \cup L_s$,. Note that we assume $2N$ ramp flows. In the case where there are less on-ramp or off-ramp, the dimension of the matrices A, B and C are reduced accordingly.

We summarize below the measurement requirements for the proposed estimation algorithm.

- The speed of connected vehicles at any segment of the highway is measured and used for computing the average segment speed $u^c_{i,j}, i = 1, \ldots, N$ and $j = 1, \ldots, M$ employed by the estimator; the assumption being that there is no systematic difference between the average speed of connected vehicles and the average speed of all vehicles in a segment.

- The position of connected vehicles at any segment of the highway is measured and used, with regard to the last reported position, to compute the lateral flow of connected vehicles $L^c_{i,j_1\to j_2}$ at segment $i = 1, \ldots, N$ from lane $j_1 = 1, \ldots, M$ to lane $j_2 = j_1 \pm 1$, employed in the model for total lateral flow.

- The number of connected vehicles at any segment of the highway is measured and used to compute the density of connected vehicles $\rho^c_{i,j}, i = 1, \ldots, N$ and $j = 1, \ldots, M$ employed in the model for total lateral flow.

- The total flow of vehicles at the entry $q_{0,j}$, and exit $q_{N,j}$, for all $j = 1, \ldots, M$, of the considered highway stretch is available via a fixed flow sensor.

- Either the total flow of vehicles at a ramp, is measured or an additional mainstream flow measurement, say $q_{i,j}$, where i, j corresponds to any highway segment of the considered highway stretch between every two consecutive unmeasured ramps is assumed to be available via a corresponding fixed flow sensor.

2.2.2 Per-lane total density estimation using Kalman filter

We employ then a Kalman filter for the estimation of the traffic state of a highway. Defining the vector $\hat{x}$ as the system state to be estimated,

$$\hat{x} = (\hat{\rho}_{1,1}, \ldots, \hat{\rho}_{N,1}, \ldots, \hat{\rho}_{1,M}, \ldots, \hat{\rho}_{N,M}, \hat{\theta}_{1,M}, \ldots, \hat{\theta}_{l_r+l_s,M})^T. \tag{2.17}$$

the estimation scheme equations are given by (see, e.g., (Bekiaris-Liberis et al., 2017))

$$\begin{aligned}\hat{x}(k+1) =& A(v^c(k), L^c(k), \rho^c(k))\hat{x}(k) + Bu(k) \\ &+ A(v^c(k), L^c(k), \rho^c(k))K(k)(z(k) - C(v^c(k))\hat{x}(k)) \qquad (2.18)\\ K(k) =& P(k)C(v^c(k))^T(C(v^c(k))P(k)C(v^c(k))^T + R)^{-1} \qquad (2.19)\\ P(k+1) =& A(v^c(k), L^c(k), \rho^c(k))(I - K(k)C(v^c(k)))P(k)A(v^c(k), L^c(k), \rho^c(k))^T + Q, \qquad (2.20)\end{aligned}$$

where z is a possibly noisy version of the measurement y defined in (2.15), $Q = Q^T$ and $R = R^T$ are tuning parameters which, in the ideal case in which there is additive, zero-mean Gaussian white noise in the state and output equations, represent the covariance matrices of the process and measurement noise, respectively. The initial

conditions of the filter described by Eq. (2.18)–(2.20) are chosen as

$$\hat{x}(k_0) = \mu \tag{2.21}$$

$$P(k_0) = H, \tag{2.22}$$

where μ and $H = H^T > 0$, which, in the ideal case in which $x(k_0)$ is a Gaussian random variable, represent the mean and auto covariance matrix of $x(k_0)$, respectively.

Chapter 3

Microscopic simulation setup

3.1 Simulation environment

The behavior of the developed estimation scheme is examined and evaluated through a microscopic simulation process using AIMSUN (Advanced Interactive Microscopic Simulator for Urban and Non-urban Networks) by Transport Simulation Systems (2014). AIMSUN is a commercial microscopic traffic simulation software widely used nowadays by transport professionals and researchers, as it provides tools (e.g. Aimsun API and microSDK), which among others offer the possibility to users to exchange information dynamically with the Aimsun module or configure the simulation parameters and modify or replace the incorporated vehicle models through user-defined applications in C++ or Python. Nevertheless, the two main critical components of a traffic simulation system are the car-following and lane-changing models. The default car-following and lane-changing model implemented in Aimsun are based on the models developed in Gipps (1981) and Gipps (1986), respectively. However, in this work the model utilized replaces the aforementioned default car-following model with the Intelligent Driver Model (IDM) (Treiber et al., 2000),

as (Wang et al., 2005) pointed to the insufficiency of (Gipps, 1981) model in reproducing capacity drop phenomena realistically. Also, some heuristic rules that were introduced in (Roncoli et al., 2014) were implemented to replace the default lane-changing model in sections where Gipps (1986) could not capture the merging behavior in a critical flow regime (Chevallier and Leclercq, 2009).

The estimation scheme (Section 2.2.2) scheme is applied on a multi-lane microscopic model that is designed and calibrated by (Perraki, 2016) with real lane-specific traffic data obtained from detectors, taken from (Schakel and Van Arem, 2014), providing us a realistic reference case. The case-study network (see, Fig. 3.1) is a stretch of the motorway A20 which is entirely located in the Dutch province South Holland. It links the N213 road of the Westland municipality with the cities of Rotterdam and Gouda, where there at the interchange of Gouwe it connects with the A12 motorway. Traffic demand is composed by traffic states with one minute time interval, same as the real data measurement interval and using a single vehicle type. Since there are no available detector data at on-ramp locations, the flow at these entrances is computed based on the measurements of the flow from the detectors placed before and after the locations that vehicles enter the mainstream network from the on-ramps. Similarly to this, due to the lack of data at off-ramp locations the percentage of vehicles exiting the network via an off-ramp is again calculated based on the flow measured after and before the off-ramp.

3.2 Network and data description

3.2.1 Network configuration

The case-study highway stretch, as shown in Fig. 3.1, constitutes a challenging test-bed for the estimation scheme proposed in the previous section, as it incorporates a non-trivial combination of lane-drops, on-ramps, and off-ramps, the congestion pattern, and the relatively closely-spaced detectors which cover anomalous but important traffic conditions. The considered stretch is about 9.3 km in length, comprises 3 homo-directional lanes until segment 8 where we have the drop of lane 1 and is space-discretised with N = 21 segments. Two on-ramps and two off-ramps are located within segments 8, 14 and 10, 16, respectively. There are 22 detectors placed rather homogeneously along the stretch, as illustrated in Fig. 3.1 and a flow detector located at off-ramp "Moordrecht".

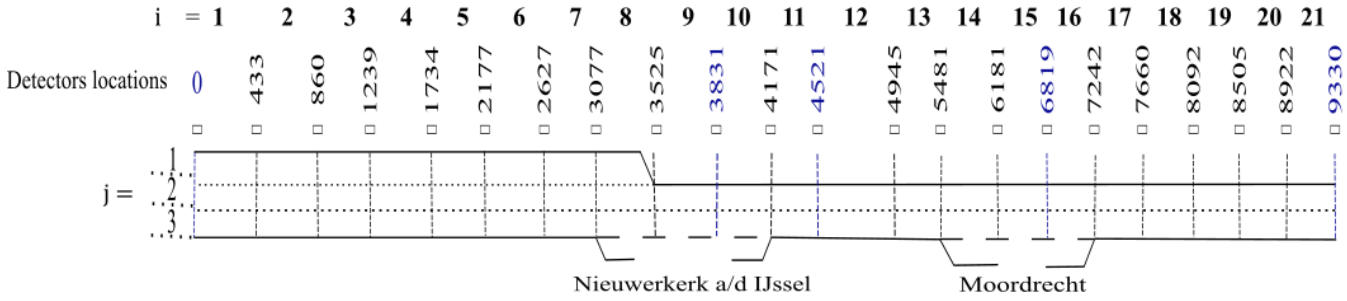


Figure 3.1: A schematic representation of the case study network. Detector positions are indicated as the distance (in m) from the network entrance. The detectors used by the estimator for obtaining flow measurements within the case study are colored blue.

The utilized network measurements are summarized in flow measurements retrieved

from a limited number of detectors according to the configuration shown in Fig. 3.1, that is sufficient to guarantee observability as explained in the previous section. Specifically, the detector located at the highway entrance (0 m) is used to obtain the input of the system $q_{0,j}$; all ramp flows are assumed unknown, therefore three additional detectors in each lane are utilized (one per lane between every pair of unmeasured ramps), choosing the ones located at 3831 m (segment i = 9), 4521 m (segment i = 11), 6819 m (segment i = 15); and, finally, one measurement per lane is taken at the network exit 9330 m (segment i = 21), employed as $q_{N,j}$. The duration of the simulation is 4 hours, same as the real data measurements duration. Hence, the traffic demand assigned in the simulation scenario is composed by 240 traffic states with one minute duration for each state. The utilized network parameters are summarized as shown in Table 3.1.

T	τ	Δ_i	v_f	N	M
$\frac{1}{360}(h)$	$\frac{1}{900}(h)$	$\simeq 0.4(km)$	$120(\frac{km}{h})$	21	3

Table 3.1: Parameters of the model

where T, which meets the stability condition defined in Eq. 2.3, denotes the detection interval of flow sensors, as well as the interval for calculating average segment speeds, τ is the simulation step and v_f is the free-flow speed.

3.2.2 Employed scenario

The employed scenario replicates traffic conditions of the morning rush hour of Wednesday, May 26, 2010, where a strong congestion is created in segment 10 around 6:30 AM because of the increased flow entering from on-ramp "Nieuwerkerk a/d

IJssel"; consequently, the congestion spills back and standing queues are also formed due to the lane drop in segment 8 and the congestion covers the stretch up to segment 3 until 7:00 AM. From 7:00 AM until 7:30 congestion continues at segment 10 and 9 while the bottlenecks appearing upstream disappear. Figure 3.2, demonstrates the speeds measured from detectors placed along each lane from 5 AM to 9 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. Positions of the off-ramps and on-ramps are specified on the right side of 3.2, while the dot line in the plot of the speed conditions in lanes 2 and 3 defines the exact location where lane 1 drops. Since lane 1 after some point merges with lane 2 the measured speeds are presented until 3:6 km where the lane drops.

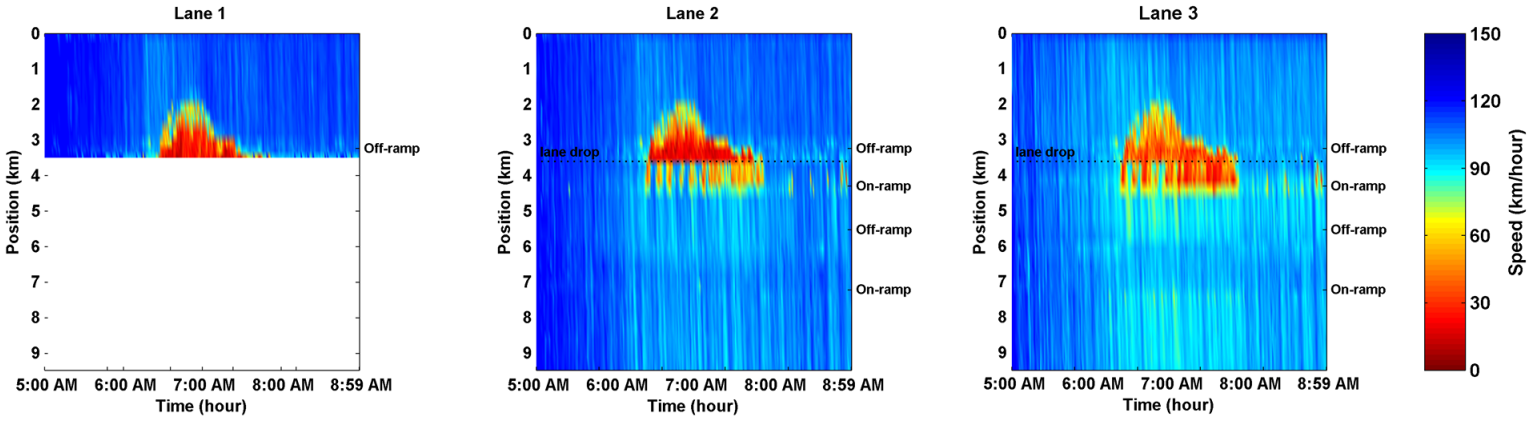


Figure 3.2: Simulated speed, Wednesday 26-05-2010,

Perraki (2016)

3.2.3 Ground truth generation

The ground truth in our experiments, used to evaluate the performance of the developed estimation scheme, is represented by the total density in each segment and the total ramp flows. The cell density $\rho_{i,j}$ and $\rho^c_{i,j}$ is computed by accumulating the number of all vehicles and of connected vehicles, respectively, that are present within segment i of lane j at a time instant kT divided by the length the segment length; whereas all flows at mainstream segments and ramp flows are computed by counting the number of vehicles that cross the corresponding detector within the time interval $[kT, (k+1)T]$. However, because lane-based density dynamics and ramp flows are very oscillatory a moving average of the last 6 measurements is considered as ground truth.

Average segment speed $v_{i,j}$ that represents ground truth is computed by averaging arithmetically at time step kT the instant speed of all vehicles present in a segment which is collected every 2 s. Finally, position information of all vehicles present at a highway are collected every 2 s; lateral flow $L_{i,j_1 \to j_2}$ is then computed comparing the reported position of an identified vehicle between every two consecutive reports and we accumulate lane changes from j_1 to j_2 within segment i every time step kT.

Chapter 4

Estimation performance evaluation via microscopic simulation

4.1 Experimental configuration

The traffic network and scenario described in Chapter 3 is considered to realize experiments within which we simulate cases featuring mixed traffic conditions (of connected and ordinary vehicles) with different penetration rates of connected vehicles. Essentially, this corresponds to evaluating the robustness of the approach to potentially less accurate measurements stemming from connected vehicles that are considered as input to the Kalman filter equations (see, Eq. 2.18–2.20). Vehicles entering the network are characterized as all of the same vehicle type, thus featuring identical overall behavior, and are randomly marked as connected, according to the assumed penetration rate based on a uniform distribution. 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, $2\%, 5\%, 10\%, 20\%$, and 50%.

To assess the overall performance level of the estimation scheme suggested under the quantitative term, we adopt a performance index as presented in Eq. 4.1, formulated as the Coefficient of Variation (CV) of the root mean square error of the estimated density $\hat{\rho}_{i,j}$ with respect to the corresponding ground truth density $\rho_{i,j}$, with simulation time horizon $K = \frac{4}{T} = 1440$, :

$$CV_\rho = \frac{\sqrt{\frac{1}{MNK}\sum_{i=1}^{N}\sum_{j=1}^{M}\sum_{k=1}^{K}[\hat{\rho}_{i,j}(k) - \rho_{i,j}(k)]^2}}{\frac{1}{MNK}\sum_{i=1}^{N}\sum_{j=1}^{M}\sum_{k=1}^{K}\rho_{i,j}(k)}. \tag{4.1}$$

Identically to (Eq. 4.1), CV of the estimated ramp flows $\hat{\theta}_{i,M}$ in the case of unmeasured ramp flows estimation, is formulated with respect to the corresponding ground truth ramp flows $\theta_{i,M}$ as :

$$CV_{r,s} = \frac{\sqrt{\frac{1}{K(l_r+l_s)}\sum_{k=1}^{K}\sum_{i=1}^{l_r+l_s}\frac{\Delta_i}{T}^2[\hat{\theta}_{i,M}(k) - \theta_{i,M}(k)]^2}}{\frac{1}{K(l_r+l_s)}\sum_{k=1}^{K}\sum_{i=1}^{l_r+l_s}\frac{\Delta_i}{T}\theta_{i,M}(k)}. \tag{4.2}$$

4.2 Computation of the measurements utilized by the estimator

Prior to the performance evaluation of the estimation scheme proposed, we address the issue related to the information provided to the system. The estimation performance is manly based on the quality of these information and thus we employ simple algorithms to ensure that these information are reliable and representative of the ground truth. As mentioned in Section 2.2.1.3, the estimation scheme is developed based on the assumption that the average speed of connected vehicles is roughly equal to the average speed of conventional vehicles. This assumption relies on the fact that, regardless of traffic conditions, there is no reason for connected vehicles to adopt a systematically different mean speed than conventional vehicles,

as all vehicles (connected or not) feature the same driving statistics. In cases of low penetration rates of connected vehicles , because of the limited number of vehicles transmitting information, two main issues may be obtained, that may degrade the estimation performance:

- Individual speed reports retrieved only from a few vehicles, that are present in the segment, may be non-representative for the ground truth average speed of the segment under certain conditions; for example, consider the case where at the time of report the driver is forced to strongly decelerate or stop due to an aggressive lane-changing manoeuvre from some other vehicle or an accidental vehicle breaking.

- With a time interval as small as $T = 10$ s, in low penetration rates there may be no connected vehicles within some segments during some time intervals. In such cases, information concerning segment speed and the number of connected vehicles remain unobserved. In our simulation experiments, considering lane-based detection, such intervals are of significant proportion even from penetration rates of 20% (see Eq. 4.1).

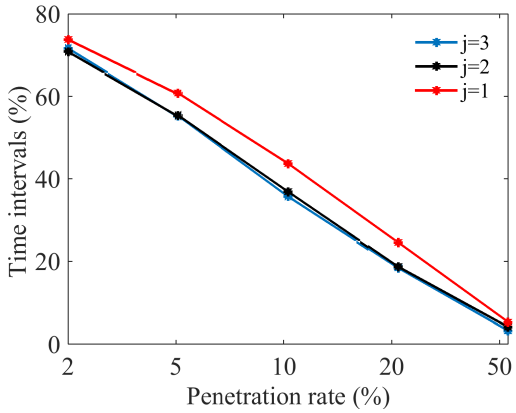


Figure 4.1: Average percentage of time intervals of T = 10 s that feature no connected vehicle report against penetration rate of connected vehicles, for all lanes $j = 1, 2, 3$.

A moving average, utilizing available measurements from previous times steps is considered to compensate for potential large errors in the measurements utilized by the estimator in such issues. Time intervals that feature no connected vehicles' reports, are addressed utilizing the last available measurements. With regard to the speed measurements $\upsilon^c_{i,j}$, for every time step k, we feed the filter with a moving average of the last n available speed measurements as :

$$\upsilon^c_{i,j}(k) = \sum_{l=0}^{n-1} \frac{\nu^c_{i,j}(k-l)}{n}, \tag{4.3}$$

where $\nu^c_{i,j}(k)$ is the average speed computed from connected vehicles reports at segment i of lane j computed from speed reports collected every 2 s and averaged arithmetically every time step kT. In cases where there are no connected vehicles reports available at cell i, j during a time interval $((k-1)T, kT]$, we replace this $\nu^c_{i,j}(k)$ with the speed reported at the previous time step as , i.e., $\nu^c_{i,j}(k) = \nu^c_{i,j}(k-1)$. In our experiments we considered values for n equal to 6 (1 m time window) and 12 (2 m time window). We choose to employ a moving average of the last $n = 12$ last

speed measurements available, as it is more representative of the overall segment speed, especially in low penetration rates.

Similarly, with regard to the measurements of $\rho^c_{i,j}$ utilized in Eq. 2.4, we employ a moving average of the last $n = 6$ (time window of 1 m) available measurements in order to anticipate time intervals where we do not have available reports from connected vehicles.

$$\rho^c_{i,j}(k) = \sum_{l=0}^{n-1} \frac{\rho^c_{i,j}(k-l)}{n}. \tag{4.4}$$

Finally, lateral flow $L_{i,j_1 \to j_2}$ of connected vehicles is observed when a successive observation of a connected vehicle occurs referring to lane j_2 subsequent to lane j_1, within a time interval of 2 s; every time instant kT, lateral flow measurements are accumulated. Since the lateral flow vector of connected vehicles, namely L^c, may contain some spiky values, mainly due to the rare appearance of connected vehicles lateral movement, we employ in the estimation scheme a filtered (exponentially smoothed) version of the lateral flows of connected vehicles rather than the original measured lateral flows. Thus for each lateral flow measurement we define for all $i = 1, \ldots, N$, $j_1 = 1, \ldots, M$, and $j_2 = j_1 \pm 1$

$$\bar{L}^c_{i,j_1 \to j_2}(k+1) = (1-a)\bar{L}^c_{i,j_1 \to j_2}(k) + aL^c_{i,j_1 \to j_2}(k), \tag{4.5}$$

where the smoothing factor $a \in [0, 1]$ will be decided based on statistical analysis (see Fig. 4.10) and $\bar{L}^c_{i,j_1 \to j_2}(0) = 0$. The problem, at some time instants, of the density of connected vehicles being equal to zero in Eq. 2.4 is resolved by setting the term $\frac{L^c_{i,j_1 \to j_2}(k)}{\rho^c_{i,j_1}(k)}$, equal to zero.

Statistical analysis was also performed in order to evaluate the magnitude of the error related to the computation of lateral flows $L_{i,j_1 \to j_2}(k)$ (see Eq. 2.4) with respect to the corresponding ground truth lateral flows, for all possible combinations

of $j_1 = 1, \ldots, M$ and $j_2 = 1, \ldots, M$ and for all $i = 1, \ldots, N$, with simulation time horizon $K = \frac{4}{T} = 1440$. Note that $\rho^c_{i,j_1}(k)$ is computed as shown in Eq. 4.4 and $\rho_{i,j_1}(k)$, within this tests, is considered to be the ground truth density. The results are presented in Fig. 4.2 for various penetration rates of connected vehicles.

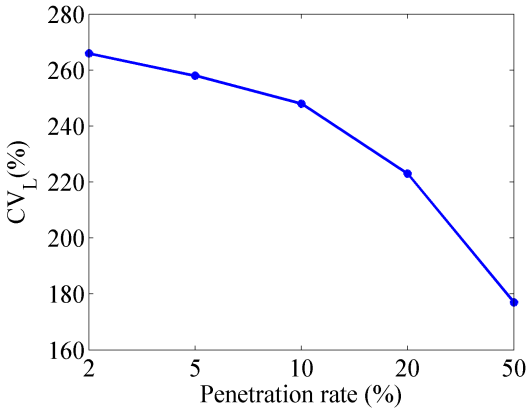


Figure 4.2: Performance comparison of the lateral flow computation via Eq. 2.4 for various penetration rate.

4.3 Mixed traffic simulation results in the presence of connected vehicles

4.3.1 Selection of the estimation scheme parameters

Prior to the operation of the filter, any tuning parameters involved should be measured. We will examine the sensitivity of the estimator with regard to variations of these parameters within a broad range of values for the case of our basic scenario, where we obtain 20% penetration rate of connected vehicles. To this end, we perform a series of experiments evaluating the sensitivity of the estimation scheme to the values of the filter parameters Q and R as shown in Fig. 4.4, where each of the

parameters σ_ρ (Q entries that correspond to density), $\sigma_{r,s}$ (Q entries that correspond to ramp flows), and R are varying within several orders of magnitude, while the other two remain constant. As it is evident for the plots, ramp flow estimation and density estimation for low values of σ_ρ is shown to be very sensitive.

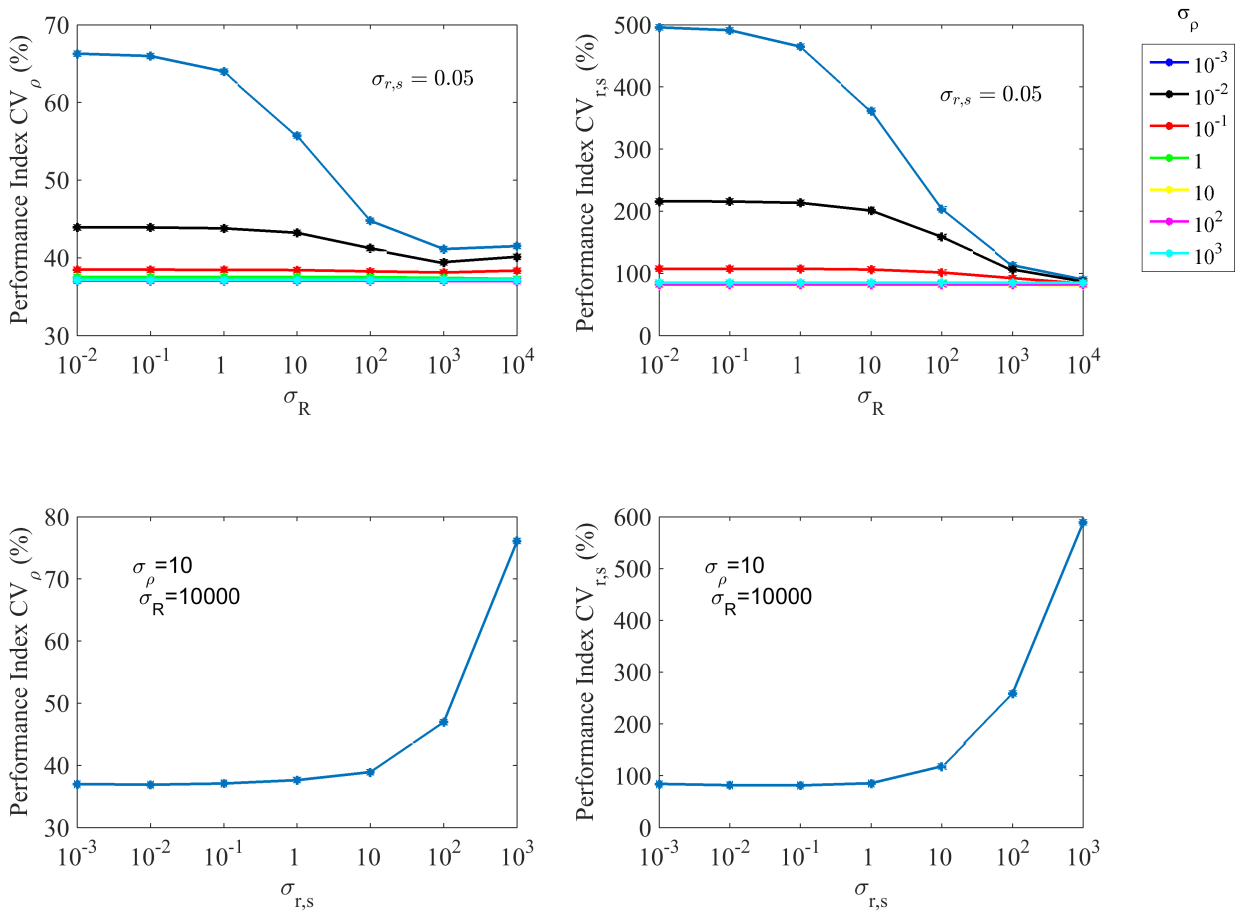


Figure 4.4: Performance comparison of the density and ramp flow estimations for different values of the parameters σ_ρ, σ_R (top) and $\sigma_{r,s}$ (bottom), for various penetration rates of connected vehicles, when the speed utilized by the estimator is calculated via (4.3) and $n = 12$.

The choice of the elements of μ and P_0 may range from infinitely large values to small values depending on the available information about the relevant states. It has further been observed that the choices of the initial values affect only the initial

part (warm-up phase) of the estimation so these are usually not very critical unless the initial estimation exceeds acceptable limits.

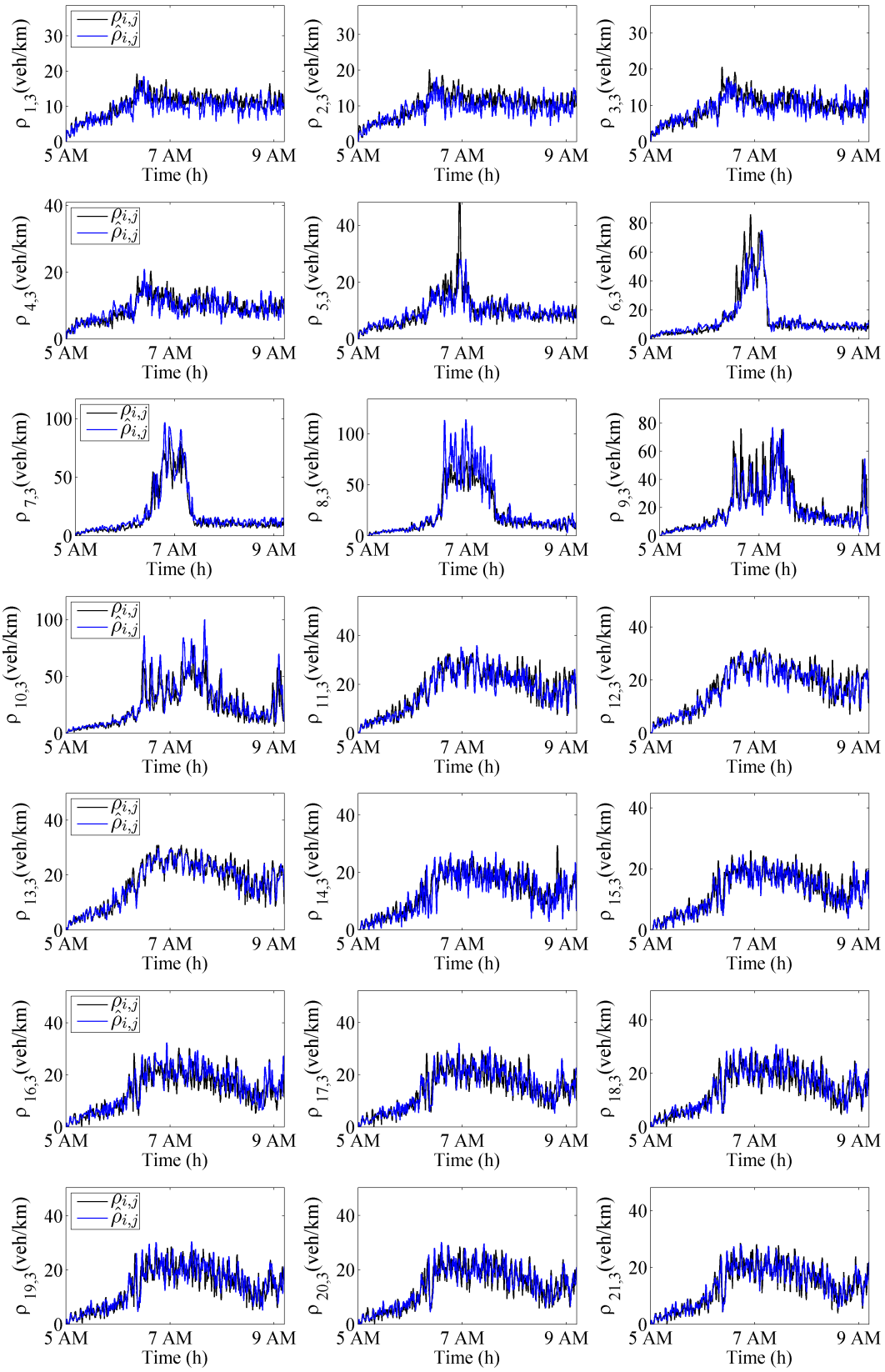
Q	σ_ρ	$\sigma_{r,s}$
$diag(\sigma_\rho, \sigma_{r,s})$	$10 \times I_{(N \times M)}$	$0.01 \times I_{((l_r+l_s)\times 1)}$
R	μ	H
$10^4 \times I_{(l_r+l_s)}$	$(5, \ldots, 5, 2, \ldots, 2)^T$	$I_{(N \times M + l_r + l_s)}$

Table 4.1: Filter parameters used in the simulation

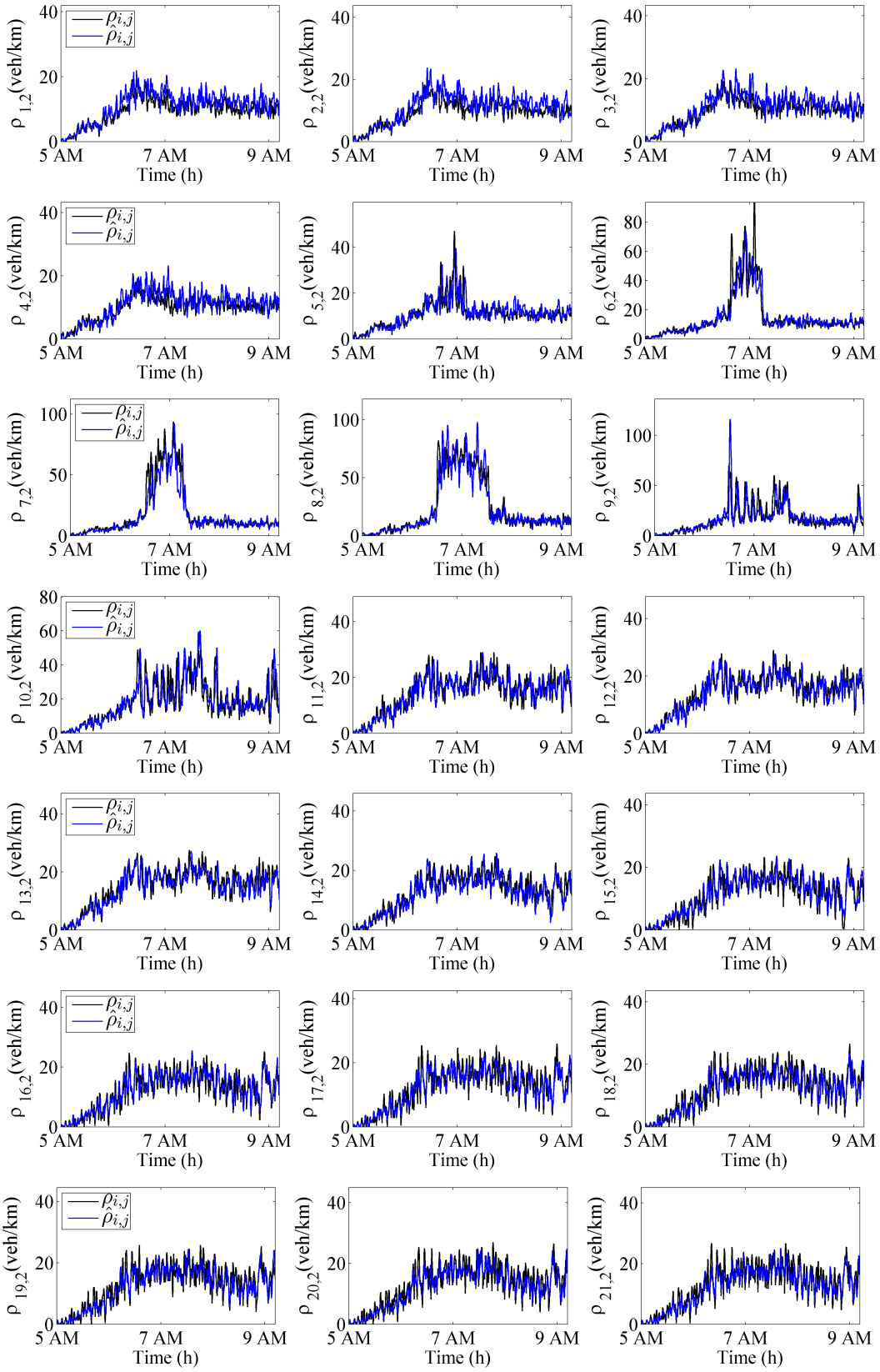
4.3.2 Performance evaluation for varying penetration rates of connected vehicles

For the sake of brevity, only the results for segment density for all lanes and ramp flow estimation related to one of these tests, namely with a 20% penetration rate of connected vehicles, is illustrated in Fig 4.5 and 4.6 respectively; the speed utilized by the filter is calculated via Eq. 4.3 with $n = 12$, lateral flow of connected vehicles is computed as in Eq. 4.5 with smoothing factor $a = 0.05$, density of connected vehicles is computes as shown in Eq. 4.4 with $n = 6$. Finally the diagonal factor p is chosen to be equal to 0.3, based on statistical analysis as shown in Section 4.3.3.

Lane 3


$\rho_{i,j}$
$\hat{\rho}_{i,j}$
ρ 1,3(veh/km)
ρ 2,3(veh/km)
ρ 3,3(veh/km)
ρ 4,3(veh/km)
ρ 5,3(veh/km)
ρ 6,3(veh/km)
ρ 7,3(veh/km)
ρ 8,3(veh/km)
ρ 9,3(veh/km)
ρ 10,3(veh/km)
ρ 11,3(veh/km)
ρ 12,3(veh/km)
ρ 13,3(veh/km)
ρ 14,3(veh/km)
ρ 15,3(veh/km)
ρ 16,3(veh/km)
ρ 17,3(veh/km)
ρ 18,3(veh/km)
ρ 19,3(veh/km)
ρ 20,3(veh/km)
ρ 21,3(veh/km)
5 AM
7 AM
9 AM
Time (h)

Lane 2


ρ 1,2(veh/km)
ρ 2,2(veh/km)
ρ 3,2(veh/km)
ρ 4,2(veh/km)
ρ 5,2(veh/km)
ρ 6,2(veh/km)
ρ 7,2(veh/km)
ρ 8,2(veh/km)
ρ 9,2(veh/km)
ρ 10,2(veh/km)
ρ 11,2(veh/km)
ρ 12,2(veh/km)
ρ 13,2(veh/km)
ρ 14,2(veh/km)
ρ 15,2(veh/km)
ρ 16,2(veh/km)
ρ 17,2(veh/km)
ρ 18,2(veh/km)
ρ 19,2(veh/km)
ρ 20,2(veh/km)
ρ 21,2(veh/km)
$\rho_{i,j}$
$\hat{\rho}_{i,j}$
5 AM
7 AM
9 AM
Time (h)

Lane 1

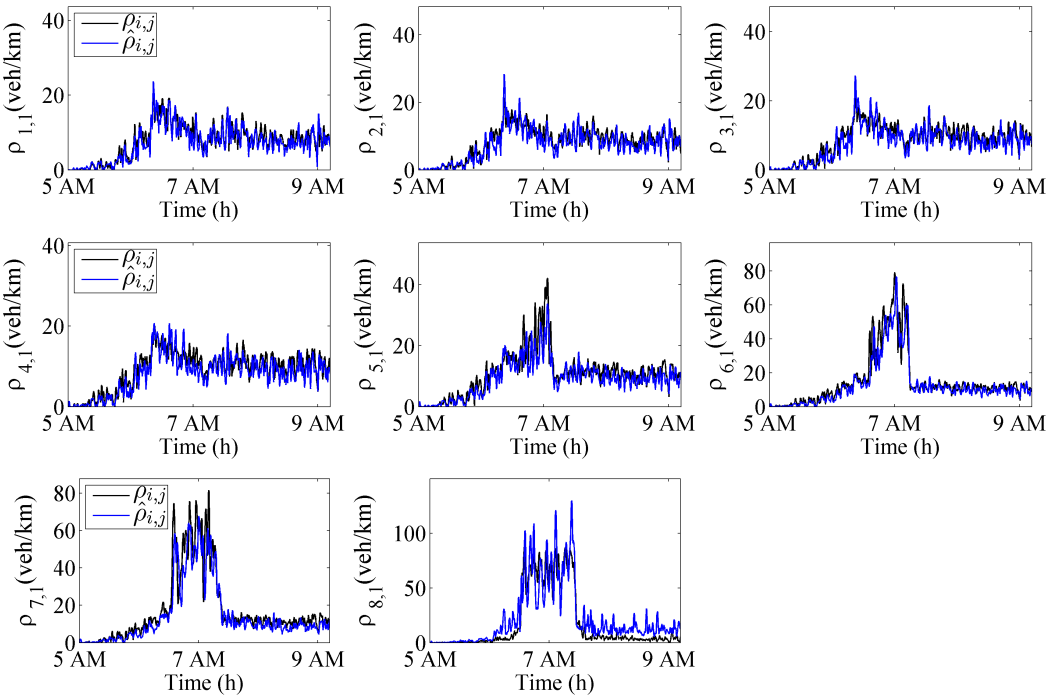


Figure 4.5: Comparison between real (black line) and estimated (blue line) densities in veh/km for all network segments for mixed traffic with a 20% penetration rate of connected vehicles.

Ramp flows

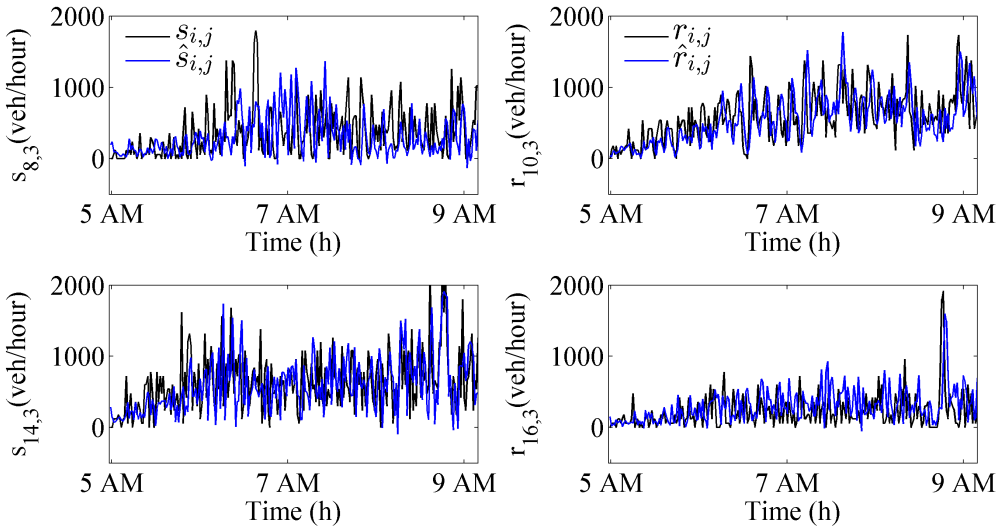


Figure 4.6: Comparison between real (black line) and estimated (blue line) ramp flows in veh/hour for all on-ramps and off ramps in the network for mixed traffic with a 20% penetration rate of connected vehicles.

It is evident from the plots that the proposed scheme successfully estimates and is capable of following the dynamics of both segment densities and ramp flows under various traffic conditions, including congested and free-flow conditions and even for time intervals where we obtain no information from connected vehicles' reports. Flow estimation results in the case of off-ramp 8 are less satisfactory. It is plausible that the length of the deceleration lane of off-ramp Nieuwerkerk, which is almost half of the length of the corresponding mainstream segment, may have influenced the results obtained in flow estimation of this off-ramp. Density estimation is characterized by a performance index $CV_p = 34.4\%$, whereas ramp flow estimation is characterized by a performance index $CV_{r,s} = 78.0\%$.

Figure 4.7 illustrates the performance comparison for density and ramp flow estimation with regard to the different penetration rates of connected vehicles con-

sidered. A high sensitivity is observed in the performance of the estimation scheme with regard to the different penetration rates of connected vehicles present at the highway, which accounts for the inadequacy of traffic information available for a large proportion of time intervals; reaching up to 70% for 2% of penetration rate (see 4.1).

Model described in Section 2.2.1.3, assumes data collected from connected vehicles' reports, in addition to data collected from stationery flow detectors placed at mainstream segments along the highway to reconstruct the local density. A conventional flow detector is placed at the entrance and exit of the the considered highway, along with additional spot sensors placed at segments 9, 11, 15 in order to guarantee observability. The assumption that such sensors deployed are capable to deliver observed data both for flow and speed, is reasonable as well as valid in most freeway environments. For this reason, the entries of the output equation (2.16) that correspond to the average cell speed as computed from connected vehicles' reports may be replaced by the total average speed measured by the detector. Moreover,when there are available occupancy detectors in a cell, the density of this cell could be directly extracted. In this case, the corresponding entries of the C matrix in (2.16) would be equal to one. Note that in this case, the model needs re-tuning, as different order of magnitude for the parameters Q, R may be more appropriate, and thus we will not include this case in the results shown in Fig. 4.7. Density estimation is shown to be little sensitive to different configurations of the output matrix.

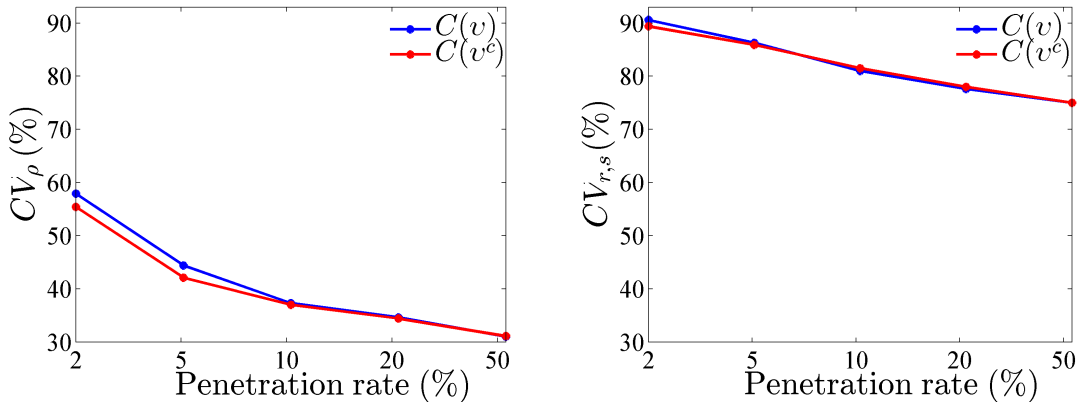


Figure 4.7: Performance comparison of density and ramp flow estimations for various penetration rates when mainstream speed measurements are available at the locations of flow detectors (blue line) and when only connected vehicles' speed measurements are available(red line).

4.3.3 Sensitivity analysis for the percentage of lane-drop diagonal flow

Experimental analysis was performed in order to obtain the optimal value of parameter p, prior to be employed in Eq. 2.6, in order to succeed an overall estimation improvement. Figure 4.8, shows the comparison of the performance of the estimation scheme for different values of parameter $p \in [0, 1]$, for various penetration rates, while Fig. 4.9 illustrates the impact of various values of diagonal factor p on the performance index that corresponds to cell 8 of lane 2, where the updated model for longitudinal flows is applied, for various penetration rates of connected vehicles. In all our experiments we use the original measurements of lateral flow ($a = 1$). From Fig. 4.8 it is evident that the overall estimation performance is quite insensitive to the different values of parameter p, whereas ramp flow estimation is more sensitive. With regard to the density estimation, a value $p = 0.3$ may best depict the diagonal

flows in the updated model for the longitudinal flows, for almost all penetration rates. Density estimation at cell 8 of lane 2, is more sensitive to changes of parameter p, as shown in Fig. 4.9. For almost all penetration rates the optimal value is $p = 0.3$, whereas for 2% penetration rate the optimal value is about $p = 0.5$. Thus, Fig. 4.8– 4.9, depict the need of imposing a non-zero value for p, especially for low penetration rates.

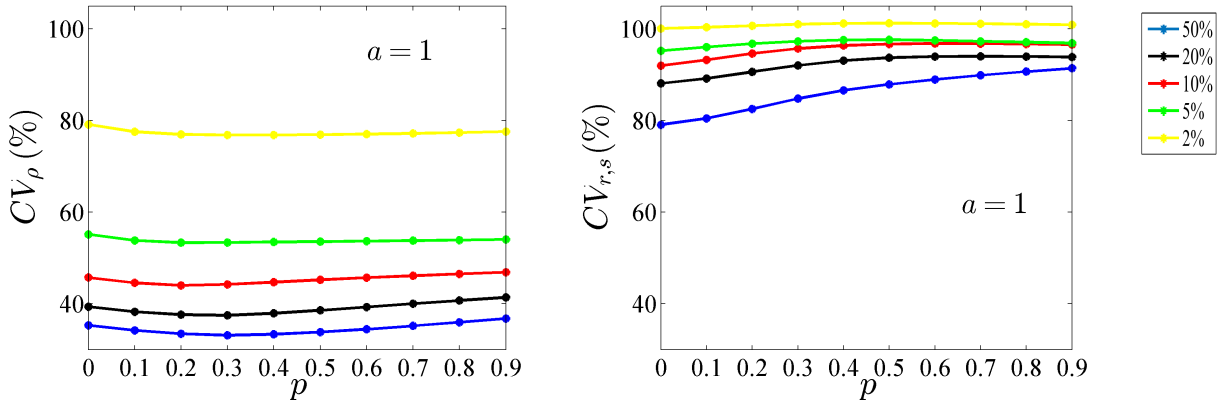


Figure 4.8: Performance comparison of the density and ramp flow estimations for different values of the parameter p and for various penetration rates of connected vehicles

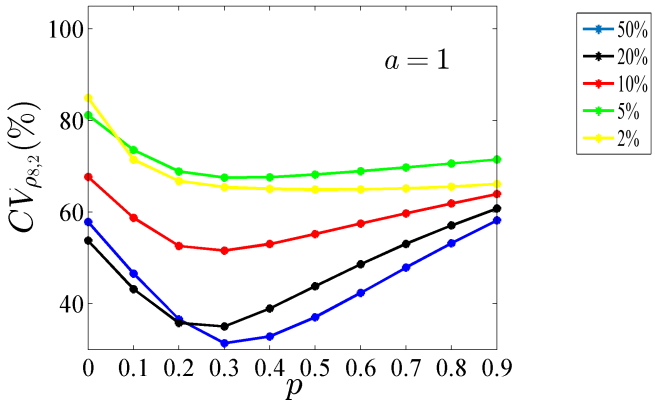


Figure 4.9: Performance comparison of the density of cell (8,2), where the lane-drop is located, for different values of the parameter p and for various penetration rates of connected vehicles.

4.3.4 Sensitivity analysis for the smoothing factor

Estimation of lateral flows is heavily based on measurements from connected vehicles. These measurements are prone to inconsistencies when we observe a rare appearance of connected vehicles. For low values of the smoothing factor a (strong smoothing) employed in Eq. 4.5, inconsistencies in measurements utilized for the estimation of lateral flows, are eliminated; However, in this case we may loose information from the time intervals where we have reliable measurements. Sensitivity of the estimation performance to the variations of the smoothing factor a is shown in Fig. 4.10, for $p = 0.3$ and for various penetration rates of connected vehicles. From Fig. 4.10 one can observe that at low penetration rates, the density estimation is more sensitive to different values of smoothing factor a and that the value of a at which the optimal performance of the model is obtained almost decreases with the penetration rate. This is consistent with the expectation that at lower penetration rates a stronger smoothing effect is needed, as the available density and lateral flow information from connected vehicles is sufficiently poor. The average value of a at which the minimum performance index is succeeded, for all penetration rates is about 0.05.

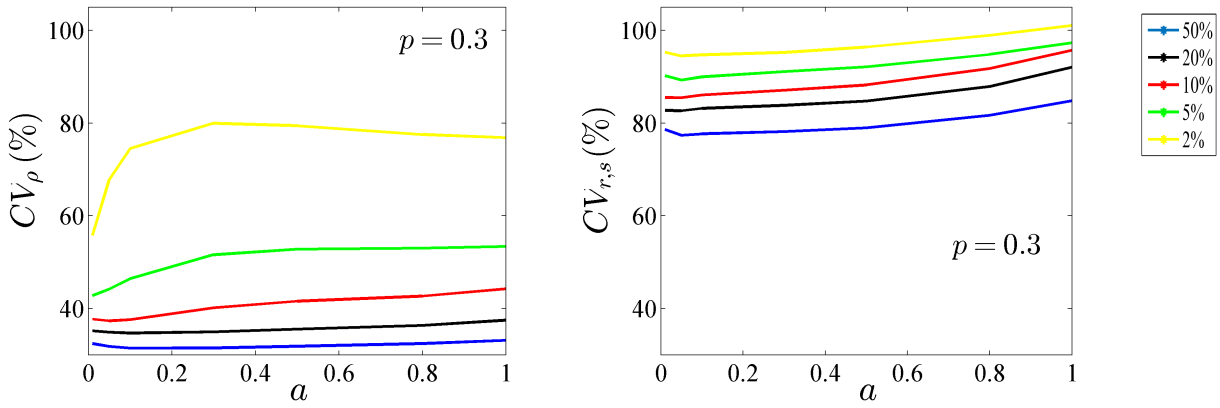


Figure 4.10: Performance comparison of the density and ramp flow estimations for different values of the smoothing factor a and for various penetration rates of connected vehicles.

Chapter 5

Conclusions

In this thesis, the problem of multi-lane traffic density as well as ramp flow estimation was addressed employing a Kalman filter that utilizes measurements retrieved from connected vehicles' reports and a limited amount of flow measurements from conventional detectors. The effectiveness of the proposed estimation scheme was thoroughly tested using the AIMSUN microscopic traffic simulator, which was calibrated with real traffic data in order to reproduce realistic traffic conditions. The test site network is a stretch of the motorway A20 from Rotterdam to Gouda in the Netherlands with two sets of on-off/ramps and a lane-drop, which constitutes it a challenging case-study. The employed simulation scenario featured 4-hour simulation horizon where both congested and free-flow conditions are observed. The complex infrastructure characteristics along with the mixed traffic conditions of the case-study network constitutes it a challenging test-bed.

The estimation performance of the proposed scheme is largely based on speed and position information obtained from connected vehicle's regular reports. The effectiveness of the methodology proposed, was examined both in in qualitative and quantitative terms. The obtained results demonstrated that the estimation scheme

captures the onset of congestion with accurate timing and at the correct location even for low penetration rates. Density estimation is satisfactory even for as low penetration rates as 2%. Finally, the estimation performance was demonstrated to be quite insensitive to the variations of the model parameters.

The developed approach has several advantages for possible future real-world applications, including the use of a limited amount of fixed sensors, and the use of connected vehicles data, which constitute a growing source of real-time traffic information. The estimated traffic density can facilitate lane-based traffic control and also provide lane advice and inputs for long-term transport planning.

Bibliography

Anand, A., G. Ramadurai, and L. Vanajakshi (2014). Data fusion-based traffic density estimation and prediction. *Journal of Intelligent Transportation Systems 18*(4), 367–378.

Bekiaris-Liberis, N., C. Roncoli, and M. Papageorgiou (2016). Highway traffic state estimation with mixed connected and conventional vehicles. *IEEE Transactions on Intelligent Transportation Systems 17*(12), 3484–3497.

Bekiaris-Liberis, N., C. Roncoli, and M. Papageorgiou (2017). Highway traffic state estimation per lane in the presence of connected vehicles. *Transportation Research Part B: Methodological (in press)*.

Bishop, R. (2005). *Intelligent Vehicle Technology and Trends*. Artech House Publishers, Norwood MA, USA.

Cayford, R. and T. Johnson (2003). Operational parameters affecting the use of anonymous cell phone tracking for generating traffic information. In *Institute of transportation studies for the 82th TRB Annual Meeting*, Volume 1, pp. 03–3865.

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

Chang, M.-F. and D. C. Gazis (1975). Traffic density estimation with consideration of lane-changing. *Transportation Science 9*(4), 308–320.

Chevallier, E. and L. Leclercq (2009). Do microscopic merging models reproduce the observed priority sharing ratio in congestion? *Transportation Research Part C: Emerging Technologies 17*(3), 328–336.

Coifman, B. (2003). Estimating density and lane inflow on a freeway segment. *Transportation Research Part A: Policy and Practice 37*(8), 689–701.

Dargay, J., D. Gately, and M. Sommer (2007). Vehicle ownership and income growth, worldwide: 1960-2030. *The Energy Journal*, 143–170.

De Fabritiis, C., R. Ragona, and G. Valenti (2008). Traffic estimation and prediction based on real time floating car data. In *Intelligent Transportation Systems, 2008. ITSC 2008. 11th International IEEE Conference on*, pp. 197–203. IEEE.

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

Diakaki, C., M. Papageorgiou, I. Papamichail, and I. Nikolos (2015). Overview and analysis of vehicle automation and communication systems from a motorway traffic management perspective. *Transportation Research Part A: Policy and Practice 75*, 147–165.

Fountoulakis, M., N. Bekiaris-Liberis, C. Roncoli, I. Papamichail, and M. Papageorgiou (2016). Highway traffic state estimation with mixed connected and conventional vehicles: Microscopic simulation-based testing. In *Intelligent Transport-*

ation Systems (ITSC), 2016 IEEE 19th International Conference on, pp. 1761–1766. IEEE.

Gipps, P. G. (1981). A behavioural car-following model for computer simulation. *Transportation Research Part B: Methodological 15*(2), 105–111.

Gipps, P. G. (1986). A model for the structure of lane-changing decisions. *Transportation Research Part B: Methodological 20*(5), 403–414.

Herrera, J. C., D. B. Work, R. Herring, X. J. Ban, Q. Jacobson, and A. M. Bayen (2010). Evaluation of traffic data obtained via gps-enabled mobile phones: The mobile century field experiment. *Transportation Research Part C: Emerging Technologies 18*(4), 568–583.

Liu, K., T. Yamamoto, and T. Morikawa (2006). An analysis of the cost efficiency of probe vehicle data at different transmission frequencies. *International Journal of ITS Research 4*(1), 21–28.

Messner, A. and M. Papageorgiou (1990). Metanet: A macroscopic simulation program for motorway networks. *Traffic Engineering & Control 31*(8-9), 466–470.

Nanthawichit, C., T. NAKATSUJI, and H. SUZUKI (2003). Dynamic estimation of traffic states on a freeway using probe vehicle data. *Doboku Gakkai Ronbunshu 2003*(730), 43–54.

Perraki, G. (2016). Evaluation of a model predictive control strategy on a calibrated multilane microscopic model. Master's thesis, School of Production Engineering and Management, Technical University of Crete.

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

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

Rahmani, M., H. N. Koutsopoulos, and A. Ranganathan (2010). Requirements and potential of gps-based floating car data for traffic management: Stockholm case study. In *Intelligent Transportation Systems (ITSC), 2010 13th International IEEE Conference on*, pp. 730–735. IEEE.

Rempe, F., P. Franeck, U. Fastenrath, and K. Bogenberger (2016). Online freeway traffic estimation with real floating car data. In *Intelligent Transportation Systems (ITSC), 2016 IEEE 19th International Conference on*, pp. 1838–1843. IEEE.

Roncoli, C., N. Bekiaris-Liberis, and M. Papageorgiou (2016). Highway traffic state estimation using speed measurements: case studies on ngsim data and highway a20 in the netherlands. *Transportation Research Record 2559*, 90–100.

Roncoli, C., I. Papamichail, and M. Papageorgiou (2014). Model predictive control for multi-lane motorways in presence of vacs. In *Intelligent Transportation Systems (ITSC), 2014 IEEE 17th International Conference on*, pp. 501–507. IEEE.

Schakel, W. J. and B. Van Arem (2014). Improving traffic flow efficiency by in-car advice on lane, speed, and headway. *IEEE Transactions on Intelligent Transportation Systems 15*(4), 1597–1606.

Schreiter, T., H. van Lint, M. Treiber, and S. Hoogendoorn (2010). Two fast implementations of the adaptive smoothing method used in highway traffic state estimation. In *Intelligent Transportation Systems (ITSC), 2010 13th International IEEE Conference on*, pp. 1202–1208. IEEE.

Seo, T. and T. Kusakabe (2015). Probe vehicle-based traffic flow estimation method without fundamental diagram. *Transportation Research Procedia 9*, 149–163.

Sheu, J.-B. (1999). A stochastic modeling approach to dynamic prediction of section-wide inter-lane and intra-lane traffic variables using point detector data. *Transportation Research Part A: Policy and Practice 33*(2), 79–100.

Singh, K. and B. Li (2012a). Discrete choice modelling for traffic densities with lane-change behaviour. *Procedia-Social and Behavioral Sciences 43*, 367–374.

Singh, K. and B. Li (2012b). Estimation of traffic densities for multilane roadways using a markov model approach. *IEEE Transactions on industrial electronics 59*(11), 4369–4376.

Stenneth, L. O. and G. A. Giurgiu (2016, May 31). Differentiation of probe reports based on quality. US Patent 9,355,560.

Tao, S., V. Manolopoulos, S. Rodriguez, A. Rusu, et al. (2012). Real-time urban traffic state estimation with a-gps mobile phones as probes. *Journal of Transportation Technologies 2*(01), 22.

Transport Simulation Systems (2014). *Aimsun 8 Dynamic Simulators Users' Manual*. Transport Simulation Systems.

Treiber, M., A. Hennecke, and D. Helbing (2000). Congested traffic states in empirical observations and microscopic simulations. *Physical review E 62*(2), 1805.

Treiber, M., A. Kesting, and R. E. Wilson (2011). Reconstructing the traffic state by fusion of heterogeneous data. *Computer-Aided Civil and Infrastructure Engineering 26*(6), 408–419.

Turksma, S. (2000). The various uses of floating car data.

van Hinsbergen, C. P., T. Schreiter, F. S. Zuurbier, J. Van Lint, and H. J. van Zuylen (2012). Localized extended kalman filter for scalable real-time traffic state estimation. *IEEE Transactions on Intelligent Transportation Systems 13*(1), 385–394.

Wang, J., R. Liu, and F. Montgomery (2005). Car-following model for motorway traffic. *Transportation Research Record: Journal of the Transportation Research Board* (1934), 33–42.

Wang, R., D. B. Work, and R. Sowers (2016). Multiple model particle filter for traffic estimation and incident detection. *IEEE Transactions on Intelligent Transportation Systems 17*(12), 3461–3470.

Waterson, B. and S. Box (2012). Quantifying the impact of probe vehicle localisation data errors on signalised junction control. *IET Intelligent Transport Systems 6*(2), 197–203.

Work, D. B., O.-P. Tossavainen, S. Blandin, A. M. Bayen, T. Iwuchukwu, and K. Tracton (2008). An ensemble kalman filtering approach to highway traffic estimation using gps enabled mobile devices. In *Decision and Control, 2008. CDC 2008. 47th IEEE Conference on*, pp. 5062–5068. IEEE.

Wright, M. and R. Horowitz (2016). Fusing loop and gps probe measurements

to estimate freeway density. *IEEE Transactions on Intelligent Transportation Systems 17*(12), 3577–3590.

Yılan, M. (2016). Multilane traffic density estimation with kde and nonlinear ls and tracking with scalar kalman filtering. Master's thesis.

Yuan, Y., J. Van Lint, R. E. Wilson, F. van Wageningen-Kessels, and S. P. Hoogendoorn (2012). Real-time lagrangian traffic state estimator for freeways. *IEEE Transactions on Intelligent Transportation Systems 13*(1), 59–70.

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

Zito, R., G. D'este, and M. Taylor (1995). Global positioning systems in the time domain: How useful a tool for intelligent vehicle-highway systems? *Transportation Research Part C: Emerging Technologies 3*(4), 193–209.