Sebastiaan Thoen (4304942)

Using relative flow data to expose the fundamental diagram

Abstract

The fundamental diagram (FD) describes the relation between the macroscopic traffic variables speed (u), flow (q), and density (k) in stationary and homogenous traffic. The FD is an important input for traffic control measures. Currently, loop detectors installed in the road are most often used to collect data to estimate the FD. Such loop detectors are expensive, can be out of order, and do not provide information when traffic is at standstill. This paper introduces a methodology to estimate points of q and k on the FD using only probe vehicles as moving observers. Areas in space-time (xt-areas) are constructed using the trajectories of moving observers as boundaries. The values of q and k in an xt-area can be calculated with Edie's definitions. Furthermore, we developed a simpler method, which we call the Corner Points Delta (CDN) method, to estimate q and k in each xt-area, using the relative flow and average speed of each boundary as input. For (nearly) stationary and homogeneous traffic conditions, which can be identified with our method, the CDN method can estimate points on the FD with high accuracy. It also provides more data about very congested states, improving on traditional methods that use loop detector data. Because it only requires data collected by vehicles with sensing equipment, the FD can be estimated for road sections where loop detectors are out of order or not installed. The paper indicates possible extensions for a wider application of the method.

1. Introduction

The fundamental diagram (FD) in traffic flow theory describes the relation between the macroscopic variables speed (u), flow (q), and density (k) when traffic is in equilibrium. The FD is an important input for traffic control measures such as ramp metering (Drake et al., 1967). Each road section has different characteristics and users, therefore it is desired to estimate the FD of each road section individually (Knoop & Daamen, 2014). The most commonly available data that can be used to estimate the FD is measured with induction loops installed in the road (Daamen et al., 2014). These induction loops are expensive, can be out of order, and do not measure traffic conditions when vehicles are at standstill (Knoop & Daamen, 2014; Van Erp et al, 2017b).

Therefore, another method to finds points of q and k that form the FD is desired. This paper proposes a method using data collected by moving observers that observe the change in cumulative vehicle number (ΔN) over paths in space-time. Van Erp et al. (2017a) have hypothesized that such data may become widely available when the sensing equipment of automated vehicles is used for more purposes than individual vehicle control.

The method consists of two mains steps: (1) constructing areas in space-time (denoted as xt-areas), and (2) estimating q and k in each constructed xt-area. For the first step, we use the trajectories of moving observers as the boundaries to xt-areas.

The two main contributions of this paper are as follows. Firstly, we construct xt-areas in such a manner that its traffic state can be calculated with Edie's definitions, and information is provided about its adequacy for FD estimation. Secondly, we developed a simpler method, that can provide a perfect estimate of traffic states on the FD in stationary and homogenous xt-areas.

This paper is organized as follows. In Section 2 a literature overview is given that serves as background information for the developed method, which is treated in Section 3. In Section 4 the simulation study to test the developed method is explained, after which the results of the study are

reported in Section 5. Conclusions and a discussion are provided in Section 6. In Section 7 we discuss possible extensions of the developed method.

2. Literature overview

This section given an overview of relevant literature for the development of our method. Two main topics are discussed. Firstly, the description of traffic conditions is treated. Secondly, the estimation of traffic states for the FD is discussed.

2.1 Describing traffic conditions

Traffic conditions are most often described with the macroscopic variables flow (q), density (k), and speed (u) (Edie, 1965). Edie's generalized definitions should be used to obtain the values of these three variables for FD estimation (Edie, 1965; Laval, 2011). These definitions are based on the size (A) of an xt-area, the total distance traveled by all vehicles in an xt-area (TD), and the total time spent by all vehicles in an xt-area (TT). The equations are as follows:

$$q = \frac{TD}{A} \tag{1}$$

$$k = \frac{TT}{A} \tag{2}$$

$$u = \frac{TD}{TT} \tag{3}$$

Another way to describe traffic, is the three-dimensional representation of Makigami et al. (1971). The three dimensions considered are: space (x), time (t), and cumulative vehicle number (N). $N(x,t)$ is thus the cumulative number of counted vehicles at a point in space and time (Van Erp et al., 2017a). This way, changes in traffic conditions over space and time can be graphically shown. If $N(x,t)$ is considered continuous and differentiable to x and t, q is the time derivative of N, shown in equation 4 (Makigami et al., 1971).

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

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

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

2.2 Estimating traffic states for the FD

This section consists of three parts. Firstly, stationarity and homogeneity of traffic, two important features for FD estimation, are discussed. Secondly, the usage of road-side detectors for FD estimation is discussed. Finally, the usage of vehicles as moving observers for FD estimation is covered.

2.2.1 Stationarity and homogeneity

The FD describes the relation of traffic state variables in equilibrium. Therefore data describing transient phases should not be used for FD estimation (Knoop & Daamen, 2014). Hence, in order to find data points which lie on the FD, traffic states in the xt-areas over which data is aggregated should stay the same over the time axis (stationarity) and over the space axis (homogeneity).

Usually, a lot of scatter is found in the relation between the three variables. It has been shown that a large amount of this scatter can be removed by only selecting those data that describe (nearly) stationary conditions (Cassidy, 1998; Laval, 2011). Cassidy (1998) states that when more than one traffic state is present in an xt-area, there is no reason to expect the traffic state variables in this area to lie on a reproducible function, as the values will be merely a weighted average of two traffic states. However, it should be noted that if both states are on the same linear function, the average will still lie on this function. To maximize the chance of finding stationary xt-areas, Laval (2011) proposes to choose boundaries to the xt-areas with free flow and wave speed.

Ideally, Edie's generalized definitions are used to estimate the traffic state inside each xt-area. However, for this purpose TD and TT of each xt-area have to be known. Hence, it is required to know the entrance and exit time and location of each vehicle inside an xt-area, which is often not possible. Therefore, other methods have been proposed that use road-side detectors or moving observers, these are discussed in the next part.

2.2.2 Road-side detectors

A road-side detector is installed in the road and observes when a vehicles passes it. The most common detector is the induction loop (Daamen et al., 2014). Depending on its configuration, it can measure the flow and time mean speed over time intervals. To arrive at the density, the formula $k = q/u$ can be used. However, to estimate the density, the space mean speed is needed instead of the time mean speed. Under stationary conditions, the space mean speed can be found by calculating the harmonic time mean speed (Gerlough & Huber, 1975; Knoop et al., 2009).

To estimate the FD, lines can be fitted on the obtained data points (e.g. Dervisoglu et al., 2009, Knoop & Daamen, 2014).

Four important problems related to using road-side detectors for FD estimation are: (1) traffic may not be stationary during time intervals, (2) detector failures and measurement errors occur, (3) the detectors do not detect traffic at standstill (Knoop & Daamen, 2014), and (4) FD estimation is limited to those road sections where detectors are installed (Sao et al., 2017).

2.2.3 Moving observers

Instead of a road-side detector, a vehicle can be used as a moving observer. It counts the number of vehicles overtaking him or being overtaken by him during a period of time (Wardrop & Charlesworth, 1954). This yields a relative flow (q_{rel}) that is dependent on the speed of the moving observer (v_{obs}) (equation 7). When v_{obs} equals the speed of all surrounding vehicles, q_{rel} equals zero. When the speed difference with the average speed increases, q_{rel} increases too.

$$q_{rel} = k * (u - v_{obs}) \tag{7}$$

Chiabaut et al. (2009) used vehicles as moving observers, but estimated the wave speed parameter of an FD instead of data points related to q and k. For this purpose, they assumed that the passing rate is constant when a moving observer traverses the opposite side of a road at wave speed. The estimated wave speed is then the moving observer speed at which the deviation in the passing rate is the lowest. A downside to this approach is that many observations of moving observers at wave speed are required. It also only estimates the congested branch of the FD.

Sao et al. (2017) used probe vehicles in a different fashion for FD estimation. Probe-traffic-states (i.e. traffic states considering only the probe vehicles) are estimated and used for the estimation of a separate probe FD. Based on the number of vehicles between the probe vehicles, the free flow speed

and critical density of the FD for the road section are estimated. For this purpose, the jam density parameter is fixed.

3. Methodology

We developed a method to estimate combinations of traffic variables q and k that lie on the FD. Only data of moving observers is used that do not need to cruise at wave speed. The data that moving observers collect is discussed in section 3.1. In section 3.2 we discuss the first step of the developed methodology: constructing xt-areas. Next, in section 3.3 and 3.4 we discuss how q and k can be found with Edie's definitions and approximated with our Corner Points Delta method respectively.

3.1 Data collection

The proposed method is developed for a road section with two opposing streams (figure 3.1). A share of the vehicles on both streams of this road is a moving observer. These moving observers are assumed to collect two types of data. The first type of data these moving observers collect is their own trajectory in space and time. The second type of data they collect is the change in cumulative vehicle number (ΔN) over their trajectory in space-time. Moving observers can collect ΔN on both their own side and opposite side of the road. Similarly, Coifman et al. (2017) equipped municipal vehicles with GPS and sensing equipment to detect vehicles on two opposing streams. When overtaking is not possible, ΔN is always 0 in the direction that the moving observer is traveling in. ΔN is only collected with respect to the right stream to ease the explanation of the methodology.

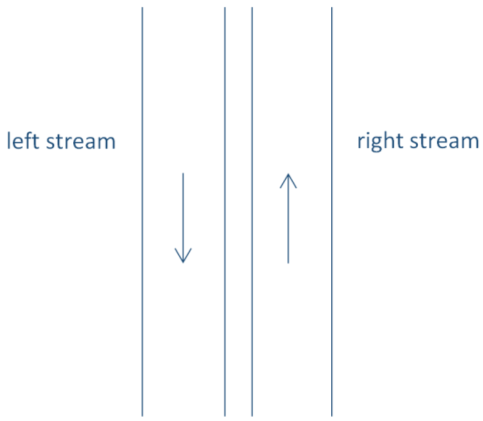


Figure 3.1. The considered road section.

3.2 Areas in space-time

In the first step of the proposed methodology, we construct xt-areas based on the data collected by the moving observers. A moving observer observes ΔN over a path in space-time (which we denote as an observation path). When two observation paths intersect, the ΔN observations over the two paths can be related to each other (Van Erp et al., 2017a). Depending on the available observation paths, xt-areas can be found that are enclosed by observation paths. The boundaries of these xt-areas are thus the observation paths that provide information on ΔN.

There are two general advantages to construct such xt-areas. Firstly, we can relate information (e.g. speed or relative flow) of boundaries to each other. This can assist in making statements about the stationarity and homogeneity of an xt-area. Secondly, we are able to use Edie's definitions to find q and k inside each xt-area.

3.3 Edie's q and k

The constructed xt-areas provide the information required to calculate Edie's generalized definitions. Based on the observation paths in space-time, the area enclosed by these paths (i.e. A) can be calculated.

TD and TT can be calculated because it is known when ΔN changes along these trajectories. Given that overtaking is not possible, the distance traveled of by the first (non-moving observer) vehicle in this xt-area, is the location at which ΔN changes for the first time at the right boundary minus the location at which ΔN changes for the first time at the left boundary. The sum of the distances for all these vehicles in the xt-area gives us TD. To deal with the fact that a vehicle is found exactly at the

upper and lower boundary of the xt-area (i.e. the moving observers), half of the distance traveled by these two vehicles is added to TD. TT can be found in a similar fashion.

3.4 Corners Delta N

We also propose a simpler approach, instead of Edie's definitions. We call this approach the Corner Delta N (CDN) method. This approach is twofold. Firstly, the relative flow and average speed of each boundary of the xt-area are used to construct lines in the qk-plane. For this purpose, equation 7 can be rewritten as equation 8. This yields several lines, namely one for every moving observer (whatever direction he is travelling at).

$$q = q_{rel} + k * v_{obs} \tag{8}$$

Secondly, a point in the qk-plane is chosen based on the points where these lines intersect. If the traffic conditions inside an xt-area are stationary and homogeneous, these lines intersect in the same point, because only one combination of values for q and k holds for this xt-area (figure 3.2). In that case, the chosen qk-point is the point where these lines intersect.

If traffic conditions change inside an xt-area, these lines do not necessarily intersect in the same point, but can form an area enclosed by intersection points in the qk-plane (figure 3.3). In this case, the mean of the four intersection points is the chosen qk-point, because there is no reason to suggest that on average one of the four intersection points is more often closest to the correct q and k.

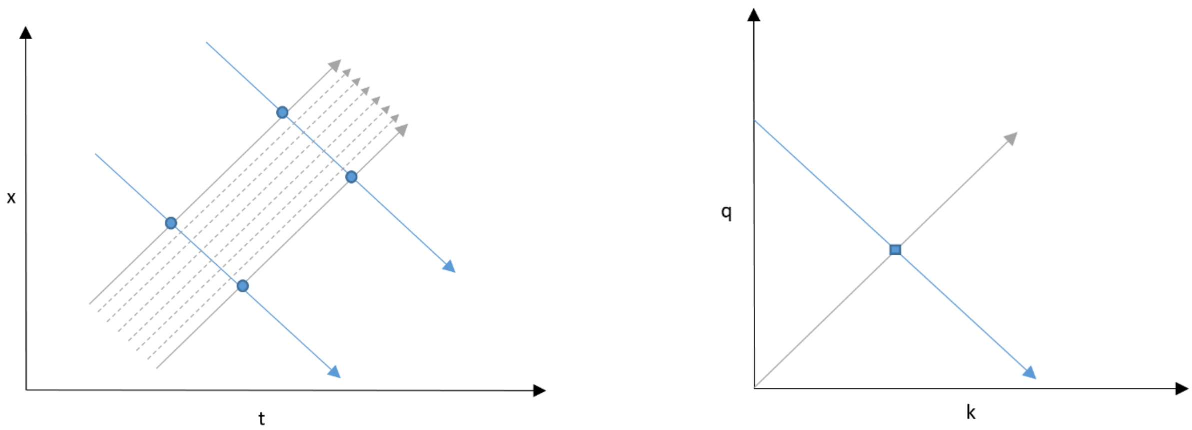


Figure 3.2. In an xt-area with homogenous and stationary traffic conditions, one intersection point is found in the qk-plane.

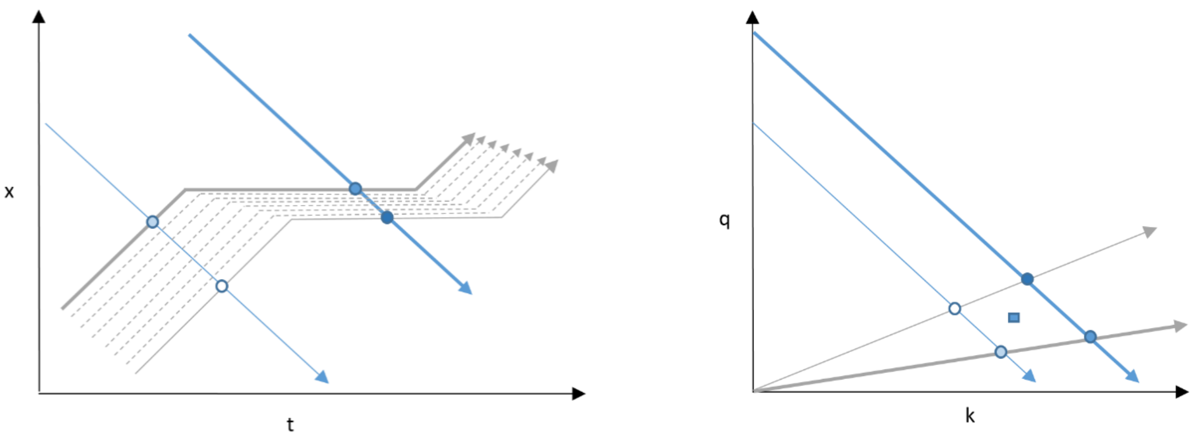


Figure 3.3. In an xt-area with heterogeneous or non-stationary traffic conditions, multiple intersection points can be found in the qk-plane.

4. Simulation study

A simulation study is performed to test the developed method that estimates realizations of q and k on the FD. The advantage of a simulation study is that the ground truth is known, and can be compared with the estimates. The study consists of two main steps. Firstly, we construct vehicle trajectories. We then determine the points in the qk-plane as the method above describes. Secondly, we quantify the measurement error of the developed method.

4.1 Microscopic simulation of traffic

The required input for the method is a set of vehicle trajectories on two opposing streams. To create vehicle trajectories, two different car following models are used: Newell's car following model and the Intelligent Driver Model (IDM). We will be using these two models to consider both stationary and homogeneous conditions (following from Newell's model), and instable traffic (from the IDM).

Newell's model (Newell, 2002) describes traffic in equilibrium, speed changes occur instantaneously and hence no transient phases are found. The IDM, developed by Treiber et al. (2000a), includes acceleration and deceleration towards a desired speed. Hence, it includes transient phases and does not only describe traffic in equilibrium conditions. The values for the model parameters are found in table 4.1, the IDM parameter values are taken from Treiber et al. (2000b).

A number of vehicle trajectories are created on the right ($n_{veh,rs}$) and left stream ($n_{veh,ls}$). At the first time step of the simulation, all vehicles move with v_{free}. Furthermore, their initial spacing equals $s_{jam} + \tau * v_{free}$ for Newell and $s_0 + l_{veh} + T * v_{free}$ for the IDM, plus a random draw of an exponential distribution with mean $\mu = 10$ m. The size of a time step (dt) equals 0.1 s.

In the simulation with Newell's model, a bottleneck is simulated on the right stream by reducing the speed of the first vehicle to 0 m/s from $t = 30$ s until $t = 149.9$ s, to 5 m/s from $t = 150$ s until $t = 299.9$ s, and increasing it back to v_{free} at $t = 300$ s until the end of the simulation. In the IDM simulation, the first vehicle decelerates with 0.8 m/s^2 to standstill starting at $t = 30$ s, and accelerates back to v_{free} with 0.8 m/s^2 starting at $t = 250$ s. Congested states are found upstream of this vehicle.

The FD that holds with Newell's model, has a triangular shape in the qk-plane. The free flow branch is represented with a line through the origin that increases linearly with a slope v_{free}. The congested branch is a line that decreases linearly with a slope ${}^{s_{jam}}/_{\tau}$, and goes through the jam density ($k_{jam} = s_{jam}{}^{-1}$). The FD of the IDM can be found for s and v, by taking ${}^{\Delta v}/_{\Delta t} = 0$ (i.e. stationary traffic), after which the FD for q and k can be calculated.

Table 4.1. Values for Newell and IDM parameters.

Newell parameters	**Value**
s_{jam}	6 m
τ	0.9 s
IDM parameter	**Value**
T	1.2 s
a	0.8 m/s^2
b	1.25 m/s^2
s_0	1 m
l_{veh}	5 m

4.2 Reference traffic state estimator

The reference traffic state estimator to compare the developed methodology to, is based on induction loop detectors installed in the road, because this is the most common way to obtain traffic data (Daamen et al, 2014). Following common practice on Dutch motorways, it is assumed these detectors have a spacing of 500 m, and collect the flow and time mean speed over time intervals of 60 s. Each detector is assumed to be located in the middle of an xt-area, i.e. a detector located 500m is used to estimate the traffic states in xt-areas with space boundaries 250m and 750m. Twenty detectors are placed, such that they estimate traffic states for xt-areas located in approximately the same range in space and time as the moving observers.

4.3 Using moving observers

A percentage of the vehicles on the right and left stream ($\%_{obs,rs}$ and $\%_{obs,ls}$) is a moving observer, randomly selected with a uniform distribution. The trajectories of these moving observers are the boundaries to construct xt-areas with. For each xt-area, q and k are determined with Edie's definitions (section 3.3), and estimated with the CDN method (section 3.4).

Table 4.2. Values for other model parameters

Parameter	Value
$v_{free} = v_0$	33.33 m/s
$n_{veh,rs}$	250
$n_{veh,ls}$	250
$\%_{obs,rs}$	5
$\%_{obs,ls}$	5
$n_{timesteps}$	6000
dt	0.1 s

4.4 Quantifying the estimation error

To quantify the performance of the developed method with respect to traffic state estimation for the FD, the Root Mean Square Error (RMSE) is determined for q. Also the bias of q (i.e. mean difference) is determined. Three types of errors are determined: (1) the q estimates with respect to Edie's q, (2) the q estimates with respect to the FD, and (3) Edie's q with respect to the FD.

To find the RMSE and bias of q with respect to the FD, we calculate the vertical difference in the qk-plane. It is determined which q holds with given FD for the value of k that is found or estimated in this xt-area. The difference with the q that is found or estimated in this xt-area, is used to calculate the RMSE and bias.

Because stochasticity is present in the initial spacing and the selection of vehicles that are moving observers, ten model runs are performed, of which the average and standard deviation are reported.

Firstly, the estimation errors are quantified with Newell's model. Secondly, these are compared with the estimation errors with the IDM. Finally, the findings with both Newell's model and the IDM are compared with the reference traffic state estimator.

5. Results

The results of the simulation study discussed in the previous section are now reported. Firstly, the estimation errors of the developed method are quantified for simulations using Newell's model. Secondly, the estimation errors for Newell's model and IDM are compared. Thirdly, the estimation

error of the developed method is compared with the reference traffic state estimator. Finally, we discuss the information each boundary provides about the adequacy of an xt-area for FD estimation.

5.1 Estimation error with Newell's model

The CDN method has a RMSE of 382 veh/h with respect to the FD in this simulation study. On average, the CDN method overestimates q on the FD by 35 veh/h. If Edie's definitions are used to determine the traffic state in each xt-area for the FD, the RMSE is approximately halved. On average, Edie's q is 60 veh/h under the FD.

Table 5.1. The performance with respect to q with Newell's model, over ten model runs.

Measure	Root Mean Square Error		Bias	
	Average	St. dev.	Average	St. dev.
CDN ←→ FD	382	48	35	23
CDN ←→ Edie	383	72	95	21
Edie ←→ FD	194	22	-60	10

Four types of data points can be distinguished:

1. Both CDN estimates and Edie's lie on (or very near) the FD.
2. Edie's lies on the FD, but CDN estimates do not.
3. Both CDN estimates and Edie's do not lie on the FD.
4. The CDN estimates lie on the FD, but Edie's do not.

The first type of data point is found when the traffic conditions in an xt-area are stationary and homogenous. Figure 3.2 graphically shows such a case. This can also occur when traffic is heterogeneous (i.e. different spacing), but only if all vehicles in the xt-area cruise at v_{free}.

The second type is found when there are two (or more) traffic states on the same branch of the FD in an xt-area. For example, if two different congested states are found in an xt-area, Edie's definitions find a weighted average of the two. Because the congested branch of the FD with Newell's model decreases linearly, this weighted average will also lie on the FD. The CDN method almost always provides an overestimated or underestimated q in these cases.

The third type of data point is found when a state on the free flow branch (below capacity) and the congested branch are present in an xt-area. Edie's q is again a weighted average of the two states, but lies below the FD. This explains the negative bias for Edie's q with respect to the FD in table 5.1, when Edie's q and k are not on the FD, they are found below the FD.

The fourth type of data point occurs only rarely and due to coincidence.

5.2 Estimation error with IDM

When traffic behaves according to the IDM (i.e. dynamically), CDN performs slightly better compared to when it behaves according to Newell's model. The average RMSE of q is 301 veh/h with respect to the FD. Using Edie's definitions yields roughly the same results with the two models, again it is only found on or below the FD.

It is mostly in xt-areas with two very different states, where CDN can have large errors. In Newell's model, traffic states change instantaneously, whereas in the IDM they change more gradually. Therefore, changes in traffic states in xt-areas are often less strong in the latter.

Table 5.2. The performance with respect to q with the IDM, over ten model runs.

Measure	Root Mean Square Error		Bias	
	Average	St. dev.	Average	St. dev.
CDN ←→ FD	301	66	11	22
CDN ←→ Edie	314	69	93	20
Edie ←→ FD	181	12	-74	11

5.3 Comparison with reference traffic state estimator

In both the Newell and IDM simulation, the RMSE of q is similar for CDN and loop detectors with respect to the FD. Using Edie's definitions in the developed method, yields a large improvement for FD estimation compared to using loop detectors. In Newell's simulation, the RMSE of q is reduced from 381 veh/h to 194 veh/h. In the IDM simulation, the RMSE of q is reduced from 278 veh/h to 181 veh/h.

The developed method also provides far more data about very congested states (i.e. states near jam density). In figure 5.1, it can be seen that detector loops provide little data about congested states, and no data about very congested states in this simulation study.

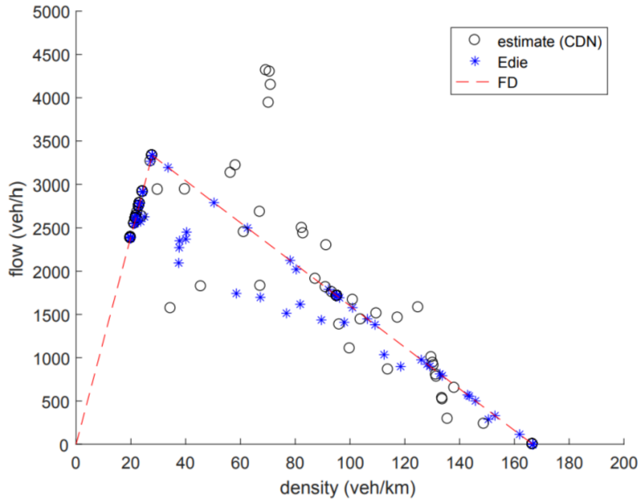
The loop detectors are better in approximating Edie's q than the CDN method. Induction loops simply count the number of vehicles that have passed it. The only errors in this simulation are made when vehicles are inside an xt-area, but are not detected, which are mostly short distances at the corners of the xt-area, it can therefore be expected that it approximates Edie's q well on average.

However, the CDN performs better when we look at Edie's q with respect to the FD. This may be caused by the form of the xt-areas with the developed methodology. It may be less prone to include multiple states, since the shape of xt-areas is more similar to the suggested xt-area shape by Laval (2011) more often (i.e. boundaries at wave speed and free flow speed). Another explanation relates to the way this simulation study is designed. With the loop detectors, some xt-areas are found that include the empty space downstream of the first vehicle. Therefore, in the congested period of the simulation, Edie's q is sometimes found in between the origin and a state on the congested branch in the qk-plane (figure 5.1b and d).

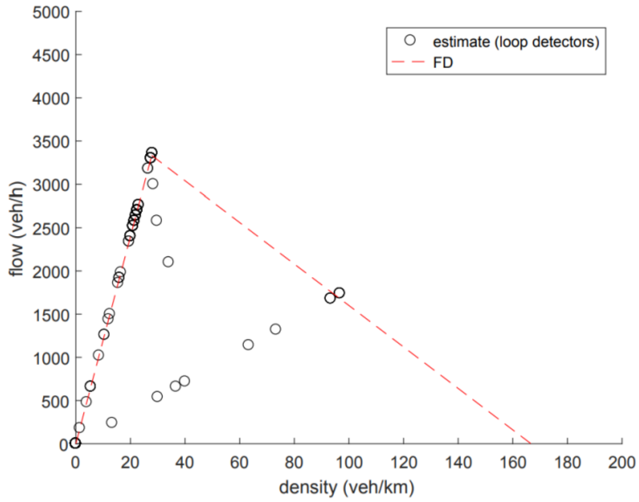
The upper boundary of the xt-planes with the developed method is always a trajectory, and hence does not include empty space in this simulation study. In reality, we may construct xt-areas with an upper boundary downstream of an empty space and a lower boundary upstream of it (e.g. at a signalized intersection). However, the developed method provides information about its boundaries that can be used cleverly to decide which xt-areas provide useful information for FD estimation. This is discussed in further detail in the next section.

Table 5.3. The performance with respect to q, using loop detectors, over ten model runs.

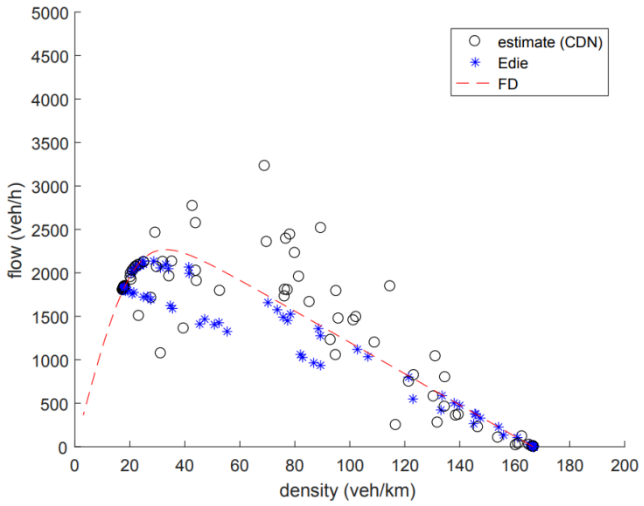
Measure	Root Mean Square Error		Bias	
	Average	St. dev.	Average	St. dev.
Newell: detector ←→ FD	381	59	-82	22
Newell: detector←→ Edie	111	16	1	1
Newell: Edie ←→ FD	485	31	-151	10
IDM: detector ←→ FD	278	21	-72	7
IDM: detector ←→ Edie	78	3	0	1
IDM: Edie ←→ FD	280	61	-81	15



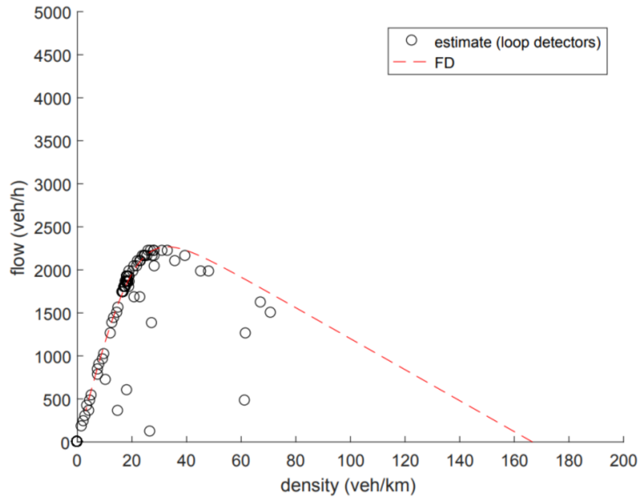
a) *Newell, developed method.*



b) *Newell, loop detectors.*



c) *IDM, developed method.*



d) *IDM, detector loops.*

Figure 5.1. Comparison with the FD. The CDN estimates and Edie's for constructed xt-areas, and detector loop estimates for one model run.

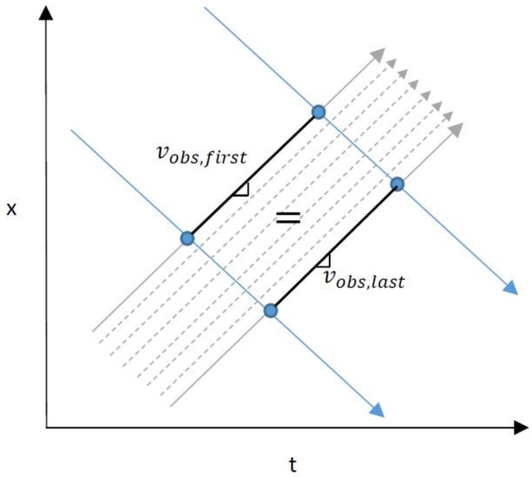
5.4 Using information of the boundaries

The xt-areas that we construct provide the information required to calculate Edie's generalized definitions. Therefore we have the possibility to decide between the CDN method and Edie's for each xt-area. For this purpose, we can make use of the distinction of four types of data points in section 5.1.

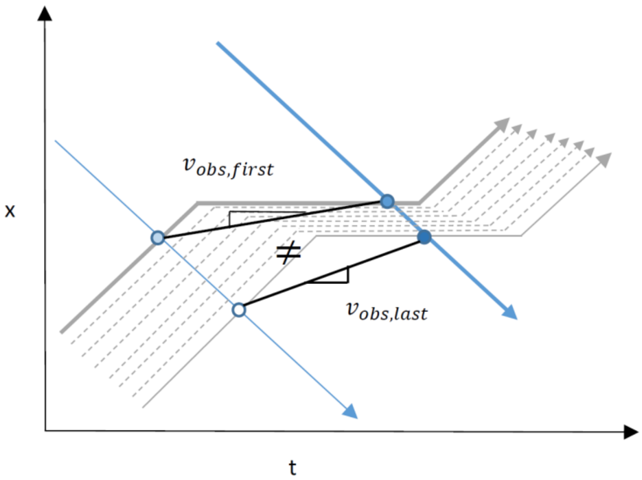
For xt-areas with data points of the first type, (i.e. both CDN estimates and Edie's lie on the FD), we could make use of the (simpler) CDN method. To distinguish these data points, we can use the fact that the first type of data point is found for stationary and homogeneous xt-areas, and for heterogeneous xt-areas in the free flow branch.

What these xt-areas have in common, is that the speeds do not differ between vehicles. For this purpose, we can use the fact that we know v_{obs} of both the first and last vehicle (i.e. both moving observers) of each xt-area. In Newell's model, we can decide to use the CDN method only for xt-areas where $v_{obs,first} = v_{obs,last}$ (figure 5.2a). For those xt-areas, CDN provides a perfect estimate of q and k on the FD, the RMSE of the q-estimate is practically zero (in the simulation with Newell's

model). After ten model runs, on average 38% of xt-areas complies with this criterion, and states both in the free flow branch and congested branch are found.



a) The average speed of the first moving observer equals the speed of the last moving observer.



b) There is a speed difference between the first and last moving observer.

Figure 5.2. Two examples of xt-areas.

When drivers behave according to the IDM, speeds are virtually never exactly equal. Therefore, instead we can set a threshold speed difference (v_{th}). If the speed difference between $v_{obs,first}$ and $v_{obs,last}$ is larger than v_{th}, we can decide not to include this xt-area. Similarly, we find a reduction in the RMSE. With $v_{th} = 1.0$, the RMSE of q with the CDN method with respect to the FD is reduced from 301 veh/h to 111 veh/h compared to having no v_{th}, while still 66% of xt-areas are included (figure 5.3 and table 5.4).

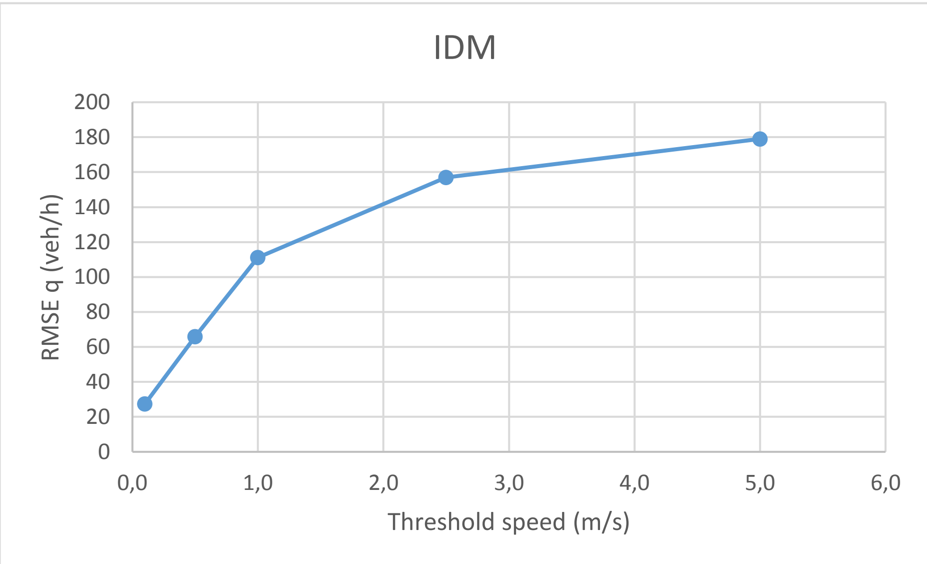


Figure 5.3. The average RMSE of q with the CDN method to the FD, when only xt-areas with a difference between $v_{obs,first}$ and $v_{obs,last}$ lower than v_{th} are included. Five IDM model runs are performed with each value of v_{th}.

Table 5.4. The percentage of xt-areas where the difference between $v_{obs,first}$ and $v_{obs,last}$ is lower than v_{th} in the IDM simulation.

v_{th}	% xt-areas
0.1	29%
0.5	58%
1.0	66%
2.5	76%
5.0	83%

6. Conclusions and discussion

In this paper, we have developed a method to estimate combinations of q and k on the FD, using only connected vehicles as moving observers. We construct xt-areas using moving observers as boundaries, q and k in each xt-area can be found with Edie's definitions or estimated with our CDN method.

If we apply CDN to all constructed xt-areas, the method has an estimation error (RMSE of q with respect to the FD) similar to the reference traffic state estimator, which makes use of loop detectors. It does, however, provide more information about the congested branch. Especially states near jam density are found more often with our method. For this purpose, of course congestion needs to occur in the observation period.

When Edie's definitions are used to find the traffic states inside the constructed xt-areas, the RMSE of q with respect to the FD is reduced considerably, in comparison to both CDN and loop detectors. Edie's definitions always provide a combination of q and k on or below the FD. Hence, if Edie's definitions are used for FD estimation, it may be desirable to place the FD higher than a least squares fit on the data points would.

The CDN method performs better for FD estimation when traffic states change dynamically (i.e. IDM) than when they change instantaneously (i.e. Newell's model). The CDN method can have large estimation errors when traffic states change very strongly inside an xt-area, which is more likely to happen if traffic states change instantaneously.

The boundaries of each xt-area provide information about its adequacy for FD estimation. For each boundary to an xt-area (i.e. trajectory of a moving observer), we know the average speed. If we use the CDN estimate only for xt-areas that have an upper and lower boundary with approximately the same average speed, the RMSE of q can be considerably reduced.

The CDN method shows promising results for FD estimation: its estimation error is relatively low, it provides a large amount of useful data about very congested states, and the data we collect for it provides information about the adequacy of each xt-area for FD estimation. Furthermore, it only requires vehicles with sensing equipment. Therefore, it can estimate the FD on road sections where loop detectors are out of order or not installed.

7. Future directions

In this section we discuss several possible extensions of the method.

We suggest to search for more clever ways to choose a point of q and k based on the intersection points in the qk-plane. Right now, we choose the middle point of four intersection points. We hypothesize that the fact that ΔN is known between any two points in each xt-area is very useful for this. Multiple lines can be constructed inside an xt-area that can provide more information about the traffic state(s) inside the xt-area.

The fact that ΔN is known between any two points in an xt-area, allows us to correct potential observation errors (van Erp et al., 2017a). Moving from one point on the boundary to another point on the boundary is possible in two directions (clock-wise and counter-clock-wise). Both should lead to the same; however, in case of observation errors, there may be a difference between the two directions. Based on this difference and characteristics of the observation paths, we may correct for errors and work with more accurate information on ΔN.

A strength of the approach is that the moving observers in the opposite direction do not need to cruise at wave speed. This paper has shown promising results when the moving observers cruise at v_{free}. While it is more likely to have many observations at v_{free} than at wave speed, in reality we will find moving observers at many different speeds. Therefore, to test the adequacy of the method for FD estimation in the field, the estimation error should be investigated when traffic at the opposite stream moves at other speeds, and when it is non-stationary and heterogeneous.

In section 5.4 we discussed choosing which xt-areas to use for FD estimation with the CDN method. For the remaining xt-areas, we can use Edie's generalized definitions. However, preferably we can distinguish which data points are of the second type (i.e. Edie is on the FD, method is not) and of the third type (both Edie and the method are not on the FD). For this purpose, it is required to know if states in an xt-area are on the same branch of the FD. One could work with a prior FD. If both lines of the left stream observers (i.e. the blue lines in figure 3.3) intersect with this prior FD in the congested branch, we could use Edie's definitions to find q and k. The remaining xt-areas would not be used for FD estimation.

References

Cassidy, M.J. (1998). *Bivariate relations in highway traffic.* Transportation Research Part B, Vol. 32, No. 1, pp. 49–59.

Chiabaut, N., Buisson, C., Leclercq, L. (2009). *Fundamental Diagram Estimation through Passing Rate Measurements in Congestion*. IEEE Transactions on Intelligent Transportation Systems, Vol. 10, No. 2, pp. 355-359.

Coifman, B., Redmill, K., Yang, R., Mashalani, R., McCord, M. (2017). *Municipal Vehicles as Sensor Platforms to Monitor Roadway Traffic*. Transportation Research Record, Vol. 2644, pp. 48-54.

Daamen, W. (Ed.), Buisson, C. (Ed.), Hoogendoorn, S. (Ed.) (2014). *Traffic Simulation and Data*. Boca Raton: CRC Press.

Dervisoglu, G., Gomes, G., Kwon, J., Muralidharan, A., Varaiya, P., Horowitz, R. (2009). *Automatic calibration of the fundamental diagram and empirical observations on capacity.* Transportation Research Board 88th Annual Meeting.

Drake, J. S., Schöfer, J.L., May, A. (1967). *A Statistical Analysis of Speed Density Hypotheses*. In Proceedings of the Third International Symposium on the Theory of Traffic Flow (Edie, L.C., Herman, R., Rothery, R., eds.). Elsevier North-Holland, New York, 1967.

Edie, L. C. (1965). *Discussion of traffic stream measurements and definitions*. In 2nd Int. Symp. on the Theory of Traffic Flow. Paris: OECD.

Gerlough, D.L., Huber, M.J. (1975). *Traffic flow theory: a monograph.* Washington: Transportation Research Board, National Research Council.

Knoop, V.L., Hoogendoorn, S.P., van Zuylen, H. (2009). *Empirical Differences Between Time Mean Speed and Space Mean Speed*. In Traffic and Granular Flow '07 (Appert-Rolland, C., Chevoir, F., Gondret, P., Lassarre, S., Lebacque, J.P., Schreckenberg, M., eds.). Springer, Berlin, Heidelberg.

Knoop, V.L., Daamen, W. (2014). *Automatic fitting procedure for the fundamental diagram.* Celebrating 50 Years of Traffic Flow Theory. Portland, Oregon. August 11-13, 2014.

Laval, J.A. (2011). *Hysteresis in traffic flow revisited: An improved measurement method*. Transportation Research Part B, Vol. 45, pp. 385–391.

Makigami, Y., Newell, G.F., Rothery, R. (1971). *Three-Dimensional Representation of Traffic Flow*. Transportation Science, Vol. 5, No. 3, pp. 302-313.

Newell, G.F. (2002). *A simplified car-following theory: a lower order model.* Transportation Research Part B, Vol. 36, pp. 195-205.

Sao, T., Kusakabe, T., Asakura, Y. (2017). *Calibration of fundamental diagram using trajectories of probe vehicles: Basic formulation and heuristic algorithm.* Transportation Research Procedia, Vol. 21, pp. 6-17.

Treiber, M., Hennecke, A., Helbing, D. (2000a). *Congested Traffic States in Empirical Observations and Microscopic Simulations*. Physical Review E, Vol. 62, pp. 1805–1824.

Treiber, M., Hennecke, A., Helbing, D. (2000b). *Microscopic Simulation of Congested Traffic*. In Traffic and Granular Flow ’99 (Helbing, D., Herrmann, H.J., Schreckenberg, M., Wolf, D.E., eds.). Springer, Berlin, Heidelberg.

Van Erp, P.B.C., Knoop, V.L., Hoogendoorn, S.P. (2017a). *Macroscopic traffic state estimation: Estimating the state based on cumulative vehicle number point-observations in space-time.* August 3, 2017.

Van Erp, P.B.C., Knoop, V.L., Hoogendoorn, S.P. (2017b). *On the value of relative flow data.* December 1, 2017.

Wardrop, J.G., Charlesworth, G. (1954). *Method of Estimating Speed and Flow of Traffic from a Moving Vehicle*. Proceedings of the Institution of Civil Engineers, Vol. 3, pp. 158-169.