Transportation Research Part B 114 (2018) 281–299

Contents lists available at ScienceDirect

Transportation Research Part B

journal homepage: www.elsevier.com/locate/trb

Macroscopic traffic state estimation using relative flows from stationary and moving observers

Paul B.C. van Erp*, Victor L. Knoop, Serge P. Hoogendoorn

Department of Transport & Planning, Faculty of Civil Engineering and Geosciences, Delft University of Technology, Stevinweg 1 Delft, 2628 CN, The Netherlands

ARTICLE INFO

Article history:
Received 16 March 2018
Revised 7 June 2018
Accepted 8 June 2018

Keywords:
Traffic state estimation
Cumulative vehicle number
Relative flow data

ABSTRACT

This article presents a procedure to estimate the macroscopic traffic state in a pre-defined space-time mesh using relative flow data collected by stationary and moving observers. The procedure consist of two consecutive and independent processes: (1) estimate point observations of the cumulative vehicle number in space-time, i.e., $N(x, t)$, based on relative flow data from the observers and (2) estimate flow and density in a pre-define space-time mesh based on the point observations of N. In this paper, the principles behind the first process are explained and a methodology (the Point-Observations N (PON) estimation methodology) is introduced for the second process. This methodology does not incorporate information in the form of a traffic flow model or historical data. To evaluate this performance and improve our understanding of the methodology, a microscopic simulation study is conducted. The estimation performance is effected by the homogeneity and stationarity of traffic in estimation area and in the sample area. In case of large changes in traffic conditions, e.g., from free-flow to congestion or stop-and-go waves, a better sampling resolution will improve localizing these changes in space and time and hence improve the estimation performance. In the simulation study, the proposed methodology is also compared with estimates based on loop-detector data. This indicates that the combination of the proposed methodology and data yields an alternative for existing combinations of methodology and data. Especially, in terms of density estimation the introduced methodology shows promising results.

1. Introduction

This paper addresses macroscopic traffic state estimation. Estimates of the macroscopic traffic flow variables, i.e., flow q, density k and speed u, can be used as input for control decisions within dynamic traffic management applications (Papageorgiou et al., 1991; Smaragdis et al., 2004).

The estimation procedure introduced in this paper allows us to estimate the macroscopic traffic flow variables within a pre-defined space-time mesh using stationary and moving observers. This procedure consists of two main (independent and consecutive) processes. These are: (1) estimate the cumulative vehicle number N for points along the observed paths in

* Corresponding author.
E-mail addresses: p.b.c.vanerp@tudelft.nl (P.B.C. van Erp), v.l.knoop@tudelft.nl (V.L. Knoop), s.p.hoogendoorn@tudelft.nl (S.P. Hoogendoorn).

https://doi.org/10.1016/j.trb.2018.06.005
0191-2615/

space-time (we will call these point-observations) using traffic sensing data from stationary and moving observers and (2) estimate q and k in a pre-defined mesh based on point-observations of N.

This paper has two important contributions. The first and more generic contribution is the full estimation procedure. We propose to use equipped and/or automated vehicles that observe the relative flow with respect to their trajectory in combination with stationary observers. The traffic sensing data from these observers can be fused on the cumulative vehicle number level (first process) and can be used to estimate the macroscopic traffic conditions (second process). Both processes are explained in this paper. The second and more specific contribution is the methodology designed for the second process. This methodology is called the Point-Observation N (PON) estimation methodology. The two processes require independent methodologies and should in our view not be discussed in detail in a single paper. Therefore, in this paper, we explain the second process in detail, while we solely explain the principles behind the first process.

This paper is organized as follows. Section 2 provides the existing methodological basis for our work and positions this work within the research field. Next, we explain the two processes in Sections 3 and 4. The PON estimation methodology (explained in Section 4) is evaluated using a simulation study in Sections 5 and 6. We conclude with the conclusions and discussion in Section 7.

2. Background on macroscopic traffic state estimation

In this section, we discuss the topic of macroscopic traffic state estimation and explain how the proposed methodology differs from existing work. First, Edie's generalized definitions of the macroscopic traffic flow variables and the three-dimensional representation of traffic flow are provided. Second, the categorization discussed by Seo et al. (2017) is used to categorize the proposed estimation approach. Third, we elaborate on different types of traffic sensing data used for macroscopic traffic state estimation and which data are used in the proposed estimation procedure. And finally, we discuss our estimation output and how this relates to existing work.

The generalized definitions of flow q, density k and speed u, for an area $\mathcal{D}$ in space-time are provided by Edie (1965):

$$q_{\mathcal{D}} = \frac{\sum_i d_i}{A_{\mathcal{D}}} \tag{1}$$

$$k_{\mathcal{D}} = \frac{\sum_i r_i}{A_{\mathcal{D}}} \tag{2}$$

$$u_{\mathcal{D}} = \frac{q_{\mathcal{D}}}{k_{\mathcal{D}}} \tag{3}$$

where d_i and r_i respectively denote the distance traveled and time spent by vehicle i within the area $\mathcal{D}$ and $A_{\mathcal{D}}$ denotes the surface of $\mathcal{D}$. $\Sigma_i d_i$ and $\Sigma_i r_i$ respectively denote the Total Travel Distance (*TTD*) and Total Time Spent (*TTS*) in $\mathcal{D}$.

Makigami et al. (1971) proposed the three-dimensional representation of traffic flow. The three dimensions are space, time and the cumulative vehicle number, where $N(x, t)$ denotes the cumulative vehicle number at location x and time instant t. As vehicles are discrete, N can be represented as a discrete variable. Here, $N(x, t)$ increases instantly by one vehicle at the time instant t when a vehicle passes location x.

We want to describe traffic flow on a macroscopic level. For this purpose, the discrete N can be smoothed (Makigami et al., 1971). For the smoothed and continuously differentiable N, the macroscopic traffic flow variables can be described based on the three dimensions. The macroscopic variables for a point in space-time, i.e., (x, t), are given by the time and space derivatives of $N(x, t)$:

$$q(x,t) = \frac{\partial N(x,t)}{\partial t} \tag{4}$$

$$k(x,t) = -\frac{\partial N(x,t)}{\partial x} \tag{5}$$

$$u(x,t) = \frac{q(x,t)}{k(x,t)} \tag{6}$$

Seo et al. (2017) categorizes the estimation approach into three categories (i.e., model-driven, data-driven and streaming-data-driven) based on information input. Following this categorization, the methodology presented in this study can be categorized as a streaming-data-driven traffic state estimation methodology. Examples of other streaming-data-driven methodologies are Wardrop and Charlesworth (1954), Seo and Kusakabe (2015) and Florin and Olariu (2017). A streaming-data-driven methodology does not depend on information in the form of a traffic flow model, fundamental diagram or historical data, but solely relies on real-time data and 'weak' assumptions such as the conservation-of-vehicles. Therefore, 'it is robust against uncertain phenomena and unpredictable incidents' (Seo et al., 2017). At the same time, Seo et al. (2017) denotes two limitations of streaming-data-driven methodologies: (1) additional information (e.g., a traffic flow model) is needed to

predict the (future) traffic state and (2) massive streaming data may be needed to obtain accurate estimates. However, the combination of sensing data and estimation methodology presented in this study can accurately estimate the traffic state without needing massive streaming data. Although we cannot solve the former limitation, the methodology results in estimates that can be used as inputs for a traffic flow model, e.g., the Cell Transmission Model (CTM) (Daganzo, 1994; 1995). Furthermore, the methodology may be extended by adding information in the form of a traffic flow model. However, in contrast to methodologies such as Newell's three-detector method (Newell, 1993a; 1993b; 1993c; Laval et al., 2012) and Claudel's approach (Claudel and Bayen, 2010a; 2010b), it does not need this information.

The fundamental input of traffic state estimation is traffic sensing data. These data are collected using sensing equipment and observe specific traffic variables for specific space-time properties. In existing work, different types of sensing data are used for macroscopic traffic state estimation. Treiber and Helbing (2002) and Wang and Papageorgiou (2005) use detectors to observe macroscopic variables, i.e., mean speed and flow. The speed can be aggregated over time (as is done by Dutch detectors), thereby yielding a time-mean speed, which differs from the desired space-mean speed. As a result, estimating the density based on the aggregated detector variables is biased (Knoop et al., 2009). Alternatively, vehicle-based sensing equipment can be used to observe microscopic traffic variables, i.e., individual vehicle speed (Nanthawichit et al., 2003; Herrera and Bayen, 2010), space-headway (Seo and Kusakabe, 2015). These data have a potential bias-problem (Seo et al., 2017) as equipped vehicles can have a largely consistent different driving behavior than the average road user. For instance, we may collect speed data for a fleet of trucks, which will a lower speed than average. Furthermore, we may collect headway data from vehicles equipped with headway sensors and automated driving technologies, which may act different than manual vehicles.

The estimation procedure proposed in this study solely uses observations of ΔN, i.e., flow data from stationary observers and relative flow data from moving observers. Stationary sensing-equipement that observe flow (e.g., detectors or cameras) are widely applied in practice (e.g., PeMS). As proposed by Redmill et al. (2011) and Florin and Olariu (2017), equipped (e.g., automated) vehicles can serve as moving observers that record the flow relative to their position over time. These vehicles observe ΔN along their trajectory. Earlier studies (e.g., Claudel and Bayen, 2010a) have used trajectory data of vehicles that cannot observe the relative flow in the same manner. In this case, they often assume that $\Delta N = 0$, i.e., no overtaking, along the vehicle trajectory. Although this assumption is expected to hold for single-lane traffic, it is likely to violated in multi-lane traffic. Therefore, these methodologies could also benefit from using moving observers that observe ΔN.

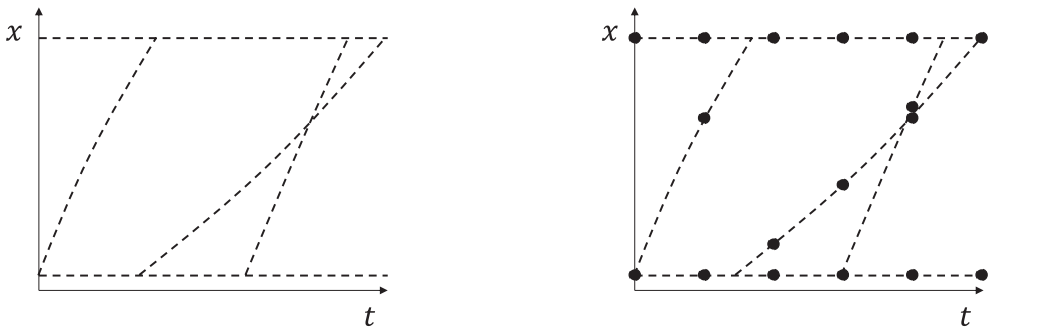
In terms of estimation output, we can differentiate various types of estimations. These can differ both in the variable types that are estimated as the spatial-temporal characteristics related to the estimates. The estimation methodology proposed in this study estimates the flow and density in a pre-defined space-time mesh. A potential space-time mesh is a discritisation of space in road-segments (cells) and time in periods, e.g., Nanthawichit et al. (2003), Wang and Papageorgiou (2005) and Herrera and Bayen (2010). In contrast to these methodologies, our methodology is free to work with any other pre-defined mesh. In the remainder of this paper, our focus lies on the two direct output of the methodology, i.e., flow and density. However, as we can obtain the speed from the flow and density, all three macroscopic variables can be estimated. Other methodologies exist that only estimate the speed using data from individual vehicles (e.g., Work et al., 2010; Del Arco et al., 2011). Newell's three-detector method (Newell, 1993a; 1993b; 1993c; Laval et al., 2012) and Claudel's approach (Claudel and Bayen, 2010a; 2010b) estimate the cumulative vehicle number for different points in space-time. Although this is related to our estimation output, comparing the outputs would require an additional step for one of the estimation methodologies.

3. Point-observations of the cumulative vehicle number

We combine traffic sensing data from stationary and moving observers to estimate N for points in space-time, i.e., $N(x, t)$. In this section, we explain the principles behind this process.

We define observations paths as paths in space-time over which we observe ΔN. Fig. 1 shows a combination of five observation paths. For stationary observers (dashed lines), the observation path is a horizontal line in space-time, i.e., a fixed location over time is observed. For moving observers (solid lines), the observation path is the trajectory of the connected automated vehicle. The relative flow observed between points along these boundaries is the change in N between these points. Individual observers thus provide the relative N along a single boundary. However, we are interested in having the relative N for all combinations of point-observations in space-time.

Combining sensing data from stationary and moving observers allows us to relate the data from different observers and to deal with potential observation errors. As stationary and moving observers move with different speeds through space-time, the observation paths intersect. Observation paths that have not yet been initialized (e.g., a moving observer entering the link) can be initialized based on their first interaction (e.g., with a stationary observer at the upstream link boundary). An intersection of two already-initialized observation paths can be used for error correction. Observation errors can lead to a discrepancy between N at the intersection point on the two observation paths. As both paths should have the same N-value at the intersection point, we can correct for observation errors based on the difference in N. This principle is simple; however, in designing an error correction methodology, we need to define how the difference in N translates in to a correction in ΔN over the observation paths. As explained in the introduction, designing this methodology is not the focus of this paper. However, we believe it is valuable to show that the full estimation procedure is robust to observation errors.



(a) Observation boundaries, stationary sensing (dashed lines) and moving observations (solid lines)

(b) Point-observations along the observation boundaries

Fig. 1. Visualization of observation boundaries and point-observations along these boundaries.

For this purpose, we design a simple methodology and evaluate the estimation performance under different observation errors in Appendix A.

To apply the PON-estimation methodology (which will be explained in the next section), all point-observations of N have to be defined in a single framework. Therefore, it should be possible to travel from an observation path to any other observation path. This holds for the combination of observation paths presented in Fig. 1. Each moving observer is initialized by intersecting with the stationary observer at the upstream boundary and the moving observers cross all stationary observers. Therefore, we have a ΔN for each combination of two points that lie on an observation path.

4. The PON estimation methodology

Three steps are taken to estimate flow q and density k in a pre-defined mesh using point-observations of N. These are (1) subdivide the space-time domain in triangular areas, where the three corners are point-observations of N, i.e., $N(x, t)$, (2) estimate q and k for each triangular area $\mathcal{T}$, i.e., $q_{\mathcal{T}}$ and $k_{\mathcal{T}}$, based on the three point-observations and (3) estimate q and k for each area $\mathcal{D}$, i.e., $q_{\mathcal{D}}$ and $k_{\mathcal{D}}$, in a pre-defined mesh based on $q_{\mathcal{T}}$ and $k_{\mathcal{T}}$. In the former step, we use an existing methodology (Delaunay triangulation). The latter two steps are derived based on the three-dimensional representation of traffic flow (Makigami et al., 1971) and Edie's generalized definitions of the macroscopic traffic flow variables (Edie, 1965).

This section contains four sub-sections. Sections 4.1 and 4.4 respectively explain the first and third steps. Sections 4.2 and 4.3 both relate to the second step (estimating $q_{\mathcal{T}}$ and $k_{\mathcal{T}}$). In Section 4.2, we derive the equations used for this step based on the three-dimensional representation of traffic flow (Makigami et al., 1971). However, there is a subtle difference between the interpretation of the estimates obtained using the derived equations and estimates for a triangular area in space-time. In Section 4.3, we use Edie's generalized definitions (Edie, 1965) to show that we can use the equations to estimate $q_{\mathcal{T}}$ and $k_{\mathcal{T}}$ are quantify the related estimation errors.

4.1. Subdivide space-time into triangular areas

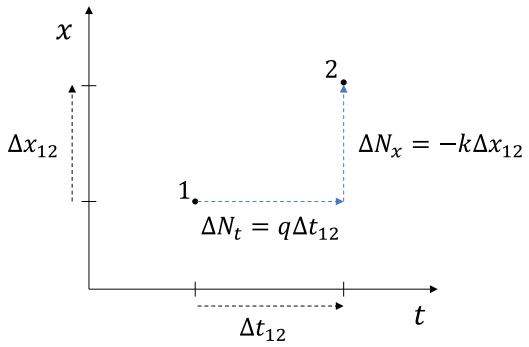
Given a set of point-observations, we want to subdivide the space-time domain into triangular areas, where the three corners are point-observations. We use Delaunay triangulation, which avoids sliver, i.e., narrow, triangular areas, to subdivide space-time based on coordinates of the corner points.

The triangular areas are defined on two different dimensions, i.e., space and time. In order to define three interior angles of triangular areas, we need to define the relation between distances in space and time. This yields the sole parameter of our traffic state estimator, i.e., the space-time ratio υ used in Delaunay triangulation. By defining υ we define a desired dimensions of the triangular areas. For instance, reducing υ yields triangular areas with a smaller space-time ratio, i.e., wider in terms of time and more sliver in terms of space. Therefore, υ influences which information is used to estimate the traffic conditions within the individual areas of the pre-defined mesh.

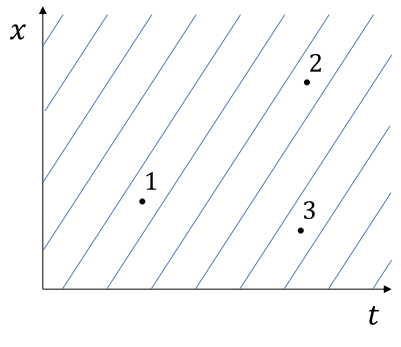
4.2. Three-point traffic state estimation

We propose equations to estimate the macroscopic traffic conditions based on three point-observations of the cumulative vehicle number in space-time, i.e., $N(x, t)$. Let us consider two points in the space-time domain, i.e., (x_1, t_1) and (x_2, t_2), for which the difference in N, i.e., $\Delta N_{12} = N(x_2, t_2) - N(x_1, t_1)$ is known.

Fig. 2a shows two points in space-time and the considered path between these points. Using (4) and (5) a relation between q and k based on ΔN_{12}, Δx_{12} and Δt_{12} is derived. To derive the relation, homogeneous and stationary traffic



(a) Two point observations in space-time and the considered path between these points



(b) Estimated homogeneous and stationary traffic conditions

Fig. 2. Visualizations related to three-point traffic state estimation.

conditions are assumed, i.e., $q(x,t) = q$ and $k(x,t) = k$ for any $(x,\ t)$ in the space-time domain. Therefore, over space and time the changes in N are respectively equal to $-k\Delta x_{12}$ and $q\Delta t_{12}$. This yields:

$$\Delta N_{12} = q\Delta t_{12} - k\Delta x_{12} \tag{7}$$

Note that any path between the two points will yield the same equation for ΔN_{12} as we assume homogeneous and stationary traffic conditions. Re-arranging (7) allow us to write q as a function of k and vice versa:

$$q = \frac{\Delta N_{12}}{\Delta t_{12}} + \frac{\Delta x_{12}}{\Delta t_{12}} k \tag{8}$$

$$k = -\frac{\Delta N_{12}}{\Delta x_{12}} + \frac{\Delta t_{12}}{\Delta x_{12}} q \tag{9}$$

An extra (third) point is added to estimate q and k as a function of Δx, Δt and ΔN. In combination with the two initial points, the third point provides ΔN_{13}, Δx_{13}, Δt_{13} and ΔN_{23}, Δx_{23}, Δt_{23}. To incorporate the information of all three points, it suffices to consider two combination of points. In this case, we consider combinations 12 and 23. For both combinations, the relations (8) and (9) are obtained. Below, we show the steps taken to derive the equation for q, i.e., (14):

$$q = \frac{\Delta N_{12}}{\Delta t_{12}} + \frac{\Delta x_{12}}{\Delta t_{12}} \left(-\frac{\Delta N_{23}}{\Delta x_{23}} + \frac{\Delta t_{23}}{\Delta x_{23}} q \right) \tag{10}$$

$$q\left(1 - \frac{\Delta x_{12}}{\Delta t_{12}} \frac{\Delta t_{23}}{\Delta x_{23}}\right) = \frac{\Delta N_{12}}{\Delta t_{12}} - \frac{\Delta N_{23}}{\Delta x_{23}} \frac{\Delta x_{12}}{\Delta t_{12}} \tag{11}$$

$$q\left(\frac{\Delta t_{12}}{\Delta t_{12}} \frac{\Delta x_{23}}{\Delta x_{23}} - \frac{\Delta x_{12}}{\Delta t_{12}} \frac{\Delta t_{23}}{\Delta x_{23}}\right) = \frac{\Delta N_{12}}{\Delta t_{12}} \frac{\Delta x_{23}}{\Delta x_{23}} - \frac{\Delta N_{23}}{\Delta x_{23}} \frac{\Delta x_{12}}{\Delta t_{12}} \tag{12}$$

$$q(\Delta t_{12}\Delta x_{23} - \Delta t_{23}\Delta x_{12}) = \Delta N_{12}\Delta x_{23} - \Delta N_{23}\Delta x_{12} \tag{13}$$

$$q = \frac{1}{\Delta t_{12}\Delta x_{23} - \Delta t_{23}\Delta x_{12}} (\Delta N_{12}\Delta x_{23} - \Delta N_{23}\Delta x_{12}) \tag{14}$$

Similar steps can be taken to derive the equation for k, which results in the following equation:

$$k = \frac{1}{\Delta t_{12}\Delta x_{23} - \Delta t_{23}\Delta x_{12}} (\Delta N_{12}\Delta t_{23} - \Delta N_{23}\Delta t_{12}) \tag{15}$$

The resulting q and k can be interpreted as the homogeneous and stationary traffic conditions that satisfy the cumulative vehicle number for three points in the space-time domain. Fig. 2b provides a visual interpretation of the estimate homogeneous and stationary conditions. Here, the blue lines denote trajectories satisfying q and k.

If the conditions are indeed homogeneous and stationary, any three observations of N are sufficient to estimate the traffic conditions for the full space-time domain. However, one conditions has to hold. The three points should not lie on a straight line in space-time, i.e., $\Delta x_{12}/\Delta t_{12} \neq \Delta x_{23}/\Delta t_{23}$. If $\Delta x_{12}/\Delta t_{12} = \Delta x_{23}/\Delta t_{23}$ the denominator will be zero. In this case the third point does not provide any additional information. It is furthermore important to note that (14) and (15) are invariant to the numbering of the three points, i.e., the same results are obtained if the same three points are numbered differently.

4.3. Space-time area traffic state estimation based on the mean boundary conditions

We consider Edie's generalized definitions (Edie, 1965) to relate the equations derived above to a triangular area in space-time and explain the related estimation errors. Edie's generalized definitions (1) and (2) can be rewritten on an aggregated level:

$$q = \frac{I\bar{d}}{A} \tag{16}$$

$$k = \frac{I\bar{r}}{A} \tag{17}$$

where $\bar{d}$ and $\bar{r}$ are respectively the mean distance traveled and time spent by the I number of vehicles. $\bar{d}$ and $\bar{r}$ can be described as function of the mean entry $(\bar{x}_{\text{in}}, \bar{t}_{\text{in}})$ and exit $(\bar{x}_{\text{out}}, \bar{t}_{\text{out}})$ points, i.e., $\bar{d} = \bar{x}_{\text{out}} - \bar{x}_{\text{in}}$ and $\bar{r} = \bar{t}_{\text{out}} - \bar{t}_{\text{in}}$.

Let us consider areas in the space-time domain with straight boundaries, e.g., a triangle or pentagon. Each boundary b has a net flow ΔN_b. Here outflow and inflow are respectively denoted by positive and negative values of ΔN_b. Furthermore, each b has a mean intersection point $(\bar{x}_b, \bar{t}_b)$, which reflects the mean point in space-time where vehicles intersect with the boundary. As we consider straight boundaries, this point lies on the boundary. In the case that traffic conditions are homogeneous and stationary and N is considered continuous, $(\bar{x}_b, \bar{t}_b)$ lie at the middle point of boundary b.

For the specific case of having an area with a finite number of straight boundaries, (16) and (17) can rewritten as a function of the boundary conditions:

$$q = \frac{\sum_b \Delta N_b \bar{x}_b}{A} \tag{18}$$

$$k = \frac{\sum_b \Delta N_b \bar{t}_b}{A} \tag{19}$$

Let us consider a triangular area for which all three ΔN_b are known. This resembles the case presented in Section 4.2, in which we know ΔN for three combinations of two points in space-time. Drawing a straight line (boundary) between the combinations of points in space-time results in a triangle.

Theorem 4.1. *For triangular areas in which the mean intersection of each boundary coincides with the middle point of the boundary, (18) and (19) are equal to respectively (14) and (15).*

Proof. We prove that (18) and (19) are equal to respectively (14) and (15) in the case a triangular area is considered and the mean intersection of each boundary coincides with the middle point of the boundary, i.e.,:

$$\frac{\sum_b \Delta N_b \bar{x}_b}{A} = \frac{1}{\Delta t_{12}\Delta x_{23} - \Delta t_{23}\Delta x_{12}}(\Delta N_{12}\Delta x_{23} - \Delta N_{23}\Delta x_{12}) \tag{20}$$

$$\frac{\sum_b \Delta N_b \bar{t}_b}{A} = \frac{1}{\Delta t_{12}\Delta x_{23} - \Delta t_{23}\Delta x_{12}}(\Delta N_{12}\Delta t_{23} - \Delta N_{23}\Delta t_{12}) \tag{21}$$

For triangular areas we have three corner points, i.e., 1, 2 and 3, and three boundaries, i.e., 12, 23 and 31. For this situation, the left part of the equation can be rewritten as

$$\frac{\sum_b \Delta N_b \bar{x}_b}{A} = \frac{1}{A}(\Delta N_{12}\bar{x}_{12} + \Delta N_{23}\bar{x}_{23} + \Delta N_{31}\bar{x}_{31}) \tag{22}$$

where the conservations of vehicles condition depicts that $\Delta N_{12} + \Delta N_{23} + \Delta N_{31} = 0$, thus $\Delta N_{31} = -(\Delta N_{12} + \Delta N_{23})$. Furthermore, for case in which the mean intersection of each boundary coincides with the middle point of the boundary, $\bar{x}_b$ can be described as a function of the corner points, e.g., $\bar{x}_{12} = (x_1 + x_2)/2$. This allows us to rewrite (22):

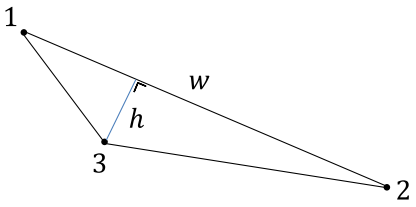
$$\frac{\sum_b \Delta N_b \bar{x}_b}{A} = \frac{1}{A}\left(\Delta N_{12}\frac{x_1 + x_2}{2} + \Delta N_{23}\frac{x_2 + x_3}{2} - (\Delta N_{12} + \Delta N_{23})\frac{x_1 + x_3}{2}\right) \tag{23}$$

$$= \frac{1}{A}\left(\Delta N_{12}\frac{x_2 - x_3}{2} + \Delta N_{23}\frac{x_2 - x_1}{2}\right) \tag{24}$$

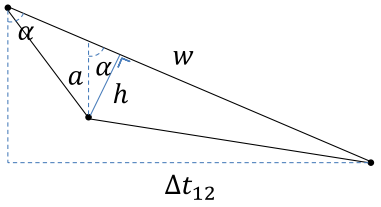
where $x_2 - x_3 = -\Delta x_{23}$ and $x_2 - x_1 = \Delta x_{12}$, thus:

$$\frac{\sum_b \Delta N_b \bar{x}_b}{A} = -\frac{1}{2A}(\Delta N_{12}\Delta x_{23} - \Delta N_{23}\Delta x_{12}) \tag{25}$$

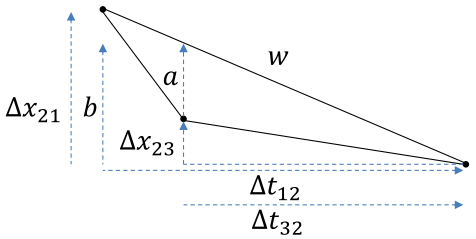
Next, we consider the surface of the triangular area, i.e., A. To calculate the surface of a triangular area we can use $A = \frac{1}{2}wh$, where w is the width and h is the height. Fig. 3a provides the triangular area and definitions of w and h which we consider to find an equation for A. Note that we want to describe A as a function of the boundaries 12 and 23 as these



(a) Height h and width w



(b) Relevant dimensions and angles to describe h



(c) Relevant dimensions to find a

Fig. 3. Important dimensions to find an equation for *A*.

boundaries are also considered for the right sides of Eqs. (20) and (21). Furthermore, note that there are multiple correct approaches to find the desired function for *A*.

To find h we first consider Fig. 3b. In this figure, we observe two triangles which have a common angle α. This allows us to relate the ratio of different known and unknown triangle sides:

$$\sin(\alpha) = \frac{h}{a} = \frac{\Delta t_{12}}{w} \tag{26}$$

Thus h is given by:

$$h = \frac{a\Delta t_{12}}{w} \tag{27}$$

This allows us to simplify the equation for *A*:

$$A = \frac{1}{2} a \Delta t_{12} \tag{28}$$

As Δt_{12} is known we only need to find a. Fig. 3c shows the relevant dimensions used to find a. This figure provides two basic relations, which we use to find a, namely:

$$a = b - \Delta x_{23} \tag{29}$$

$$\frac{b}{\Delta t_{32}} = \frac{\Delta x_{21}}{\Delta t_{12}} \tag{30}$$

where $\Delta t_{32} = -\Delta t_{23}$ and $\Delta x_{21} = -\Delta x_{12}$, thus:

$$b = \frac{\Delta t_{23} \Delta x_{12}}{\Delta t_{12}} \tag{31}$$

$$a = \frac{\Delta t_{23}\Delta x_{12}}{\Delta t_{12}} - \Delta x_{23} \tag{32}$$

This allows us to find A given (28):

$$A = \frac{1}{2}\left(\frac{\Delta t_{23}\Delta x_{12}}{\Delta t_{12}} - \Delta x_{23}\right)\Delta t_{12} \tag{33}$$

$$= \frac{1}{2}(\Delta t_{23}\Delta x_{12} - \Delta t_{12}\Delta x_{23}) \tag{34}$$

Substituting this equation into (25) yields:

$$\frac{\sum_b \Delta N_b \bar{x}_b}{A} = -\frac{1}{2A}(\Delta N_{12}\Delta x_{23} - \Delta N_{23}\Delta x_{12}) \tag{35}$$

$$= -\frac{1}{2 \cdot \frac{1}{2}(\Delta t_{23}\Delta x_{12} - \Delta x_{23}\Delta t_{12})}(\Delta N_{12}\Delta x_{23} - \Delta N_{23}\Delta x_{12}) \tag{36}$$

$$= \frac{1}{\Delta x_{23}\Delta t_{12} - \Delta t_{23}\Delta x_{12}}(\Delta N_{12}\Delta x_{23} - \Delta N_{23}\Delta x_{12}) \tag{37}$$

Similar steps can be taken to prove equality for k, i.e., (21).

□

The comparison made in Theorem 4.1 is important for two reasons. Firstly, it shows that (14) and (15) provide a proxy for the *TTS* and *TTD* within the triangular area enclosed by the three points. Secondly, it allows us to describe the estimation error in (14) and (15) based on the mean level of inhomogeneity and non-stationarity of the traffic conditions over the three boundaries. The mean level of inhomogeneity and non-stationarity of the traffic conditions over b affects $(\bar{x}_b, \bar{t}_b)$. The difference between this point and the middle point is denoted as $\Delta\bar{x}_b$ and $\Delta\bar{t}_b$. As both the mean intersection point and middle point lie on the boundary, the inhomogeneous and non-stationary traffic conditions can be described by a single variable, i.e., the fractional difference between the two points $\Delta\bar{\mu}_b$, where $\Delta\bar{x}_b = \Delta\bar{\mu}_b x_b$ and $\Delta\bar{t}_b = \Delta\bar{\mu}_b t_b$. This allows us to quantify the error in (14) and (15), i.e., ε_q and ε_k:

$$\varepsilon_q = \frac{\sum_b \Delta N_b \Delta\bar{\mu}_b x_b}{A} \tag{38}$$

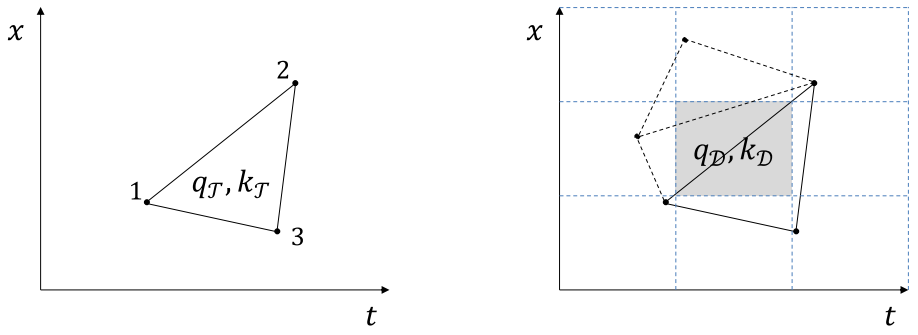
$$\varepsilon_k = \frac{\sum_b \Delta N_b \Delta\bar{\mu}_b t_b}{A} \tag{39}$$

These equations provide insight in the estimation error dependent on the area dimensions, i.e., A, x_b and t_b, and traffic conditions, i.e., N_b and $\Delta\bar{\mu}_b$. They show that the estimator estimates the traffic conditions perfectly, i.e., $\varepsilon_q = 0$ and $\varepsilon_k = 0$, in homogeneous and stationary conditions, i.e., $\Delta\bar{\mu}_b = 0$ for all b. Furthermore, the equations provide insight into the estimation errors of two triangular areas which share a boundary, i.e., adjacent triangular areas. Inhomogeneous and non-stationary traffic conditions over the shared boundary result in an estimation error in both triangular areas. There are, however, two differences, i.e., (1) the flow is opposite (inflow vs. outflow) and (2) the error is scaled by the related triangular area surface. Therefore, the error induced by a shared boundary is inversely proportional to the surfaces of the adjacent triangular areas.

4.4. Traffic state estimation in a pre-defined mesh

The relations introduced in previous sections allow us to estimate q and k for triangular areas in space-time, see Fig. 4a. However, depending on the application, we may want to select the dimensions of the estimation area, i.e., define the estimation mesh. As an example, in this research, we consider a mesh which subdivides space-time in rectangular areas, see Fig. 4b. The estimates for the rectangular areas can be interpreted as the mean traffic conditions for a road segment (cell) during a time period. The selected mesh, i.e., discretizing space and time, allows for model-based prediction with the Cell Transmission Model (CTM) (Daganzo, 1994; 1995). Therefore, our estimates can be used as input for traffic state estimation using the CTM. Prior proposed estimators (Wang and Papageorgiou, 2005; Herrera and Bayen, 2010) and control algorithms (Smaragdis et al., 2004) also consider a discrete space and time mesh.

Let us consider the case in which the complete space-time domain is subdivided in triangular areas. For each triangular area $\mathcal{T}$ we know the flow $q_{\mathcal{T}}$ and density $k_{\mathcal{T}}$. To estimate the traffic conditions for a defined area in space-time, the conditions within $\mathcal{T}$ are assumed to be homogeneous and stationary. This means that the *TTS* and *TTD* within a subarea of $\mathcal{T}$



(a) Three-point estimation

(b) Estimation in desired space-time mesh, in gray: the triangle subareas which are relevant for the rectangular area

Fig. 4. Visualisation of three-point traffic state estimation and estimation in a defined mesh.

is proportional to the relative size of the subarea. Based on Edie (1965), we known that $TTD_{\mathcal{T}} = q_{\mathcal{T}} A_{\mathcal{T}}$ and $TTS_{\mathcal{T}} = k_{\mathcal{T}} A_{\mathcal{T}}$. Combined with the assumption stated above, the *TTD* and *TTS* for a subarea $\mathcal{S}$ of the triangular area $\mathcal{T}$ are given by:

$$TTD_{\mathcal{S}} = TTD_{\mathcal{T}} \frac{A_{\mathcal{S}}}{A_{\mathcal{T}}} = q_{\mathcal{T}} A_{\mathcal{S}} \tag{40}$$

$$TTS_{\mathcal{S}} = TTS_{\mathcal{T}} \frac{A_{\mathcal{S}}}{A_{\mathcal{T}}} = k_{\mathcal{T}} A_{\mathcal{S}} \tag{41}$$

Let us consider a rectangular area $\mathcal{D}$ with a surface $A_{\mathcal{D}}$. The surface of the relevant subarea $A_{\mathcal{S}}$ is the surface of triangle $\mathcal{T}$, i.e., $A_{\mathcal{T}}$, in the desired area $\mathcal{D}$, i.e., $A_{\mathcal{T}} \in A_{\mathcal{D}}$. Based on the traffic conditions in all relevant areas, i.e., $q_{\mathcal{T}}$ and $k_{\mathcal{T}}$, and their contribution, i.e., $(A_{\mathcal{T}} \in A_{\mathcal{D}})/A_{\mathcal{D}}$, the traffic conditions in the rectangular area can be estimated:

$$q_{\mathcal{D}} = \frac{\sum_{\mathcal{T}} (A_{\mathcal{T}} \in A_{\mathcal{D}}) q_{\mathcal{T}}}{A_{\mathcal{D}}} \tag{42}$$

$$k_{\mathcal{D}} = \frac{\sum_{\mathcal{T}} (A_{\mathcal{T}} \in A_{\mathcal{D}}) k_{\mathcal{T}}}{A_{\mathcal{D}}} \tag{43}$$

Fig. 4b provides a visualization of the relevant triangle subareas for a single rectangular area.

The accuracy of estimates obtained with (42) and (43) depend on two factors. Firstly, if $q_{\mathcal{T}}$ and $k_{\mathcal{T}}$ are estimates, they may contain estimation errors. If these estimates are obtained using three-point estimation, the estimation errors of adjacent areas induced by the shared boundary are negatively correlated. In this case the errors are partially corrected by combining adjacent areas. Secondly, we assumed that conditions within $\mathcal{T}$ are homogeneous and stationary. If $\mathcal{T}$ partly falls outside the desired area $\mathcal{D}$, inhomogeneity and non-stationarity within $\mathcal{T}$ can induce estimation error for $\mathcal{D}$.

5. Simulation study

As explained before, the estimation procedure consists of two processes: (1) estimating $N(x, t)$ based on measurements from stationary and moving observers and (2) estimating q and k in a pre-defined space-time mesh based on $N(x, t)$. These processes were consecutively discussed in Sections 3 and 4. In the simulation study we focus on the second process, for which the PON estimation methodology is designed. The simulation study is explained in this section and the results are discussed in Section 6.

To investigate and explain the working of the methodology in more detail, it is assumed that the first process is perfect, so we correctly know $N(x, t)$ for the point-observations along the observations paths of the available stationary and moving observers. This requires that the observation paths are correctly initialized and there are no (or fully corrected) observation errors. In practice, there will be count-errors in stationary (and moving) observers. Therefore, error correction is required for methodologies that only require cumulative counts, e.g., Van Lint et al. (2014). In our study, as explained in Section 3, the cumulative drift problem is less of an issue. We will therefore further elaborate on the second process, assuming no errors, in this and the next section. Additionally, Appendix A shows how observations errors might affect the performance.

5.1. Microscopic simulation in FOSIM

The simulation study is conducted with the microscopic simulation program FOSIM (Dijker and Knoppers, 2006). FOSIM is calibrated and validated for Dutch freeways (Minderhoud and Kirwan, 2001; Henkens et al., 2017). However, the simulated

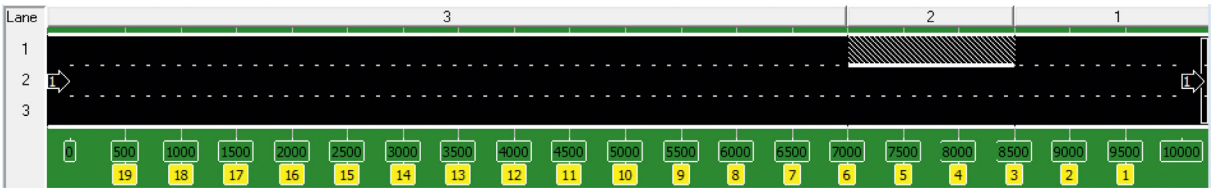


Fig. 5. Considered road lay-out in the microscopic simulation program FOSIM.

traffic conditions may still deviate from real traffic conditions for similar road infrastructure properties and traffic demand. In estimation we want to reconstruct the conditions based on a limited data-set. In our case study, this means that we want to reconstruct the simulated traffic conditions using the proposed methodologies and data. It therefore is essential that the full ground truth, which should be reconstructed, is know, which is the case for simulation. Therefore, we collect two synthesized data-sets; (1) a data-set of all vehicle trajectories to obtain the true traffic conditions, i.e., the ground truth, and (2) a limited data-set of traffic sensing data to reconstruct the true traffic conditions.

We consider a 10 km freeway link with a bottleneck at $x = 7$ km. The number of lanes drops from three to two at this location and goes back to three at $x = 8.5$ km. This road lay-out, as visualized in FOSIM, is shown in Fig. 5.

The traffic conditions are described by the macroscopic traffic flow variables in a rectangular mesh, i.e., discrete space and time. Here, the freeway is subdivided in 500 m long road segments and 15 s time-periods. Given the speed limit of 120 km/h this combination of road segment length and time-period duration satisfies the Courant–Friedrichs–Lewy conditions (Courant et al., 1928). Although this condition is not a prerequisite for the proposed method, it is important for numerical stability in model-based estimation and prediction. As the estimates may be used as initial conditions for model-based traffic state prediction, we decided to evaluate the methodology in this mesh.

Traffic is simulated for two one-hour periods, one with solely free-flow conditions and one with free-flow and congested conditions. In the remainder of this paper these two cases are respectively referred to as the free-flow and congested case. Fig. 6a and b show the traffic demand that is used as input into the microscopic simulation program is shown. In both cases traffic is composed of 90% passenger cars and 10% trucks. More details about the microscopic models and parameters used in FOSIM can be found in the users manual (Dijker and Knoppers, 2006). The true $q_{\mathcal{D}}$, $k_{\mathcal{D}}$ and $u_{\mathcal{D}}$, are obtained from the trajectory information for all vehicles using (1), (2) and (3) (Edie, 1965). This ground truth is shown in Fig. 6.

Fig. 6d, f and h, i.e., the congested case, show patterns in the congested area. To understand these patterns, we determine the ground truth for a finer mesh, i.e., smaller road segments and shorter time periods. Fig. 7 shows the ground truth in terms of density and speed for road segments of 100 m and 5 s time-periods. This figure shows that there are multiple waves which move upstream with an approximate speed of 25 km/h. The speed within these waves is close to 0 km/h, while the speed between the waves reaches 40 km/h. These patterns are thus stop-and-go waves in congestion. As explained above, our objective is to reconstruct the true simulated traffic conditions. Therefore, we will not address the realism of the observed patterns. However, we are interested in the ability of our estimator to reconstruct these patterns.

In addition to the free-flow and the congested case, we simulated and evaluated the estimation performance for an incident situation. As the results show large similarities with the congested case, we shortly discuss the estimation performance, but do not provide detailed information in this and the next section.

5.2. The reference traffic state estimator

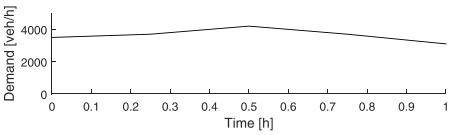
Existing types of traffic sensing data and estimation methodologies can be used to estimate the macroscopic traffic flow variables. To evaluate the added value of our work, we compare the estimation performance with a reference estimator which uses loop-detector data. It is assumed that a loop-detector is installed in the middle of each cell, i.e., the loop-detector spacing is equal to 500 m (thereby following the Dutch state-of-the-art). The choice of considering loop-detectors installed in the middle of each cell is beneficial for its estimation performance compared to other locations, e.g., upstream or downstream boundaries.

The loop-detector data characteristics are based on the loop-detectors installed on the Dutch freeways, i.e., lane-specific one-minute aggregated speed u_l^T and flow q_l. The flow for all lanes q is obtained by summing the lane-specific flows q_l. To approximate the mean speed, the following equations are considered (Knoop and Hoogendoorn, 2012):

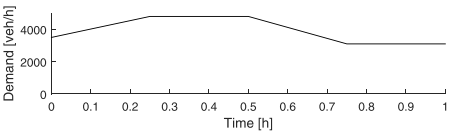
$$q = \sum_{l=1}^{\lambda} q_l \tag{44}$$

$$u = \frac{\sum_{l=1}^{\lambda} q_l}{\sum_{l=1}^{\lambda} \frac{q_l}{u_l^T}} \tag{45}$$

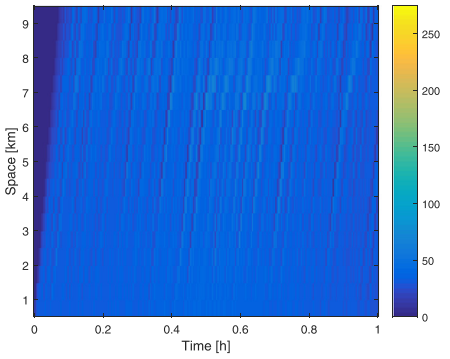
$$k = \frac{q}{u} \tag{46}$$



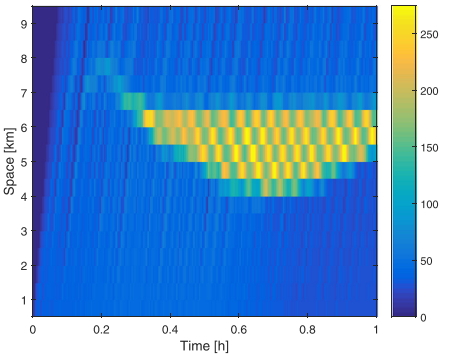
(a) Demand, free-flow case



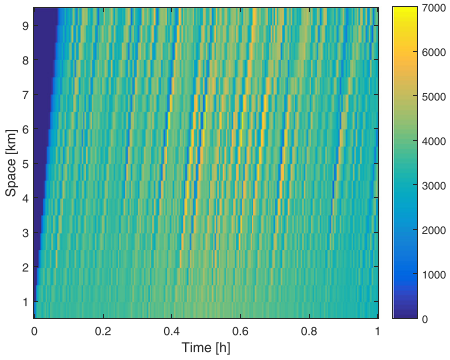
(b) Demand, congested case



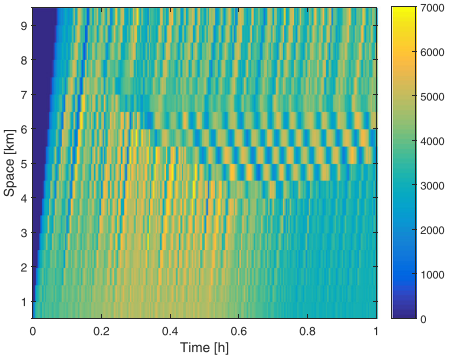
(c) Density, free-flow case



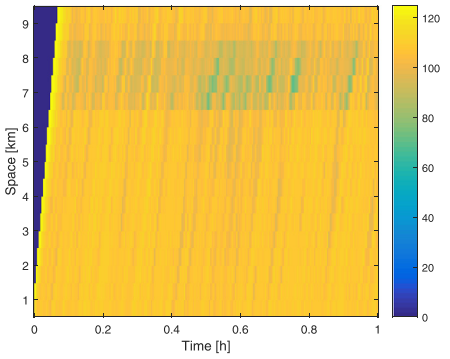
(d) Density, congested case



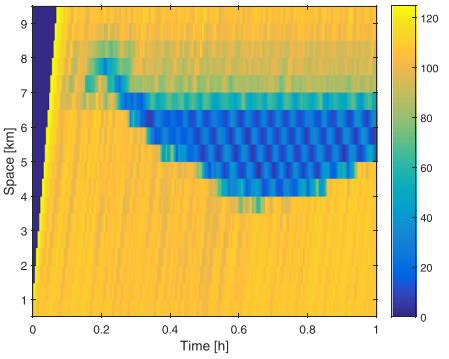
(e) Flow, free-flow case



(f) Flow, congested case



(g) Speed, free-flow case

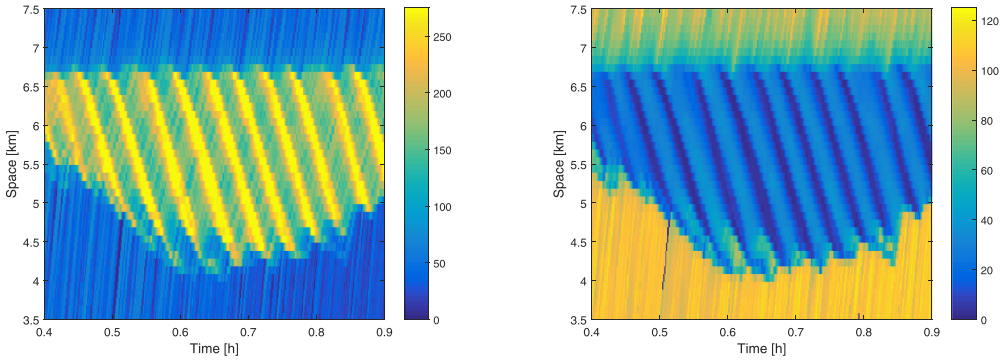


(h) Speed, congested case

Fig. 6. Demand and ground truth for free-flow and congested cases.

where λ denotes the number of lanes.

The resulting estimates relate to the one-minute loop-detector data period, i.e., four consecutive 15 s periods in the considered estimation mesh. The estimates are assigned to each 15 s period with the one-minute period. As the conditions obtained using Edie (1965) differ within this one-minute period, the one-minute data aggregation period is a cause of the estimation errors.



(a) Density (b) Speed

Fig. 7. Ground truth of the congested case for a finer mesh.

Loop-detector data characteristics differ across the world. A downside of the Dutch state-of-the-art is that time-mean speeds are observed. These data overestimate the speed (and thus underestimate the density) is congested conditions (Knoop et al., 2009). This bias-problem can be addressed by collecting harmonic mean speeds instead of time-mean speed. As some systems outside the Netherlands can collect harmonic mean speeds, we will also discuss this type of mean speeds in the results section.

5.3. The PON traffic state estimator

The data used by the proposed traffic state estimator is retrieved from sensing equipment installed on fixed locations (stationary observers) and vehicle-based sensing equipment (moving observers). In our study, all observers have a fixed sampling period of 15 s. Every 15 s, i.e., $t = 0$ s, 15 s, 30 s, ..., the observers share their position, which yields a set of point-observations along the observation path. For each period between these points ΔN is observed, i.e., the number of vehicles that passed the observer minus the number of vehicles that are passed by the observer within this period.

We fix three observation paths, which are thus used for all evaluated PON estimators. These are: the upstream, i.e., $x = 0$ km, and downstream, i.e., $x = 10$ km, link boundaries, and the first vehicle entering the link. Additional to this fixed information, a fraction of the vehicles serve as moving observers. In this set-up all observation paths can be connected, thereby yielding a ΔN for each combination of two point-observations. Furthermore, as we observe the upstream and downstream boundaries, all corner points of space-time domain until the current time T are observed, i.e., (0,0), (10,0), (0, T) and (10, T). Observing the corner points is important as this means that the complete space-time domain will be subdivided into triangular areas.

Initial evaluations have shown that the space-time ratio used in Delaunay triangulation can affect the estimation performance. However, these evaluations also indicated that the relations are not straightforward. Therefore, to keep this article concise and to-the-point, we have decided to leave additional analysis related to subdivision of space-time in triangular areas for future research. This also means that there still is room for improvement in the results shown in the next section. In this study, the space-time ratio used in Delaunay triangulation is selected to be equal to the space-time ratio of the defined estimation mesh, i.e., $\upsilon = 120$ km/h. When the triangular area dimensions are in line with estimation mesh dimensions, the PON estimator is expected to use the observations nearest (in space-time) to the desired estimation area.

The PON estimator is used at the same time interval as the reference estimator, i.e., every minute. For both estimators it is assumed that there is no data availability latency. Therefore, at the end of each minute both estimators estimate the traffic conditions for the four 15 s periods within this minute. In contrast to the reference estimator, the PON estimator is able to utilize information prior to this minute.

5.4. Evaluation of the estimators

To evaluate the overall performance of the different estimators we consider the bias (where a positive bias indicates an underestimation) and Root Mean Squared Error (RMSE). These statistics are used to evaluate the performance in terms of the flow and density, and for the two cases, i.e., the free-flow and the congested case. Although we focus on flow and density (which are the direct outputs of the PON estimation methodology), we will briefly discuss the estimation performance in terms of speed. We use a warm up period of 15 min, which ensures all observation paths are connected.

At first we will focus solely on the PON estimator. As explained in Section 4, the homogeneity and stationarity of the traffic conditions is expected to influence the estimation accuracy. The two cases shown in Fig. 6, i.e., free-flow and congested, allow us to evaluate the estimation performance for different levels of inhomogeneity and non-stationarity. Furthermore, as the constructed triangular areas depend on the data-availability it is interesting to vary this factor and discuss the related

estimation performance for the two cases. In terms of the data-availability we solely vary the penetration rate. Here, we will depict the estimation performance for seven different penetration rates, i.e., $p = 0.10, 0.25, 0.50, 1.0, 2.5, 5.0, 10.0\%$.

The main goal of the simulation study is to show the working of the proposed methodology. Additionally, to discuss the added value of the PON methodology and considered data, the PON estimator is compared with the reference estimator. Here, it is specifically interesting to see how the relative estimation performance differs between the two cases and the macroscopic traffic flow variables.

6. Results

Fig. 8 shows the estimation performance of the estimators in terms of the bias and RMSE. The eight sub-figures show the macroscopic traffic flow variables flow and density for the two cases, i.e., free-flow and congested. Each sub-figure shows two lines, i.e., one for the PON estimator (solid blue line) and one for the reference estimator (dashed black line). For all penetration rates p the upstream and downstream boundary are observed, i.e., we have more information than solely traffic sensing data from the moving observers.

6.1. Estimation performance of the PON estimator

The estimation performance for the different penetration rates depends on the homogeneity and stationarity of the traffic conditions. The lower the penetration rate, the larger the constructed triangular areas. Therefore, for lower penetration rates these triangular areas are expected to individually be part of more rectangular areas in the defined estimation mesh. If the traffic conditions are inhomogeneous and non-stationary within the triangular area, the macroscopic traffic conditions are not representative for all related rectangular areas yielding estimation errors. This explains the clear difference between the RMSEs in the free-flow and the congested case observed in Fig. 8. At low penetration rates, the RMSEs are larger in the congested case; however, at a penetration rate of 10.0% the performance in terms of RMSE is similar for both cases. In terms of flow estimates, the bias improves with the penetration rate (similar to the RMSE). However, in terms of density estimates, we observe a near-zero bias for all penetration rates.

To illustrate the importance of localizing changes in traffic conditions over space-time, we look at Fig. 9. This figure shows that the estimates for PON estimator at three different penetration rates, i.e., $p = 0.10$, 1.0 and 10.0%, and the reference estimator. These estimates can be compared to the ground truth given in Fig. 6d. The pattern at $p = 0.10\%$ is caused by three individual vehicles of which the trajectories are shown using the black lines. These vehicles do not provide sufficient information to localize the congestion (delay). Although the congested pattern is more in line with the ground truth at $p = 1.0\%$, the identification of this pattern clearly improves when we move to $p = 10.0\%$. When estimates for $p = 10.0\%$, i.e., Fig. 9c, with the ground truth, i.e., Fig. 6d, we see that the estimator is able to approximate both the correct values of density and the pattern caused by the stop-and-go waves in congestion.

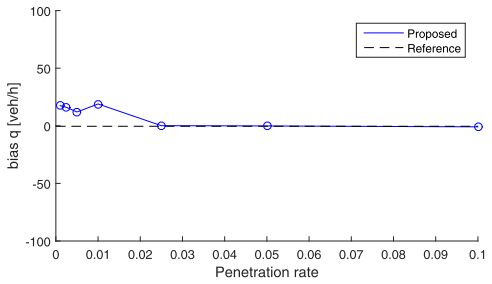
In addition to the free-flow and congested case, the PON estimator was applied to an incident case. This represents a form of non-recurrent congestion and a situation in which the traffic flow behavior temporarily changes due to a lane closure. The resulting estimation performance was similar to the congested case and therefore we do not present the detailed results in this paper. Similar results to the congested case were expected as the PON estimator does not use any information in the form of a traffic flow model or historical information. Therefore, the performance of the PON estimator should and did not differ between recurrent and non-recurrent congestion, or be effected by using an inaccurate description of the traffic flow behavior.

6.2. Comparison between the PON and reference estimator

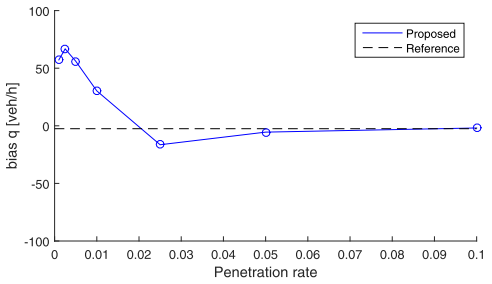
The simulation study allows us to compare the estimation performance of the PON estimator with the reference (loop-detector data) estimator. We discuss two elements of the relative performance, i.e., (1) the relation to the inhomogeneity and non-stationarity in traffic conditions and (2) the differences between the macroscopic variables.

The relative performance of the two estimator depends on the level of inhomogeneity and non-stationarity of the traffic conditions. As discussed above, the PON estimator requires a higher data-availability for traffic phase identification if the inhomogeneity and non-stationarity of traffic conditions increases. The reference estimator uses observations from the middle of each cell, i.e., sensing equipment is installed every 500 m. As a result, the reference estimator accurately locates the different traffic phases in space-time for the congested case, see Fig. 9. This explains the larger difference between the PON and reference estimators in the congested with respect to the free-flow case at low penetration rates.

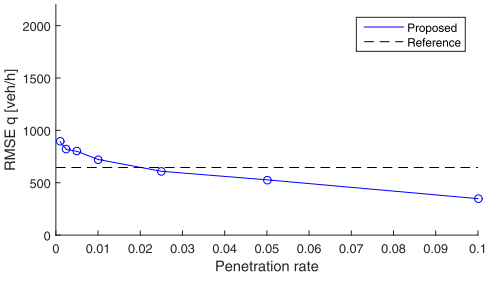
With respect to the reference estimator, we observe a better relative performance of the PON estimator for density estimates than for flow estimates. In terms of flow, the relative estimation performance is better in the free-flow case than the congested case, i.e., the PON estimator outperforms the reference estimator in terms of RMSE respectively at penetration rates higher than approximately 2.5% and 5.0%. The bias of the PON estimator at penetration rates higher than 5% and the reference estimator are both (near) zero. In terms of density, the PON estimator outperforms the reference estimator in terms of RMSE starting from a penetration between 1.0% and 2.5% for both cases. As shown in Fig. 8f and by comparing Figs. 6d and 9 d, the reference estimator has a large bias in congested conditions. This bias is reduced if the loop-detectors would observe harmonic mean speeds instead of lane-specific time-mean speeds (ceteris paribus). In this case the PON



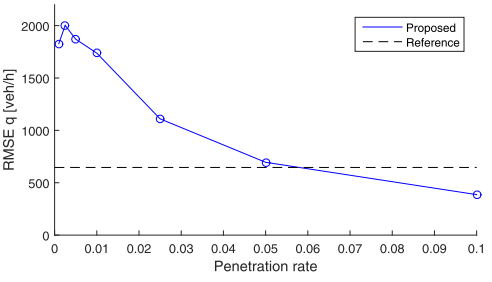
(a) Bias flow, free-flow case



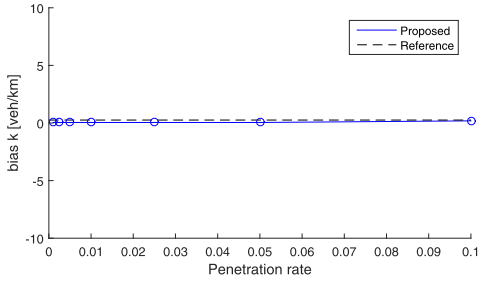
(b) Bias flow, congested case



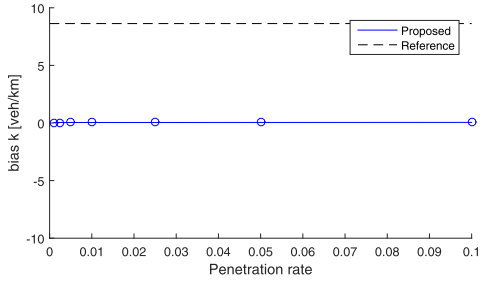
(c) RMSE flow, free-flow case



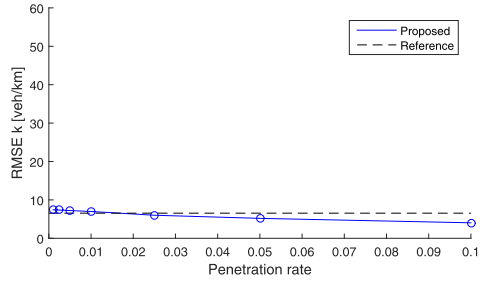
(d) RMSE flow, congested case



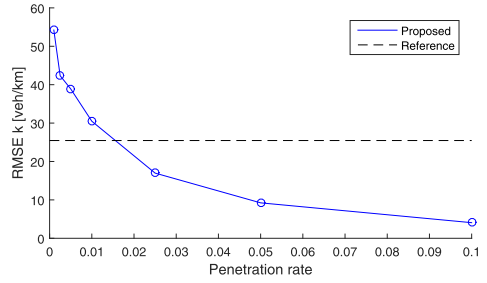
(e) Bias density, free-flow case



(f) Bias density, congested case



(g) RMSE density, free-flow case



(h) RMSE density, congested case

Fig. 8. Estimation performance (bias and RMSE) of the PON and reference (loop-detector) estimator for the free-flow and congested case.

estimator outperforms the reference estimator in terms of density RMSE starting from a penetration rate between 2.5% and 5.0% for the congested case.

In contrast to density and flow estimation, we do not recommend using the PON estimator for speed estimation. Existing methodologies that use trajectory data or probe speed data (e.g., Work et al., 2010; Del Arco et al., 2011) are expected to outperform speed estimates with the PON estimator. However, our approach to combine stationary and moving observers is valuable for (the speed-related) dynamic link travel-time estimation. This approach allows us to estimate the cumulative curves at the upstream and downstream link boundaries, which in turn can be used to estimate the link travel-time, e.g., Van Lint et al. (2014).

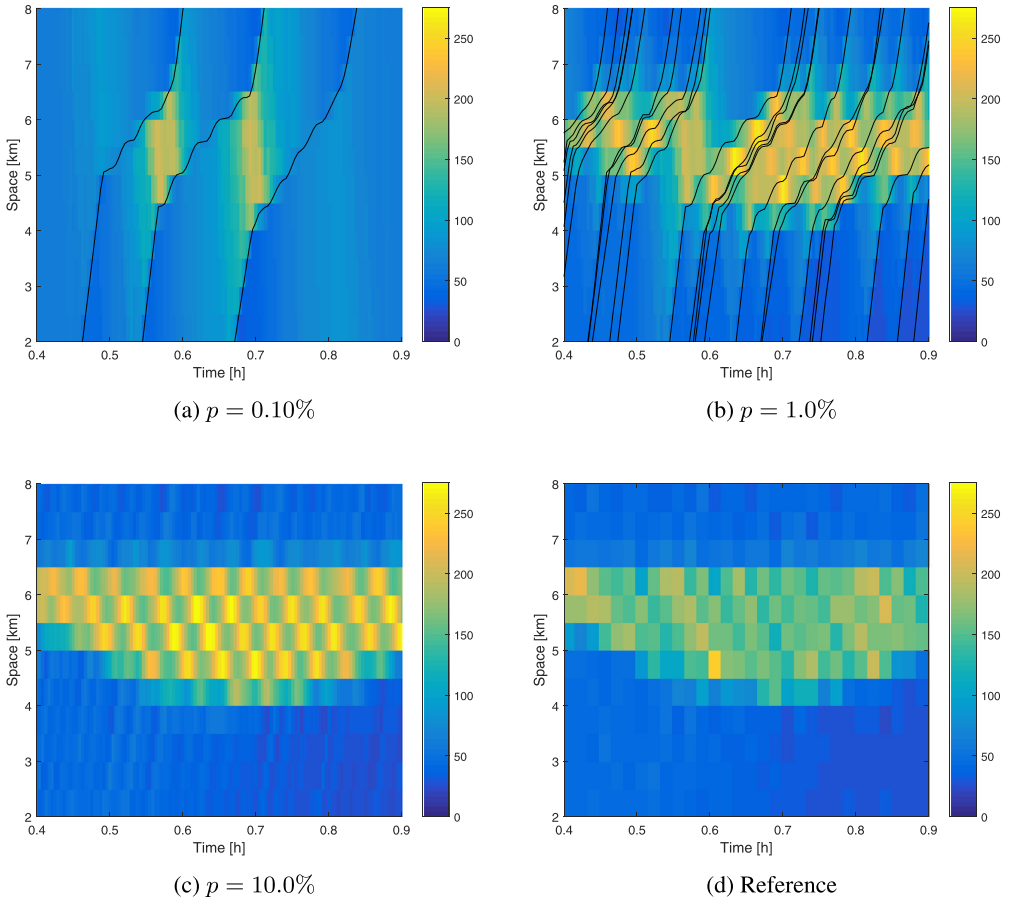


Fig. 9. Density estimation in space-time for the PON estimator at $p = 0.10$, 1.0 and 10.0%, and the reference estimator. The moving observers trajectories (black lines) are shown for $p = 0.10\%$ and 1.0%.

7. Conclusions and discussion

In this study, we propose an estimation procedure to estimate the macroscopic traffic conditions in a pre-defined space-time mesh using traffic sensing data collected by stationary and moving observers. This procedure consists of two processes: (1) obtain point-observations of the cumulative vehicle number N using stationary and moving observers and (2) estimate the macroscopic traffic conditions based on the point-observations of N. For the second, we designed a full methodology, which we denote as the Point-Observations N (PON) methodology. This methodology does not use any information related to traffic behavior (e.g., a fundamental diagram) or historical data and is thus a streaming-data-driven methodology.

The PON estimation methodology assumes homogeneous and stationary traffic conditions. If this assumption holds the methodology perfectly estimates the macroscopic traffic conditions in space-time. If this assumption is violated, errors are induced. Here, we make a difference between errors induced when estimating the conditions for the triangular areas that form the basic estimation unit and the errors induced when going from the triangular areas to the desired areas. The relation between these assumptions and the estimation errors explain that a positive relation exists between the level of inhomogeneity and non-stationarity of the traffic conditions and the required data-availability to reach a similar estimation performance. If the traffic conditions change highly over space and time, e.g., from free-flow to congestion or stop-and-go waves, having sufficient point-observations is important to localize the different traffic phases. Nevertheless, in the conducted case-study we still only needed to observe 1.0–2.5% of the vehicles in combination with the upstream and downstream 10 km link boundary to reach the same density estimation performance in terms of RMSE as having loop-detectors installed every 500 m. In this case, the flow estimation performance is similar for a penetration rate of approximately 5%. The estimation performance for the same link was also evaluated for solely free-flow conditions. Here, the PON estimation methodology outperforms loop-detector data-based estimated in terms of both flow and density estimates at a penetration rate of 1.0–2.5%. At lower penetration rates the relative estimation performance is clearly better than in the congested case. The PON estimation methodology does not have advantages when estimating speed in a discrete space-time mesh. However, the approach to combine observations of stationary and moving observers can be used to accurately estimate the cumulative

vehicle number at the upstream and downstream link boundaries, which in turn can be used to accurately estimate the link travel-times.

The estimation performance seems to be largely determined by the ability to localize the changes in traffic conditions in space and time. Especially, miss-localizing large changes in traffic conditions, i.e., from free-flow to congestion or stop-and-go waves, results in large errors. The ability to localize the changes in traffic conditions depends on the sampling (in space and time) of the point-observations of N. In the simulation study, the influence of the penetration rate (defined as the fraction of vehicles that is observed) on the estimation performance was discussed. However, the penetration rate is not the only factor affecting the sampling of the moving observers. The number of moving observers in space and time respectively depend on the penetration rate in combination with the density and flow. Furthermore, a fixed number of available observers can be spread differently in space and time, which can (depending on the changes in traffic conditions) affect the estimation performance.

In addition to the PON estimation methodology, the more general estimation procedure is an important contribution of this paper. Equipped and/or automated vehicles can be used to collect traffic sensing data on the relative flow with respect to their trajectories. This describes the change in the cumulative vehicle number ΔN over a path in space-time. These data (together with data from stationary observers) can be fused on the N-level using the principles discussed in this paper. This is highly valuable information as N is the core macroscopic traffic flow variable. Knowing N over space and time allows for deriving all three macroscopic traffic flow variables. The PON estimation methodology can be used to estimate the flow and density in a pre-defined space-time mesh; however, different (existing) methodologies may also be used to estimate macroscopic traffic conditions, e.g., N at non-observed points in space-time or the travel-time.

Acknowledgments

We thank the Netherlands Organisation for Scientific Research, i.e., NWO in Dutch, for providing the funding used to perform this research. The grant-number assigned to this project is 022.005.030.

Appendix A. Performance under observation errors

This paper proposes and evaluates new methodology, i.e., the PON estimation methodology. Up till now, we assumed that the data allow us to obtain error-free point-observations of N. However, in reality, it is likely that the data is not perfect. To investigate the effects of observation errors, we relax the prior assumptions, i.e., observation errors in N over the observation paths are introduced. As explained in Section 3 the data, i.e., a combination of sensing data from stationary and moving observers, allows us to deal with observation errors; however, we need to design an error correction methodology to do so.

The objective of this appendix is to show that the combination of sensing data and the PON estimator also yields a good estimation output when having to deal with observation errors. For this purpose, a simple error-correction methodology is designed. We opt to show that even a simple methodology suffices to deal with observation errors. Similar to the TSE methodology discussed in Section 4, we do not want to rely on parameters that have to be calibrated, as this is a way of incorporating additional information in the form of historical data.

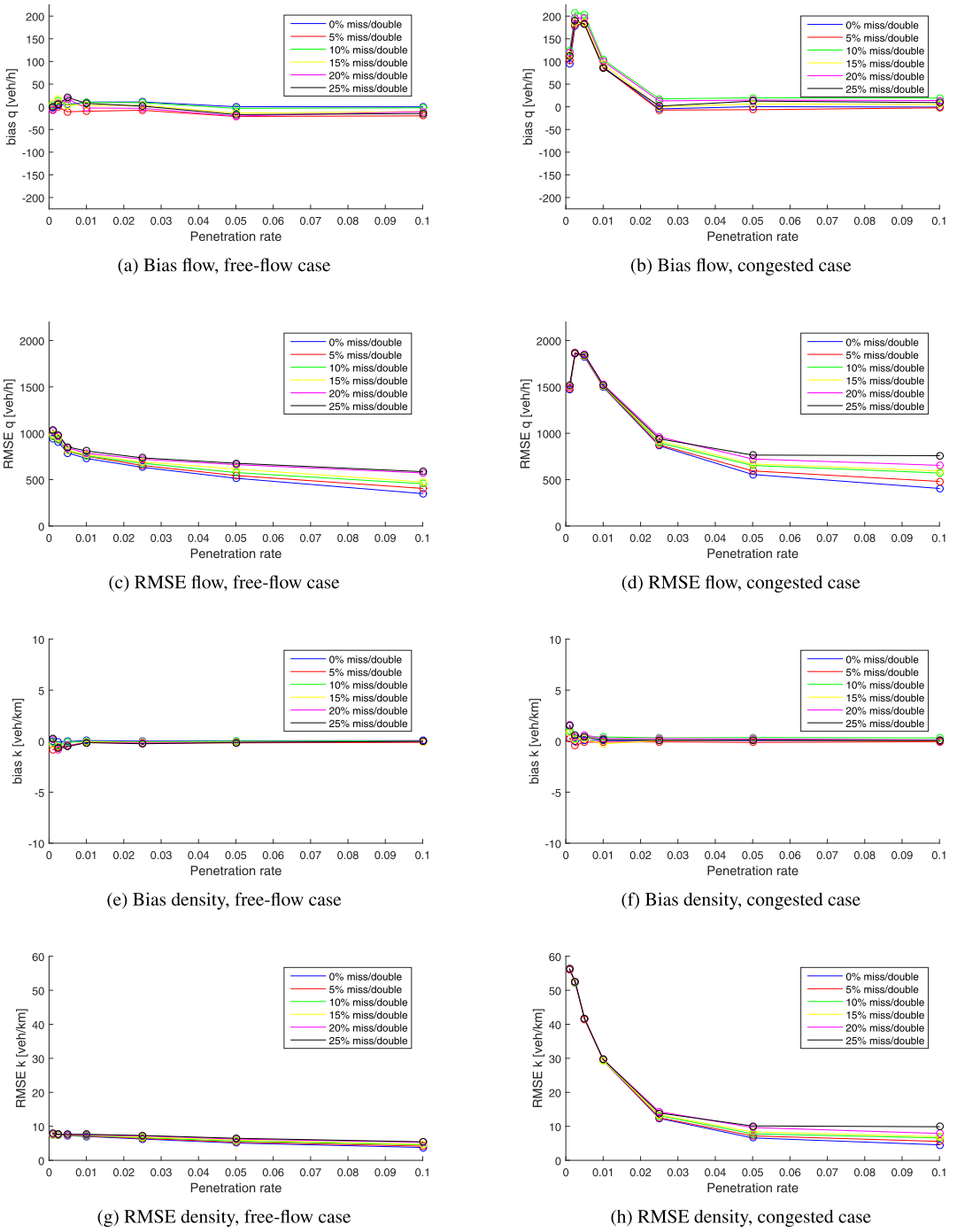
A1. Methodology

Section 3 explains the basic concept behind the error correction methodology: *An intersection of two already-initialized observation paths can be used for error correction. Observation errors can lead to a discrepancy between N at the intersection point on the two observation paths. As both paths should have the same N-value at the intersection point, we can correct for observation errors based on the difference in N. This principle is simple; however, in designing an error correction methodology, we need to define how the difference in N translates in to a correction in ΔN over the observation paths.*

Let us consider (as we do in the simulation study) a link with stationary observers at the upstream and downstream link boundaries, and more than one moving observer. To initialize N over the observation paths, the upstream link boundary ($x = 0$ m) is taken as the reference point and $N(0, 0) = 0$. Upon entering the link, thus when the moving observer interacts with the upstream boundary, each moving observer is assigned an initial N. The downstream link boundary ($x = L$) is initialized by the first moving observer exiting the link. Initialization of the downstream boundary does not occur at $t = 0$ s. Therefore, in the simulation study, a warm up period of 15 min is used.

Any other intersection of observation paths is *an intersection of two already-initialized observation paths* and can thus *be used to for error correction* (see above). In our simple methodology, for each of these intersections, we define a leading and following path. The leading path has the most recent information from the upstream link boundary. In case of interaction between the downstream link boundary and a moving observer, the moving observer is leading. In case of an interaction between two moving observers, the overtaking (i.e., fastest) moving observer is leading.

A difference in N in the intersection point between the two observers is accounted for by the following observer in the period between its last and current interaction. In this way every observation is at most corrected one time. To account for the difference, which is the error, the ΔN are altered. the difference is spread out evenly over the absolute number of overtakings registered by the observer. For instance, if there is an error of +4 veh and the following vehicle registered 16



(a) Bias flow, free-flow case

(b) Bias flow, congested case

(c) RMSE flow, free-flow case

(d) RMSE flow, congested case

(e) Bias density, free-flow case

(f) Bias density, congested case

(g) RMSE density, free-flow case

(h) RMSE density, congested case

Fig. 10. Estimation performance (bias and RMSE) of the PON estimator for the free-flow and congested case under observation errors.

overtakings, we add -0.25 veh to each of the 16 registered overtakings. Here we assume that the connected vehicle provides information on each overtaking it observes. For instance, if the connected vehicle registers that it overtakes one vehicle and is overtaken by another vehicle in one data-period, we want to know that two overtakings were registered and that $\Delta N = 0$ veh within this period. If the following observer did not register any overtakings, the difference is spread out evenly over the time periods.

This simple error correction methodology does not require any parameters. Furthermore, for each observed point in space-time there are maximally two values of N, that is, before and after correction. If the intersecting of the two relevant observations paths (which are used for correction) falls within the available data period, we use the corrected value.

A2. Implementation in the simulation study

The performance under observation errors is evaluated for the cases that were discussed in Sections 5 and 6. The difference between the simulation study in these sections and this appendix lies in the observations and implementation of an error correction methodology (see above). Instead of assuming that we know N for specific points in space-time, we consider potentially erroneous observations of ΔN and estimate N based on these observations.

Stationary and moving observers should observe vehicles that they pass or that pass them. However, these observers can miss or incorrectly define vehicle-passings. In this study, we consider the following errors: each passing can either be (1) correctly observed, (2) missed or (3) double-counted. Furthermore, we assign equal probabilities to events (2) and (3).

In theory, it is possible that with one data-period (5 s) a vehicle overtakes another vehicle and is directly overtaken again by the same vehicle. These overtakings are missed in our study, as we expose the passings based on the difference between the vehicle positions in consecutive time-instances, which are 5 s apart. However, as vehicle speeds are not expected to show large short-term fluctuations, we believe that we only miss a very small fraction of the overtakings.

In line with the main simulation study of this article, we evaluate the RMSE of density, flow and speed estimates for different penetration rates. However, in this simulation study, we also consider different probabilities for missing or double-counting overtakings, i.e., 0, 5, 10, 15, 20, 25%. A miss/double-count probability of 10% means that the probabilities of correct, missed and double-counted observations are respectively 90%, 5% and 5%.

A3. Results

Fig. 10 shows the estimation performance for different miss/double-count probabilities. As expected, observation errors lead to a lower estimation performance. However, even with large observation errors (e.g., 25.0%), we observe a good estimation performance.

The estimation performance statistics, i.e., bias and RMSE, under observation errors (Fig. 10) have a similar shape to those without observation errors. The RMSE error reduces with increasing penetration rate in both cases (free-flow and congested) and for both variables (flow and density). However, the influence of the observation errors, i.e., the difference between the lines with different miss/double-count probabilities, is smallest for low penetration rates. At these low penetration rate, the influence of incorrect ΔN values between points in space-time may be of lower significance compared to the existing estimation errors (which are discussed in Section 6). Furthermore, the errors caused by incorrect observations of ΔN can become more local, e.g., an underestimation and overestimation of the density respectively upstream and downstream of a moving observer can come hand-in-hand and both contribute to a larger RMSE.

For each observation error probability the bias moves towards a level that depends on the errors in the upstream stationary detector. However, depending on these errors this level lies at a certain non-zero value. This makes sense, as our error correction methodology does not try to correct all errors (e.g., no corrections are performed on the upstream link boundary), but opts to limit the errors between point observations in near proximity in space and time. This may be a reason to design a better error correction methodology; however, a stated before, this is out of the scope of this appendix. In contrast to the RMSE, the bias of one observation error probability can be better (i.e., closer to zero) than for a lower observation error probability. Again, this makes sense, as we consider a zero-mean error distribution. Therefore, additional errors may actually compensate earlier errors.

We discussed that the intersection between observation paths is important for initialization and error correction. However, the observers considered in this study, do not directly observe these intersection point. Instead, in the error correction methodology we needed to estimate the intersection points based on the trajectory observation points that have an interval of 15 s. Therefore, it would be beneficial if the stationary and moving observers observe the exact intersection points. However, as this puts an extra strain on the required data characteristics, in this study, we want to keep this as a recommendation for practice.

Supplementary material

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.trb.2018.06.005.

References

Claudel, C.G., Bayen, A.M., 2010. Lax-Hopf based incorporation of internal boundary conditions into Hamilton-Jacobi equation. Part I: theory. IEEE Trans. Automat. Contr. 55 (5), 1142–1157.
Claudel, C.G., Bayen, A.M., 2010. Lax-Hopf based incorporation of internal boundary conditions into Hamilton-Jacobi equation. Part II: computational methods. IEEE Trans. Automat. Contr. 55 (5), 1158–1174.
Courant, R., Friedrichs, K., Lewy, H., 1928. Uber die partiellen Differenzengleiehungen der mathematischen Physik. Math. Ann. 100 (1), 32–74.
Daganzo, C.F., 1994. The cell transmission model: a dynamic representation of highway traffic consistent with the hydrodynamic theory. Transp. Res. Part B 28B (4), 269–287.
Daganzo, C.F., 1995. The cell transmission model, Part II: network traffic. Transp. Res. Part B 29B (2), 79–93.
Del Arco, E., Morgado, E., Ramiro-Barguene, J., More-Jimenez, I., Caamano, A., 2011. Vehicular sensor networks in congested traffic: Linking STV field reconstruction and communications channel. In: 14th International IEEE Conference on Intelligent Transportation Systems. Washington, DC, USA, pp. 606–613.
Dijker, T., Knoppers, P., 2006. FOSIM 5.1 Gebruikershandleiding (Users Manual). Technical Report. Technische Universiteit Delft.
Edie, L.C., 1965. Discussion of traffic stream measurements and definitions. 2nd Int. Symp. On the Theory of Traffic Flow. OECD, Paris.

Florin, R., Olariu, S., 2017. On a variant of the mobile observer method. IEEE Trans. Intell. Transp. Syst. 18 (2), 441–449.
Henkens, N., Mieras, W., Bonnema, D., 2017. Validatie FOSIM. Technical Report. Sweco, De Bilt.
Herrera, J.C., Bayen, A.M., 2010. Incorporation of lagrangian measurements in freeway traffic state estimation. Transp. Res. Part B 44 (4), 460–481.
Knoop, V.L., Hoogendoorn, S.P., 2012. Empirics of a generalized macroscopic fundamental diagram for urban freeways. In: 92th Annual Meeting of the Transportation Research Board, pp. 1–16.
Knoop, V.L., Hoogendoorn, S.P., Van Zuylen, H., 2009. Empirical differences between time mean speed and space mean speed. In: Appert-Rolland, C., Chevois, F., Gondret, P., Lassarre, S., Lebacque, J.-P., Schreckenberg, M. (Eds.), Proceedings of Traffic and Granular Flow 07. Springer, New York, pp. 351–356.
Laval, J.A., He, Z., Castrillon, F., 2012. Stochastic extension of Newells three-detector method. Transp. Res. Record 2315, 73–80.
Makigami, Y., Newell, G.F., Rothery, R., 1971. Three-dimensional representation of traffic flow. Transp. Sci. 5 (3), 302–313.
Minderhoud, M.M., Kirwan, K., 2001. Validatie FOSIM voor asymmetrische weekvakken - CAPWEEF fase 1. Technical Report. Delft: Laboratorium voor Verkeerskunde, Faculteit Civiele Techniek en Geowetenschappen, Technische Universiteit Delft.
Nanthawichit, C., Nakatsuji, T., Suzuki, H., 2003. Application of probe-vehicle data for real time traffic state estimation and short term travel time prediction on a freeway. Transp. Res. Rec. 5890 (1855), 49–59.
Newell, G.F., 1993. A simplified theory of kinematic waves in highway traffic, Part I: general theory. Transp. Res. Part B 27 (4), 281–287.
Newell, G.F., 1993. A simplified theory of kinematic waves in highway traffic, Part II: queueing at freeway bottlenecks. Transp. Res. Part B 27 (4), 289–303.
Newell, G.F., 1993. A simplified theory of kinematic waves in highway traffic, Part III: multi-destination flows. Transp. Res. Part B 27 (4), 305–313.
Papageorgiou, M., Hadj-salem, H., Blosseville, J.-m., 1991. ALINEA : A local feedback control law for on-Ramp metering. Transp. Res. Rec. 1320, 58–64.
Redmill, K.A., Coifman, B., Mccord, M., Mishalani, R.G., 2011. Using transit or municipal vehicles as moving observer platforms for large scale collection of traffic and transportation system information. In: 14th International IEEE Conference on Intelligent Transportation Systems.
Seo, T., Bayen, A.M., Kusakabe, T., Asakura, Y., 2017. Annual reviews in control traffic state estimation on highway : a comprehensive survey. Annu. Rev. Control 43, 128–151.
Seo, T., Kusakabe, T., 2015. Probe vehicle-based traffic state estimation method with spacing information and conservation law. Transp. Res. Part C 59, 391–403.
Smaragdis, E., Papageorgiou, M., Kosmatopoulos, E., 2004. A flow-maximizing adaptive local ramp metering strategy. Transp. Res. Part B 38 (3), 251–270.
Treiber, M., Helbing, D., 2002. Reconstructing the spatio-temporal traffic dynamics from stationary detector data. Cooper@tive Tr@nsport@tion Dyn@mics 1, 3.1–3.24.
Van Lint, H., Bertini, R.L., Hoogendoorn, S.P., 2014. Data fusion solutions to compute performance measures for urban arterials. In: Celebrating 50 Years of Traffic Flow Theory, pp. 1–5. Portland, Oregon
Wang, Y., Papageorgiou, M., 2005. Real-time freeway traffic state estimation based on extended Kalman filter: a general approach. Transp. Res. Part B 39 (2), 141–167.
Wardrop, J., Charlesworth, G., 1954. A method of estimating speed and flow of traffic from a moving vehicle. Proc. Inst. Civ. Eng. 3 (1), 158–171.
Work, D.B., Blandin, S., Tossavainen, O.-P., Piccoli, B., Bayen, A.M., 2010. A traffic model for velocity data assimilation. Appl. Math. Res. eXpress 1, 1–35.