DELFT UNIVERSITY OF TECHNOLOGY

Incorporating the information and uncertainties of loop-detector and floating car data in freeway traffic state estimation

by

Paul B.C. van Erp

A thesis submitted in partial fulfillment for the
degree of Master of Science in Civil Engineering

in the
Faculty of Civil Engineering and Geosciences
Department Transport and Planning

May 29, 2015

Summary

Traffic management measures may be implemented to make use of the road traffic network as efficient as possible. Therefore, it is an important field of research for traffic researchers. Potential traffic management measures first have to be proven theoretically. However, before the decision is made to implement these measures in real-life, their effectiveness has to be tested.

Road traffic involves many different actors. For instance, different authorities are responsible for different parts of the road traffic network and traffic data is available at multiple, both public and private, organizations. A collaboration between these actors yields advantages. Not only for the actors, but also for the road users. In Amsterdam, the Netherlands, a unique project was set-up in which different public and private organizations work together to test traffic management systems in a field operational test. In this research this will be denoted by FOT Amsterdam.

The FOT Amsterdam consist of multiple phases of which the first has been successfully completed. The traffic management system tested in this first phase aims to prevent the congestion at the freeway as this negatively affects the maximum flow (capacity). In short, the traffic management system observes the traffic conditions on the roads, evaluates if breakdowns are likely to occur and if needed takes control action. This research focuses on a single (monitoring) component of the entire system, namely the freeway traffic state estimator. As the name states, this component estimates the traffic conditions on the freeway using the available data.

In the first phase of the FOT Amsterdam the estimator only Loop-Detector Data (LDD) is utilized. LDD is gathered by sensors (loops) which have fixed road locations and is therefore denoted as Eulerian sensing data. At the location where loop-detectors are installed, for each individual lane, the one-minute aggregated speed and flow are available. In the second phase of the FOT Amsterdam a new data-type becomes available, namely Floating Car Data (FCD). Potential sources of FCD are mobile phones and navigation systems. In contrast to Eulerian sensors, which have a fixed road location, these sensors follow individual vehicles and FCD is therefore denoted as Lagrangian sensing data. The potential availability of this data is a consequence of the cooperative nature of the FOT Amsterdam. Potential FCD-providers are private companies. In a meeting with FCD-providers the characteristics of this data-type were discussed. It is expected that the individual vehicle speed and location may be available for up to 20 % of the road users. The percentage of road users which is observed by means of these sensors is called the penetration rate.

This research investigates the added value of FCD for freeway traffic state estimation. Potentially this may lead to an improved estimation performance. However, the objective is not to achieve the highest possible accuracy. The freeway traffic state estimator should yield estimations which are sufficiently accurate for the entire traffic management system to function properly. In this way involving FCD may lead to a reduction of the required LDD. This is interesting for the road authorities as installing and maintaining loop-detectors comes with certain costs, which may be lowered using FCD.

The output of the freeway traffic state estimator should consist of estimates for the space-mean speed, flow and density. These estimates are used in other components of the traffic management system. The data, both LDD and FCD, provides information related to the space-mean speed, flow and density. However, the data does not directly provide all the correct variables. For instance, the space-mean speed is desired, but the LDD provides the time-mean speed and the FCD provides individual vehicles speeds of limited number of road users. Furthermore, the LDD provides one-minute aggregates while shorter periods may be desired, which may have a different value.

In short, estimating the space-mean speed, flow and density directly from the data comes with uncertainties. This holds both for the estimates based on LDD and FCD. However, it is expected that uncertainties of the FCD are highly dependent on the fraction of observed vehicles and traffic conditions. There are differences in driving behavior of road users, which is denoted as driver heterogeneity, the mean speed for a select group of road users may not be representative for the true mean speed. This difference, the driver heterogeneity, depends on the traffic conditions. As one may imagine when there is less traffic on the road the fast drivers are not restricted and larger speed differences are observed.

The freeway traffic state estimator implemented in the first phase of the FOT Amsterdam is based on the Adaptive Smoothing Method (ASM). This methodology uses knowledge about the way traffic conditions, for instance congestion, propagates over time and space. However, it does not implement any physical laws like the conservation of vehicles, which means that vehicles cannot disappear or originate out of nothing. The ASM is thus not based on a traffic flow model and therefore not denoted as a model-based methodology. For the FOT Amsterdam, it may however be of added value to use a model-based methodology. This would allow to integrate the freeway traffic state estimator with other components of the traffic management system and may lead to an improved performance of the system as a whole. For instance, the traffic state estimator may be extended to make short-term prediction for the traffic conditions or breakdown probability. For this reason and other benefits explained below, it is decided to design a model-based traffic state estimation methodology.

Model-based traffic state estimation consists of three components, namely a traffic flow model, an observation model and an assimilation technique. In the observation model, the desired traffic flow variables are estimated based on the LDD and FCD. The traffic flow model describes the dynamics of the traffic. Based on the current traffic conditions, the traffic conditions may be predicted using such a model. It is chosen to use the second-order traffic flow model METANET. This model provides equations to make predictions for the density and space-mean speed. To predict the density it incorporates the physical law of conservation of vehicles. Similarly to the data-based estimations of the space-mean speed and flow, the model-based predictions have a degree of uncertainty.

The assimilation technique should be able to combine the information of the observation model and traffic flow model. For this purpose an Extended Kalman Filter (EKF) will be applied. A Kalman Filter estimates the traffic conditions in two steps. Firstly, the traffic conditions are predicted using the traffic flow model. Secondly, these predictions are corrected using the estimation from the observation model. This correction is dependent on the uncertainties in both models. As one may imagine, estimations with a large uncertainty are trusted less than those with a small uncertainty. The estimations which are trusted the most, thus with the smallest uncertainty should be assigned the largest weight in the correction step.

Using a model-based traffic state estimator with an EKF is not new in the field of traffic state estimation. Therefore an existing methodology is used as a starting point for the functional design of the proposed traffic state estimation methodology. The existing methodology only involves LDD. Two important changes are proposed with respect to the existing methodology. Firstly, in the existing methodology the two-steps denoted above, prediction and correction, are taken together. These are separated in the proposed methodology as the predicted traffic conditions will be used to determine the uncertainties of the FCD-based space-mean speed estimations. This leads to the second change. As explained above it is expected that the uncertainty of FCD-based estimates depends on the driver heterogeneity and the penetration rate. These may both be estimated based on the predicted traffic conditions, namely respectively from the space-mean speed and density estimates. This is used in combination with the measurements to determine the FCD-based estimation uncertainty. In line with the existing methodology, the uncertainties of the estimations obtained from the traffic flow model and LDD are fixed. If both LDD and FCD are available for a single cell, the combined estimate and uncertainty is determined using the individual estimates and uncertainties.

To evaluate the proposed methodology and the added value of FCD for traffic state estimation, experiments are conducted with synthesized data. In addition to the proposed methodology an ASM-based estimator which is able to use the FCD is designed. This serves two purposes. Firstly, the ASM-based estimator combined with the LDD availability in the first phase of the FOT Amsterdam serves as a reference. This yields the estimation performance which is accurately enough for the entire traffic management system to work properly. Based on this reference performance, the added value of FCD and potential reduction in required LDD may be evaluated. Secondly, it is an alternative approach to involve FCD. The performances of both methodologies may be compared for equal data availability.

The experiments show that FCD has a large added value to the proposed methodology, while the added value is limited for the ASM-based methodology. In the less complex ASM-based estimator the FCD only affects the space-mean speed estimates directly and thereby indirectly the density estimates. The flow estimates are unaffected. Furthermore, it already performs relatively well for space-mean speed estimation without FCD. Therefore, the potential of FCD in this estimation methodology is limited. FCD improves the estimation performance of the ASM-based estimator in congestion, while it may even lead to a decreased performance in free-flow estimation. The ASM-based estimator provides insight into the errors of estimation based on individual data-types. For FCD-based estimation, low penetration rates yield inaccurate estimates due to driver heterogeneity. This inaccuracy is smaller in congested conditions than in free-flow conditions. Furthermore, the inaccuracy decreases when the penetration rate increases. Both observations were expected and these dynamics are included in the proposed methodology in the FCD-based estimation uncertainty.

The proposed methodology benefits highly from incorporating FCD. This has four main reasons. Firstly, stability issues may occur when only fixed locations are observed, thus only LDD. The FCD which observes all locations over time is able to overcome this problem yielding more accurate estimations. Secondly, all interesting variables, thus space-mean speed, flow and density, may be affected by the FCD-based estimates through the traffic flow model. Thirdly, other cells and future estimations are affected through the traffic model. This may be explained intuitively. If the accuracy of the current estimations is improved the predictions which use these estimations as input are likely to be more accurate. Fourthly, the proposed methodology involves the knowledge of the uncertainties of the FCD-based estimates. It is shown that the uncertainties are indeed dependent on the penetration rate and driver heterogeneity. However, it seems that there are more, unexplained factors of uncertainty, which may be included in future research.

The proposed methodology shows promising results. In the experimental set-up a low penetration rate of 4 % was sufficient to reduce the number of required loop-detectors by more than half, while having a accurately enough estimation performance. In road stretches with less discontinuities, like ramps, the number of loop-detectors may be reduced even further. Furthermore, there is still room for improvement in the proposed methodology without having to involve more data.

In the meeting with potential FCD-providers, it was stated that a penetration rate of up to 20 % may be expected. If this is truly the case, in time, this may lead to a reduction in the installed loop-detectors at the Dutch freeways. Therefore, this research may be taken into account when making the long-term decisions related to the installment of loop-detectors.

Preface

In front of you lies my Master thesis and thus the final product of my Master at Delft University of Technology. This is the direct result of nine months of research and my previous experiences in Delft, Rotterdam and Berkeley. During this time I gained many insights and learned a lot from a large number of individuals, which unfortunately I cannot all thank personally.

During the summer of 2014, I approached Prof. Serge Hoogendoorn and Prof. Hans van Lint to discuss potential subjects for my Master thesis. I just returned from the University of California: Berkeley. Here I worked with different traffic data-types, however did not fully utilize the potential of these data-types. Therefore, for my Master thesis, I really wanted to exploit the potential of different traffic data-types. The result was the subject of this Master thesis. So thank you both.

First of all, I want to thank my committee consisting of Prof. Serge Hoogendoorn, dr. Victor Knoop, dr. Yufei Yuan, ir. Kai Yuan, ir. Paul Wiggenraad and dr. Mathijs Spaan, for there supervision and insights during our meetings. Serge Hoogendoorn, your connections and high level insights have made it possible to perform this research. Victor Knoop, you have challenged me to think like a researcher and perform this research as such. Yufei Yuan, your in-depth knowledge on traffic state estimation and Kalman Filters has helped me a lot while developing the proposed methodology. Kai Yuan, your curious mind and comments have helped to structure the report and explain my thoughts clearer. Paul Wiggenraad, thank you for helping me in the steps of the graduation procedure. Finally, thank you Mathijs Spaan for being my external reader.

Alongside the research on this thesis, I have been working part-time at MIcompany. In the end this combination was sometimes challenging and has lead to some evenings and weekend working on my thesis. For this I apologize to my room-mates, other friends and family.

To test the understandability and use of the English language, my sister Simone van Erp and dear friend Tom Armbrust have read important parts of this research. Thank you both for your contribution during this thesis and in earlier stages of my life. Finally and certainly not least, I would like to thank my parents for supporting me during my studies.

Paul van Erp

List of symbols

The list on the next page shows an overview of the important symbols used in this thesis.

Traffic flow symbols

ρ	Density	veh/km/lane
q	Flow	veh/h
r	On-ramp inflow	veh/h
s	Exiting flow	veh/h
β	Exiting rate	-
u^S	Space-mean speed	km/h
u^T	Time-mean speed	km/h
v	Individual vehicle speed	km/h

Estimation domain symbols

x	Location	km
i	Cell	-
t	Time	h
k	Period	-
Δ	Cell length	km
T	Period duration	h
λ	Number of lanes	-

Traffic flow model METANET symbols

v_f	Free-flow speed	km/h
rho_{cr}	Critical density	veh/km/lane
a	Parameter of fundamental diagram	-
τ	Time constant of the METANET speed relaxation term	-
κ	METANET speed anticipation term parameter	veh/km/lane
ν	METANET speed anticipation term parameter	km^2/h
δ	METANET parameter for the speed drop term caused by merging at an on-ramp	-
ξ	Model error	-

Observation model symbols

u^S_{fcd}	FCD-based space-mean speed estimate	km/h
u^S_{ldd}	LDD-based space-mean speed estimate	km/h
$\mu_{1,2}$	Mean of combined u^S PDF	km/h
$\sigma^2_{1,2}$	Variance of combined u^S PDF	$(km/h)^2$
η	Measurement error	-

Assimiliation technique symbols

$\mathbf{x}$	State vector	-
$\mathbf{y}$	Measurement vector	-
$\mathbf{P}$	State error covariance matrix	-
$\mathbf{Q}$	Model error covariance matrix	-
$\mathbf{R}$	Measurement error covariance matrix	-
$\mathbf{K}$	Kalman gain	-

Contents

Chapter 1

Introduction

In theory, traffic management measures have a lot of potential. However, the road towards empirically testing and implementing these measures is long and in many cases does not even lead to an implementation of the measure. In Amsterdam, the Netherlands, a unique traffic management project was set-up (Hoogendoorn et al., 2013). This public-private cooperation is called the The Field Operational Test (FOT) Integrated Network Management (INM) Amsterdam (Dutch: *Praktijkproef Amsterdam* or PPA for short). In the remainder of this report it will be denoted as the FOT Amsterdam.

1.1 Freeway state estimation in the Field Operation Test Amsterdam

The integrated element of the FOT Amsterdam makes this project unique. Parties responsible for different elements of the road-network, data providers, researchers and companies with other potential applicable contributions all come together. For instance, different authorities are responsible for national roads and the Amsterdam network. Furthermore, different types of traffic data are available at both public and private cooperation, which may be used together for traffic management. By working together and stepping over the individual boundaries, the FOT Amsterdam opts to show the potential of integrated network management.

The FOT Amsterdam aims to introduce traffic management measures to maintain or increase the capacity in the traffic network. These measures are introduced step-by-step over multiple consecutive phases. Currently, the first phase is finished and the research related to the second phase has started. In this research, the focus lies on one specific component of the hierarchical control scheme of the FOT Amsterdam, namely the freeway state estimator.

With respect to the first phase innovations related to the measurement data input will be incorporated in the second phase. The measurement inputs of the freeway state estimator in the first phase are purely from loop-detectors. In the second phase another data-type, namely Floating Car Data (FCD), will

be added. FCD is data which is obtained from vehicle specific data-sources, like navigation systems or mobile phones. Normally it is owned by private parties, which also determine the characteristics of the data they are willing to share. In contrast to the Loop-Detector Data (LDD), the road authorities have no investment and maintenance costs for FCD. This makes, in combination with the difference in data-characteristics, FCD an interesting data-type to be added.

For effective control by the hierarchical control system, the freeway state estimator output has to fulfill certain requirements. These are denoted in the required variable(s), variable accuracy and granularity in time and space. It is important to note that the objective is not to get the most accurate estimations, but to get estimations which are accurate enough to fulfill the requirements, while also keeping the cost limited.

In summary, the objective of this research is twofold. Firstly, we want to evaluate how the freeway traffic conditions may be estimated using a combined input of LDD and FCD. Secondly, we want to test if it is possible to maintain a desirable accuracy for the freeway state estimator when reducing the LDD availability. The first and second objectives are respectively methodology-driven and data-driven.

To fulfill these objective, the following main research question with corresponding sub questions are formulated. The main research question is:

How can we estimate the freeway traffic state accurately enough by balancing LDD and FCD?

The term *accurately enough* may seem vague, but actually is key in this research. The freeway state estimator is one element within the FOT Amsterdam traffic management system. The performance of the system as a whole is important. For effective control, a certain level of estimation performance is required of the freeway traffic state estimator. In order words, the freeway traffic state estimations have to be *accurately enough*.

Before this question can be answered, some steps have to be taken which follow the following sub research questions:

1. What is the role of the freeway traffic state estimator in the FOT Amsterdam? (estimator output)
2. What are the characteristics of the existing LDD and potentially available FCD? (estimator input)
3. How do the data-characteristics relate to the macroscopic traffic flow variables? (variable relations)
4. Which existing methodologies are suitable for freeway traffic state estimation with heterogeneous data-types? (existing methodologies)
5. How can we utilize and potentially improve the existing traffic state estimation methodologies to combine the information and uncertainties of the LDD and FCD? (methodology development)
6. How can the different features of the proposed methodology be evaluated? (experimental set-up)
7. Which combinations of the data input characteristics should be tested in order to answer the main research question? (experimental set-up)

1.2 Research contributions

This research has four main contributions.

- **New method to estimate traffic state using heterogeneous data-types**: In this research a methodology is proposed to utilize the information of loop-detector and floating car measurements. The key principle is expressing the macroscopic traffic variables space-mean speed and flow as a Probability Density Function (PDF) based on the measurements. If there are multiple PDFs for the same variable in a segment and period, these are fused to a single PDF. Although these two data-types are used, the methodology allows any other data-type to be added. Therefore, this is interesting for traffic state estimation application for which other data-types are available. *(Scientific contribution)*
- **Determining the FCD-based measurement error variance on the measurements and model-based state estimates:** The proposed methodology applies the Extended Kalman Filter (EKF). In short, this assimilation technique has two steps, namely the model-based prediction step and the correction step. The correction factor used in the correction step depends on the expected measurement and model errors. In the proposed model the uncertainties of the FCD-based estimations are based on the actual measurements. For the FCD, important parameters for the accuracy of the space-mean speed estimates are the penetration rate and traffic conditions. Both are unobserved, however these are estimated by means of the model-based predictions. *(Scientific contribution)*
- **Providing a valuable estimation of the final estimation uncertainty:** The state error covariance matrix, which contains the uncertainties of the state vector estimations, may be an interesting input variable for other components of the FOT Amsterdam hierarchical control scheme. The FCD-based measurement error variance is based on empirically studied traffic relations and the other error variances are scaled with respect to this error variance. Therefore, the values in the state error covariance matrix may show uncertainties in the order of magnitude of the true estimation uncertainties. This makes the values easier to interpret and may provide better information as input for other applications. *(Scientific/practical contribution)*
- **Developing a model-based freeway traffic state estimator for the FOT Amsterdam**: The developed freeway traffic state estimator satisfies the objective to include FCD and is a model-based approach. This makes an integral network approach (estimation, prediction and control optimization) possible with other components of the FOT Amsterdam hierarchical control scheme. *(Practical contribution)*

1.3 Outline of this thesis

The outline of this thesis is visualized in Figure 1.1. The core-components are included in the dashed box. In the literature review (Chapter 2), the role of the traffic state estimator within the FOT Amsterdam will be discussed together with the expected data characteristics of LDD and FCD. This information serves as

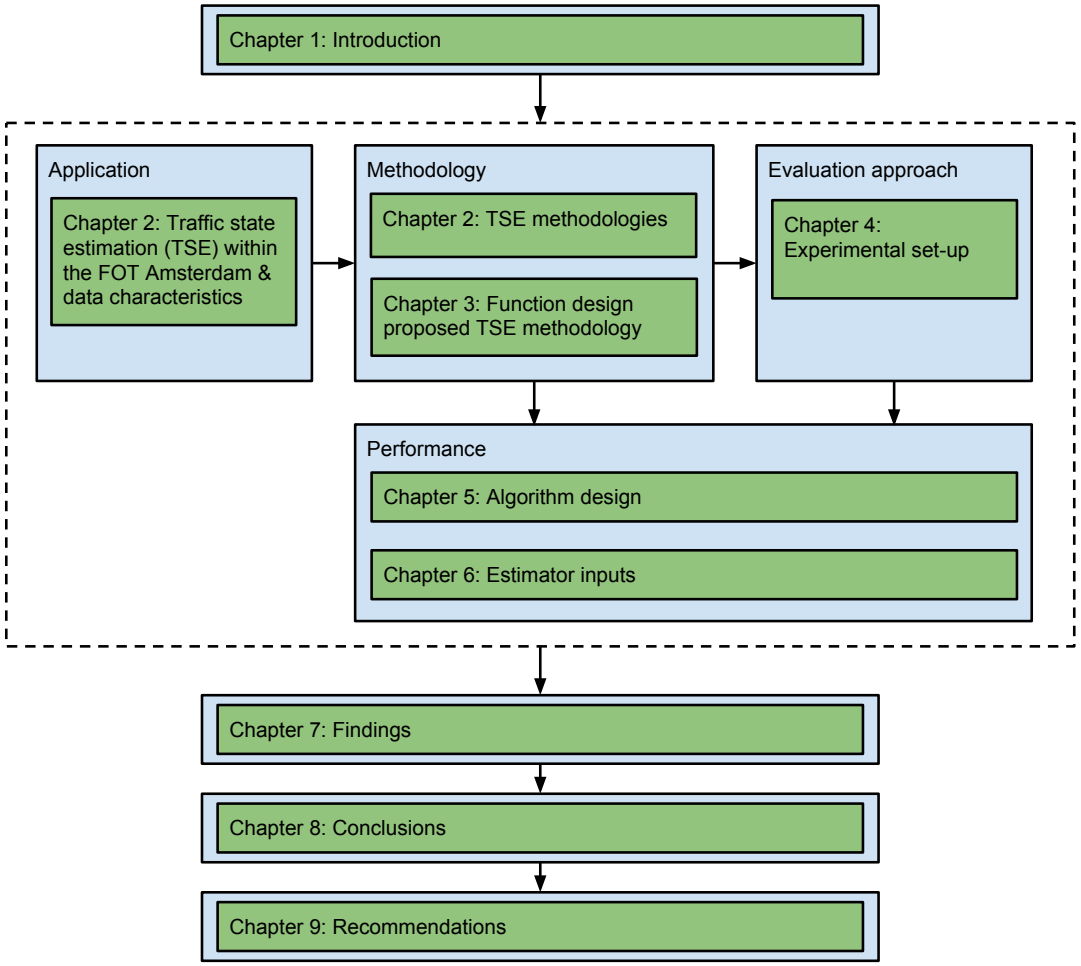


FIGURE 1.1: Visualization of the outline of this thesis.

an input for the functional design of the proposed traffic state estimation methodology which is provided in Chapter 3. An algorithm will be designed in Chapter 5 based on the proposed methodology and the experimental set-up (evaluation approach) discussed in Chapter 4. The algorithm design is based on the experimental set-up as it should be flexible in doing the experiments. In Chapter 6, the choices related to estimator inputs are made. These involve the evaluated freeway and estimator parameters. This is the final step before the experiments are conducted. The findings based on the experimental results are discussed in Chapter 7. These serve as input to answer the main research question. The conclusions of this researchare provided in Chapters 8. Finally, in Chapter 9, some recommendations are given for the application and future research.

Chapter 2

Literature review

By means of a literature review important information in three main fields is collected. Firstly, the required estimator output follow from the review. Secondly, the characteristics and the (theoretical) relations of the expected estimator input (data) with the desired output are found. Finally, different traffic state estimation and data fusion methodologies are discussed. Combined, this information leads to an initial selection of the traffic state estimation and data fusion methodology.

In Section 2.1, the role of the freeway state estimator in the FOT Amsterdam is discussed. Traffic state estimation is addressed in Section 2.2. In this section a three-way split is made into the observation (data) model, traffic flow model and assimilation technique. Finally, the initial methodology selection is made in Section 2.3.

2.1 The role of the freeway state estimator in the FOT Amsterdam

In this research a traffic state estimator is designed to function as a component in a traffic management system, namely within the traffic management system of the FOT Amsterdam. The objective of this system is provided by Hoogendoorn et al. (2013).

1. To prevent the capacity drop at the freeway.
2. To prevent spill back at the freeway and urban arterials.
3. To use spare storage capacity in the network given the prevailing Level of Service (LoS) of the subnetwork and the archetype situation (traffic problem) that requires a specific control action.
4. To solve problems locally if possible, and to coordinate (escalate) only when needed.

In Figure 2.1 the monitoring components in the hierarchical control system of the FOT Amsterdam are shown. The freeway state estimator is one of the traffic monitoring systems in the FOT Amsterdam.

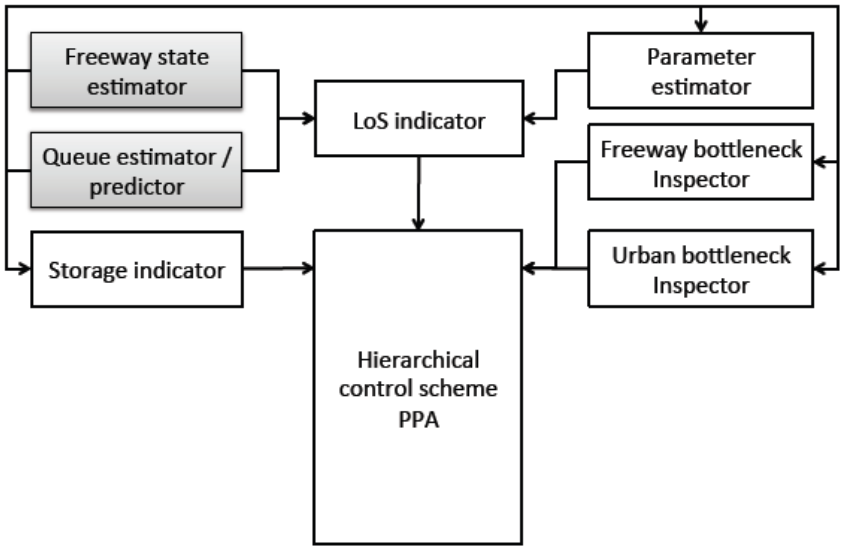


FIGURE 2.1: Simplified scheme of the monitoring components in the hierarchical control system of the FOT Amsterdam (H. Van Lint et al., 2014)

H. Van Lint et al. (2014) provide the components and how they work. The outputs of the freeway state estimator in the first phase of the FOT Amsterdam were speed, flow and density contours. The objective is to make continuous and unbiased state estimators for these three variables. These estimates are made with the Adaptive Smoothing Method (ASM) (Treiber & Helbing, 2002), (J. Van Lint & Hoogendoorn, 2010).

The freeway state estimator outputs are used by the Level of Service (LoS) indicator, the parameter estimator and the Freeway Bottleneck Inspector (FBI). The LoS indicator also utilizes the urban queue estimator to provide a network LoS indicator based on the average density and its spatial variance (Knoop & Hoogendoorn, 2012). In the parameter estimator, the critical density is estimated. This parameter is the key parameter for ALINEA ramp metering control scheme proposed by Papageorgiou et al. (1991). The last element, the FBI, should predict the breakdown probabilities at locations where breakdowns are historically likely. Additionally, the FBI is able to estimate the head, body and tail location of (already set in) congestion. These elements and those related to the urban queue estimator, provide the inputs of the hierarchical control scheme of the FOT Amsterdam.

It is clear that the freeway state estimator plays an important role with the FOT Amsterdam. The speed, flow and density are all used as inputs for the other elements in the FOT traffic management system. Therefore, the freeway state estimator should provide estimates for all three variables. The current state estimator is able to provide continuous state estimates. However, discrete state estimates may also be used. In the methodologies currently used in the LoS indicator (Knoop & Hoogendoorn, 2012) and parameter estimator/control scheme (Smaragdis et al., 2004) the road is divided in discrete road segments.

The performance of the traffic management system is dependent on the performance of all individual components and the communication between components. The performance of a traffic state estimation may thus be expressed on multiple fields. Firstly, by comparing the estimated values with the true values

using an accuracy statistic. Secondly, as stated by H. Van Lint et al. (2014) it would be of added value to have a traffic state estimation methodology allowing an integral network approach. Compared with the current system, the communication from the traffic state estimator to other components may even be enhanced, for instance by considering the uncertainties of our estimations.

As the total system performance is decisive, assigning a specific value to the performance of the freeway state estimator which suffices is difficult if not impossible. However, it is possible to compare the accuracy of the current traffic state estimator methodology and inputs with the new methodology and inputs. Here the current estimator is based on the ASM and uses LDD with a certain availability (or spacing). Given the new methodology and new inputs (FCD), the amount of LDD required to achieve the current accuracy may be determined.

2.2 Model-based traffic state estimation

In addition to the added value of model-based estimation within the FOT Amsterdam, there is another data-related potential advantage. As will be explained in Section 2.2.1, FCD only provides speed information. If a model-based traffic state estimation methodology is used all macroscopic variables can be affected by the FCD[1].

Model-based traffic state estimation consists of three components, namely a dynamic traffic model, an observation model and an assimilation technique (Yuan, 2013). These components may be categorized in respectively the X- Y- and Z-dimensions (Ou, 2011).

2.2.1 The observation model

In this research, we opt to develop a traffic state estimator which may be used in practice in the near future. Therefore, only potentially available data-types will be considered. As the name denotes, the FOT Amsterdam freeway state estimator will be designed for a Dutch freeway. Therefore, the loop-detectors installed on these roads are considered. The FCD characteristics are based on a meeting with the parties involved in the FOT Amsterdam. A comparison between the two data-types is shown in Table 2.1.

Loop-detector and other observations from sensors with fixed locations is denoted as Eulerian sensing. The double loop-detectors installed on the Dutch freeways should observe the time at and with which speed every vehicle passes. This raw data is automatically processed to compute time averaged speeds and flows with embedded routines (H. Van Lint et al., 2014). As results, the LDD consists of one-minute lane-specific aggregated speeds and flows. The average speeds is the time-mean speed u^T, which is the average value of all measured individual vehicle speed within the considered period. The flow is equal to the number of vehicles passing the loop-detector within the considered period. Note that the LDD does not contain any information related to the vehicle IDs. As a consequence, separate observations of consecutive loop-detectors cannot directly be linked.

[1] In this research the proposed methodology will be compared with a ASM-based methodology, which is discussed in Appendix C. In the latter methodology FCD only affects the u^S estimations and indirectly ρ, while q is unaffected.

TABLE 2.1: Characteristics of the expected LDD and FCD. These characteristics are based on H. Van Lint et al. (2014) and a meeting with potential FCD-providers.

	Loop-detector data	Floating car data
Traffic variables	Time-mean speed u^T Flow q	Vehicle speed v_n
Time/period	One-minute aggregates	Point observations
Location	Fixed	Variable
Lane-information	Yes	No
Vehicle ID	No	Yes
Penetration rate	100 %	0-20 %

Observations from sensors following vehicle trajectories, like FCD, is denoted as Lagrangian sensing data. The available FCD provides map-matched location (based on GPS-sensing) and speed estimates for a fraction of the vehicles. This fraction is denoted as the penetration rate. The term 'map-matched' means that the location is given in terms of the road location, in contrast to original GPS location[2]. Due to inaccuracies in the GPS location, the FCD-provider is not able to provide the lane number of the observation. Here the FCD thus differs from the LDD. Another difference is that the floating car observations contain a vehicle ID, thereby making it possible to link individual observations. Depending on the vehicle (equipment) the update-frequency differ.

The traffic variables available in the LDD and FCD are directly related to the macroscopic traffic variables space-mean speed u^S, flow q and density ρ. For LDD, assigning the LDD to periods smaller than the one-minute aggregation period comes with inaccuracies. This hold for both q and u^T. Both LDD and FCD contain information related to u^S, however this variable is not available in the raw data. Due to driver and vehicle heterogeneity, thus different individual vehicle speeds, faster vehicles are over-represented in the u^T with respect to the u^S. As the u^S is desired the LDD provides a biased speed estimation. To reduce to error u^S may be derived from the lane-specific u^T and q (Knoop & Hoogendoorn, 2012). For the FCD speed estimate, it is a problem that only a fraction of the vehicles is observed. Due to driver and vehicle heterogeneity the mean of this limited set of individual speeds v does not have to equal u^S. It is not possible to improve this estimation by using lane information, as this is not available for the FCD.

In addition to inaccuracies in the estimation of the macroscopic variables based on the raw data, there may be measurement errors in sensing data. This also holds for the LDD and FCD. There may be errors in the individual vehicle speed measurements by double loop-detectors or mobile phones (FCD). Furthermore, loop-detectors may have missing or double observations, for instance when a vehicle is switching lanes at the loop-detector location. Another possible error is the FCD-location error which was already briefly discussed and making it impossible to assign a lane to the floating car observation.

As a result of potential measurement errors and uncertainty in the relation between the observations and local macroscopic traffic variables, the observation actually lead to a Probability Density Function (PDF) of the macroscopic variables instead of a single value estimate. For a given data-type, either LDD or FCD, the parameters of the PDF depend on the measurement equipment and the observations. For

[2]GPS provides the location in longitude and latitude. The FCD-provider derives the location on the Dutch road network from this raw GPS-data.

instance, consider a road stretch with a certain length at a moment in time. At this moment there are N vehicles present on this road stretch of which we observe n vehicles via floating car measurements. In this case the penetration rate is n/N. The individual speed estimates have a potential error which is based on the quality of the measurement system. As one can imagine, a higher penetration rate[3] of the FCD and higher accuracy of the measurement system leads to a more certain estimate of the space-mean speed.

The q is only estimated by LDD, while u^S estimates may be based on both LDD and FCD. Individually, these data-types may provide a u^S PDF. If both types are available for a given time and space, two separate PDFs are available. These may be combined to form a single PDF (Bromiley, 2003).

2.2.2 Assimilation technique: the Extended Kalman Filter

H. Van Lint et al. (2014) states that there are many benefits to a model-based control approach for the FOT Amsterdam. The most important benefit is the possibility of an integral network approach to estimation, prediction and control optimization. Examples of such an approach are Wang & Papageorgiou (2005) and Van Hinsbergen et al. (2012), which both use an Extended Kalman Filter (EKF) to combine the information in traffic flow models and sensor data.

Filters, like the different versions of the KF, are capable of weighing the estimates obtained from traffic flow models and sensor data based on their uncertainties. This capability is already valuable when only LDD is included. However, it becomes more important when the uncertainties are more dependent on the data characteristics. As argued in Section 2.2.1, the penetration rate is an important data characteristic which influences the uncertainties of FCD-based estimates. Therefore, the assimilation technique should correctly incorporate the data-dependent uncertainties.

Yuan (2013) compared different recursive assimilation techniques with existing traffic state estimation applications. In this overview[4], he compares four versions of the Kalman Filter (KF), namely the original KF, the Extended KF (EKF), Ensemble KF (EnsKF) and Unscented KF (UKF), and the Particle Filter (PF). The EKF is argued to be most suitable for real-time traffic state estimation. He states that there are two classical arguments for using more involved and theoretically superior assimilation methods. These are that the EKF uses an inaccurate linearisation for highly nonlinear problems and that the Gaussian assumptions for the noise terms (inaccuracies) is invalid. The former may hold and this could be a problem for traffic flow models with mode (free-flowing or congested) switching, which are highly nonlinear around capacity. However, other (continuous) models may be chosen, such as the METANET model (Papageorgiou et al., 1990), to overcome this problem. in reality the second classical argument may be valid, however Yuan (2013) list earlier applications which have shown that the Gaussian assumption provides good approximations. Based on this earlier research it is therefore concluded that the Gaussian may be used. Furthermore, a downside of the more involved and theoretically superior assimilation

[3] Although this fraction is unknown based on the FCD, it may be estimated using the methodology proposed in the next chapter.

[4] Van Hinsbergen et al. (2012) also provides a short discussion on the different Kalman Filters. Based on similar arguments as Yuan (2013), they also chose the EKF. This research follows their comparison. Although the choice is made to involve the EKF, components of the proposed methodology may also be used in combination with other filters.

methods (EnsKF, UKF, PF) is the computationally inefficiency compared to the EKF. This makes them less suitable for real-time traffic state estimation[5].

2.2.3 The Eulerian coordinate system

The freeway monitoring components in the current hierarchical control scheme of the FOT Amsterdam utilize the Eulerian coordinate system (H. Van Lint et al., 2014). Using the ASM (Treiber & Helbing, 2002) the traffic state $z(t, x)$ may be estimated for any $\{t, x\}$. Similarly, the approach proposed by Smaragdis et al. (2004), on which the parameter estimator is based, considers Eulerian coordinates. The important parameters, critical density and capacity, are determined for the freeway bottleneck locations, thus corresponding to the Eulerian coordinate system. Similarly, the FBI does not move with the traffic stream, but is fixed in space. For integration purposes of the different FOT Amsterdam component, a choice of the Eulerian coordinate system for traffic state estimation is thus most straightforward.

Yuan (2013) argues that traffic state estimation using a first-order traffic flow model in the Lagrangian coordinates may be formulated and solved more efficiently than in Eulerian coordinates. Furthermore, it is argued that estimation using the EKF performs much better near capacity with a Lagrangian model due to the mode-switching between congestion and free-flow. This problem, however, does not directly apply to this application as we opt for using a second-order traffic model with a continuous fundamental diagram.

Due to the choice for the Eulerian coordinate system, the FCD may not be used to its full potential. If the floating car observations in a segment are used to determine the PDF of space-mean speed, the vehicle IDs are not required[6].

2.2.4 Macroscopic traffic flow model: METANET

As explained in Section 2.1 the traffic state estimator should provide estimates for the macroscopic traffic variables space-mean speed u^S, flow q and density ρ. The reason for our interest in u^S instead of u^T, which is provided by the LDD, lies in the simple relation it has with q and ρ (Wardrop, 1952).

$$u^S = \frac{q}{\rho} \tag{2.1}$$

The traffic dynamics may be explained using the macroscopic traffic flow variables in a so-called macroscopic traffic model. Lighthill & Whitham (1955) and Richards (1956) were the first to explain these dynamics in such a model. The results is the LighthillWhithamRichard (LWR) model, in which the

[5]For future research the proposed methodology may implemented with other filters. This is discussed in more detail in the recommendation, see Chapter 9.

[6]Except for removing double observations. Although this will not be included in the first set-up, the IDs may be used to obtain information about the vehicle and driver behavior.

fundamental diagram describes the relation between each set of two macroscopic variables. In this diagram a distinction is made between the free-flow and congestion state. The potential traffic conditions in these states are expected to be located on the two straight lines (Daganzo, 2005). Over the years multiple interpretations of the fundamental diagram are proposed. Instead of two separate straight lines, a continuous function may be used to represent the fundamental diagram (Papageorgiou et al., 1990).

The LWR model is a first-order macroscopic traffic model as it has one partial-derivative equation (PDE), namely the conservation equation. The METANET model (Papageorgiou et al., 1990) is a second-order macroscopic traffic model in which a second PDE, namely the dynamic speed equation, is added. This model is selected to be implemented in the proposed methodology. With respect to a first-order model[7], like implemented by Van Hinsbergen et al. (2012), it is expected to be able to describe the traffic dynamics more realistically. However, there is still much discussion among traffic research related to the choice of traffic flow models Daganzo (1995).

2.3 Initial methodology selection

The freeway traffic state estimator should provide estimates for the space-mean speed, flow and density. Although the currently used traffic state estimator is not able to allow an integral network approach (estimation - prediction - control), this ability is desired to improve the performance of the traffic management system. In order to allow an integral network approach, the Eulerian coordinate system is selected. The main advantage of this coordinate system is that the monitoring system of the FOT Amsterdam wants to identify the (Eulerian) bottleneck locations.

A traffic state estimator consists of three components, namely the observation model, traffic flow model and assimilation technique component. In the observation model two data-types will be available, namely LDD and FCD. The information within the data differs between the two types. However, both are related to the desired traffic state estimator output variable space-mean speed. The estimates of the macroscopic variables based on the LDD and FCD come with an uncertainty. Therefore, the data actually leads to a PDF of the macroscopic variables instead of a single value estimate. This PDF depends on the measurements available in the data.

Based on a traffic flow model the traffic conditions may be predicted. The complexity, expressed in the number of Partial Differential Equations (PDEs) incorporated, of such a model may be selected. In this research, the second-order traffic model METANET model is selected as it is expected to depict the traffic dynamics more realistically than commonly used first-order traffic models. Furthermore, it incorporates a continuous stationary speed equation, which may lead to better performance of the assimilation technique. Similarly to the observation model, the predictions made with the traffic flow model come with an uncertainty[8] and may thus be described by a PDF. As both traffic state estimation components provide

[7] In this research different order macroscopic traffic models will not be compared. However, for future research, it will be interesting to evaluate the proposed methodology with different traffic flow models, like the one implemented by Van Hinsbergen et al. (2012). This will be further addressed in Chapter 9.

[8] Furthermore, similar to the monitoring component, these inaccuracies may be related to the base traffic conditions and we may want to include this knowledge in our methodology/model. However, for the moment this step will only be evaluated for the monitoring component

estimations in combination with the uncertainty of this estimations, a assimilation technique able to utilize all this information is desired. The Extended Kalman Filter fulfills this requirement and is relatively computationally efficient, such that it may be used for real-time applications.

Chapter 3

Functional design of the traffic state estimator

In the previous chapter, an initial selection of the methodologies and models used for the traffic state estimator was made. It was chosen to estimate the traffic conditions in the Eulerian coordinate system using an EKF in combination with the second-order traffic flow model METANET. Based on these choices, the function design of the traffic state estimator is made in this chapter.

The freeway traffic state estimation methodology based on EKF proposed by Wang & Papageorgiou (2005) will be used for traffic state estimation and inference is considered as a basis for our methodology. However, some changes are proposed and implemented with respect to Wang & Papageorgiou (2005) methodology. Before the functional design of the traffic state estimator is described, the methodology developments with respect to Wang & Papageorgiou (2005) methodology are discussed.

Figure 3.1 visualizes the proposed traffic state estimation methodology. Three methodology developments are proposed. Firstly, multiple (heterogeneous) data-types are used. In Figure 3.1 j data-types are shown. Although only two data-types (LDD and FCD) will be considered more may be added if they are available without having to alter the methodology. As a result of having multiple (heterogeneous) data-types, a data-fusion methodology is required. Secondly, the measurement error covariance matrix in period k will be based on the measurements available for k. Based on the measurement of individual data-types a PDF of the macroscopic variables u^S or q is determined. For FCD the penetration rate and traffic conditions are used to determine the u^S PDF. As this information is not available in the FCD, the model based (prior) state vector estimate $\hat{\mathbf{x}}(k|k-1)$ is used. This results in an extra (one-direction) link in the methodology schematized in Figure 3.1. Finally, the EKF steps used Wang & Papageorgiou (2005) differ from the standard steps in a Kalman filter (Kalman, 1960). Wang & Papageorgiou (2005) take the prediction and correction step simultaneously, thereby correcting the state estimation of period $k+1$ with the measurement data from k. In other applications of an EKF, like Van Hinsbergen et al. (2012), every period two separate steps are taken. In the first step a prior estimate (prediction) $\hat{\mathbf{x}}(k|k-1)$ is made based on a model. This estimate is corrected using the measurement information on that period leading to a posterior estimate $\hat{\mathbf{x}}(k|k)$. The final estimate in k of the proposed methodology thus involves the

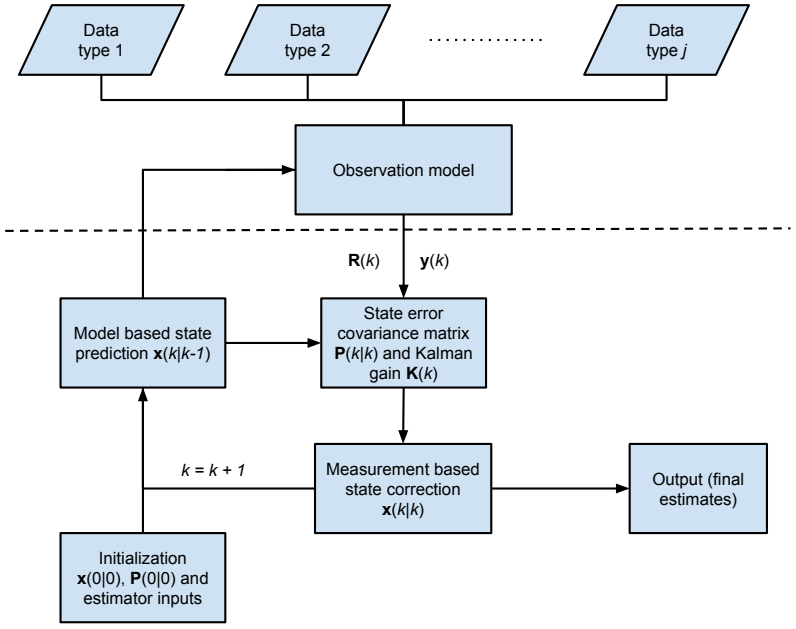


FIGURE 3.1: Visualization of the functional design of the proposed methodology. The information of heterogeneous data-types is fused in the observation model.

measurement information till k, while Wang & Papageorgiou (2005) involves the measurement information till $k-1$. It is expected that the proposed methodology provides more accurate estimations as more information is utilized. However, this hypothesis is not evaluated in this research[1].

The implication of these changes for the traffic state estimation methodology will be discussed in the related sections. In total, this chapter provides the complete function design of the proposed traffic state estimation methodology.

The proposed model-based estimation methodology requires a discretisation of time and space. Therefore, in Section 3.1 the related requirements are denoted. Next, the state vector is defined in Section 3.2. As discussed in the Literature Review (Chapter 2), model-based traffic state estimation consists of three components. Each of these components is discussed separately. The traffic flow model, observation model and assimilation technique are reviewed in Sections 3.3, 3.4 and 3.5 respectively.

3.1 Discretization

Van Hinsbergen et al. (2012) state that the length of the time and space discretizations, respectively T and Δ_i, should be based on the Courant-Friedrichs-Lewy condition (Courant et al., 1928) in order to be numerical stable.

[1] Due to the complexity of the methodology and thus the algorithm it is not feasible to evaluate the proposed methodology with the recursive equation of Wang & Papageorgiou (2005). This equation expresses that the final update of k uses the measurement information till $k-1$. However, as will be discussed in Chapter 9 it may be interesting to evaluate the difference in performance of different methodologies. The proposed methodology to involve FCD in the estimations may be incorporated in existing EKF methodology such as Wang & Papageorgiou (2005) and Van Hinsbergen et al. (2012).

$$\Delta_i \geq v_f T \tag{3.1}$$

This equation may be explained intuitively. In the used methodology it is assumed that conditions are homogenous within each individual cell during the entire period. The number of vehicles $N_i(k)$ in cell i during period k is equal to the total density $\lambda_i \rho_i(k)$ times the cell length Δ_i. As ρ_i expressed the density per lane, the total density equals ρ_i times the number of lanes λ_i. Furthermore, the rate of vehicles leaving i in k is denoted by the flow $q_i(k)$, which is equal to $\lambda_i \rho_i(k) u_i^S(k)$. The maximum number of vehicles leaving i at this rate is $N_i(k)$. Therefore, the time at which the flow may be sustained and thus the maximum duration of a period is equal to the cell length divided by the space-mean speed. The duration of the period should thus comply with the following condition.

$$T \leq \frac{\Delta_i \lambda_i \rho_i(k)}{\lambda_i \rho_i(k) u_i^S(k)} = \frac{\Delta_i}{u_i^S(k)} \tag{3.2}$$

While the speed may vary a single and constant period duration should be sized on the speed-limit.

3.2 The state vector

In a EKF a state vector $\mathbf{x}$ is recursively updated based on a model and measurement information. The variables included in $\mathbf{x}$ should thus be chosen based on the selected models. Wang & Papageorgiou (2005) makes a three-way split in $\mathbf{x}$. As explained in Chapter 2 the traffic flow model METANET will be used. This model predicts the freeway macroscopic traffic variables density ρ_i and space-mean speed u_i^S for all cells $i = 1, \ldots, N$. Therefore, these variables have to be included in $\mathbf{x}$, which combined form the model state $\mathbf{z}$.

$$\mathbf{z} = \begin{bmatrix} \rho_1 & v_1 & \cdots & \rho_N & v_N \end{bmatrix}^T$$

Additional to the freeway conditions, the boundary conditions are required inputs for the traffic flow model. These boundary conditions consist of the upstream flow q_0 and space-mean speed u_0^S, the downstream density ρ_{N+1}, the on-ramp inflows r_i and exiting rates β_i. A ramp variable, r_i and β_i only has to be included if a ramp is located in cell i. Combined these variables form the boundary variables $\mathbf{d}$.

$$\mathbf{d} = \begin{bmatrix} q_0 & v_0 & \rho_{N+1} & r_1 & \cdots & r_N & \beta_1 & \cdots & \beta_N \end{bmatrix}^T$$

Wang & Papageorgiou (2005) includes the stationary speed equation variables free-flow speed v_f, critical density ρ_{cr} and a in the state vector as these are allowed to change during the recursive updating. These three variables form the unknown stationary speed equation parameters $\mathbf{p}$.

$$\mathbf{p} = \begin{bmatrix} v_f & \rho_{cr} & a \end{bmatrix}^T$$

In total, the state vector $\mathbf{x}$ is build-up of the above shown vectors, namely $\mathbf{x} = [\mathbf{z}^T \ \mathbf{d}^T \ \mathbf{p}^T]^T$.

3.3 The traffic flow model

To predict the freeway state, expressed in the freeway variables ρ and u^S, the second-order macroscopic traffic flow model METANET (Papageorgiou et al., 1990) is used. This model incorporates the conservation equation (3.3), dynamic speed equation (3.5), stationary speed equation (3.6) and flow equation (3.7). Note that the space-mean speed is expressed by v instead of u^S. In this section, the original variables of Papageorgiou et al. (1990) are used. However, in the remainder of this report, u^S will be used to denote the space-mean speed.

$$\rho_i(k+1) = \rho_i(k) + \frac{T}{\Delta_i \lambda_i}\left[q_{i-1}(k) - q_i(k) + r_i(k) - s_i(k)\right] \tag{3.3}$$

$$s_i(k) = \beta_i(k) q_{i-1}(k) \tag{3.4}$$

$$\begin{aligned} v_i(k+1) = v_i(k) &+ \frac{T}{\tau}\left[V(\rho_i(k)) - v_i(k)\right] + \frac{T}{\Delta_i} v_i(k)\left[v_{i-1}(k) - v_i(k)\right] \\ &- \frac{\nu T}{\tau \Delta_i}\frac{\left[\rho_{i+1}(k) - \rho_i(k)\right]}{\rho_i(k) + \kappa} - \frac{\delta T}{\Delta_i \lambda_i}\frac{r_i(k) v_i(k)}{\rho_i(k) + \kappa} + \xi_i^v(k) \end{aligned} \tag{3.5}$$

$$V(\rho) = v_f exp\left[-\frac{1}{a}\left(\frac{\rho}{\rho_{cr}}\right)^a\right] \tag{3.6}$$

$$q_i(k) = \rho_i(k) v_i(k) \lambda_i + \xi_i^q(k) \tag{3.7}$$

The density is predicted using the conservation equation, see (3.3). As the name states, this is based on the physical law that vehicles cannot exist out of nothing or disappear and are thus conserved. The change in ρ_i is equal to the residual inflow divided by the length of the cell Δ_i and number of lanes λ_i. Here, the residual inflow of a road stretch during a period is given by the inflow subtracted by the outflow times T. Inflow may come from the upstream cell q_{i-1} and an on-ramp r_i. The outflow is equal to the flow leaving the cell to the downstream cell q_i and via an off-ramp s_i. Following Wang & Papageorgiou (2005), the off-ramps are placed at the most upstream position of a cell. The off-ramp flow s_i can therefore be expressed as a fraction β_i of the inflow from the upstream cell q_{i-1}, see (3.4). As the flows are not in the state vector, these are determined from the other macroscopic variables via the flow equation. This equation states that the outflow of cell i with homogenous conditions is equal to the density per lane ρ_i times the number of lanes λ_i times the space-mean speed v_i.

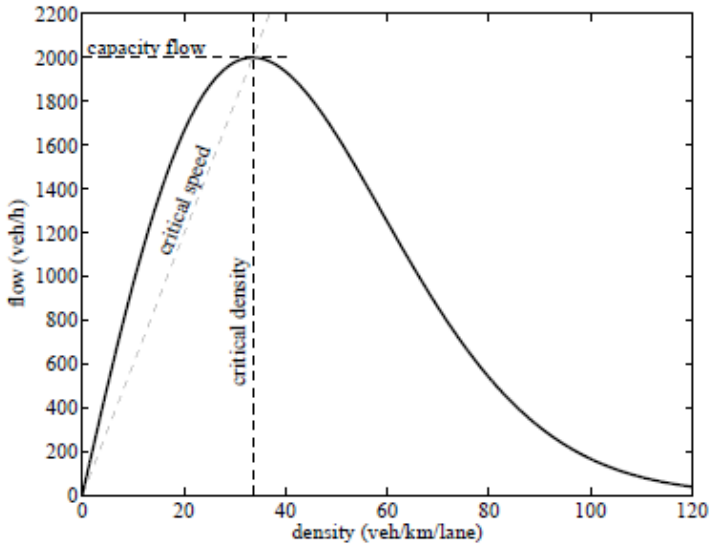


FIGURE 3.2: Example of the fundamental diagram (3.6) with parameters $a = 1.867$, $v_f = 102$ km/h and $\rho_{cr} = 33.5$ veh/km/lane (Hegyi, 2004).

The space-mean speed is predicted using the dynamic speed equation. The first term of this equation is the speed in the previous period $v_i(k)$. The second term involves the stationary speed equation. The stationary speed equation $V(\rho)$ is a continuous fundamental diagram which describes the speed-density relation (Hegyi, 2004). This relation expresses that the speed decreases for an increasing density ρ. This is intuitive as the speed is likely to decrease for increasing restrictions. The unrestricted or free speed is denoted by v_f and a denotes a parameter which indicates to which extend the speed is affected by an increasing ρ. Based on the speed-density relation the flow-density relation may be expressed, see Figure 3.2. In this fundamental diagram, two traffic states may be distinguished, namely free-flow and congestion[2]. The maximum capacity, expressed in q, is reached at the boundary of these states, namely at ρ_{cr}. The speed is corrected based on the difference between the stationary speed equation and $v_i(k)$. In the case of a negative difference, thus when $v_i(k)$ is lower then expressed by $V(\rho_i(k))$, a positive correction is expected. This is correctly implemented by the dynamic speed equation. The time constant of the speed relaxation term τ is the parameter which in combination with T determines the extend to which the speed estimate is corrected to $V(\rho_i(k))$. This will initially be based on the research of Hegyi (2004).

The third term of (3.5) is the convection term which expresses the speed change caused by the inflow of vehicles (Hegyi, 2004). This term is dependents on the speed difference between the inflowing vehicles and vehicles originally in the cell, thus $v_{i-1}(k) - v_i(k)$. As expected, if the inflowing vehicles v_{i-1} had a higher speed in k, this term results in a positive speed change. The influence of the inflowing vehicles is dependent on the fraction of vehicles which has left the cell, which is expressed by the term Tv_i/Δ_i. The fourth term of (3.5) is the anticipation term. In this term the effect of a drivers experience of a downstream

[2]These states will be considered in a later phase of this research and are therefore expressed. Among traffic engineers there is an ongoing debate on the traffic phases/states which may be distinguished in the traffic conditions. For instance, Kerner (1999) states that the there are three traffic phases. In this research, this will not be discussed into more depth. However, for future research it may be interesting to incorporate these relations and the differences in uncertainties which come with it.

density change, $\rho_{i+1}(k) - \rho_i(k)$, is described. In case of a higher ρ_{i+1} then ρ_i a driver is expected to reduce its speed by anticipating on the increased density. Three METANET parameters are included in this term namely κ, ν and τ. Hegyi (2004) denotes the first two as METANET speed anticipation term parameters and the latter as the time constant of the speed relaxation term. An increasing κ and ν respectively reduce and increase the effect of the anticipation term. These parameters will be selected based on Hegyi (2004). The fifth and final term of (3.5) describes the speed drop caused by the merging phenomena in case of an on-ramp. This term is dependent on the density influx $Tr_i/\Delta_i\lambda_i$. As would be intuitively expected a larger influx results results in a larger speed drop. To involve the relative influx the freeway density ρ_i is involved. The METANET parameter for the speed drop term caused by merging at an on-ramp δ and κ will be selected based on Hegyi (2004).

In the flow and dynamic speed equation zero-mean Gaussian errors are included. These are depicted by ξ_i^q and ξ_i^q respectively and combined form the model error vector ξ_1.

$$\xi_1 = \begin{bmatrix} \xi_1^q & \xi_1^{u^S} & \cdots & \xi_N^q & \xi_N^{u^S} \end{bmatrix}^T \tag{3.8}$$

The traffic flow model equations are used to make a one-period prediction of the model state vector $\hat{\mathbf{z}}(k+1|k)$. Additionally the boundary variables $\mathbf{d}$ and stationary speed equation parameters $\mathbf{p}$ need to be predicted. Wang & Papageorgiou (2005) introduces a random walk for this purpose.

$$\mathbf{d}(k+1) = \mathbf{d}(k) + \xi_2(k) \tag{3.9}$$

$$\mathbf{p}(k+1) = \mathbf{p}(k) + \xi_3(k) \tag{3.10}$$

the zero-mean Gaussian errors of these random walks are denoted by ξ_2 and ξ_3. The traffic flow model combined with the random walks form the set of nonlinear differentiable functions is denoted as $\mathbf{f}[\cdot]$. These describe the state vector in period $k+1$, $\mathbf{x}(k+1)$ as a function of the previous state vector $\mathbf{x}(k)$ and white noise $\xi(k) = [\xi_1^T \ \xi_2^T \ \xi_3^T]^T$.

$$\mathbf{x}(k+1) = \mathbf{f}\left[\mathbf{x}(k), \xi(k)\right] \tag{3.11}$$

As the white noise is unknown, the predicted state variable $\hat{\mathbf{x}}(k+1|k)$ based on the period state variable estimation $\hat{\mathbf{x}}$ is given by

$$\hat{\mathbf{x}}(k+1|k) = \mathbf{f}\left[\hat{\mathbf{x}}(k|k), \mathbf{0}\right] \tag{3.12}$$

This function represents the state vector prediction.

3.4 Observation model

The raw measurement data (indirectly) provides information on u^S and q for the observed cells in k. These variables may be estimation from the measurements yielding the measurement vector $\mathbf{y}(k)$ and measurement error covariance matrix $\mathbf{R}(k)$. In the correction step of the EKF, $\mathbf{y}(k)$ will be compared with estimations of the same variables based on the model-based (prior) state vector estimate $\hat{\mathbf{x}}(k|k-1)$, namely $\hat{\mathbf{y}}(k)$. In this section the equation used to determine $\mathbf{y}(k)$, $\mathbf{R}(k)$ and $\hat{\mathbf{y}}(k)$ will be discussed. The methodology and equations used to correct the state vector based on $\mathbf{y}(k)$, $\mathbf{R}(k)$ and $\hat{\mathbf{y}}(k)$ will be discussed in Section 3.5.

3.4.1 Measurement-based Probability Density Functions

Two data-types are considered, namely LDD and FCD. All measurements are assigned to a single cell. The estimates of u^S and q based on the raw measurement data have an uncertainty and may thus be expressed as a Probability Density Function (PDF). An assumption of the EKF is that all errors are zero-mean Gaussian distributed. Therefore, the estimated value and uncertainty will be expressed as the mean and variances of the Gaussian PDF. The mean value is an element of $\mathbf{y}(k)$, while the variance is assigned to the respective diagonal element in $\mathbf{R}(k)$. Based on the LDD both u^S and q may be estimated, while FCD only provides a u^S estimate. Independently, for both data-types, these PDFs may be estimated. If both the LDD and FCD provides a PDF for an $\{i, k\}$ combination, the PDFs should be fused to form a single PDF.

In this research it is investigated if and how the measurement error variance, diagonal elements in $\mathbf{R}(k)$, may be described based on the measurements and prior state vector estimate. As stated before, it is expected that the FCD-based estimation errors are highly dependent on the raw data characteristics and traffic conditions. Therefore, this PDF is based on the data and prior state vector estimate. For LDD it is investigated if this approach also leads to an accurate estimation of the LDD-based estimation errors. The purpose of assigning measurement-based error variances to observed cell is twofold. Firstly, a measurement-based PDF may provide a better representation of the potential errors in the measurement vector $\mathbf{y}(k)$. This may lead to a better comparison between model-based predictions and measurement-based estimations. Secondly, the weight assigned to the different data-types is based on the expected uncertainties of the raw data to the measurement vector variables.

3.4.1.1 Floating car measurements

In the proposed methodology, homogenous traffic conditions are assumed within cells during a single period. Vehicles observed[3] within this area of the time-space domain provide information on u^S.

[3]Note that a vehicle may be observed more than once during a period. In this research, data with these properties will not be considered. However, if the data would have these properties it has to be dealt with correctly. For instance, an aggregated value with the same weight as a single vehicle observation may be taken. This will be discussed in Chapter 9.

The estimated space-mean speed using FCD u^S_{fcd} is taken as the mean of all individual speeds v_n of all observed vehicles with i and k. To keep the equation and variables clear no subscripts with i and k are included.

$$u^S_{fcd} = \frac{1}{n}\sum_{n=1}^{n} v_n \tag{3.13}$$

To describe the uncertainty of this estimation two factors are considered, namely the penetration rate and individual speed measurement errors. The former provides information of the representativeness of the estimated space-mean speed with respect to the mean speed of all vehicles. In the case a larger fraction of the vehicles is observed, thus the penetration rate is higher, a smaller error is expected. Although the penetration rate is not observed, it may be estimated using ρ in prior state vector $\hat{\mathbf{x}}(k|k-1)$ and the number of observed vehicles n. The estimated total number of vehicles N on a given cell in k is equal to the prior estimate of the density $\rho(k|k-1)$ times the cell length Δ.

$$N = \rho(k|k-1)\Delta \tag{3.14}$$

It is assumed that the individual speed measurement errors v_{me} are uncorrelated and may independently be described by a zero-mean Gaussian distribution. Due to the independence of the measurement errors, the error variance of the mean speed reduces when the number of observed vehicles increases. Both factors are taken into account yielding a combined FCD error variance estimate. The derivations leading to this error variance estimate are shown in Appendix A.

$$\sigma^2_{u^S_{fcd}} = \frac{1}{n}\sigma^2_{v_{me}} + max\left(\frac{N-n}{n(N-1)}\sigma^2_{u^S}, 0\right) \tag{3.15}$$

The max-function is included as the variance can never be negative. If this function would not be included and N is underestimated while all vehicles are observed, the equation may lead to a negative value.

Equation (3.15) requires two inputs, namely the individual speed measurement error variance $\sigma_{v_{me}}$ and the space-mean speed variance $\sigma^2_{u^S}$. The former will be based on the literature. The latter will be based on the empirical relation between u^T and $\sigma^2_{u^S}$ found by J. W. C. Van Lint (2004). Here the time-mean speed is denoted as u_L and the variance of the instantaneous speed distribution is denoted by $\hat{\sigma}_M$. The latter is the desired variable. However, as discussed before the u^T and u^S differ. Still the relation is considered to provide an accurate enough approximation of the $\sigma^2_{u^S}$ based on the traffic conditions. For this purpose the u^S is inserted in (3.16) as u_L.

$$\hat{\sigma}_M = \begin{cases} 0.5u_L - 34 & u_L \leq 74 \\ 0.02u_L + 5 & \text{else} \end{cases} \tag{3.16}$$

3.4.1.2 Loop-detector measurements

The space-mean speed u^S may be approximated from the lane-specific time-mean speed u_l^T and flow measurements q_l (Knoop & Hoogendoorn, 2012) where l expresses the lane.

$$u_{ldd}^S = \frac{\sum_{l=1}^{\lambda} q_l}{\sum_{l=1}^{\lambda} \frac{q_l}{u_l^T}} \tag{3.17}$$

Here λ denotes the number of lanes. Errors in u_{ldd}^S may have multiple factors. The factor which are considered are related to measurement and assumption errors. Measurement errors lead to errors in the aggregated time-mean speed. Here individual speed measurement errors and miss/double observations are considered. The considered assumption errors are the constant speeds within lanes and constant speeds during the (one-minute) aggregation period[4]. The former assumption is related to (3.17). If this assumption is violated, the space-mean speed estimation is non-Gaussian distributed and biased.

It is decided to make the LDD space-mean speed error variances independent of the observations due to two main reasons. Firstly, the considered errors lead to complex, non-Gaussian and potentially biased PDFs. Secondly, there is not a clear relation between the measured speed level and the error variance. This means that the variance within the one-minute period is unobserved by the data and difficult to describe.

The flow q at the downstream boundary of i in period k is directly estimated by taken the sum of the lane-specific flows q_l.

$$q = \sum_{l=1}^{\lambda} q_l \tag{3.18}$$

Two factors of uncertainty of this estimation are considered, namely missing/double observations and variations in flow during the one-minute period. A relation was found for the error variance expressed a function of the probability that a single vehicle is missed or double observed[5]. However, the resulting

[4]The one-minute aggregated space-mean speed will be assigned to all shorter periods within this minute. There are some methods to capture differences within the one-minute periods, for instance due to platooning (Schreiter et al., 2010), however the effect on this research is expected to be non-decisive. Therefore, it is decided not to stray away from the research and incorporate this methodology.

[5]In this relation some simplifications were made. For instance, the a single value was taken for the miss/double observation probability. Furthermore, this probability may dependent on vehicle characteristics, for instance trucks with double trailers have a larger probability to be double observed then a passenger car. However, to keep the complexity limited this vehicle-type dependency was not included

error was found to be small with respect to the influence of the variation within the one-period aggregation period. The variation within this aggregation period is expected to have a limited relation with the observed flow. The relative variation is expected to reduce with increasing flow, due to increasing restrictions by other vehicles. However, still the variation may increase due to increasing flow.

As these derivations and expected relations do not lead to an accurate measurement data-based q uncertainty, it is decided to keep the LDD flow error variances independent of the observations.

3.4.1.3 Loop-detector and floating car measurements

LDD and FCD both provide estimates for the space-mean speed u^S. If only one of the two data-types observe cell i in period k, the resulting PDF parameters can directly be used as input for the measurement vector $\mathbf{y}(k)$ and measurement error covariance matrix $\mathbf{R}(k)$. However, if both types provide a PDF, an extra step has to be taken. The combined PDF is the product of the two individual PDFs. The product of two Gaussian distributions is a Gaussian distribution (Bromiley, 2003).

$$\mu_{1,2} = \frac{\mu_2\sigma_1^2 + \mu_1\sigma_2^2}{\sigma_1^2 + \sigma_2^2} \tag{3.19}$$

$$\sigma_{1,2}^2 = \sqrt{\frac{\sigma_1^2\sigma_2^2}{\sigma_1^2 + \sigma_2^2}} \tag{3.20}$$

In these equations the subscripts 1 and 2 denote the parameters for the two independent Gaussian distributions. Intuitively more emphasis should be placed on the estimate with the lowest uncertainty, thus with the lowest error variance σ^2. In (3.19), the mean μ is weight with the error variance of the other data-type. Therefore, the estimate with the lowest error variance is assigned the highest weight. The combined error variance (3.20), is lower than the lowest error variance. This is also initiative, in the way that more information yields a lower uncertainty.

3.4.2 Model-based measurement vector estimations

In the measurement vector four types of variables may be included. These are flow q, in-flow r, out-flow s and space-mean speed u^S. Only two variables, namely r and u^S, are elements of the state vector $\mathbf{x}$.

The flow q is obtained by multiplying the density per lane ρ with the space-mean speed u^S and number of lanes λ. Here the state noise ξ and measurement noise γ denote the uncertainties for the measurement flow estimate m^q. In $\mathbf{x}$, the out-flow is included as a fraction β of the freeway flow q. To estimate the m^s based on the state vector, the freeway flow q is multiplied with β. The uncertainties are a dependent on β, ξ and γ.

$$m_i^q(k) = q_i(k) + \gamma_i^q(k) = \rho_i(k)v_i(k)\lambda_i + \xi_i^q(k) + \gamma_i^q(k) \tag{3.21}$$
$$m_i^v(k) = v_i(k) + \gamma_i^v(k) \tag{3.22}$$
$$m_i^r(k) = r_i(k) + \gamma_i^r(k) \tag{3.23}$$
$$m_i^s(k) = s_i(k) + \gamma_i^s(k) = \beta_i(k)\left(\rho_{i-1}(k)v_{i-1}(k)\lambda_{i-1} + \xi_{i-1}^q(k)\right) + \gamma_i^s(k) \tag{3.24}$$

These equations are used to estimate the measurement vector $\hat{\mathbf{y}}(k)$. $\mathbf{g}[\cdot]$ is a nonlinear differentiable vector function which contains the equations stated above related to the observed variables. The true values of the measurement vector are thus

$$\mathbf{y}(k) = \mathbf{g}[\mathbf{x}(k), \eta(k)] \tag{3.25}$$

where $\eta(k)$ is the output noise vector, which is a function of the measurement noise vector $\gamma(k)$ and the state noise vector $\xi(k)$ (Wang & Papageorgiou, 2005). The estimated measurement vector $\hat{\mathbf{y}}(k)$ obtain from the prior state estimate $\hat{\mathbf{x}}(k|k-1)$ is given by

$$\hat{\mathbf{y}}(k) = \mathbf{g}[\hat{\mathbf{x}}(k|k-1), \mathbf{0}] \tag{3.26}$$

3.5 Assimilation technique

The EKF is a recursive updating assimilation technique. It combines the predictions from the traffic flow model with the estimations from the observation model. These predictions and estimations have an uncertainty.

Every period, two consecutive steps are taken to obtain the final state vector estimate $\hat{\mathbf{x}}(k|k)$. In the first step, a model-based state prediction $\hat{\mathbf{x}}(k|k-1)$ using the nonlinear differentiable vector function $\mathbf{f}$ described by the equations in Section 3.3.

$$\hat{\mathbf{x}}(k+1|k) = \mathbf{f}\left[\hat{\mathbf{x}}(k|k), \mathbf{0}\right]. \tag{3.27}$$

The true model errors $\xi(k)$ are unknown, so their expected value $E[\xi(k)] = \mathbf{0}$[6] is included in the model-based state prediction. In the second step, the state vector prediction is corrected using the measurement data and observation model, leading to the final state vector estimate $\hat{\mathbf{x}}(k|k)$. This correction follows the recursive equation.

[6] The errors are expected to be Gaussian distributed with a zero mean. Therefore, $E[\xi(k)] = \mathbf{0}$.

$$\hat{\mathbf{x}}(k|k) = \hat{\mathbf{x}}(k|k-1) + \mathbf{K}(k)\left[\mathbf{y}(k) - \hat{\mathbf{y}}(k)\right] \tag{3.28}$$

The correction part of the recursive equation $\mathbf{K}(k)\left[\mathbf{y}(k) - \hat{\mathbf{y}}(k)\right]$ is dependent on a correction factor, the Kalman Gain $\mathbf{K}(k)$, and the difference between the measurement vector $\mathbf{y}(k)$ and estimated measurement vector $\hat{\mathbf{y}}(k)$. The former, $\mathbf{y}(k)$, is based on the measurements, while the latter, $\hat{\mathbf{y}}(k)$, is based on the prior estimate of the state vector and relation with the measurement variables. The equations used to obtain these vectors are respectively depicted in Sections 3.4.1 and 3.4.2.

$\mathbf{K}(k)$ determines the effect which the difference $\mathbf{y}(k) - \hat{\mathbf{y}}(k)$ has on the state vector. This factor depends on the uncertainties in the state vector and observation model. These uncertainties are balanced to determine the weight which is given to the state prediction and the difference with the observation model. For a given state vector uncertainty, the weight given to observations should increase with a decreasing uncertainty. This principle is exploited by the Kalman Filter.

$\mathbf{P}$ denotes the state error covariance matrix. Every step $\hat{\mathbf{x}}$ is updated, $\mathbf{P}$ is updated. The $\mathbf{P}(k|k-1)$ and $\mathbf{P}(k|k)$ are respectively the error covariance matrices of $\hat{\mathbf{x}}(k|k-1)$ and $\hat{\mathbf{x}}(k|k)$.

$$\mathbf{P}(k|k-1) = E\left\{[\mathbf{x}(k) - \hat{\mathbf{x}}(k|k-1)] \cdot [\mathbf{x}(k) - \hat{\mathbf{x}}(k|k-1)]^T\right\} \tag{3.29}$$

$$\mathbf{P}(k|k) = E\left\{[\mathbf{x}(k) - \hat{\mathbf{x}}(k|k)] \cdot [\mathbf{x}(k) - \hat{\mathbf{x}}(k|k)]^T\right\} \tag{3.30}$$

It is the objective to minimize the error covariance matrix $\mathbf{P}(k|k)$ (Wang & Papageorgiou, 2005), thus minimize (3.30). The Kalman Gain $\mathbf{K}(k)$ may be derived accordingly. The derivations shown in Appendix B.1, lead to equations for $\mathbf{P}(k|k-1)$, $\mathbf{K}(k)$ and $\mathbf{P}(k|k)$. These equations are different than those used by Wang & Papageorgiou (2005). This has to do with a difference in the recursive equations. Wang & Papageorgiou correct the model prediction for $k+1$ with the measurements in k. As this deviates from the standard Kalman Filter recursive equation, like the equation used by Van Hinsbergen et al. (2012), it was chosen not to use the recursive equation used by Wang & Papageorgiou (2005). Instead, the recursive equation used by Van Hinsbergen et al. (2012) is used.

The prior error covariance matrix $\mathbf{P}(k|k-1)$ is given by

$$\mathbf{P}(k|k-1) = \mathbf{A}(k-1)\mathbf{P}(k-1|k-1)\mathbf{A}^T(k-1) + \mathbf{\Gamma}(k-1)\mathbf{Q}(k-1)\mathbf{\Gamma}^T(k-1) \tag{3.31}$$

where

$$\mathbf{A}(k-1) = \frac{\partial \mathbf{f}}{\partial \mathbf{x}}\left(\hat{\mathbf{x}}(k-1|k-1), \mathbf{0}\right) \tag{3.32}$$

$$\mathbf{\Gamma}(k-1) = \frac{\partial \mathbf{f}}{\partial \xi}\left(\hat{\mathbf{x}}(k-1|k-1), \mathbf{0}\right) \tag{3.33}$$

$$\mathbf{Q}(k-1) = E\left\{\xi(k-1)\xi^T(k-1)\right\} \tag{3.34}$$

$\mathbf{P}(k|k-1)$ is dependent on two factors. Firstly, errors in the model-based prediction input $\hat{\mathbf{x}}(k-1|k-1)$ lead to errors in $\hat{\mathbf{x}}(k|k-1)$. This is covered by $\mathbf{A}(k-1)\mathbf{P}(k-1|k-1)\mathbf{A}^T(k-1)$. Secondly, the model error term $\xi(k-1)$ causes extra uncertainty for the state vector estimate. This factor is described by $\mathbf{\Gamma}(k-1)\mathbf{Q}(k-1)\mathbf{\Gamma}^T(k-1)$.

The Kalman Gain $\mathbf{K}(k)$ is given by

$$\mathbf{K}(k) = \frac{\mathbf{P}(k|k-1)\mathbf{C}^T(k)}{\mathbf{C}(k)\mathbf{P}(k|k-1)\mathbf{C}^T(k) + \mathbf{\Sigma}(k)\mathbf{R}(k)\mathbf{\Sigma}^T(k)} \tag{3.35}$$

where

$$\mathbf{C}(k) = \frac{\partial \mathbf{g}}{\partial \mathbf{x}}\left(\hat{\mathbf{x}}(k|k-1), \mathbf{0}\right) \tag{3.36}$$

$$\mathbf{\Sigma}(k) = \frac{\partial \mathbf{g}}{\partial \eta}\left(\hat{\mathbf{x}}(k|k-1), \mathbf{0}\right) \tag{3.37}$$

$$\mathbf{R}(k) = E\left\{\eta(k)\eta^T(k)\right\} \tag{3.38}$$

As stated before, the Kalman Gain $\mathbf{K}(k)$ balances the prior state vector uncertainty, expressed by $\mathbf{P}(k|k-1)$, and the observation model uncertainty, expressed by $\mathbf{R}(k)$. An increase in absolute value of an element in $\mathbf{K}(k)$ means a larger correction, thus a higher weight to measurement data. This should be the case if the uncertainty in the measurement vector estimate based on the prior state vector estimate increases with respect to the measurement vector uncertainty. The former is given by $\mathbf{C}(k)\mathbf{P}(k|k-1)\mathbf{C}^T(k)$ and the latter by $\mathbf{\Sigma}(k)\mathbf{R}(k)\mathbf{\Sigma}^T(k)$. These two factors form the denominator, while the numerator is dependent on the state vector error covariance matrix and its relation with the measurement vector. The equation given by (3.35) thus correctly incorporate the desired relation. Increases in $\mathbf{P}(k|k-1)$, ceteris paribus, lead to larger absolute values in $\mathbf{K}(k)$ which results in a stronger correction.

The error covariance matrix of $\hat{\mathbf{x}}(k|k)$ is given by

$$\mathbf{P}(k|k) = \left[\mathbf{I} - \mathbf{K}(k)\mathbf{C}(k)\right]\mathbf{P}(k|k-1) \tag{3.39}$$

The information included in the correction step reduced the uncertainty. Therefore the prior error covariance matrix $\mathbf{P}(k|k-1)$ is reduced by factor. This leads to $\mathbf{I}-\mathbf{K}(k)\mathbf{C}(k)$.

In (3.31), (3.35) and (3.39) the derivative matrices $\mathbf{A}(k-1)$, $\mathbf{\Gamma}(k-1)$, $\mathbf{C}(k)$ and $\mathbf{\Sigma}(k)$ are used. These are derived in Appendix B.2.

3.6 Functional design and next steps

Due to the many important equations and relations which are described in this chapter, the functional design is very technical. To concisely denote the important decisions made for the functional design, they are summarized in this section.

The proposed methodology is based on Wang & Papageorgiou (2005) methodology. No changes are made with respect to the traffic flow model and the state vector $\mathbf{x}$. The $\mathbf{x}$ contains the macroscopic traffic variables density ρ and space-mean speed u^S. All macroscopic traffic variables, where q is obtained from ρ and u^S, are thus estimated. Furthermore boundary condition variables and parameters of the fundamental diagram are included in $\mathbf{x}$. For model-based prediction of ρ and u^S the second-order traffic flow model METANET is used. The remaining state vector variables are assumed to be random walks.

Due to the introduction of a new data-type, namely FCD, changes are proposed for the observation model. This new data-type only provides information related to u^S, while LDD provides information related to both u^S and q. The output of the observation model is the measurement vector $\mathbf{y}$ and measurement error covariance matrix $\mathbf{R}$. In multiple EKF traffic state estimation applications, like Wang & Papageorgiou (2005) and Van Hinsbergen et al. (2012), $\mathbf{y}$ is based on the raw LDD and the elements within $\mathbf{R}$ are fixed values which have to be initialized. To incorporate the LDD a similar approach is taken[7]. However, a new approach is proposed to determine the FCD-based estimates of u^S and the related measurement error variance. The mean of the observed vehicle speeds is taken as the estimated u^S. The uncertainty is based on expected measurement errors and a combination of the estimated penetration rate and driver heterogeneity. The penetration rate is estimated using the ρ-estimate and the number of FCD-observations, while the driver heterogeneity is estimated from the u^S-estimate.

Additional to the observation model also a change to the assimilation technique is proposed. In this research, the model-based prior estimations serve as input for the FCD-based measurement error variance. Furthermore, the measurement-corrected posterior estimates will be evaluated. To allow this the recursive equation of Wang & Papageorgiou (2005) is changed by the traditional recursive equation proposed by Kalman (1960). This equation is also used by Van Hinsbergen et al. (2012). However, as they use a

[7] The possibility to base the error variances of LDD-based estimated on the measurement. However, it was found that it was not feasible to express the uncertainties based on the measurement. A major cause for the uncertainties is that the aggregation period (60 s) is larger the discretisation period (15 s). Other factors for uncertainties are measurement errors and model assumption errors. For u^S, basing the uncertainties on the measurement and assumption error leads to inaccurate and complex relation, which is not expected to perform better than a fixed error variance. For q, the main contribution to the uncertainty is expected to be the aggregation period. It is expected that this uncertainty may both positively and negatively correlate with the measured q a fix error. For instance, the variation in q respectively to q may vary more at low demand as drivers are less restricted. However, the absolute variation will be small as q is small. For a higher q, for instance near capacity, the relative variation will be lower as drivers are more restricted. However, the absolute may still be equal or higher than for less restricted conditions. Although there may still be a relation it is difficult to accuracy describe and its variation is considered to be less important than that of the FCD-based estimates.

first-order instead of a second-order traffic flow model, there updating equations cannot directly be used. For this reason the new equations are derived and presented.

The functional design of the proposed methodology is based on literature, the theoretical understanding of traffic dynamics and traffic state estimation, and knowledge about the potential data-types. However, this does not directly mean that the proposed methodology has a good traffic state estimation performance. In order to evaluate its performance and be able to answer the research questions an experimental set-up is discussed in the next Chapter.

Chapter 4

Experimental set-up

The previous chapter provides the functional design of the proposed traffic state estimator. In this chapter the experimental set-up is explained.

The objective of the experimental set-up is to evaluate the choice for a model-based estimator with respect the currently implemented methodology (ASM)[1] and choices related to the incorporation of the FCD. This should provide insight in the added value of the proposed methodology for the FOT Amsterdam traffic management system.

In order to quantitatively compare different methodologies and the inclusion of FCD, a performance statistic is required. Section 4.2 explains the performance statistics which will be used and how these are obtained. To conduct the interesting experiments, four steps will be taken. Firstly, the ASM is assessed for a varying LDD and FCD availability. Secondly, the proposed methodology is assessed for varying data availability and other varying features to evaluate the functional design choices. Thirdly, the estimation performance of two methodologies is compared. In this step, the complexity of the methodologies is also discussed. Finally, the added value of the proposed methodology to the FOT Amsterdam is evaluated. Important features for these experiments are explained in more detail in Section 4.3.

4.1 Data collection

In this research only synthesized data will be used. This approach allows to collect the data required to obtain the true conditions and select a fraction of the data for traffic state estimation. In this way it is possible to vary the data availability and determine the estimation performance of the resulting traffic state estimations.

[1] The ability to make model-based traffic state predictions and breakdown probability predictions are also important advantage. Furthermore, the possibility to include the estimation uncertainties $\mathbf{P}$ in other elements of system may be interesting in future applications. The accuracy of such predictions will, however, not be evaluated by means of an experimental set-up.

Two potential microscopic simulation programs were considered, namely VISSIM (Fellendorf & Vortisch, 2001) and FOSIM (Dijker et al., 1997). Out of these options FOSIM was selected for three main reasons. Firstly, the program is specifically designed for freeway traffic (Minderhoud & Kirwan, 2001). As this research only considers freeway traffic and traffic state estimation, extra features and thereby extra complexity are not needed. Secondly, FOSIM is a Dutch program and aims to describe the Dutch freeway traffic, which is in line with the choices made during this research. Finally, a working FOSIM model of an important freeway in the Netherlands, namely the A13-S is available from other research. This model may directly be implemented for this research.

4.2 Traffic state estimator performance

The output of EKF with data-fusion is collected in the posterior estimates of the state vector $\hat{\mathbf{x}}(k|k)$. These contain space-mean speed u^S and density ρ estimates for each cell over the entire estimation period. To indicate the performance of the algorithm with specific model, assimilation technique and data parameters, the estimates have to compared with the true conditions. The true conditions are denotes as the ground truth (Section 4.2.2) and comparing them to the estimations yields a certain performance statistic (Section 4.2.3). A three-step evaluation approach is used to assess the performance of the estimators, see Section 4.2.1.

4.2.1 Evaluation approach

To assess the performance and indicate potential causes of difference between estimators, two approaches may be taken[2]. Firstly, the overall performance is assessed using a performance statistic. The performance statistics yield an initial comparison between different estimators (settings). However, these statistics only provide limited insight in the traffic dynamics which are missed by the estimator. Therefore, secondly, the contour plots of the ground truth and different estimators may be evaluated. These plots may show whether the estimator is able to describe the traffic dynamics correctly. It may for instance be the case that the propagation of a traffic jam is missed. A contour plot does not have to depict the estimations, but may also depict the estimation error. The latter contour is used multiple in this research to assess the features of the traffic state estimators.

These approaches provide insight in the differences based on methodology and data-availability. Additional to the objective of evaluating the best or accurately enough methodology and data-availability, it also serve as input for potential improvements.

[2]It is possible to evaluate the estimation performance in more ways. However, in this research the states evaluation approaches suffice to gain insight in the important features of the estimators and assess the decisions made. In earlier phases of this research also time-series plots at individual cells of the estimations and the ground truth were considered.

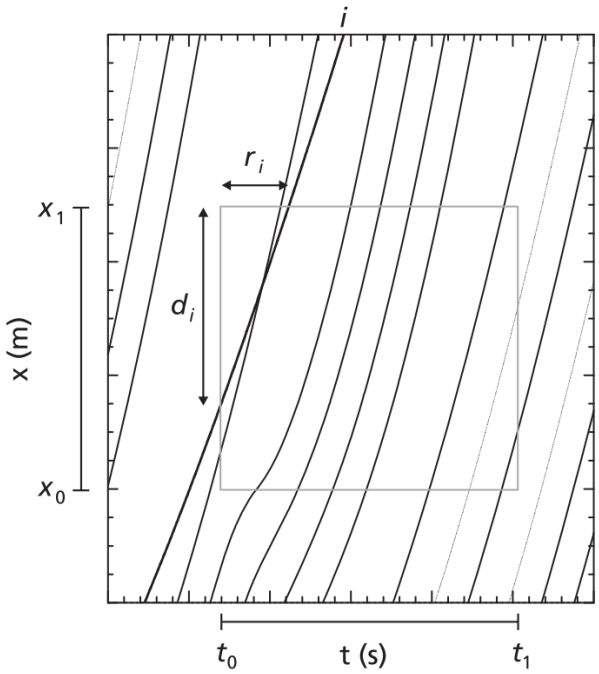


FIGURE 4.1: Generalization according to Edie (1965)

4.2.2 Ground truth

Edie (1965) provided generalized definitions of flow density and speed (Hoogendoorn & Knoop, 2013). Based on the generalization shown in Figure 4.1, the conditions within a rectangular region in time and space may be determined based on the boundary observations.

The distance traveled and time spent by vehicle i in the rectangular region are respectively denoted by d_i and r_i. The dimensions of the region are given by the period duration T and cell length X. In the raw data, the intersects with the space (LDD) and time (FCD) borders are provided. This may be used to determine the ground truth of the space-mean speed, flow and density. According to Edie (1965), the density is given by the total travel time, which is the sum over the travel time r_i of all vehicles crossing the region, divided by the area of the region, thus:

$$\rho = \frac{\sum_i r_i}{XT} \tag{4.1}$$

Similarly the flow may be determined by dividing the total distance traveled by the area of the region:

$$q = \frac{\sum_i d_i}{XT} \tag{4.2}$$

From these variables, the space-mean speed is obtained by dividing the flow with the density:

$$u^S = \frac{q}{\rho} \tag{4.3}$$

Edie (1965) thus yields a ground truth estimation for all macroscopic traffic variables. Therefore, it is possible to determine the estimator performance for each of these variables.

4.2.3 Performance statistic

Performance statistics provide insight in the overall performance of the estimator. It compares the true u and estimated $\hat{u}$ value of the variable yielding a single performance statistic. The single statistic makes comparisons between different estimators easy and clear.

Estimator performance may be indicated by multiple statistics. Yuan (2013) uses the common performance statistics are the Root Mean Square Error (RMSE). Another performance statistic is the Mean Absolute Error (MAE). To keep comparison clear a single performance statistic will be used. It is chosen to use the MAE as it provides an intuitive result which has a direct meaning. The MAE is obtained using the following equation.

$$MAE = \frac{\sum (abs\,(u - \hat{u}))}{NN} \tag{4.4}$$

Here NN denotes the number of estimates which are compared. The state vector $\mathbf{x}$ contains the density ρ and space-mean speed u^S estimates for the freeway. The final estimates of these variables, thus in $\hat{\mathbf{x}}(k|k)$, will be compared with the ground truth. As discussed in Section 2.1 the ρ estimates are the most important macroscopic traffic variable for control in the FOT Amsterdam. Therefore, the estimation performance of this variable will be assessed. Furthermore, as the new data-type, FCD, is related to u^S this variable will be assessed.

Additional to the overall estimation performance, the performance in free-flow and congested conditions will be assessed separately. This provides extra insight in the performance of the estimators in different traffic conditions. As some of the decisions made in the functional design are related to the expected performance in different traffic conditions. For instance, the driver heterogeneity is expected to be lower in congestion and for that reason a lower error variance is assigned to FCD-based u^S estimates in congested conditions.

4.3 Experiments

As stated above, three experiments will be performed. Firstly, the proposed methodology is compared with the traffic state estimation methodology used in the first phase of the FOT Amsterdam. Secondly,

the effect of the data densities on the traffic state estimator performance is investigated. Thirdly, the effect of the choices and assumptions made related to the measurement error covariance matrix $\mathbf{R}$ will be evaluated.

4.3.1 Comparison with the ASM

Additional to the proposed methodology the Adaptive Smoothing Method (ASM) (Treiber & Helbing, 2002) is applied. This methodology is extended following the original idea of Yuan & Hoogendoorn (2014) combined with some new (own) inputs. These expansions allow to incorporate point floating car observations (FCD). To assign different weights to individual LDD and FCD observations a fixed weighing factor is introduced. In the ASM, the FCD only affects the u^S estimations and thereby indirectly the ρ estimations. However, the q estimations are not affected by the introduction of this new data-type. This is limitation which is not overcome in the extension of the ASM. The methodology and important decisions are discussed in more detail in Appendix C.

Comparison with the ASM serves three objectives. Firstly, assessing the performance of the ASM-estimator itself and the contribution of the individual data-types to these estimations. Secondly, evaluating the added value of a model-based methodology. Finally, providing a reference performance for further analysis. The latter is required to evaluate the main research question. According to the researchers involved in the FOT Amsterdam, the traffic management system in the first-phase provided good results. Therefore, the ASM only using LDD provides freeway state estimates which are accurately enough to yield the desired affect. Therefore, the performance of the ASM-based estimator with the data-availability of the Dutch freeway A10-W will serve as reference performance. On this freeway the freeway loop-detector spacing is approximately $500m$ (Hoogendoorn et al., 2013).

The proposed model-based estimator requires a discretisation of space and time. This discretisation is more restrictive than for an estimator based on the ASM. Nevertheless, the comparison will be done on the level of discretisation of the proposed methodology. In the FOT Amsterdam, the ASM-based estimates are used as input for the Freeway Bottleneck Inspector (FBI). The FBI estimates the probability of a breakdown a certain discritised locations. In the end the ASM-based estimates will thus be used for discretised space. Therefore, it is assumed that in this application it is fair to compare the ASM-based estimates and the proposed methodology-based estimates on the same level.

The ASM does not involve flow measurements of the ramps. However, these measurements, r and s, are valuable for model-based methodologies. To have a fair comparison these measurement are not included[3]. Only if it is required from a stability perspective, thus if more than one ramp is located in a single cell, the ramp flow r or s is included.

[3]It would be interesting to assess the added value of ramp flow data, thus inflow r and outflow fractions β. Furthermore, if the loop-detector spacing increases and these detectors are not placed at convenient locations, an ASM-based estimator may miss breakdowns (Yuan & Van Lint, 2015). As this advantages of model-based traffic state estimators are already shown in previous work, we want to keep the number of experiments limited and the research focus lies on other elements this will not be evaluated in this research. However, for future research it is interesting and may even lead to improved performance with respect to the (relative) performance shown in this research.

4.3.2 Data-availability

For the two data-types, the data-availability may be expressed using different variables. The LDD-availability may be expressed by the loop-detector spacing. This denotes the distance between two consecutive loop-detectors. In reality, there are some variations in the loop-detector spacing based on discontinuities in the road[4]. In such a case, one may denote the LDD-availability in terms of the average loop-detector spacing. The FCD-availability may be expressed by the penetration rate and the update frequency. The former is the percentage of vehicles which is observed. The latter is the frequency at which individual observations are shared with the FCD-provider. For instance, individual vehicles may provide their location and speed every 15 s.

In this research, loop-detectors are placed at the downstream boundary of each cell. These cells, used for the model-based estimator, will vary in length based on the road discontinuities. Furthermore, every ramp and the upstream boundary conditions have loop-detectors The LDD-availability will be varied by selecting which loop-detectors are included. To keep the complexity of the comparisons between the estimation using different data-availability limited, the FCD-availability will only be varied using the penetration rate. An update frequency with the same duration as the period duration will be selected. Furthermore, the FCD-observations are all in the middle of the period. For instance, if a period is considered from $t = 0.0$ s to $t = 15.0$ s, the observed vehicles provide their information at $t = 7.5$ s.

When varying the data-availability, it is important that the data which is included for a certain availability setting is always included in higher availability settings. This criterion is introduced to exclude the possibility that a vehicle or loop-detector provides more informative data which yields better estimation results irrespective of the data-availability. For both data-types this effect may be expected in the case this criterion is not implement. This may lead to biased results in the data-availability experimental set-up, thereby making it difficult to observe the effect of the data-availability variation.

4.3.3 Estimator input parameters

The proposed methodology requires inputs related to the traffic flow model and the uncertainties of model and LDD-based estimations. To evaluate how these parameters affect the traffic state estimation, these will be varied.

By varying the traffic flow model parameters the influence of the terms in the METANET model dependent on the respectively parameter may be evaluated. Varying the uncertainty in the model provides an insight in the effect of assigning more weight (or trust) to a model-based or LDD-bases estimate.

[4] An example of such a discontinuity is an on-ramp. It may be convenient to retrieve data before and after the ramp. This may provide information related to the on-flow. Furthermore, it may be a bottleneck location for which it is thus desired to retrieve real-time information for traffic management purposes.

4.3.4 Data-fusion schemes

In this research, an approach is proposed to fuse the LDD and FCD based u^S estimation according to there uncertainties and use the resulting combined estimate and uncertainty in the EKF. Here the FCD-based u^S estimate uncertainty is dependent on the measurement characteristics and estimated traffic conditions. To assess the validity of the choices made to come to the proposed approach two alternative approaches will be evaluated.

The first alternative approach is the simplest alternative. In this alternative a fixed u^S measurement error variance will be used which is thus independent of the data-type(s) used for the estimation. Furthermore, if both LDD and FCD are available for the respective estimates are weight with a fixed factor. This leads to the following relation.

$$u^S = w_{fcd} u^S_{fcd} + (1 - w_{fcd})\, u^S_{ldd} \tag{4.5}$$

Where w_{fcd} is used to weigh the FCD-based estimation u^S_{fcd} and LDD-based estimation u^S_{ldd}. This alternative provides insight in the validity to assign a different uncertainty to u^S estimations based on the information which is available.

In the second alternative approach a fixed FCD-based u^S measurement error variance is selected. If a single data-type is available it is assigned its respective error variance, while if both data-types are available they are weight with the individual error variances and assigned a combined, lower error variance. In this alternative the only difference with the proposed approach is that the FCD measurement error variance is not based on the estimated penetration rate and predicted traffic conditions. This validates the choices made related to the dependent FCD-based measurement error variance.

4.4 Experiments and next steps

The microscopic traffic simulation program FOSIM provides the LDD and FCD and data to determine the ground truth. Furthermore, experiments are proposed to evaluate the performance of the proposed algorithm. The next step is design an algorithm which makes it possible to conduct these experiments in this chapter. This algorithm should be flexible while taking into account the condition expressed in this chapter, like those provided for varying data availability.

Chapter 5

Algorithm design

In order to test the performance of the methodology proposed in Chapter 3 an algorithm has to be designed. This algorithm is designed with the objective to perform the experimental set-up. As explained in Chapter 4, there are multiple parameters and methodologies which we want to vary. The algorithm is designed such that it is easy to vary these parameters and methodologies.

5.1 Algorithm build-up

The MATLAB algorithm is designed to perform two tasks, namely traffic state estimation and performance estimation. In this thesis, all steps taken to come from the raw (FOSIM) data to the estimations of the macroscopic variables are considered to be part of the traffic state estimation. This thus also entails determining the input data based on data-density and measurement error parameters.

For this purpose the algorithm is build-up according to the scheme shown in Figure 5.1.

5.1.1 Estimator inputs

The estimator inputs are divided into three categories, namely estimation domain, estimator settings and measurements.

5.1.1.1 Estimation domain

The estimation domain is related to the time and space domain for which the estimation will be performed. Firstly, the road lay-out has to be defined. For each cell i the length Δ_i and number of lanes λ_i should be defined. Furthermore, the location of all the ramps are expressed in the cell number. Secondly, the time domain has to be defined. This entails the start and end time and the period length T.

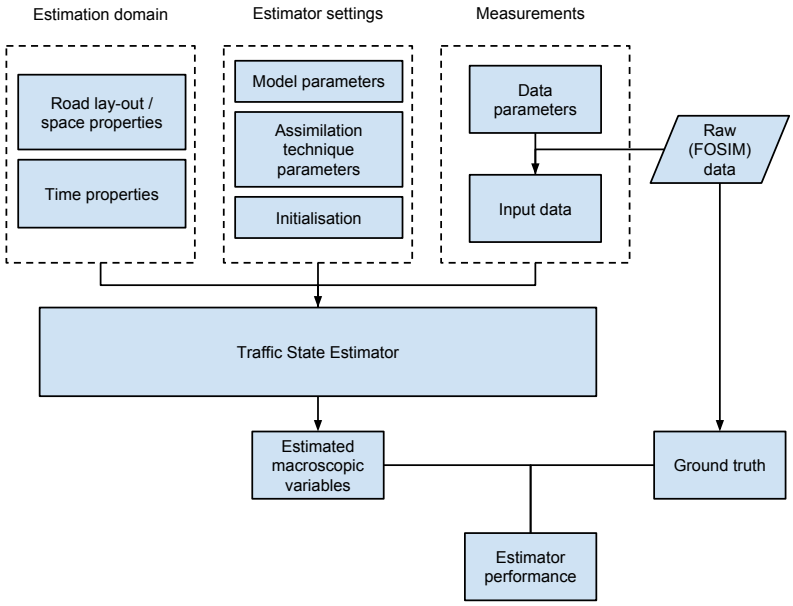


FIGURE 5.1: Visualization of the MATLAB algorithm for freeway traffic state estimation based on the proposed methodology

5.1.1.2 Estimator settings

The estimator requires input parameters and some variables have to be initialized. The required input parameters are related to the traffic flow model METANET and the error variances of the observation and traffic flow model. For the METANET model, the parameters of the dynamic speed equation have to be provided. For the error variances, the model error covariance matrix $\mathbf{Q}$ and parameters for the measurement error covariance matrix $\mathbf{R}(k)$ are needed. Additional to these parameters the state vector $\mathbf{x}$ and state error covariance matrix $\mathbf{P}$ should be initialized, thus $\mathbf{x}(0|0)$ and $\mathbf{P}(0|0)$ are defined.

5.1.1.3 Measurements

The data parameters are related to the data-density and the expected real data characteristics. The raw data retrieved from FOSIM cannot directly be used for traffic state estimation as it has another format as the real data. For instance, FOSIM registers all individual vehicle passes of the loop-detectors. Therefore, the raw LDD should be transformed to the one-minute lane-specific time-mean speeds which are available in reality.

For the data-density the available loop-detectors and penetration rate has to be defined. As stated in Section 4.3.2, it important that data included in lower densities settings is always included in higher density settings. The algorithm is build in such a way that this condition holds. The LDD density is varied by selecting the available loop-detectors. The FCD density varied based on a random draw from a uniform distribution between 0 and 1 assigned to each individual vehicle. Every vehicle for which the random draw yield a number lower than the penetration rate, where 1 denotes a penetration rate of

100%, is observed during its whole trajectory. The different FCD-densities are drawn from a single draw such that the condition stated in Section 4.3.2 holds.

5.1.2 Traffic state estimator

The mentioned estimator settings are used for the traffic state estimator based on the proposed methodology. This methodology is a recursive updating filter. Each time-step the state vector is estimated based on the traffic flow model and observation model. For period k four steps are taken starting with the previous period posterior estimates of the state vector $\mathbf{x}(k-1|k-1)$ and state error covariance matrix $\mathbf{P}(k-1|k-1)$.

1. A priori estimates of the state vector $\mathbf{x}(k|k-1)$ and error covariance matrix $\mathbf{P}(k|k-1)$: The former is estimated based on the macroscopic traffic flow model and $\mathbf{x}(k-1|k-1)$. The later is determined using the derivative matrices $\mathbf{A}(k-1)$ and $\mathbf{\Gamma}(k-1)$, the model error covariance matrix $\mathbf{Q}$ and $\mathbf{P}(k-1|k-1)$.
2. Measurement vector $\mathbf{y}(k)$ and measurement error covariance matrix $\mathbf{R}(k)$: These are based on the data-input and $\mathbf{x}(k|k-1)$. If both data-types provide information for a certain element in $\mathbf{y}(k)$ the information will be fused by means their uncertainties.
3. Kalman Gain $\mathbf{K}(k)$: This is determined using the derivative matrices $\mathbf{C}(k)$ and $\mathbf{\Sigma}(k)$, $\mathbf{P}(k|k-1)$ and $\mathbf{R}(k)$.
4. Posterior estimates of the state vector $\mathbf{x}(k|k)$ and error covariance matrix $\mathbf{P}(k|k)$: In the former $\mathbf{x}(k|k-1)$ is corrected based on $\mathbf{K}(k)$ and the difference between $\mathbf{y}(k)$ and the model based estimate of the measurement vector.

5.1.3 Estimator performance

As stated in Section 4.2, the traffic state estimator performance is determined by comparing the estimates with the ground truth. The estimates of the variables which are compared, ρ and u^S, are direct outputs of the traffic state estimator based on the proposed methodology. The ground truth is derived following Edie (1965) using the raw FOSIM data.

5.2 Algorithm design and next steps

The algorithm design is important to conduct fair experiments. Therefore, in this research emphasis was placed on the algorithm design. Due to the complexity of the algorithm, especially the proposed methodology-based estimator, error may occur. For instance, errors may occur while programming or in the earlier derivation stages. To solve such error the approach provided in Appendix E proved valuable.

Using the designed algorithm all experiments proposed in the previous chapter may be conducted. In the next chapter, the estimator inputs will be defined. This is the last step before the experiments are conducted and the related findings are presented.

Chapter 6

Estimator inputs

The performance of the proposed traffic state estimator is tested by means of the experimental set-up presented in Chapter 4. Before the experiments may be conducted, the estimator inputs have to be defined. As depicted in Figure 5.1, the traffic state estimator requires certain inputs. This section shows all parameters and choices related to these inputs. The denoted three-way split, namely estimation domain (Section 6.1), estimator settings (Section 6.2) and measurements (Section 6.3), will be followed in this chapter.

6.1 Estimation domain

A model of a real freeway stretch is considered to evaluate the traffic state estimator performance. This is the Dutch freeway A13 from The Hague to Rotterdam. The FOSIM model of this freeway stretch is shown in Figure 6.1.

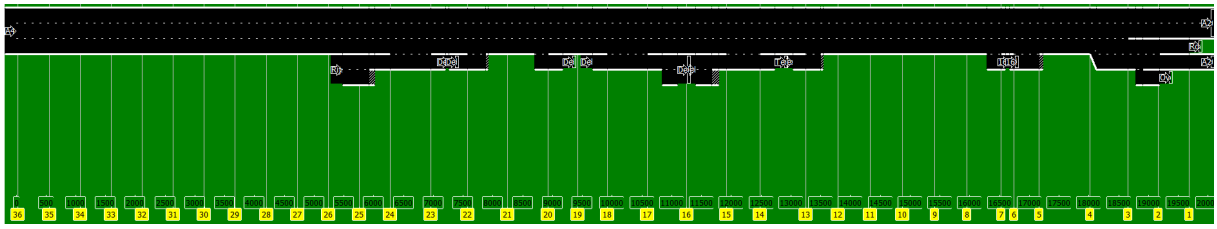
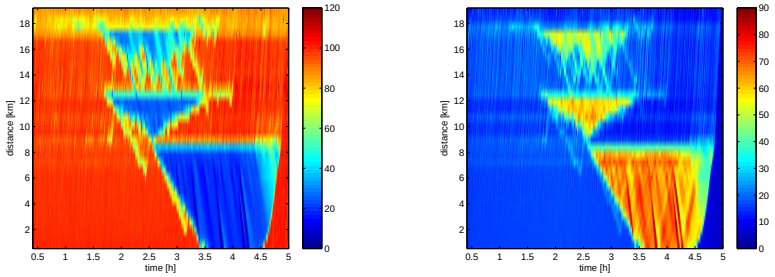


FIGURE 6.1: The A13 freeway stretch from The Hague to Rotterdam as FOSIM model.



(A) Space-mean speed (B) Density

FIGURE 6.2: Contour plots of the ground truth for the space-mean speed $u^S[km/h]$ and density per lane $\rho[veh/km/lane]$

6.1.1 Discretization

A period duration of $15s$ is chosen. It is expected that this duration leads to a relatively stable ground truth[1]. Following the Courant-Friedrichs-Lewy condition (3.1), the minimum cell length may be calculated as a function of the period duration and speed-limit. This leads to a minimum cell length of $500m$.

The total length of the model road stretch is $19,683m$. It is divided in 34 segments with lengths varying between $510m$ and $860m$. The segment-lengths are chosen based on the locations of the road discontinuities, thus the ramps and lane number changes. Following Wang & Papageorgiou (2005), the ramps are placed at the upstream boundary of the segments. This desired design is only violated at one instance, namely for on-ramp 6, as off-ramp 5 and that on-ramp are located to close to each other. The characteristics of each segment, which include the length, number of lanes and ramps, are provided in Appendix D.1.

6.1.2 Ground truth

The ground truth is obtained from the raw FOSIM data. Figure 6.2 shows the contour plots for the space mean speed u^S and density ρ. These plots are useful for comparisons with the estimated contour plots shown in the following section.

The contour plots show that both free-flow and congested conditions occur in the space-time domain. This makes it possible to evaluate the estimation performance of the estimators separately for free-flow and congested conditions. Furthermore, it may show what the effect of changes in traffic conditions, like a breakdown, on the stability of the proposed estimator[2].

[1] During earlier performance estimation shorter periods and cell lengths were chosen. The short period, namely $6s$, resulted in relatively large short-term fluctuation of the ground truth, namely in the order of 10%. This may be caused by platoons of vehicles. To reduce this problem a larger period is selected.

[2] Yuan & Van Lint (2015) showed that in some conditions the ASM may miss congestion while it is observed by model-based EKF methodology. In this research, the combination of road lay-out, traffic conditions and available loop-detectors does not show this downside of the ASM. However, their research provides an extra indication that a model-based estimation methodology may lead to improved estimation performance. On the other side, it is expected that the observed downsides of the ASM are reduced when FCD is involved.

TABLE 6.1: METANET parameters

Parameter	Value	
v_f	102	km/h
ρ_{cr}	33.5	veh/km
a	1.867	-
τ	1200	s
ν	60	km^2/h
κ	40	veh/km/lane
δ	0.0122	-
ρ_{max}	180	veh/km/lane

6.2 Estimator settings

A three-way split is made in the estimator setting, namely the traffic flow model parameters, assimilation technique parameters and data-densities.

6.2.1 Traffic flow model parameters

The traffic flow model METANET is dependent on a number of parameters which have to be defined. Firstly, the parameters of the stationary speed equation, namely the free-flow speed v_f, critical density ρ_{cr} and a, are defined. These are selected based on the work of Hegyi (2004). Secondly, the dynamic speed equation parameters, namely τ, ν, κ and δ are chosen. Here only τ differs from the parameters chosen by Hegyi (2004). In estimation runs with $\tau = 20s$ the filter became unstable. It was found that this problem could be solved by increasing the value of τ to $1200s$. Effectively this means that the effect of the fundamental diagram and the anticipation term within the METANET dynamic speed equation are reduced.

6.2.2 Assimilation technique parameters

The error covariance matrices $\mathbf{Q}$ and $\mathbf{R}$ need to be defined. The model error covariance matrix $\mathbf{Q}$ is fixed, while the measurement error covariance matrix $\mathbf{R}(k)$ is dynamic when FCD is included. The diagonal elements of both matrices provide the error variances of the corresponding variables. The off-diagonal elements denote error covariances between variables. In $\mathbf{Q}$ and $\mathbf{R}(k)$, the off-diagonal elements are set to zero.

The element-values selected for both matrices influence the performance of the estimator. For further analysis and comparison with other methodologies it is therefore important that these parameters are correctly selected. This will be based on a selection procedure.

In Section 3.4.1.1, the error variance of FCD-based u^S estimates is derived. As this error variance is given, it is taken as the starting point to determine the other error variances. In Appendix D.3 the procedure

TABLE 6.2: Assimilation technique parameters selected based on the procedure shown in Appendix D.3.

Parameter	Value	
$E[(\eta^{u^S}_{i,LDD})^2]$	10^2	$(km/h)^2$
$E[(\eta^{q}_{i,LDD})^2]$	250^2	$(veh/h)^2$
$E[(\xi^{u^S}_{i,LDD})^2]$	25^2	$(km/h)^2$
$E[(\xi^{q}_{i,LDD})^2]$	100^2	$(veh/h)^2$
$E[(\xi^{r}_{i,LDD})^2]$	ramp-specific	-
$E[(\xi^{\beta}_{i,LDD})^2]$	ramp-specific	-
$E[(\xi^{u^S_0}_{i,LDD})^2]$	25^2	$(km/h)^2$
$E[(\xi^{q_0}_{i,LDD})^2]$	100^2	$(veh/h)^2$
$E[(\xi^{\rho_N}_{i,LDD})^2]$	1.5^2	$(veh/km)^2$
$E[(\xi^{v_f}_{i,LDD})^2]$	0.5^2	$(km/h)^2$
$E[(\xi^{\rho_{cr}}_{i,LDD})^2]$	0.1^2	$(veh/km)^2$
$E[(\xi^{a}_{i,LDD})^2]$	0.01^2	-

taken to select the parameters is shown. In this procedure the selection of all parameters is based on the performance of the algorithm. Here a single data setting will be used, namely LDD-availability setting 2 and a FCD-availability of 10%. The parameters are selected consecutively. Firstly, the error variance of LDD-based u^S estimates, $E[(\eta_i^{u^S})^2]$, are selected. Secondly, the error variance of LDD-based q estimates, $E[(\eta_i^q)^2]$, are selected. In the third and fourth step, the model error variances are considered, namely respectively $E[(\xi_i^{u^S})^2]$ and $E[(\xi_i^q)^2]$. Finally, the remaining parameters are selected. The last step will be extensive, as it is expected that the influence on the estimator performance of these parameters is less than of the other parameters. The resulting parameter settings are depicted in Table 6.2.

6.2.3 Initialization

The state error covariance matrix $\mathbf{P}$ and state vector $\hat{\mathbf{x}}$ are updated each step. The initial value of the state error covariance matrix, $\mathbf{P}(0|0)$, is taken to be $\sqrt{\mathbf{Q}}$. For the initial value of the state vector, $\hat{\mathbf{x}}(0|0)$[3], each cell is assigned the same initial speed. This is $u^S = 100$ km/h. For the initial densities it is taken into account that the road is not completely filled in the beginning of the simulation. Therefore, the first cells have a higher initial rho, namely 14 veh/km/lane, than the other cells, namely 4 veh/km/lane. The boundary condition variables are based on the settings of the FOSIM simulation.

The choices for $\mathbf{P}(0|0)$ and $\hat{\mathbf{x}}(0|0)$ are important for the estimates of $\mathbf{P}$ and $\hat{\mathbf{x}}$. However, this effect fades out with time. To make sure the effect of the chosen initialization on the performance is negligible a fade-in period of eighty periods, thus $80T$, is taken. This period is thus not considered when determining the performance of an estimator.

[3]Although the initial state is denoted as $\hat{\mathbf{x}}(0|0)$, the estimation does not start at the $t = 0s$ of the simulation. In this case no vehicles would be initially present on the road. Therefore it is chosen to start the estimation at period $k = 20$.

6.3 Data availability

As an input of the traffic state estimator, the available data needs to be defined. For the LDD, the availability is expressed in the loop-detectors which are used. This yields an average spacing. For the FCD, the availability is expressed by the penetration rate, which is the percentage of vehicles which is observed.

In order for the proposed model-based methodology to be stable, at least one loop-detector location has to be available between consecutive ramps or the ramp flow has to be observed.

As the ASM-based traffic state estimator does not utilize ramp flow information, in the comparison between the ASM and proposed methodology no ramp flow information will be used. Therefore, the minimal LDD-availability setting has at least one freeway loop-detector between two consecutive ramps. Only between off-ramp 5 and on-ramp 6 no loop-detector is available, as these ramps are located in the same cell. Therefore, as an exception on-ramp 6 is observed. Three different LDD-availability setting are considered. The available loop-detectors in these settings are depicted in Table D.2 in Appendix D.2. The average loop-detector spacing in these settings is respectively $579m$, $984m$ and $1312m$[4].

During a meeting at the Dutch road authorities (Rijkswaterstaat), a penetration rate of 10% to 20% was said to be possible by a potential FCD-provider. Therefore, the penetration rate will be varied between 0% and 20%. This leads to six FCD-densities which will be evaluated, namely 0%, 4%, 8%, 12%, 16% and 20%.

6.4 Experiments and next steps

Combined the experimental set-up, estimation algorithms, estimator inputs and ground truth provide all the required preconditions to perform the experiments and evaluate the performance of the proposed methodology. The results and related findings will be discussed in Chapter 7.

[4]If a road would have been considered with less discontinuities (like ramps), the average loop-detector spacing could have been increased further.

Chapter 7

Findings

In the previous chapters the experiments and considered traffic situation are presented. These experiments are set-up to evaluate the decisions made in this research. However, there is an overlap between the findings from individual experiments. To keep the discussion of the results and associated findings clear, it is decided to present the findings one-by-one and discuss the related experimental results.

In total, six findings are discussed in this chapter. The related field and findings are depicted below. Firstly, the estimation performance of the proposed methodology is compared with the ASM-based alternative. This is a high level comparison in which the main focus lies on the added value of FCD for the model-based EKF and the non-model based ASM. This comparison is shown in Section 7.1. Next, four more specific findings for the proposed methodology are discussed. In Sections 7.2, 7.3, 7.4 and 7.5 the estimator stability, dependent estimation uncertainties, methodology complexity and the loop-detector cell locations are discussed. Finally, the main research question is addressed in Section 7.6. Here the potential of the proposed methodology in combination with FCD for the FOT Amsterdam is discussed.

1. **Estimation performance of methodologies**: FCD has a larger added value to the proposed methodology than the ASM-based methodology.
2. **Stability proposed methodology**: FCD improves the stability of the proposed methodology-based estimator which involves (Eulerian) LDD.
3. **Dependent estimation uncertainty**: FCD-based u^S error variance depends on the penetration rate and traffic conditions. However, it seems that not all uncertainties are described. The choice to use a fixed error variance for all non FCD-based estimations may still be improved. For instance, it seems that the LDD-based estimation uncertainty is dependent on the traffic conditions.
4. **Model complexity**: The proposed methodology is complex. However, this does not have to be a hurdle when implementing it in practice.
5. **Loop-detector cell locations**: The cell location of loop-detectors (downstream boundary) is not optimized for the u^S estimates.

6. **Proposed methodology in FOT Amsterdam**: Balancing LDD and FCD in the proposed methodology yields accurately enough estimations, while the number of required loop-detectors may be reduced.

7.1 Estimation performance of methodologies

The original ASM is not able to incorporate FCD for traffic state estimation purposes. However, in Appendix C an extension of the ASM is proposed which makes it possible to combine LDD and FCD for ASM-based estimations. This methodology is thus used to obtain the results discussed in this section. The estimation performance is expressed in the MAE of density ρ and space-mean speed u^S. The performance of the ASM-based estimator for varying data availability and different true traffic conditions is shown in Figure 7.1.

The MAE for estimations based on the proposed methodology is evaluated in a similar way as the ASM-based estimations. Figure 7.2 shows the MAE of ρ and u^S for varying data available in different traffic conditions. With respect to the ASM-based estimation performance plots a reference (black) line is plotted. In this section the estimation performance is not compared to the reference line, however it will be considered in the discussion in Section 7.6.

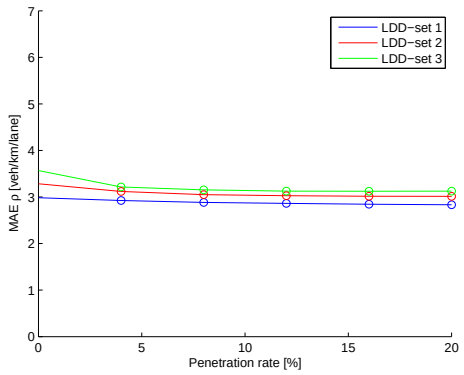
7.1.1 Adaptive smoothing method

The ASM-based estimator directly estimates u^S and q from the data. These estimates are used to estimate ρ. Involving FCD only directly influences the u^S estimates and thereby indirectly the ρ estimates. The q estimates are not affected by the FCD.

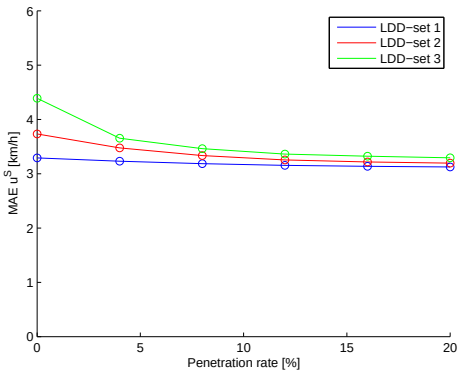
Figure 7.1a shows that the overall density estimation improves when more FCD becomes available and is involved in the estimation. However, the improvement in performance is limited. For instance, in the case of LDD availability setting 1, the difference in MEA ρ between no FCD and FCD with a penetration rate of 20 % is approximately than 5 %, namely respectively 2.98 veh/km/lane and 2.83 veh/km/lane. The overall speed estimation, shown in Figure 7.1b, also shows a slight improvement due to the involvement of FCD. Especially the performance in the case of the lowest LDD availability improves. The smallest non-zero penetration rate, namely 4 %, already leads to an MAE reduction of more than 15 %, namely from 4.39 km/h to 3.65 km/h.

The estimation performance during free-flow conditions is shown in Figures 7.1c and 7.1d for respectively ρ and u^S. For the lowest two LDD availability settings the density estimates do not seem to improve when involving FCD. The performance for the lowest LDD even slightly decreases, the MAE increases from 2.18 veh/km/lane to 2.19 veh/km/lane, respectively for 0 % and 20 % penetration rate of FCD. From Figure 7.1d it is clear that the speed estimation in free-flow do not have to benefit from additional data in the form of FCD.

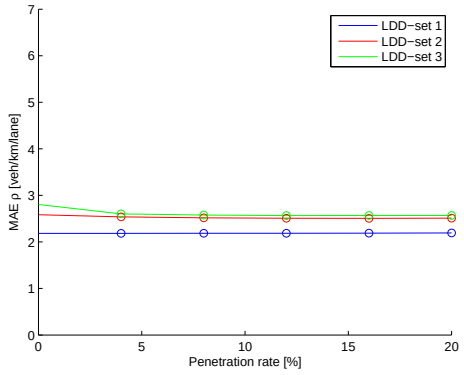
The estimation performance in congestion is shown in Figures 7.1e and 7.1f. Three features directly draw attention. Firstly, both the ρ and u^S estimations improve due to increasing FCD availability. It is



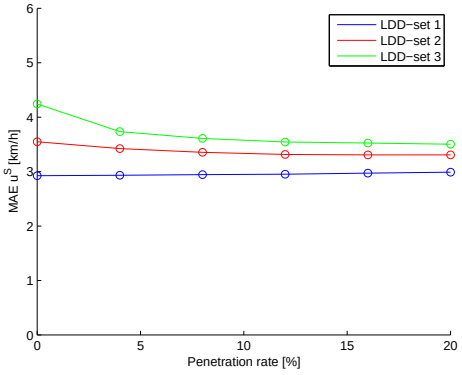
(A) Density all conditions



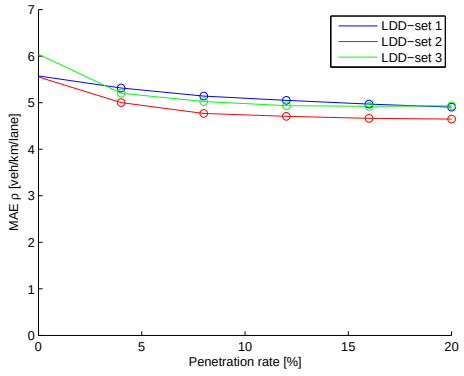
(B) Speed all conditions



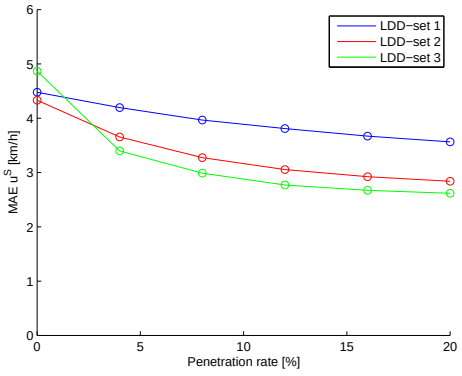
(C) Density free-flow conditions



(D) Speed free-flow conditions



(E) Density congested conditions



(F) Speed congested conditions

FIGURE 7.1: Mean Absolute Error (MAE) for the density ρ and space-mean speed u^S estimations based on the ASM for varying data availability. For LDD three data availability are considered, where LDD-den 1 denotes the highest data availability, which are depicted by the three lines. For each of these LDD availabilities the FCD availability is varied from a penetration rate of 0% to 20% with steps of 4%. To evaluate the estimator performance in free-flow, congestion and both conditions the MAE is depicted separately for these conditions.

clear that the u^S estimations benefit more from the FCD in congestion than in free-flow. Secondly, the estimation performance no longer follows the hierarchy observed in free-flow conditions in which higher LDD availability leads to better performance. The u^S estimation even shows the opposite relation, see Figure 7.1f, when FCD is involved in the estimation. The causes for this effect will also be further investigated by means of the contour plots. Thirdly, the difference in performance in free-flow and congested conditions is larger for ρ estimations than for u^S estimation, namely respectively in the order of 2.0-2.5 and 0.75-1.25. This may indicate that the q estimation drives the decreased performance in congested conditions. As explained, involving FCD in the ASM-based estimation does not overcome this problem.

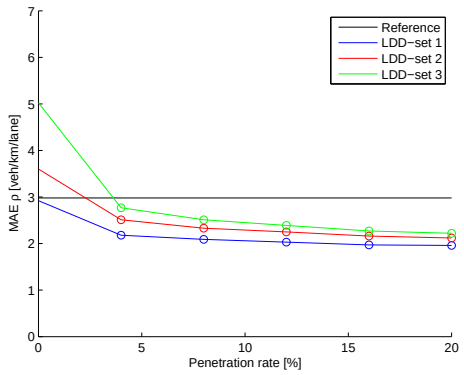
Based on the ASM-based with LDD, to added value of FCD was expected to be limited for two reasons. Firstly, the FCD only affects the u^S estimates and not the q estimates. Secondly, without FCD the u^S estimates are already relatively well, especially in free-flow. As expected adding FCD in the ASM-based estimator yields only a small increase in estimation performance. Due to driver behavior heterogeneity, FCD may even decrease the estimation performance in free-flow conditions. This will be further explained in Section 7.3. In congestion, FCD is of clear added value for u^S estimation performance. This may be explained by the lower driver behavior heterogeneity and the lower performance of LDD-based estimation. Despite the improvement in u^S estimation performance, only a limited ρ estimation performance increase is observed as the q estimation is not affected.

7.1.2 Proposed methodology

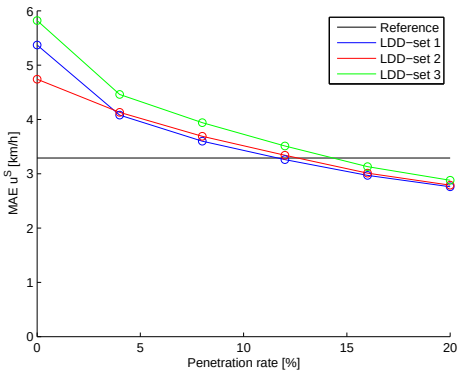
The proposed methodology is designed to utilize the information from both the available LDD and FCD. In the proposed methodology all macroscopic variables may be affected, through the model-based predictions and corrections.

The estimation performance independent of traffic conditions is depicted in Figures 7.2a and 7.2b for the ρ and u^S estimations respectively. In all LDD availability settings, both ρ and u^S are estimated more accurately for increasing FCD availability. For the ρ estimations FCD with a penetration rate of 4 % leads to relatively large performance improvements, namely respectively 25 %, 30 % and 45 % for LDD availability settings 1, 2 and 3. For higher FCD penetration rates the ρ estimations keep improving, but with smaller steps. The u^S estimates show a more continues accuracy improvement. This may have three explanations. Firstly, more information related to u^S becomes available which may lead to a more accurate u^S estimation of the observation model. Secondly, for individual cells an increase in observations leads to a decrease of the u^S measurement error variance of these cells. As a result of the decreased u^S measurement error variance more emphasis is placed on the speed observations with respect to the flow observations and traffic model predictions. Thirdly, more cells may be observed which affects the speed estimations in these cells.

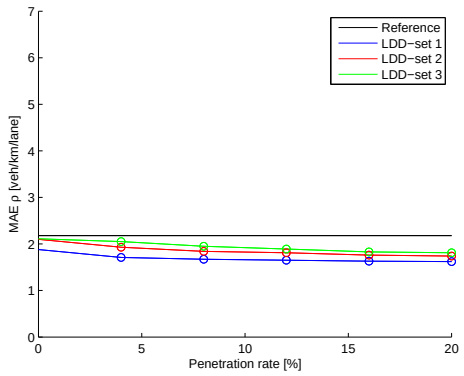
In addition to the direct positive effect caused by an increase in data in k has on the estimations in k, it also affects future estimations through the traffic flow model. The model-based predictions use the estimated variables in the state vector. Assuming that the traffic flow model provides unbiased predictions, estimations with a higher accuracy in k should lead to more accurate predictions for $k+1$.



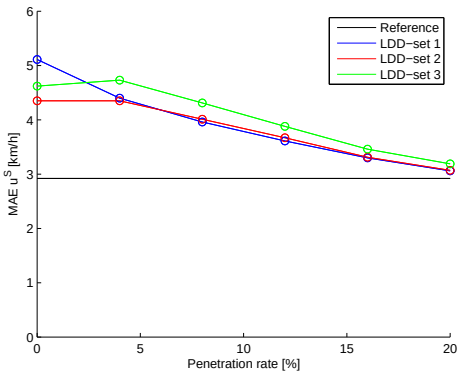
(A) Density all conditions



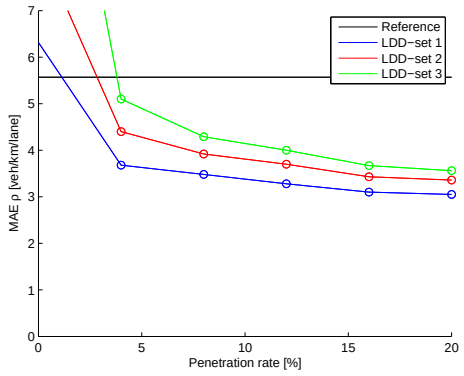
(B) Speed all conditions



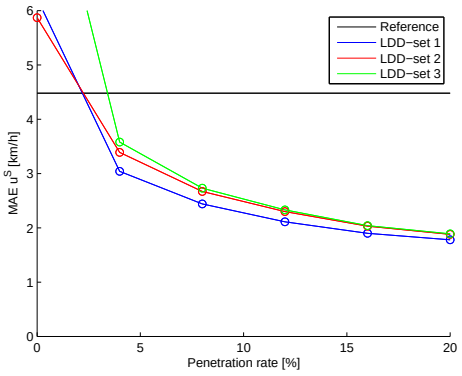
(C) Density free-flow conditions



(D) Speed free-flow conditions



(E) Density congested conditions



(F) Speed congested conditions

FIGURE 7.2: Mean Absolute Error (MAE) for the density ρ and space-mean speed u^S estimations based on the proposed methodology for varying data availability. For LDD three data availability are considered, where LDD-den 1 denotes the highest data availability, which are depicted by the three lines. For each of these LDD availabilities the FCD availability is varied from a penetration rate of 0% to 20% with steps of 4%. To evaluate the estimator performance in free-flow, congestion and both conditions the MAE is depicted separately for these conditions. The reference (black) line depicts the estimation performance of the ASM-based estimator in LDD availability setting 1 without FCD.

In this way, the increased information on u^S is able to improved accuracy for all macroscopic traffic variables.

For the proposed methodology stability may be an issue. This one of the causes of the low estimation performance in the case no FCD is available. This issue will be addressed in more detail in Section 7.2.

7.1.3 Comparison

In the previous two sections the estimation performance of an ASM-based and proposed methodology-based estimator are evaluated. The estimation methodologies incorporated by the estimators differ which leads to differences in estimation performance. These differences and the underlying dynamics are discussed in this section.

The ASM-based estimator is simple and stable, while the proposed methodology-based estimator is more complex and may become unstable if no FCD is available. Without FCD, the ASM-based estimator is able to match or even exceed the performance of the proposed methodology-based estimator. This may be explained by the low performance and instability of the latter estimator in congested conditions. In contrast to the ASM-based estimator the low performance in congestion also negatively affects future estimations through the recursive updating. However, this does not have to be seen as a negative property of the proposed methodology as it also works in favor of the estimates when individual estimates are improved.

The ASM-based estimator does not clearly benefit from the FCD. In the performance evaluation of the ASM-based estimated a number of FCD-related limitations were exposed. Firstly, the FCD is not able to affect the q estimations. As the u^S estimations are already relatively accurate, the potential added value of FCD is limited. Secondly, for low penetration rates the individual speeds available in the FCD even lead to a lower estimator performance in free-flow due to vehicle driver heterogeneity. However, the added value of FCD to the estimation performance in congested conditions is clear.

The limitations of the ASM were expected and also served as a basis for proposed methodology. The proposed methodology opts to describe the uncertainty in the FCD-based estimates. These should not only be based on the data availability, but also on the observed traffic conditions. For instance, the difference in vehicle heterogeneity dependent on the traffic conditions is incorporated by (speed variance based on speed). Furthermore, the (unknown) penetration rate is included by means of the estimated penetration rate. Combined this forms the proposed measurement based FCD uncertainty. This design has a direct effect on the u^S estimations, but also affects future estimations and the ρ estimations. Indirectly, the q estimates are also affected.

In contrast to the ASM, the model-based proposed methodology shows a large positive affect from involving FCD. Both the ρ and u^S estimation performance keeps improving for an increasing FCD penetration rate. This may be assigned to four main reasons. Firstly, there is more room for improvements as the performance without FCD is low and the u^S estimations are still less accurate in low penetration rate situation than the ASM-based estimations. This low performance without FCD is also caused by the instability of the estimator. This is discussed in more detail in Section 7.2. Secondly, all macroscopic

variables variables are affected in the proposed model-based estimator. Thirdly, improved estimations have a larger positive affect on future estimation. Incorporating new data thus has a larger influence. Finally, the proposed methodology involves the knowledge of the uncertainties of the FCD-based estimates. This will be further discussed in Section 7.3.

In this section, the dynamics behind the estimation performance of the proposed methodology are not discussed in depth. In the next sections, this will be discussed in more detail with the related findings.

7.2 Stability proposed methodology

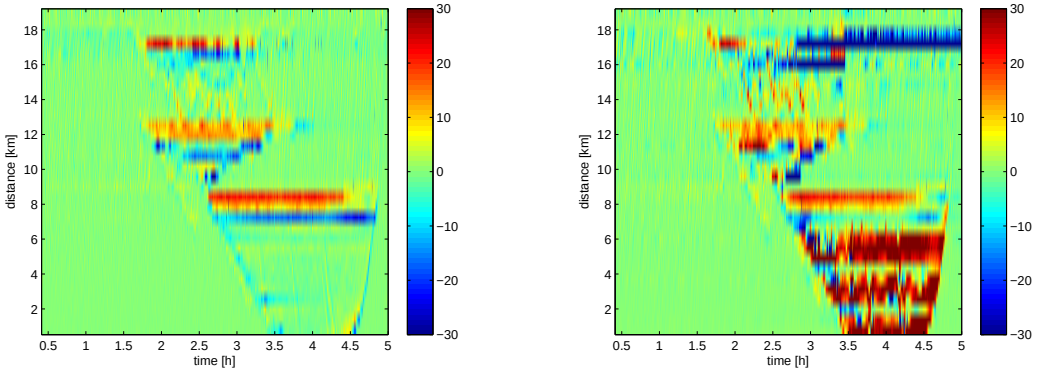
The proposed methodology-based estimator is unstable[1] for the LDD availability settings 2 and 3. This instability occurs in the correction step. To ensure stability of the estimator, τ is set to a higher value as in other applications like Hegyi (2004). A side effect is that this parameter setting directly influences the model-based prediction, more specifically the terms in the dynamic speed equation of the traffic flow model, which involve τ. Thereby, the influence of the fundamental diagram and the anticipation term is limited. As a result, the influence of ν and κ, which are involved in the anticipation term is also limited. Furthermore, the effect of the speed drop term caused by merging at an on-ramp, which involves δ and κ is also limited as the δ is small, following Hegyi (2004), and it only applies to a small number of cells. As a result of these settings the effect of varying the parameters is expected to be small. This expectations was validated by independently varying each parameter with a factor of 0.5 and 2.0. This experiment did not show any changes in the performance on the two decimal level and is thus negligible.

The density estimation errors for the lowest and highest LDD availability settings are shown in Figure 7.3. These figures indicate that the estimator has problems estimating the traffic conditions in and around congested areas in the space-time domain. For the LDD availability setting 1, this is primarily indicated by the errors at the downstream boundary of the congestion[2]. For the LDD availability setting 3 the estimator becomes unstable resulting in the most downstream congested area. This instability propagates over time till the end of the estimated period, thus $t = 5.0$ h.

A cause of the instability of the proposed methodology-based estimator may be found in the data-characteristics of the LDD. The loop-detectors are placed at fixed locations and thus only observe a fixed set of cells. In the observed cell, estimation errors may occur which propagate over time. In the example depicted in Figure 7.3b, vehicles accumulate based on the traffic flow model relations. This erroneous estimation is never corrected as these cells are never observed. FCD provides a solution for this problem. It is able to reduce the estimation errors in these previous unobserved cells, leading to a more stable estimator. This improves both the ρ and u^S estimates. The FCD thus has a positive effect on the stability of the estimator.

[1] This holds in this specific case of data-availability, road lay-out and traffic demand. Although the instability is caused by the correction step, it was found that it could be prevented by altering the traffic model parameters, more specifically τ. However, this means that the influence of the fundamental diagram and anticipation term within the METANET dynamic speed equation are reduced. As this changes the performance of the traffic flow model prediction, this is not desired. However, for the performance and stability of the estimator it is required.

[2] In Section 7.5 it will be shown that one of the factors behind the estimation errors at these congested area boundaries can be found in the loop-detector locations.



(A) LDD availability setting 1

(B) LDD availability setting 3

FIGURE 7.3: Density estimation error $\rho - \hat{\rho}$ in veh/km/lane in the lowest and highest considered LDD availability settings without FCD.

7.3 Dependent estimation uncertainty

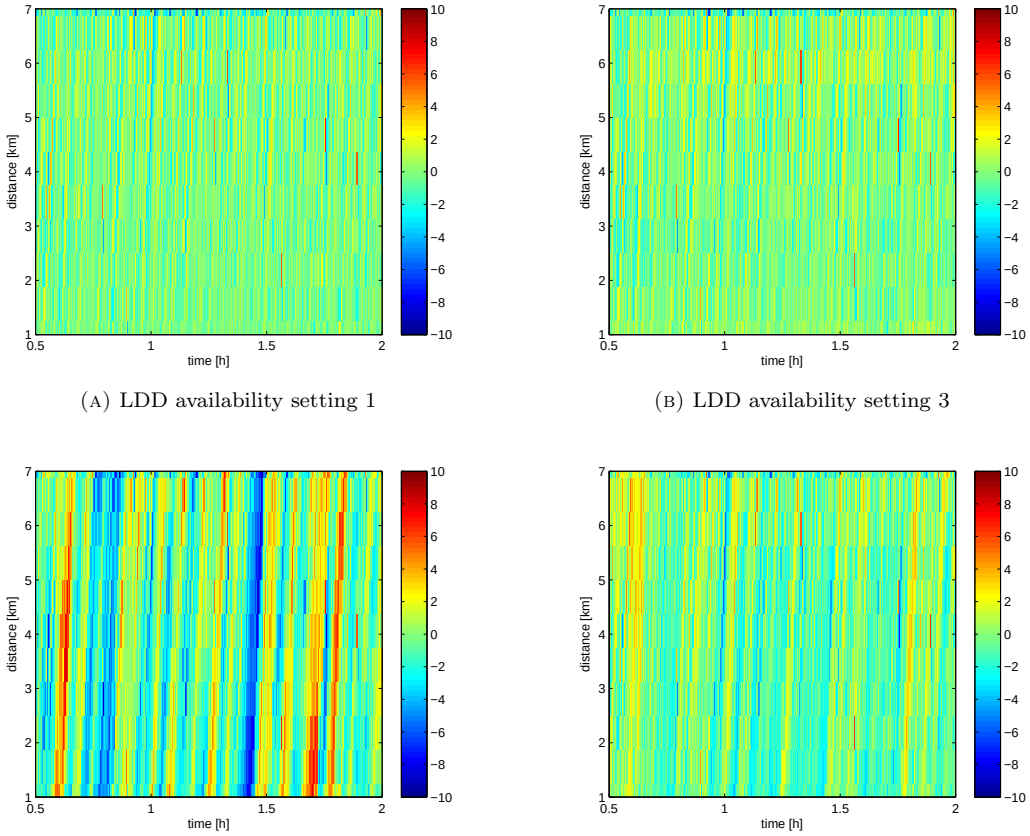
In this research a dependent FCD-based u^S estimation uncertainty is proposed and implemented in the freeway traffic state estimation methodology. In this section the findings related to these estimation uncertainties are discussed. Furthermore, findings related to the uncertainties in other, thus model- and LDD-based, estimations are discussed. Even though the proposed methodology implements fixed uncertainties for these estimates, the experiments provides evidence which may be interesting for future research.

The uncertainties of the estimations are evaluated using both the experiments conducted with the ASM and proposed methodology-based estimators. After these results are presented, the findings for the dependent estimation uncertainties are discussed.

The ASM-based estimator allows to investigate LDD- and FCD-based estimates independently. This analysis does not only provides insight in the uncertainties in these estimations, but may also explain the observations in Section 7.1. Furthermore, alternative schemes of involving FCD in the proposed methodology as assessed to evaluate the choices made related to the dependent FCD-based u^S estimation uncertainty.

7.3.1 Evaluation of ASM-based estimates

The initial analysis based on the results depicted in Figure 7.1 showed that the estimations during free-flow do not always benefit from additional data in the form of FCD. This (expected) effect may be investigated in more detail. For this purpose the estimations based on the LDD and FCD may be evaluated separately. In Figure 7.4 the u^S estimation errors based on single data-types are shown. Appendix C explains the methodology used to obtain these plots and how they may be interpreted.



(A) LDD availability setting 1

(B) LDD availability setting 3

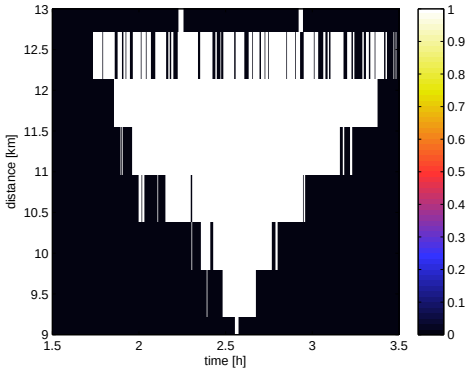
(C) 4 % FCD penetration rate

(D) 20 % FCD penetration rate

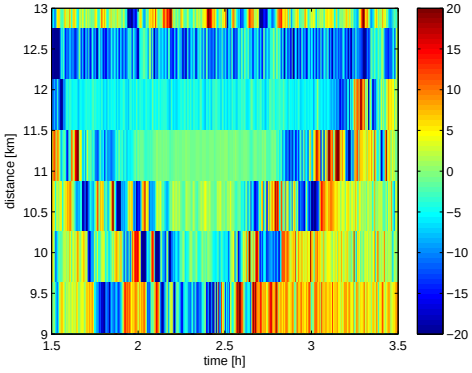
FIGURE 7.4: Space-mean speed estimation error $u^S - \hat{u}^S$ in km/h in free-flow conditions solely based on LDD or FCD for ASM-based estimations. The combination of the period from $t = 0.5$ h till $t = 2.0$ h and road-stretch from $x = 1.0$ km till $x = 7.0$ km is selected as this entire area is in free-flow conditions. For both data-types the lowest and highest evaluated data availabilities are shown.

To observe which features cause the decrease in u^S free-flow estimation performance an area within the space-time plot which is purely has free-flow conditions is selected. Within this area, the estimations solely based LDD result in errors mostly within the range of -2 to 2 km/h. Furthermore, it may be observed that the errors for the (lower) data availability setting 3, see Figure 7.4b, are larger than for data availability setting 1, see Figure 7.4a. The estimations solely based on FCD show the importance of individual vehicle speeds in the estimations and the resulting errors. Figure 7.4c clearly shows the effect of vehicle heterogeneity on the FCD-based speed estimation. Due to the low penetration rate, namely 4 %, individual vehicles have a major effect on the estimations. Vehicles which drive at a relative slow and high speed result respectively in a negative (blue) and positive (red) estimation error. These errors propagate downstream following the vehicle trajectory. If one looks closely, it may be seen that the positive estimation errors propagate with a higher speed. For a higher penetration rate, namely 20 %, the FCD-based estimation errors decrease, see Figure 7.4d. The effect is vehicle heterogeneity is still visible, but has a smaller effect on the estimation errors.

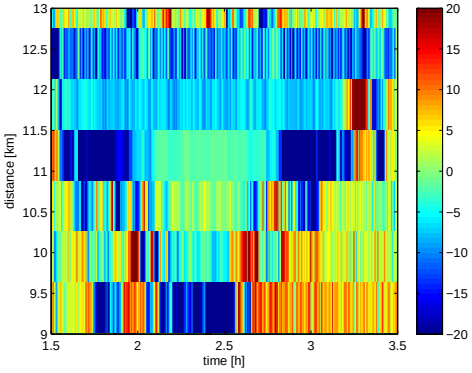
In addition to the estimation errors in free-flow conditions, the estimation errors in congested conditions



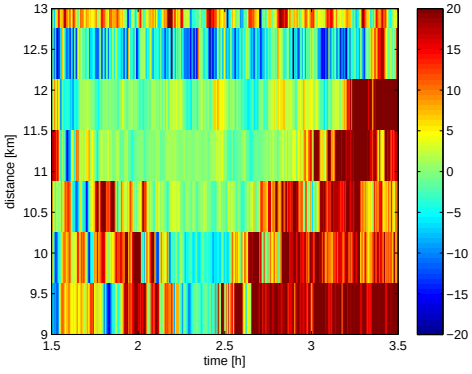
(A) Congested area



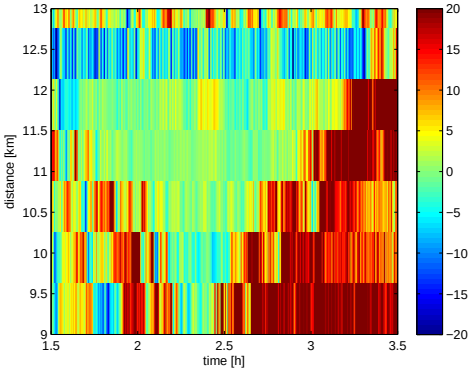
(B) LDD availability setting 1



(C) LDD availability setting 3



(D) 4 % FCD penetration rate



(E) 20 % FCD penetration rate

FIGURE 7.5: Congested area and space-mean speed estimation error $u^S - \hat{u}^S$ in km/h in congested conditions solely based on LDD or FCD. The combination of the period from $t = 1.5$ h till $t = 3.5$ h and road-stretch from $x = 9.0$ km till $x = 13.0$ km is selected as a large part of this area is in congested conditions. The congested area is represented by a binary variable where congested is 1 and free-flow is 0. For both data-types the lowest and highest evaluated data availabilities are shown.

are evaluated. Therefore, a congested area within the u^S contour plot is evaluated, see Figure 7.5. Only a part of this space-time area is in congestion. This is indicated by a binary variable in Figure 7.5a,

where white and black respectively denote congestion and free-flow.

The LDD-based estimation errors are shown in Figures 7.5b and 7.5c. These figures are relatively blue, indicating negative error and thus that the LDD-based estimations overestimate the u^S in this congested area. The lowest LDD availability setting results in larger (more/darker blue) negative errors, which is in line with the 0% FCD performance shown in Figure 7.1f. Figures 7.5d and 7.5e respectively show the FCD-based u^S estimation errors for penetration rates of 4 % and 20 %. Although some difference may be observed for between the FCD-based estimation errors, these are less clear than in free-flow conditions. In these high density conditions, the vehicles are restricted by the upstream vehicle. This leads to a lower vehicle heterogeneity between vehicles. A well-known relation between traffic conditions and vehicle heterogeneity, which also serves as a basis for the proposed methodology. The increase in u^S estimation performance for increasing FCD penetration rate, observed in Section 7.1, may be explained by means of the weight assigned to the FCD-based estimations. This weight increases when more FCD-points become available, ceteris paribus, which is the case for increasing FCD penetration rate.

7.3.2 Evaluation of proposed methodology-based estimates

In addition to the ASM-based estimator, the proposed methodology-based estimator proves important insights. To investigate the choices related to the error variances, thus estimation uncertainties, two experiments are conducted. Firstly, the fixed error variance parameters are varied to study their effect on the estimator performance. Secondly, alternative approaches are applied to incorporate the FCD in the proposed model-based estimator.

7.3.2.1 Fixed error variance parameters

The assimilation technique parameters which are varied are the LDD-based q and u^S measurement error variances and the q and u^S model error variances. For these error variances fixed values were selected in Section 6.2. Independently, a factor of 0.5^2 and 2.0^2 is assigned to each of these error variances. Reducing the error variance leads to a higher weight which is assigned to the related estimation and vice versa.

Table 7.1 shows the performance of the estimation for different assimilation technique parameters. When more weight is assigned to the u^S estimates, thus a factor 0.5^2 is assigned to the error variance, the estimation performance of u^S increases. The inverse relation is observed when the weight assigned to the q estimates is increased. As an example the q measurement error variance, $E[(\eta^q_{i,LDD})^2]$, with a factor of 0.5^2 and 2.0^2 is taken. For a higher error variance, thus less weight assigned to the q measurements, the overall MAEs of ρ and u^S are respectively 2.31 veh/km/lane and 3.18 km/h. For a lower error variance, the overall MAEs of ρ and u^S are respectively 2.13 veh/km/lane and 3.48 km/h.

Only the model u^S error variance, $E[(\xi^{u^S}_{i,LDD})^2]$, shows a better performance on both variables when it is changed. When a factor of 0.5^2 is assigned to this error variance, the ρ and u^S estimates improve in all traffic conditions. This indicates that the original choice of $E[(\xi^{u^S}_{i,LDD})^2]$ is not optimal in for this data-availability. However, as only one combination of data availability settings is compared, it is not clear how the performance is affected in other settings.

TABLE 7.1: MAE for the ASM-based and proposed methodology-based traffic state estimator with varying error variances. For this experiment the LDD-availability setting 2 and a FCD penetration rate of 12 % are considered. As no changes are proposed to the FCD-based error variances in this experiment, the FCD penetration rate is not varied. To investigate the estimation performance in different traffic conditions, the performance is assessed separately for free-flow and congested conditions.

error variance	Factor	MAE	Traffic conditions overall	free-flow	congestion
Reference		ρ [veh/km/lane]	2.16	1.73	3.55
		u^S [km/h]	3.28	3.61	2.23
$E[(\xi^{u^S}_{i,LDD})^2]$	0.5^2	ρ [veh/km/lane]	2.15	1.72	3.54
	0.5^2	u^S [km/h]	3.11	3.39	2.19
	2.0^2	ρ [veh/km/lane]	2.18	1.75	3.56
	2.0^2	u^S [km/h]	3.39	3.74	2.25
$E[(\eta^{u^S}_{i,LDD})^2]$	0.5^2	ρ [veh/km/lane]	2.20	1.77	3.61
	0.5^2	u^S [km/h]	3.14	3.32	2.53
	2.0^2	ρ [veh/km/lane]	2.19	1.75	3.60
	2.0^2	u^S [km/h]	3.57	3.99	2.20
$E[(\xi^{q}_{i,LDD})^2]$	0.5^2	ρ [veh/km/lane]	2.16	1.73	3.54
	0.5^2	u^S [km/h]	3.31	3.64	2.23
	2.0^2	ρ [veh/km/lane]	2.18	1.74	3.60
	2.0^2	u^S [km/h]	3.22	3.52	2.23
$E[(\eta^{q}_{i,LDD})^2]$	0.5^2	ρ [veh/km/lane]	2.13	1.72	3.43
	0.5^2	u^S [km/h]	3.48	3.87	2.24
	2.0^2	ρ [veh/km/lane]	2.31	1.83	3.88
	2.0^2	u^S [km/h]	3.18	3.46	2.28

For the u^S measurement error variance, $E[(\eta^{u^S}_{i,LDD})^2]$, the estimation performance in free-flow and congestion show an opposite relation. The free-flow estimates benefit from a decreased error variance, while the estimation performance in congestion decreases. This indicates that data or traffic state dependent measurement error variance may lead to a better estimator. It is also in line with the observations from the ASM-based estimator. Here it was shown that the LDD-based u^S estimates have a better performance in free-flow then in congestion. For the other error variances the difference in uncertainties dependent on traffic conditions in not clear from Table 7.1.

7.3.2.2 Alternative data-fusion schemes

The proposed approach to involve the FCD-based u^S estimations may be evaluated. It is proposed to base the FCD-based measurement error variance on the measurements and estimated traffic conditions. In the case both data-types are available for a single cell in k, the u^S estimate and related error variance are determined based on the individual estimates and related error variances. This approach is compared with a simpler one. In this simple approach, if both data-types are available for a single cell in k, their individual estimates are weight with a fixed factor. Furthermore, a fixed measurement error variance, $E[(\eta_i^{u^S})^2]$, independent of the data-type(s) which are used to come to the estimate.

TABLE 7.2: Estimation performance of the proposed methodology-based estimator with an alternative weighing scheme for FCD inclusion. The estimated value of u^S is equal to $w_{fcd}u^S_{fcd} + (1 - w_{fcd})\, u^S_{ldd}$. A fixed value is taken for the u^S error variance for all measurement-based u^S estimates, thus dependent on LDD, FCD and both. For this experiment the LDD-availability setting 2 and a FCD penetration rate of 12 % are considered. The penetration rate is not varied as this would lead to a more complex visualization of the results. To investigate the estimation performance in different traffic conditions, the performance is assessed separately for free-flow and congested conditions.

Weight FCD	Error variance	MAE	Traffic conditions overall	free-flow	congestion
Reference		ρ [veh/km/lane]	2.11	1.74	3.31
		u^S [km/h]	3.28	3.61	2.19
0.75	2.5^2	ρ [veh/km/lane]	2.18	1.83	3.33
	2.5^2	u^S [km/h]	3.43	3.78	2.30
0.50	2.5^2	ρ [veh/km/lane]	2.24	1.84	3.54
	2.5^2	u^S [km/h]	3.39	3.59	2.74
0.25	2.5^2	ρ [veh/km/lane]	2.39	1.88	4.06
	2.5^2	u^S [km/h]	3.57	3.64	3.34
0.75	5^2	ρ [veh/km/lane]	2.09	1.75	3.21
	5^2	u^S [km/h]	3.51	3.86	2.37
0.50	5^2	ρ [veh/km/lane]	2.17	1.77	3.49
	5^2	u^S [km/h]	3.54	3.76	2.81
0.25	5^2	ρ [veh/km/lane]	2.34	1.81	4.07
	5^2	u^S [km/h]	3.75	3.87	3.39
0.75	10^2	ρ [veh/km/lane]	2.15	1.74	3.46
	10^2	u^S [km/h]	3.74	4.13	2.47
0.50	10^2	ρ [veh/km/lane]	2.25	1.75	3.89
	10^2	u^S [km/h]	3.79	4.08	2.87
0.25	10^2	ρ [veh/km/lane]	2.46	1.78	4.67
	10^2	u^S [km/h]	4.01	4.17	3.46

Table 7.2 depicts the estimation performance with the proposed (reference) and alternative weighting scheme. Note that the weight factor w_{fcd} is only applied when both data-types are available. In other cells, the u^S estimate is purely based on the FCD. Note these results are provided for a fixed data-availability, namely LDD-availability setting 2 and a penetration rate of 12 %. It would be interesting to vary the penetration rate, however this would lead to a more complex table. Therefore, it is chosen to consider the estimation performance in different, free-flow and congested, conditions in this experiment and consider a varying penetration rate in the next experiment.

In these data-availability settings the best estimation performance is achieved with $w_{fcd} = 0.75$ and $E[(\eta_i^{u^S})^2] = 5^2$. The estimator even outperforms the reference estimator in ρ estimation performance.

Three features are observed in Table 7.2. Firstly, the u^S estimation performance benefits from a lower error variance. Independent of w_{fcd}, the u^S estimates show lower errors for the smallest $E[(\eta_i^{u^S})^2]$ compared with the other alternatives. Secondly, the best estimation performance in congestion on both variables is achieved when the highest weight is assigned to FCD, $w_{fcd} = 0.75$. In free-flow conditions the u^S estimates improve if more emphasis is placed on the LDD-based u^S estimates. Thirdly, the best

TABLE 7.3: Performance statistics for different FCD error variances. The considered LDD-availability is setting 2. The FCD penetration rate is varied to observe the relation between error variances and the penetration rate.

		FCD penetration rate				
$E[(\eta_{i,FCD}^{u^S})^2]$	MAE	4%	8%	12%	16%	20%
Reference	ρ [veh/km/lane]	2.49	2.25	2.11	2.07	2.04
	u^S [km/h]	4.13	3.67	3.28	3.00	2.80
5^2 (km/h)^2	ρ [veh/km/lane]	2.58	2.27	2.09	2.04	2.00
	u^S [km/h]	4.83	4.09	3.51	3.17	2.94
10^2 (km/h)^2	ρ [veh/km/lane]	2.51	2.29	2.18	2.15	2.13
	u^S [km/h]	4.36	3.88	3.51	3.28	3.13
25^2 (km/h)^2	ρ [veh/km/lane]	2.79	2.69	2.66	2.66	2.65
	u^S [km/h]	4.05	3.81	3.66	3.57	3.52

ρ estimation performance is achieved for $E[(\eta_i^{u^S})^2] = 5^2$. Despite the lower u^S estimation performance in this setting compared to settings with lower $E[(\eta_i^{u^S})^2]$, it performs better in density estimations.

These observation provide extra evidence that the FCD-based free-flow u^S estimates have a larger uncertainty than those in congestion. Furthermore, in congestion the FCD is of higher added value than in free-flow. Therefore, a higher weight relative to the LDD-based u^S estimates should be assigned to FCD-based u^S estimates in congestion with respect to free-flow. However, a better u^S estimation performance does not always yield a better ρ estimation performance. It may be that to much emphasis is placed on the u^S estimates. This holds for the alternative approach with a small fixed error variance, but also for the proposed approach. It may thus be the case that not all uncertainties are included in the proposed approach.

This analysis does not provide insight in the effect of a varying penetration rate. It is expected that this would effect the performance of the alternative approach more than the proposed approach. Therefore, this will be evaluated in the next alternative data-fusion approach.

A second, more advanced, data-fusion scheme is also evaluated. In this scheme, FCD and LDD is weight according to the error variance of the individual data-types. These error variances are fixed, which still results in a fixed weighting factor between the observations. However, the measurement error variance may differ for LDD and FCD-based estimates and is lowest when both data-types are available. This relation is depicted by (3.20) in Section 3.4.

Table 7.3 shows the estimation performance for different FCD-based measurement error variances. To keep the analysis of these results clear, the already widely studied difference between performance in free-flow and congested conditions is not depicted in this table.

Two main findings are observed in Table 7.3. Firstly, at a low penetration rate of 4 % the an error variance of 10^2 (km/h)^2 outperforms the error variance of 5^2 (km/h)^2. For higher penetration rates the estimator the smallest fixed FCD-based u^S estimation error variance ($E[(\eta_{i,FCD}^{u^S})^2]$) has the best estimation performance. Furthermore, the improvement steps for both the ρ and u^S are larger for a lower fixed $E[(\eta_{i,FCD}^{u^S})^2]$. These observations indicate that the uncertainty of the FCD-based u^S decreases

with an increasing penetration rate. Secondly, the estimator with $E[(\eta_{i,FCD}^{u^S})^2] = 5^2$ outperforms the estimator with a dependent $E[(\eta_{i,FCD}^{u^S})^2]$ (reference) for ρ estimates for penetration rates higher than 12 %. In these cases the dependent $E[(\eta_{i,FCD}^{u^S})^2]$ still shows a better u^S estimation performance. It seems that the proposed approach places to much emphasis on the u^S estimates. This may indicate that not all errors are included in the equation used to determine the FCD-based measurement error variance. A potential cause of this unobserved white noise is the time discretization.

In reality the penetration rate is not known may vary more than in these individual experiments. Therefore, the (improved) proposed approach may yield the best results. Improving the proposed approach may be as simple as adding a fixed term in (3.15) to include unexplained white noise.

7.4 Complexity of the methodologies

The proposed methodology has a higher complexity than the ASM. This reduces the understanding of the estimator and (in this case) also leads to a higher computation time. The computation time of the current algorithm based on the proposed methodology exceeds that of the ASM. However, in this research no emphasis was put on improving the efficiency and computation time of the estimator. This choice was purely made due to time restrictions and as other research was more applicable to the research department (Transport & Planning). Therefore, the computation times will not be evaluated quantitatively.

The general idea behind the proposed methodology is not difficult. The different elements of the proposed methodology may be explained intuitively and may even be more intuitive than the ASM. Firstly, the traffic flow model is based on relations and physical laws which one may understand. Furthermore, the idea of weighing estimations or a prediction and estimation based on there respective uncertainties is intuitive. Whether one does this consciously or not, this holds for many applications in life and thus comes natural to people. In the methodology this principle is used for weighing the information within the observation model and between the measurement-based estimations and model-based predictions. The traffic flow model and this principle are the blueprint for the proposed methodology.

Although the methodology may thus be explained intuitively, the exact computations will not be understandable to many individuals. The equations explained in Chapter 3 are difficult and it requires significant knowledge on the EKF and traffic flow theory to understand what they mean. This may be an extra hurdle for implementing this methodology in practice. On the other side, Kalman Filters are widely applied in different technological solutions used in every-day life, for instance GPS. This may indicate that methodologies with this level of complexity can still be used in real-life applications. For the proposed methodology there certainly is room for improvements, however, it is expected that it may be used in practice in the future.

7.5 Loop-detector cell locations

Figure 7.3a shows proposed methodology ρ estimation errors in the situation that all cells are observed by loop-detectors. At the boundaries of the congested areas some obvious, but explainable errors are

TABLE 7.4: MAE for the ASM-based traffic state estimator without FCD. In the original estimator the middle of the cell is taken as the reference-location for the entire cell. As alternative, the reference location used in the proposed methodology, namely the downstream boundary, is considered.

LDD setting	MAE	ASM original	ASM alternative
1	ρ [veh/km/lane]	3.12	3.20
	u^S [km/h]	3.29	3.95
2	ρ [veh/km/lane]	3.26	3.40
	u^S [km/h]	3.60	4.37
3	ρ [veh/km/lane]	3.55	3.69
	u^S [km/h]	4.27	4.91

observed. At the downstream boundaries positive (red) errors, which means that ρ is underestimated, are observed. At the upstream boundaries the opposite errors are observed, which means that ρ is overestimated. This may be explained by the placement of the loop-detectors, which is at the downstream boundary of the cells. The cells located at the boundaries of the congestion may be partly in congested and free-flow conditions. Due to this variation in traffic conditions the LDD gives a biased estimate for the traffic conditions at the boundaries of congestion. At the upstream boundary the LDD measures congestion conditions while a part of the cell is in free-flow conditions. This results in an overestimated ρ, thus negative (blue) estimation errors. For the downstream cells ρ may be underestimated, resulting in positive (red) errors.

In the current set-up of the two estimators, the ASM-based estimator is better able to incorporate the locations of the loop-detectors in estimating the conditions for the entire cell in a given period. This estimator assigns the observations to it true location[3], namely at the downstream boundary of the cell. To estimate the conditions for the entire cell in a given period, the middle of the cell and period is taken as reference. If two consecutive cells are observed using loop-detectors, the data of these loop-detectors may influence the cell estimations with approximately the same weight. This differs from the proposed methodology in which the observation model LDD estimations a completely based on the loop-detector at the downstream cell boundary. Placing the loop-detectors at this location is based on Wang & Papageorgiou (2005). As q is defined as the outflow of the cell, which occurs at the downstream boundary, this loop-detector location is desired. To examine the effect of taking the downstream boundary as reference for the entire cell, an ASM-based estimator is designed in with these cell reference locations. The performance of this estimator is compared with the original ASM-estimator for the three LDD availability settings. In all three settings the estimation performance for both ρ and u^S deteriorated, see Table 7.4. This is evidence that the downstream boundary condition is not the best representative location for the entire cell.

[3]Additional to the location x, the time t to which the estimation is assigned may also be better. In the ASM-based estimator the observation is assigned to a single time, namely the middle of the one-minute aggregation period. This may result in more accurate than assuming, which is done in the proposed methodology-based estimator, that these conditions hold for the entire minute. However, as this is not tested by means of an experiment, for the moment, no recommendations will follow from it.

7.6 Proposed methodology in FOT Amsterdam

In this research it is opted to develop a freeway traffic state estimator involving both LDD and FCD which may be used in practice in the near future. The traffic management application in which it may be used is the FOT Amsterdam. As explained before, the proposed methodology will be compared with currently implemented traffic state estimation methodology and data availability. Therefore, as reference, the performance of the ASM-based estimator at LDD availability setting 1 is considered. This performance is expected to be *accurately enough* for the entire traffic management system to work properly.

Figure 7.2 shows the estimation performance of proposed methodology and the reference performance. As discussed in Section 7.1, involving FCD has a clear positive effect on the performance of the proposed methodology-based estimator. In all traffic conditions the ρ estimates with the proposed methodology outperform the ASM with a FCD penetration rate equal or higher than 4%. The potential of FCD in traffic state estimation becomes especially clear when comparing LDD-availability setting 1 without FCD with the LDD-availability setting 3 combined with a FCD penetration rate of 4%. The latter data-combination shows a better estimation performance for both assessed variables, while the available loop-detectors is reduced from 35 to 16. This number may even be reduced further when a road-stretch with less discontinuities (ramps) is selected[4].

The u^S reference performance is not directly outperformed by the proposed methodology with FCD. However, as explained in Section 2.1 ρ is the most important variable for different elements within the traffic management system of the FOT Amsterdam. Therefore, the estimation performance for ρ is assumed to be decisive. Furthermore, in previous sections it was shown that the u^S estimations may be improved, at the expensive of the ρ estimations, by choosing lower u^S error variances. When the proposed methodology is implemented in the FOT Amsterdam traffic management system, these parameters may still be selected based on the performance of the entire system.

[4] As explained in Section 6.1 the presence of discontinuities limits the reduction in loop-detectors which are currently selected for the different LDD-availability settings. In this research, no loop-detectors were placed on the ramps as the ASM is not able to benefit from these loop-detectors and this may lead to an unfair comparison. However, this may also reduce the number of required freeway loop-detectors in the proposed methodology.

Chapter 8

Conclusions

This research proposes a freeway traffic state estimation methodology which is able to use a combined input of Loop-Detector Data (LDD) and Floating Car Data (FCD). LDD is a common input for traffic state estimators. However, in recent years, the rise of navigations systems and smart-phones have increased the potential availability of FCD. Therefore, the added value of FCD in traffic state estimation may be investigated. For a given traffic management application, the required traffic state estimation input may be expressed as a combination of LDD and FCD. Potentially, the LDD required to obtain the desired performance may be reduced due to the involvement of FCD. This would lead to a cost reduction for the road authorities.

Although the proposed methodology may be used for a wider range of traffic management applications, it is designed to function within the traffic management system in the FOT Amsterdam. This means that the data characteristics, thus traffic state estimator inputs, are based on the data that is (potentially) available within this application. The considered traffic management system aims to prevent or postpone the capacity drop. The capacity drop is the negative effect on the capacity in congestion. The system consists of different components which have different purposes, for instance related to estimation, prediction and control.

In order to evaluate the complete traffic management system, all components have to be designed and programmed. As this research focuses on the freeway traffic state estimator, this time expensive endeavor is not undertaken. An alternative evaluation approach is used. As a reference, the first phase of the FOT Amsterdam is considered. During this phase, the complete system was tested on a segment of the A10-N in Amsterdam. According to the involved researchers the system showed a sufficient performance during this field operational test. Therefore, the individual components may be taken as a reference. It is assumed that only one component, namely the freeway traffic state estimator is altered. The output, which consist of macroscopic variable estimates and their level in time and space, should yield at least the same accuracy as the reference traffic state estimator. The reference traffic state estimator should involve the methodology and input-data-characteristics which were available in the first phase of the FOT Amsterdam. In this phase the traffic state estimator was based on the Adaptive Smoothing Method (ASM) and loop-detector data was used with an average spacing of 500 m.

The Dutch freeways are equipped with double loop-detectors, which are able to observe individual vehicle passings and measure individual vehicle speeds for each lane independently. However, before these measurement are made available these raw measurements are aggregated over one-minute periods. This results in two variables, namely the lane-specific time-mean speed u^T and flow q. The data-characteristics of the potentially available FCD are based on a meeting at the Dutch road authorities. The potential FCD-provider stated that individual speed v measurements with a vehicle ID, time-stamp and road location (without lane-information) would be available up to a penetration rate of 20 %.

The available measurement data differs from the desired traffic state estimator output in both the level in time and space and in the variable types. The estimator should be able to provide estimations of the macroscopic variables density ρ, space-mean speed u^S and flow q for periods smaller than one minute. In the data only q is directly available, but its aggregation period is too long. In both the LDD and the FCD, u^S is not directly available. Based on the lane-specific u^T and q, u^S may be approximated. However, the approximated u^S is also still aggregated over a one-minute period. The individual speeds v are related to u^S. For a given road segment, the distribution of the instantaneous individual speeds may be approximated by a Gaussian distribution. The mean of this distribution is u^S. Literature shows that the variance of this distribution depends on the mean speed. Based on this relation u^S may be estimated from the FCD along with an uncertainty for this estimation.

It is decided to use a model-based traffic state estimation methodology. For the FOT Amsterdam this is interesting as it allows for an integral, estimation - prediction - control, approach. Furthermore, the new information, FCD, may positively affect the all the macroscopic traffic variables through the traffic flow model. Model-based traffic state estimation consists of three parts, namely a traffic flow model, an observation model and an assimilation technique.

As there are uncertainties in the relations between the measurement data and the desired estimator output, a traffic state estimation methodology able to incorporate these uncertainties is desired. It is chosen to use an Extended Kalman Filter (EKF). This assimilation technique is able to combine the observations (data) with the second-order traffic flow model METANET. Within the observation model, the uncertainty of the FCD-based u^S estimate is based on the available estimations and model-based predictions of the u^S and ρ. These predicted variables are used to estimate the traffic conditions and penetration rate. This uncertainty thus varies depending on the available data. Next, a combined LDD- and FCD-based u^S estimate uses the uncertainties of the u^S estimates based on the individual data-types.

The decisions made in the functional design of the proposed methodology are evaluated by means of various experiments. These experiments are conducted to observe the added value of combining the two data-types for traffic state estimation and evaluate how these data-types should be involved in the proposed methodology.

Additional to the proposed traffic state estimation methodology, the ASM is evaluated. To assess the potential added value of involving FCD for both models, the ASM is extended to be able to incorporate FCD. It is expected that individual floating car observations are less representative than the aggregated loop-detector observations. Therefore, a fixed factor is used to determine the respective weight of these observations.

In the evaluation of the ASM-based estimator two main limitations were found when combining LDD and FCD as input. Firstly, the ASM-based estimator already performs relatively well for u^S when only using LDD. As FCD only directly affects the u^S estimation, the potential for improvements due to involving this data-type is limited. Any improvement in the u^S estimation performance, affects the ρ estimation. However, as the q estimations are not affected, the (positive) effect on ρ remains limited. Secondly, due to driver heterogeneity, in free-flow the u^S (and indirectly ρ) estimations may be negatively affected by involving FCD. This error in FCD-based u^S estimations reduces with an increasing penetration rate. In congestion, the FCD does provide valuable information leading to improvements in ρ and u^S estimations. This positive effect on the estimation performance increases for a higher penetration rate. This finding indicates that the original decision to base the measurement error variance in the proposed methodology on the penetration rate seems valid. However, the positive estimation performance effect of an increasing penetration rate in congestion is lower than that in free-flow. In combination with the larger error for free-flow due to driver heterogeneity, this validates the decision to use the relation between mean speed and speed variance for the FCD u^S estimation error variance.

The limitations of the ASM-based estimator were expected and also served as input for the methodology design decisions. In the proposed methodology, the uncertainty assigned to the FCD-based u^S estimation is based on the estimated traffic conditions and the penetration rate. This follows the relations found in the FCD-based ASM estimations. The estimated traffic conditions and penetration rate are respectively based on the prior (model-based) u^S and ρ estimations. To determine the combined LDD- and FCD-based u^S estimate and its error variance, the individual estimations and error variances are weight based on the uncertainties, thus error variances. The experiments show that the estimation errors are dependent on the available data and traffic conditions. However, it is found that the currently implemented factors, driver heterogeneity and penetration rate, do not cover all potential factors of uncertainties.

The evaluation of the proposed methodology-based estimator performance using both LDD and FCD, showed that FCD has a large positive effect on the estimation performance. This positive effect has four main reasons. Firstly, without FCD some stability issues occur in the proposed methodology-based estimator. In these data conditions, the instability results in a relatively low performance of the estimator compared with the ASM-based estimator. The instability occurs in unobserved cells, which remain unobserved during the entire estimation period. When FCD is involved, these cells will also be observed in certain periods. These observations are able to prevent the instability and thereby have a positive effect on the performance. Secondly, in the proposed methodology FCD may affect the estimations of all macroscopic variables trough the traffic flow model. This may result in a larger (positive) effect on the ρ estimation performance than in the ASM-based estimator. Thirdly, estimation improvements do not only have local, in time and space, effects. Future estimations are also improved through the traffic flow model predictions. These should be more accurate when the implemented explanatory variables, the estimated state variables, are more accurate. Although the positive effect diminishes over time, it may propagate over the entire future estimation domain. Fourthly, the proposed methodology is designed to incorporate the uncertainties of the FCD-based estimations better than the ASM.

Although the proposed methodology shows promising results with respect to the ASM, there are still some limitations. Firstly, the instability issues, which led to alternative traffic flow model parameter choices. It is expected that there still is an error in the proposed methodology-based algorithm, which

causes this specific instability. This may further be investigated in future research. If this is indeed the case and the error is corrected, it is expected that the estimation performance will improve. Nevertheless, the estimator already provides better results than the ASM-based estimator. Secondly, it is shown that the location of the loop-detectors and how these are used to estimate u^S based on the LDD may be improved. In the proposed methodology the downstream boundary of a cell is taken as reference for the entire cell. This location is chosen as this allows to measure the outflow of the cell. However, in the ASM-based estimator the middle of the cell is considered as reference location for the entire cell. The u^S is estimated based on the observations near the estimation position in the space-time domain. The effect of taking the downstream boundary as reference for the entire cell is also clearly visible in the estimation performance at the boundaries of a congested area. Within these cells the traffic condition are expected to vary, with a fraction being in free-flow and a fraction in congestion. Therefore, taking the downstream boundary as reference results in biased estimates of the cell true traffic conditions.

There are thus still potential improvements for the proposed methodology. However, when balancing LDD and FCD it is able to achieve a accurately enough estimations. For the evaluated road, the A13-S, a penetration rate of 4 % is sufficient to yield a accurately enough estimations while the number of required loop-detectors is cut in half. A accurately enough estimation performance is defined as an estimation performance which is at least as good as in the first phase of the FOT Amsterdam. For the FOT Amsterdam it is thus expected that the proposed methodology in combination with the FCD made available by FCD-providers yield accurate enough estimations. In time, this may lead to a reduction of the required loop-detectors on the Dutch freeways.

Choices related to the number of loop-detectors installed on the freeway are long-term decisions. This research provides an insight in the potential of FCD for traffic state purposes. It shows that the number of loop-detectors may be reduced when the considered FCD becomes available without having to compromise on the estimation performance. This research may thus be seen as an input for future decisions on the number of loop-detectors installed on freeways and the locations where these are placed.

The work related to this subject is not finished. For this purpose recommendations related to the application and future research are provided in the next chapter.

Chapter 9

Recommendations

In this section the recommendations following from this research are discussed. These are split in recommendations for application, thus the traffic management system within the FOT Amsterdam, and recommendations for future research.

The recommendations provided in this chapter may benefit from a more extensive explanation. However, due to the large number of recommendations, it is chosen to keep the explanations of the individual recommendations limited. In some cases more research has gone into the recommendations that depicted here. Therefore, for a more in-depth discussion related to these topics, the author may be contacted.

9.1 Recommendations for the application

Recommendation for the application are steps/actions which are recommended to be taken for second phase the FOT Amsterdam and are related to the research performed in this thesis.

1. **Investigate causes of observed estimator instability:** The proposed methodology-based estimator without FCD becomes unstable in some settings. Even with the algorithm, the proposed methodology shows promising results. Therefore, the causes behind the instability may still be investigated in more depth. Potentially there is a programming error, which is not found yet with the procedure depicted in Appendix E. In the case of a programming error it should be solved. It is not expected that the problem lies in the methodology, as the LDD-based estimator is mainly based on Wang & Papageorgiou (2005). However, if there is an error in the methodology, changes should be made or a different methodology with Extended Kalman Filter (EKF) should be chosen.
2. **Investigate traffic state dependent LDD-based measurement error variance:** In this research the choice was made to keep the LDD-based measurement error variance independent of the measurements. This choice was based on complex and two-sided relation which the measurement and assumption errors have with u^S and the importance of the one-minute aggregation period for the q and u^S errors. However, when varying the assigned values to the error variances is was found

that there is a difference between traffic conditions. Therefore, it may still be beneficial to involve a dependent LDD-based measurement error variance. For the u^S error variances this may be as simple as having different error variances for free-flow and congested traffic conditions, which may be dependent on the prior estimate or measurements. To find how the q measurement error variance dependents on the traffic conditions and measurements extra research is necessary.

3. **Link the traffic state estimator with other components of the traffic management system:** One of the added values of having a model-based traffic state estimator is that it allows for an integral network approach (estimation - prediction - control). For instance, the Freeway Bottleneck Inspector (FBI) which is used to predict the breakdown probabilities at certain locations may be combined with the traffic state estimator.
4. **Investigate potential improvements to other components of the traffic management system:** Other components like the parameter estimator and the freeway bottleneck estimator may be improved. The conducted research also provides new opportunities which may be implemented to improve these other components. The proposed methodology-based estimator yields a state vector $\mathbf{x}$ and a state error covariance matrix $\mathbf{P}$. The former contains the estimated macroscopic variables u^S and ρ, while the latter contains error variances related to these estimations. The uncertainties of the FCD-based estimations are based on empirical observation and the other uncertainties are chosen accordingly. Therefore, the uncertainties available in $\mathbf{P}$ are expected to be in the order of magnitude of the true estimation uncertainties. However, this expectation is not tested and should therefore still be evaluated. In the expectation holds, these estimation uncertainties may be valuable information within the traffic management system of the FOT Amsterdam.
5. **Improve estimates based on LDD:** The LDD contains lane-specific one-minute aggregates for speed and flow, from which one-minute estimates for u^S and q are determined. However, in the estimator LDD-based estimations for shorter periods are desired. To obtain estimates for shorter periods it is assumed that traffic conditions are constant during this period. There are some methods to capture differences within the one-minute periods, for instance due to platooning (Schreiter et al., 2010). These may be implemented to improve the estimates based on the LDD for periods shorter than one minute.
6. **Apply with real-data**: At the moment the methodology is only tested with synthesized data. This was important to evaluate the estimation performance of the proposed estimator, as it possible to vary the data availability and compare the estimation with the true values. However, this data may differ from real data and therefore the estimator should be tested with real data before applying it in a real traffic management system.
7. **Dealing with (different) data availability lags:** In the proposed methodology the data is used to correct the estimates in the period which it describes. However, in reality there will be a certain time-lag depending on the data-source. This lag may differ for LDD and FCD. An example may be considered in which the LDD and FCD respectively have a lag of one and two periods. In period k, the FCD is thus available till $k-2$ and the LDD till $k-1$. In this case for every period till $k-2$ both data-types are available, in $k-1$ only LDD is available and for later periods no data is available. To incorporate all the information the methodology should be extended. This requires extra research. However, an approach is proposed which yields to desired effect. If an estimation is desired for k with the denoted data-availability, the following approach may be taken. Every period

k, the estimations for periods $k-2$ till k are determined. Here only $k-2$ yields the final estimates as in k all the data is available for that period. The estimation of $k-1$ is determined using the traffic flow model and LDD, while the estimation of k is determined using the traffic flow model.

8. **Different FCD-update frequencies:** In this research it is assumed that there are no differences in FCD-update frequencies. Furthermore, it is assumed that there is one observation per vehicle per period. However, in reality difference may occur and vehicles may be observed more than once per period. Dealing with these data-characteristics provides an extra challenge. Therefore, extensions to the methodology on this field have to be designed. In this approach, it is for instance important that vehicles observed twice in a single period are not counted twice. This would yield a biased representation of the u^S estimate and error variance through a biased penetration rate estimation.
9. **Algorithm efficiency:** The current algorithm is not programmed as efficiently as possible. Currently, many loops are used in the recursive updating steps. Replacing these, where possible, by matrix manipulations leads to a more efficient algorithm. This would reduce the computation time significantly. For the moment, the computation speed of the algorithm is not quantitatively assessed. However, as it is designed for real-time applications it is important to conduct this experiment before it may be implemented in the FOT Amsterdam.

9.2 Recommendations for future research

The recommendations for future research are more general recommendations than those presented in the previous section. However, in the future these may still be interesting for traffic state estimation and other applications within the FOT Amsterdam.

1. **Test the performance of simpler and cheaper loop-detectors**: The currently considered FCD only provides information related to u^S. At low penetration rates the estimations in free-flow conditions do not yield the accuracy of LDD-based estimates. However, if enough data is available it may even replace LDD speed data instead of being a supplement. In the proposed methodology, the current placement of the loop-detectors is chosen based on desired location for which we want to observe q. For the LDD-based u^S estimations this may lead to biased estimations. As the current loop-detectors have a dual observation purpose, thus are used to observe both speeds and flows, this conflict for the proposed methodology will remain. Alternatively, it may be investigated if it suffices to use loop-detectors (or other Eulerian sensing equipment) to flow measurements and FCD for speed measurements. This may potentially lead to decreased cost for loop-detectors.
2. **Implement a different traffic flow model:** It was not feasible to evaluate all choices made during this research by means of a quantitative analysis. For instance, the second-order traffic flow model METANET is not compared to other traffic flow models, due to time-constraints. The current proposed methodology-based estimator becomes unstable. To ensure stability a METANET parameter is changes, which reduced influences of elements within the dynamic speed equation. If the instability is not caused by errors in the algorithm, a simpler model may lead to more stable results. However, stability is not the only reason to test other traffic flow models. As stated in Chapter 2 there is an ongoing debate related to performance of different traffic flow models.

Therefore, it will be interesting to evaluate the performance with other traffic flow models. For instance, the first-order traffic flow model used by Van Hinsbergen et al. (2012). The proposed methodology of involving the FCD may be combined with a different traffic flow model with limited work.

3. **Implement a different filter:** In this research it is chosen to use the EKF. This choice is made based on comparison in literature. A important factor in this decision is the computation speed of the EKF (Van Hinsbergen et al., 2012), which should allow a real-time application with the proposed methodology. The EKF, however, has the restriction that in only allows Gaussian distributed errors. Yuan (2013) refers to various researches which have shown that the Gaussian assumption is not a bad choice. However, it may be interesting to compare the performance of this assimilation technique with other techniques. For instance, a Particle Filter (PF) relaxes the Gaussian assumption and may therefore lead to a more realistic representation of the true errors. The proposed methodology to involve FCD for traffic state estimation may also be incorporated in an estimator using a PF. Without the Gaussian assumption there is more freedom in describing the errors dependent on the measurements or traffic conditions. When different filters are compared both the estimation performance and the computation times should be evaluated quantitatively.
4. **Involve information about driver characteristics:** The floating car observations are treated as individual observations. However, in the FCD the vehicle ID is known. This may be used to learn about the driving behavior of vehicles. As was shown in Figure 7.4c, the driver behavior can be consistent over time. In other words, a driver which drives relatively fast is likely to drive relatively fast in the future. Although the data was collected from a simulation program, it is expected that drivers show consistent driving behavior[1].
5. **Opportunities in the case of improved FCD-characteristics:** The FCD-characteristics which followed from meetings at the Dutch road authorities and which are therefore considered in this research may be improved in the feature. For instance, the current individual speed observations may become vehicle class-specific and lane-specific. This would allow for multi-class or multi-lane traffic flow estimation using FCD.
6. **Involve a similar approach to Lagrangian coordinate traffic state estimation:** In an early stage of this research, alternative traffic state estimation methodologies were considered. Before the choice was made to estimate the traffic conditions in an Eulerian coordinate system, the methodology proposed by Kuwahara et al. (2013) was considered. He proposes to combine LDD and FCD for traffic state estimation in a Lagrangian coordinate system. Besides the difference in coordinate system, the main reason why this methodology was not evaluated is an inaccurate decision which it makes. It assumes that vehicles do not overtake each other, leading to the First-In-First-Out (FIFO) condition. However, in reality in multi-lane traffic, for instance freeway traffic, this condition may be violated. Some potential improvements were conceived. For instance, a similar probabilistic approach may be considered for the number of vehicles which are overtaken by an individual vehicle for which FCD is available. For a more in-depth discussion related to this topic, the author may be contacted.

[1] A similar approach may be adopted to learn the characteristics of other sensors, like individual loop-detectors. These may also show consistencies in the errors, like the miss probability.

7. **Incorporate additional data-types:** The proposed methodology is designed to incorporate additional data-types. This may both be Eulerian sensing data, for instance cameras, and Lagrangian sensing data, for instance headway estimations[2]. When a new data-type is added, the relation with the macroscopic variables should be described. Similar to the evaluated FCD, the data-characteristics of the new data-type should be evaluated. Based on these characteristics and the relation with the macroscopic traffic variables the data-based estimates and uncertainties may be determined and included in the methodology accordingly.

[2]A PhD-application has been submitted to investigate the potential of other types of FCD for traffic management purposes. An example of such a data-type is individual headway estimation collected from vehicles equipped with radar systems for Adaptive Cruise Control (ACC).

References

Bromiley, P. a. (2003). Products and Convolutions of Gaussian Distributions. *Biomedical Engineering*, *M*(2003).

Courant, R., Friedrichs, K., & Lewy, H. (1928). Uber die partiellen Differenzengleiehungen der mathematischen Physik. *Math. Ann.*, *100*(1), 32–74.

Daganzo, C. F. (1995). Requiem for second-order fluid approximations of traffic flow. *Transportation Research Part B: Methodological*, *29*(4), 277–286.

Daganzo, C. F. (2005, February). A variational formulation of kinematic waves: basic theory and complex boundary conditions. *Transportation Research Part B: Methodological*, *39*(2), 187–196. doi: 10.1016/j.trb.2004.04.003

Dijker, T., Bovy, P. H. L., & Vermijs, R. G. (1997). *Car-Following under Non-congested and Congested conditions.*

Edie, L. C. (1965). Discussion of traffic stream measurements and definitions. In *2nd int. symp. on the theory of traffic flow.* Paris: OECD.

Fellendorf, M., & Vortisch, P. (2001). Validation of the Microscopic Traffic Flow Model VISSIM in Different Real-World Situations. *Transportation Research Board 80th Annual Meeting*, 1–9.

Hegyi, A. (2004). *Model Predictive Control for Integrating Traffic Control Measures.* TRAIL Research School.

Hoogendoorn, S., & Knoop, V. (2013). Traffic flow theory and modeliing. In *The transport system and transport policy* (pp. 125–159). Delft.

Hoogendoorn, S., Landman, R., Van Kooten, J., & Schreuder, M. (2013). Integrated Network Management Amsterdam: Control approach and test results. *16th International IEEE Conference on Intelligent Transportation Systems (ITSC 2013)*(Itsc), 474–479. doi: 10.1109/ITSC.2013.6728276

Kalman, R. E. (1960). A New Approach to Linear Filtering and Prediction Problems. *Journal of Basic Engineering*(March), 35–45.

Kerner, B. S. (1999). The physics of traffic. *Physics World*(August), 25–30.

Knoop, V., & Hoogendoorn, S. (2012). Empirics of a Generalized Macroscopic Fundamental Diagram for Urban Freeways. In *92th annual meeting of the transportation research board* (pp. 1–16).

Kuwahara, M., Ohata, T., & Takigawa, T. (2013). Estimating Vehicle Trajectories on a Signalized Urban Arterial and a Motorway by Data Fusion of Probe and Detector Data. In *Optimum 2013 - international symposium on recent advances in transport modelling* (pp. 1–11).

Lighthill, M. J., & Whitham, G. B. (1955, May). On Kinematic Waves. II. A Theory of Traffic Flow on Long Crowded Roads. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *229*(1178), 317–345. doi: 10.1098/rspa.1955.0089

Minderhoud, M. M., & Kirwan, K. (2001). *Validatie FOSUM voor asymmetrische weefvlakken - CAP-WEEF fase 1.*

Ou, Q. (2011). *Fusing Heterogeneous Traffic Data : Parsimonious Approaches using Data-Data Consistency.*

Papageorgiou, M., Blosseville, J.-M., & Haj-Salem, H. (1990). Modelling and real-time control of traffic flow on the southern part of Boulevard Peripherique in Paris: Part II: Coordinated on-ramp metering. *Transportation Research Part A: General*, *24*(5), 361–370. doi: 10.1016/0191-2607(90)90048-B

Papageorgiou, M., Hadj-Salem, H., & Blosseville, J. (1991). ALINEA: A Local Feedback Control Law for On-ramp Metering; A Real-life Study. *Transportation Research Record: Journal of the Transportation Research Board*, *1320*, 58–64.

Richards, P. I. . (1956). Shock Waves on the Highway. *Operations Research*, *4*(1), 42–51.

Schreiter, T., van Hinsbergen, C., Zuurbier, F., van Lint, H., & Hoogendoorn, S. (2010). Data-model synchronization in extended Kalman filters for accurate online traffic state estimation. *2010 Proceedings of the Traffic Flow Theory Conference, Annecy, France*.

Smaragdis, E., Papageorgiou, M., & Kosmatopoulos, E. (2004, March). A flow-maximizing adaptive local ramp metering strategy. *Transportation Research Part B: Methodological*, *38*(3), 251–270. doi: 10.1016/S0191-2615(03)00012-2

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 Hinsbergen, C. P. I. J., Schreiter, T., Zuurbier, F. S., Van Lint, J. W. C., & Van Zuylen, H. J. (2012, March). Localized Extended Kalman Filter for Scalable Real-Time Traffic State Estimation. *IEEE Transactions on Intelligent Transportation Systems*, *13*(1), 385–394. doi: 10.1109/TITS.2011.2175728

Van Lint, H., Landman, R., Yuan, Y., Van Hinsbergen, C., & Hoogendoorn, S. (2014). Traffic monitoring for coordinated traffic management experiences from the field trial integrated traffic management in Amsterdam. (May).

Van Lint, J., & Hoogendoorn, S. P. (2010, November). A Robust and Efficient Method for Fusing Heterogeneous Data from Traffic Sensors on Freeways. *Computer-Aided Civil and Infrastructure Engineering*, *25*(8), 596–612. doi: 10.1111/j.1467-8667.2009.00617.x

Van Lint, J. W. C. (2004). *Reliable Travel Time Prediction for Freeways.*

Wang, Y., & Papageorgiou, M. (2005, February). Real-time freeway traffic state estimation based on extended Kalman filter: a general approach. *Transportation Research Part B: Methodological*, *39*(2), 141–167. doi: 10.1016/j.trb.2004.03.003

Wardrop, J. G. (1952). Some Theoretical Aspects of Road Traffic Research. *Road engineering division meeting*, 325–362.

Yuan, Y. (2013). *Lagrangian multi-class traffic state estimation.* Retrieved from _2013.pdf

Yuan, Y., & Hoogendoorn, S. (2014). Functionele specificatie Fileschatter 2.0. , *15*.

Yuan, Y., & Van Lint, J. W. C. (2015). *Innovations in Dynamic Traffic Management (CIE5804): Traffic monitoring / state estimation.* Delft University of Technology.

Appendix A

Derivations for the FCD-based measurement Probability Density Functions

In this research the FCD-based u^S measurement error variance is dependent on the measurements and estimated traffic conditions. The factors of uncertainty are taken into account, which should describe the u^S estimation errors. Here errors due to individual speed measurement error and errors due to a combination of driver heterogeneity and limited penetration rate.

A.1 Derivation of the uncertainty due to individual measurement errors

The estimated mean speed $\hat{\bar{v}}$ for N individual vehicle speeds is calculated with the following equation.

$$\hat{\bar{v}} = \frac{1}{N}\sum_{n=1}^{N} v_n + \varepsilon_n \tag{A.1}$$

$$= \bar{v} + \frac{\sum_{n=1}^{N} \varepsilon_n}{N} \tag{A.2}$$

Here v_n denotes the true individual vehicle speed and $\bar{v}$ the true mean speed. The individual measurement errors are denoted by ε_n. The individual measurement errors are assumed independently distributed, thus $E[\varepsilon_n \varepsilon_m] = 0$ for $n \neq m$ and $E[\varepsilon_n^2] = \sigma^2$

$$var(\hat{\bar{v}}) = E\left[(\hat{\bar{v}} - E[\hat{\bar{v}}])^2\right] \tag{A.3}$$

$$= E\left[(\bar{v} - \bar{v} + \frac{\sum_{n=1}^{N} \varepsilon_n}{N})^2\right] \tag{A.4}$$

$$= E\left[(\frac{\sum_{n=1}^{N} \varepsilon_n}{N})^2\right] \tag{A.5}$$

$$= E\left[\frac{1}{N^2}\sum_{n=1}^{N} \varepsilon_n^2\right] \tag{A.6}$$

$$= \frac{1}{N^2} N\sigma^2 = \frac{1}{N}\sigma^2 \tag{A.7}$$

This equation shows that the $\bar{v}$ estimation error due to individual speed measurement error decreases when the number of observed vehicles increases.

A.2 Derivation of the uncertainty due to driver heterogeneity and penetration rate

For a given road stretch and time the individual vehicle speed are assumed to be Gaussian distribution. This difference is driver behavior is denoted by the driver heterogeneity. As discussed in Section 3.4.1 the driver heterogeneity is dependent on the traffic conditions. A relation between the mean speed and instantaneous speed variance is provided in this section. Given this variance, the uncertainty based on the penetration rate may be derived.

The individual vehicle speeds may be described by the following equations.

$$v_n = u^S + \varepsilon_n \tag{A.8}$$

The speed difference of vehicle n with respect to u^S is denoted by ε_n. Furthermore, the mean-speed u^S estimated based on j vehicles is equal to

$$\hat{u}^S = \frac{\sum_{n=1}^{j} v_n}{j} \tag{A.9}$$

$$= \frac{\sum_{n=1}^{j} u^S + \varepsilon_n}{j} \tag{A.10}$$

$$= u^S + \frac{\sum_{n=1}^{j} \varepsilon_n}{j} \tag{A.11}$$

As stated above the speed differences ε_n are Gaussian distributed with zero-mean, where $E[\varepsilon_i^2] = \sigma^2$. The variance σ^2 denotes the driver heterogeneity given the traffic conditions. In contrast to the individual speed measurements, the speed differences ε_n are not independently distributed. It is assumed that the expected correlation between two random speed differences is a fixed value, namely $E[\varepsilon_n \varepsilon_m] = \sigma_{12}^2$ where n and m are two random vehicles.

If j vehicles are observed the error variance may be calculated with

$$var(\hat{u}^S) = E\left[\left(\hat{u}^S - E[\hat{u}^S]\right)^2\right] \tag{A.12}$$

$$= E\left[\left(u^S - u^S + \frac{\sum_{n=1}^{j} \varepsilon_n}{j}\right)^2\right] \tag{A.13}$$

$$= E\left[\left(\frac{\sum_{n=1}^{j} \varepsilon_n}{j}\right)^2\right] \tag{A.14}$$

$$= \frac{1}{j^2}\left(j\sigma^2 + (j^2 - j)\sigma_{12}^2\right) \tag{A.15}$$

σ_{12}^2 is assumed to be a fixed value and may be derived. If all N vehicles are observed the estimation error and thus $var(\hat{u}^S)$ is equal to zero. This allow us to solve for σ_{12}^2.

$$\frac{1}{N^2}\left(N\sigma^2 + (N^2 - N)\sigma_{12}^2\right) = 0 \tag{A.16}$$

$$\frac{1}{N}\sigma^2 + \frac{N-1}{N}\sigma_{12}^2 = 0 \tag{A.17}$$

$$(N-1)\,\sigma_{12}^2 = -\sigma^2 \tag{A.18}$$

$$\sigma_{12}^2 = -\frac{1}{N-1}\sigma^2 \tag{A.19}$$

This may be substituted in equation above yielding

$$var(u^S - \hat{u}^S) = \frac{1}{j}\sigma^2 - \frac{j-1}{j}\frac{1}{N-1}\sigma^2 \tag{A.20}$$

$$= \frac{N-1}{j(N-1)}\sigma^2 - \frac{j-1}{j(N-1)}\sigma^2 \tag{A.21}$$

$$= \frac{N-j}{j(N-1)}\sigma^2 \tag{A.22}$$

Intuitively this equation may be explained. As expected if there is one observation, thus $j = 1$, the error variance due to a limited penetration rate is equal to σ^2. Furthermore, if all vehicles are observed $j = N$, the error variance due to a limited penetration rate is equal to zero.

Appendix B

Derivations for the EKF

To obtain the important equation for the EKF, derivations are needed. As stated in Chapter 3, the recursive equation which is used differs from the one proposed by Wang & Papageorgiou (2005). Therefore new updating equations have to be derived, which are provided in Section B.1. These equations use derivative matrices, for which the individual elements are derived in Section B.2.

B.1 Derivation of the Kalman gain and error covariance matrix

Wang & Papageorgiou (2005) method will be adopted. However, we will include an extra step in the process. Wang & Papageorgiou (2005) uses:

$$\hat{\mathbf{x}}(k+1|k) = \mathbf{f}\left[\hat{\mathbf{x}}(k|k-1), \mathbf{0}\right] + \mathbf{K}(k)\left[\mathbf{y}(k) - \mathbf{g}\left(\hat{\mathbf{x}}(k|k-1), \mathbf{0}\right)\right] \tag{B.1}$$

In this equation the (final) state vector estimation $\hat{\mathbf{x}}$ for period $k+1$ is determined with the measurement information $\mathbf{y}$ on period k. Other applications, such as Van Hinsbergen et al. (2012), correct the estimate for k with the measurement information on k. For the author the latter approach seems more logical. In this set-up the updating follows two steps. Firstly, a model-based prediction:

$$\hat{\mathbf{x}}(k+1|k) = \mathbf{f}\left[\mathbf{x}(k|k), \mathbf{0}\right]. \tag{B.2}$$

Secondly, the correction step.

$$\mathbf{K}(k+1)\left[\mathbf{y}(k+1) - \mathbf{g}\left(\hat{\mathbf{x}}(k+1|k), \mathbf{0}\right)\right] \tag{B.3}$$

Combined this leads to the following recursive updating equation.

$$\hat{\mathbf{x}}(k+1|k+1) = \hat{\mathbf{x}}(k+1|k) + \mathbf{K}(k+1)\left[\mathbf{y}(k+1) - \mathbf{g}\left(\hat{\mathbf{x}}(k+1|k), \mathbf{0}\right)\right] \tag{B.4}$$

The extended Kalman filter (EKF) should minimize the covariance of the estimation error, where the estimation error is defined as $\mathbf{x}(k+1) - \hat{\mathbf{x}}(k+1|k+1)$.

$$\mathbf{P}(k+1|k+1) = E\left\{\left[\mathbf{x}(k+1) - \hat{\mathbf{x}}(k+1|k+1)\right] \cdot \left[\mathbf{x}(k+1) - \hat{\mathbf{x}}(k+1|k+1)\right]^T\right\} \tag{B.5}$$

To derive the updating equations, (B.5) may be expanded.

$$\begin{aligned} &E\left\{\left[\mathbf{x}(k+1) - \hat{\mathbf{x}}(k+1|k+1)\right] \cdot \left[\mathbf{x}(k+1) - \hat{\mathbf{x}}(k+1|k+1)\right]^T\right\} \\ &= E\{[\mathbf{x}(k+1) - \hat{\mathbf{x}}(k+1|k) - \mathbf{K}(k+1)\left[\mathbf{y}(k+1) - \hat{\mathbf{y}}(k+1|k+1)\right]] \cdot \\ &\quad [\mathbf{x}(k+1) - \hat{\mathbf{x}}(k+1|k) - \mathbf{K}(k+1)\left[\mathbf{y}(k+1) - \hat{\mathbf{y}}(k+1|k+1)\right]]^T\} \end{aligned}$$

We may consider the inner part separately for a second. The state and measurement vector (estimates) may be described by

$$\hat{\mathbf{x}}(k+1|k) = \mathbf{f}\left[\hat{\mathbf{x}}(k|k), \mathbf{0}\right] = \mathbf{A}(k)\hat{\mathbf{x}}(k|k) \tag{B.6}$$

$$\mathbf{x}(k+1) = \mathbf{f}\left[\mathbf{x}(k), \xi(k)\right] = \mathbf{A}(k)\mathbf{x}(k) + \mathbf{\Gamma}(k)\xi(k) \tag{B.7}$$

$$\hat{\mathbf{y}}(k+1|k+1) = \mathbf{g}\left[\hat{\mathbf{x}}(k+1|k), \mathbf{0}\right] = \mathbf{C}(k+1)\hat{\mathbf{x}}(k+1|k) \tag{B.8}$$

$$\mathbf{y}(k+1) = \mathbf{g}\left[\mathbf{x}(k+1), \eta(k+1)\right] = \mathbf{C}(k+1)\mathbf{x}(k+1) + \mathbf{\Sigma}(k+1)\eta(k+1) \tag{B.9}$$

with

$$\mathbf{A}(k) = \frac{\partial \mathbf{f}}{\partial \mathbf{x}}\left(\hat{\mathbf{x}}(k|k), \mathbf{0}\right) \tag{B.10}$$

$$\mathbf{\Gamma}(k) = \frac{\partial \mathbf{f}}{\partial \xi}\left(\hat{\mathbf{x}}(k|k), \mathbf{0}\right) \tag{B.11}$$

$$\mathbf{C}(k+1) = \frac{\partial \mathbf{g}}{\partial \mathbf{x}}\left(\hat{\mathbf{x}}(k+1|k), \mathbf{0}\right) \tag{B.12}$$

$$\mathbf{\Sigma}(k+1) = \frac{\partial \mathbf{g}}{\partial \eta}\left(\hat{\mathbf{x}}(k+1|k), \mathbf{0}\right) \tag{B.13}$$

Firstly $\left[\mathbf{x}(k+1) - \hat{\mathbf{x}}(k+1|k+1)\right]$ may be expanded

$$
\begin{aligned}
&[\mathbf{x}(k+1) - \hat{\mathbf{x}}(k+1|k+1)] = \\
&[\mathbf{x}(k+1) - \hat{\mathbf{x}}(k+1|k) - \mathbf{K}(k+1)\left[\mathbf{y}(k+1) - \mathbf{g}\left(\hat{\mathbf{x}}(k+1|k), \mathbf{0}\right)\right]] = \\
&\mathbf{f}\left[\mathbf{x}(k), \xi(k)\right] - \mathbf{f}\left[\hat{\mathbf{x}}(k|k), \mathbf{0}\right] - \mathbf{K}(k+1)\left[\mathbf{g}\left[\mathbf{x}(k+1), \eta(k+1)\right] - \mathbf{g}\left(\mathbf{f}\left[\hat{\mathbf{x}}(k|k), \mathbf{0}\right], \mathbf{0}\right)\right] = \\
&\mathbf{A}(k)\mathbf{x}(k) + \mathbf{\Gamma}(k)\xi(k) - \mathbf{A}(k)\hat{\mathbf{x}}(k|k) - \\
&\mathbf{K}(k+1)\left[\mathbf{C}(k+1)\left[\mathbf{A}(k)\mathbf{x}(k) + \mathbf{\Gamma}(k)\xi(k)\right] + \mathbf{\Sigma}(k+1)\eta(k+1) - \mathbf{C}(k+1)\mathbf{A}(k)\hat{\mathbf{x}}(k|k)\right] = \\
&\left(\mathbf{I} - \mathbf{K}(k+1)\mathbf{C}(k+1)\right)\mathbf{A}(k)\left[\mathbf{x}(k) - \hat{\mathbf{x}}(k|k)\right] + \\
&\left(\mathbf{I} - \mathbf{K}(k+1)\mathbf{C}(k+1)\right)\mathbf{\Gamma}(k)\xi(k) - \mathbf{K}(k+1)\mathbf{\Sigma}(k+1)\eta(k+1)
\end{aligned}
$$

Determine $\mathbf{P}(k+1|k+1)$:

$$
\begin{aligned}
&E\left\{\left[\mathbf{x}(k+1) - \hat{\mathbf{x}}(k+1|k+1)\right] \cdot \left[\mathbf{x}(k+1) - \hat{\mathbf{x}}(k+1|k+1)\right]^T\right\} = \\
&E\{\left(\mathbf{I} - \mathbf{K}(k+1)\mathbf{C}(k+1)\right)\mathbf{A}(k)\left[\mathbf{x}(k) - \hat{\mathbf{x}}(k|k)\right] \cdot \\
&\left[\mathbf{x}(k) - \hat{\mathbf{x}}(k|k)\right]^T \mathbf{A}^T(k)\left(\mathbf{I} - \mathbf{K}(k+1)\mathbf{C}(k+1)\right)^T \\
&+ \left(\mathbf{I} - \mathbf{K}(k+1)\mathbf{C}(k+1)\right)\mathbf{A}(k)\left[\mathbf{x}(k) - \hat{\mathbf{x}}(k|k)\right]\xi^T(k)\mathbf{\Gamma}^T(k)\left(\mathbf{I} - \mathbf{K}(k+1)\mathbf{C}(k+1)\right)^T \\
&- \left(\mathbf{I} - \mathbf{K}(k+1)\mathbf{C}(k+1)\right)\mathbf{A}(k)\left[\mathbf{x}(k) - \hat{\mathbf{x}}(k|k)\right]\eta^T(k+1)\mathbf{\Sigma}^T(k+1)\mathbf{K}^T(k+1) \\
&+ \left(\mathbf{I} - \mathbf{K}(k+1)\mathbf{C}(k+1)\right)\mathbf{\Gamma}(k)\xi(k)\left[\mathbf{x}(k) - \hat{\mathbf{x}}(k|k)\right]^T \mathbf{A}^T(k)\left(\mathbf{I} - \mathbf{K}(k+1)\mathbf{C}(k+1)\right)^T \\
&+ \left(\mathbf{I} - \mathbf{K}(k+1)\mathbf{C}(k+1)\right)\mathbf{\Gamma}(k)\xi(k)\xi^T(k)\mathbf{\Gamma}^T(k)\left(\mathbf{I} - \mathbf{K}(k+1)\mathbf{C}(k+1)\right)^T \\
&- \left(\mathbf{I} - \mathbf{K}(k+1)\mathbf{C}(k+1)\right)\mathbf{\Gamma}(k)\xi(k)\eta^T(k+1)\mathbf{\Sigma}^T(k+1)\mathbf{K}^T(k+1) \\
&-\mathbf{K}(k+1)\mathbf{\Sigma}(k+1)\eta(k+1)\left[\mathbf{x}(k) - \hat{\mathbf{x}}(k|k)\right]^T \mathbf{A}^T(k)\left(\mathbf{I} - \mathbf{K}(k+1)\mathbf{C}(k+1)\right)^T \\
&-\mathbf{K}(k+1)\mathbf{\Sigma}(k+1)\eta(k+1)\xi^T(k)\mathbf{\Gamma}^T(k)\left(\mathbf{I} - \mathbf{K}(k+1)\mathbf{C}(k+1)\right)^T \\
&+\mathbf{K}(k+1)\mathbf{\Sigma}(k+1)\eta(k+1)\eta^T(k+1)\mathbf{\Sigma}^T(k+1)\mathbf{K}^T(k+1)\}
\end{aligned}
$$

This may be simplified using the following relations. The measurement error covariance matrix is given by $\mathbf{R}(k)$. Furthermore, the model error covariance matrix is given by $\mathbf{Q}(k)$.

$$\mathbf{P}(k|k) = E\left\{\left[\mathbf{x}(k) - \hat{\mathbf{x}}(k|k)\right] \cdot \left[\mathbf{x}(k) - \hat{\mathbf{x}}(k|k)\right]^T\right\} \tag{B.14}$$

$$\mathbf{Q}(k) = E\left\{\xi(k)\xi^T(k)\right\} \tag{B.15}$$

$$\mathbf{R}(k) = E\left\{\eta(k)\eta^T(k)\right\} \tag{B.16}$$

Other cross-products involving a single term of the model ($\xi(k)$) or measurement ($\eta(k)$) error are expected to be zero ($\mathbf{0}$). The remaining equation for $\mathbf{P}(k+1|k+1)$ is:

$$\begin{aligned}\mathbf{P}(k+1|k+1) = \\ (\mathbf{I}-\mathbf{K}(k+1)\mathbf{C}(k+1))\,\mathbf{A}(k)\mathbf{P}(k|k)\mathbf{A}^T(k)\,(\mathbf{I}-\mathbf{K}(k+1)\mathbf{C}(k+1))^T \\ +(\mathbf{I}-\mathbf{K}(k+1)\mathbf{C}(k+1))\,\mathbf{\Gamma}(k)\mathbf{Q}(k)\mathbf{\Gamma}^T(k)\,(\mathbf{I}-\mathbf{K}(k+1)\mathbf{C}(k+1))^T \\ +\mathbf{K}(k+1)\mathbf{\Sigma}(k+1)\mathbf{R}(k+1)\mathbf{\Sigma}^T(k+1)\mathbf{K}^T(k+1)\end{aligned}$$

We want to find the Kalman gain $\mathbf{K}(k+1)$ which minimizes the error covariance matrix $\mathbf{P}(k+1|k+1)$. This may be done by setting the derivative to $\mathbf{K}(k+1)$ of the trace of $\mathbf{P}(k+1|k+1)$ to zero and solving it for $\mathbf{K}(k+1)$. Firstly we expand $\mathbf{P}(k+1|k+1)$ again:

$$\begin{aligned}\mathbf{P}(k+1|k+1) = \\ \mathbf{A}(k)\mathbf{P}(k|k)\mathbf{A}^T(k)+\mathbf{K}(k+1)\mathbf{C}(k+1)\mathbf{A}(k)\mathbf{P}(k|k)\mathbf{A}^T(k)\mathbf{C}^T(k+1)\mathbf{K}^T(k+1) \\ -\mathbf{A}(k)\mathbf{P}(k|k)\mathbf{A}^T(k)\mathbf{C}^T(k+1)\mathbf{K}^T(k+1)-\mathbf{K}(k+1)\mathbf{C}(k+1)\mathbf{A}(k)\mathbf{P}(k|k)\mathbf{A}^T(k) \\ +\mathbf{\Gamma}(k)\mathbf{Q}(k)\mathbf{\Gamma}^T(k)+\mathbf{K}(k+1)\mathbf{C}(k+1)\mathbf{\Gamma}(k)\mathbf{Q}(k)\mathbf{\Gamma}^T(k)\mathbf{C}^T(k+1)\mathbf{K}^T(k+1) \\ -\mathbf{\Gamma}(k)\mathbf{Q}(k)\mathbf{\Gamma}^T(k)\mathbf{C}^T(k+1)\mathbf{K}^T(k+1)-\mathbf{K}(k+1)\mathbf{C}(k+1)\mathbf{\Gamma}(k)\mathbf{Q}(k)\mathbf{\Gamma}^T(k) \\ +\mathbf{K}(k+1)\mathbf{\Sigma}(k+1)\mathbf{R}(k+1)\mathbf{\Sigma}^T(k+1)\mathbf{K}^T(k+1)\end{aligned}$$

The derivative:

$$\begin{aligned}\frac{\partial tr\left(\mathbf{P}(k+1|k+1)\right)}{\partial \mathbf{K}(k+1)} = \\ -2\mathbf{A}(k)\mathbf{P}(k|k)\mathbf{A}^T(k)\mathbf{C}^T(k+1)-2\mathbf{\Gamma}(k)\mathbf{Q}(k)\mathbf{\Gamma}^T(k)\mathbf{C}^T(k+1) \\ +2\mathbf{K}(k+1)\mathbf{C}(k+1)\mathbf{A}(k)\mathbf{P}(k|k)\mathbf{A}^T(k)\mathbf{C}^T(k+1) \\ +2\mathbf{K}(k+1)\mathbf{C}(k+1)\mathbf{\Gamma}(k)\mathbf{Q}(k)\mathbf{\Gamma}^T(k)\mathbf{C}^T(k+1) \\ +2\mathbf{K}(k+1)\mathbf{\Sigma}(k+1)\mathbf{R}(k+1)\mathbf{\Sigma}^T(k+1)=\mathbf{0}\end{aligned}$$

Solving this for $\mathbf{K}(k+1)$:

$$\begin{aligned}\mathbf{K}(k+1) = \\ \frac{\left[\mathbf{A}(k)\mathbf{P}(k|k)\mathbf{A}^T(k)+\mathbf{\Gamma}(k)\mathbf{Q}(k)\mathbf{\Gamma}^T(k)\right]\mathbf{C}^T(k+1)}{\mathbf{C}(k+1)\left[\mathbf{A}(k)\mathbf{P}(k|k)\mathbf{A}^T(k)+\mathbf{\Gamma}(k)\mathbf{Q}(k)\mathbf{\Gamma}^T(k)\right]\mathbf{C}^T(k+1)+\mathbf{\Sigma}(k+1)\mathbf{R}(k+1)\mathbf{\Sigma}^T(k+1)}\end{aligned}$$

This equation can be used to simplify the earlier obtain equation for $\mathbf{P}(k+1|k+1)$. This yields:

$$\mathbf{P}(k+1|k+1) = \\ [\mathbf{I} - \mathbf{K}(k+1)\mathbf{C}(k+1)] \left[\mathbf{A}(k)\mathbf{P}(k|k)\mathbf{A}^T(k) + \mathbf{\Gamma}(k)\mathbf{Q}(k)\mathbf{\Gamma}^T(k)\right]$$

This equation and the Kalman Gain may be separated by denoting the prior error covariance matrix $\mathbf{P}(k|k-1)$ separately. This is given by

$$\mathbf{P}(k+1|k) = \mathbf{A}(k)\mathbf{P}(k|k)\mathbf{A}^T(k) + \mathbf{\Gamma}(k)\mathbf{Q}(k)\mathbf{\Gamma}^T(k) \tag{B.17}$$

The Kalman Gain $\mathbf{K}(k+1)$ then simplifies to

$$\mathbf{K}(k+1) = \frac{\mathbf{P}(k+1|k)\mathbf{C}^T(k+1)}{\mathbf{C}(k+1)\mathbf{P}(k+1|k)\mathbf{C}^T(k+1) + \mathbf{\Sigma}(k+1)\mathbf{R}(k+1)\mathbf{\Sigma}^T(k+1)} \tag{B.18}$$

The posterior error covariance matrix of $\hat{\mathbf{x}}(k|k)$ then simplifies to

$$\mathbf{P}(k+1|k+1) = [\mathbf{I} - \mathbf{K}(k+1)\mathbf{C}(k+1)]\,\mathbf{P}(k+1|k) \tag{B.19}$$

The equations are explained intuitively in Section 3.5.

B.2 Derivation of $\mathbf{A}(k)$, $\mathbf{\Gamma}(k)$, $\mathbf{C}(k+1)$ and $\mathbf{\Sigma}(k+1)$

In the previous section four derivative matrices were introduced. Before implementing these in an algorithm the individual elements have to be derived. This section shows these derivations. The following derivative matrices are used.

$$\mathbf{A}(k) = \frac{\partial \mathbf{f}}{\partial \mathbf{x}}\left(\hat{\mathbf{x}}(k|k), \mathbf{0}\right) \tag{B.20}$$

$$\mathbf{\Gamma}(k) = \frac{\partial \mathbf{f}}{\partial \xi}\left(\hat{\mathbf{x}}(k|k), \mathbf{0}\right) \tag{B.21}$$

$$\mathbf{C}(k+1) = \frac{\partial \mathbf{g}}{\partial \mathbf{x}}\left(\hat{\mathbf{x}}(k+1|k), \mathbf{0}\right) \tag{B.22}$$

$$\mathbf{\Sigma}(k+1) = \frac{\partial \mathbf{g}}{\partial \eta}\left(\hat{\mathbf{x}}(k+1|k), \mathbf{0}\right) \tag{B.23}$$

B.2.1 Derivative of traffic model to the state vector $\mathbf{A}(k)$

$$\mathbf{A}(k) = \frac{\partial \mathbf{f}}{\partial \mathbf{x}}\left(\hat{\mathbf{x}}(k|k), \mathbf{0}\right) \tag{B.24}$$

$$\mathbf{A}(k) = \begin{bmatrix} \frac{\partial f_1}{\partial x_1}\left(\hat{\mathbf{x}}(k|k), \mathbf{0}\right) & \frac{\partial f_1}{\partial x_2}\left(\hat{\mathbf{x}}(k|k), \mathbf{0}\right) & \cdots & \frac{\partial f_1}{\partial x_n}\left(\hat{\mathbf{x}}(k|k), \mathbf{0}\right) \\ \vdots & \vdots & \vdots & \vdots \\ \frac{\partial f_n}{\partial x_1}\left(\hat{\mathbf{x}}(k|k), \mathbf{0}\right) & \cdots & \cdots & \frac{\partial f_n}{\partial x_n}\left(\hat{\mathbf{x}}(k|k), \mathbf{0}\right) \end{bmatrix}_{n \times n} . \tag{B.25}$$

The elements of this matrix can be determined separately. Firstly we may denote the derivative of density in cell i, $\rho_i(k)$. This is given by

$$\rho_i(k+1) = \rho_i(k) + \frac{T}{\Delta_i \lambda_i}\left[q_{i-1}(k) - q_i(k) + r_i(k) - s_i(k)\right] \tag{B.26}$$

$$= \rho_i(k) + \frac{T}{\Delta_i \lambda_i}\left[q_{i-1}(k) - q_i(k) + r_i(k) - \beta_i(k) q_{i-1}(k)\right] \tag{B.27}$$

$$= \rho_i(k) + \frac{T}{\Delta_i \lambda_i}[(1 - \beta_i(k))\rho_{i-1}(k) v_{i-1}(k) \lambda_{i-1}$$
$$- \rho_i(k) v_i(k) \lambda_i + r_i(k)] \tag{B.28}$$

For $i = 2, \ldots, N$, the non-zero elements are the derivatives to ρ_{i-1}, v_{i-1}, ρ_i, v_i, β_i, r_i. The position within the state vector of $\rho_i(k)$ is $2 \times i - 1$. (below is based on no ramps)

$$\frac{\partial \rho_i(k+1)}{\partial \rho_{i-1}} = \frac{T\lambda_{i-1}}{\Delta_i \lambda_i} v_{i-1}(k) \tag{B.29}$$

$$\frac{\partial \rho_i(k+1)}{\partial v_{i-1}} = \frac{T\lambda_{i-1}}{\Delta_i \lambda_i} \rho_{i-1}(k) \tag{B.30}$$

$$\frac{\partial \rho_i(k+1)}{\partial \rho_i} = 1 - \frac{T}{\Delta_i} v_i(k) \tag{B.31}$$

$$\frac{\partial \rho_i(k+1)}{\partial v_i} = -\frac{T}{\Delta_i} \rho_i(k) \tag{B.32}$$

In the case of an on-ramp:

$$\frac{\partial \rho_i(k+1)}{\partial r_i} = \frac{T}{\Delta_i \lambda_i} \tag{B.33}$$

In the case of an off-ramp:

$$\frac{\partial \rho_i(k+1)}{\partial \rho_{i-1}} = \frac{T\lambda_{i-1}}{\Delta_i \lambda_i} (1 - \beta_i(k)) v_{i-1}(k) \tag{B.34}$$

$$\frac{\partial \rho_i(k+1)}{\partial v_{i-1}} = \frac{T\lambda_{i-1}}{\Delta_i \lambda_i} (1 - \beta_i(k)) \rho_{i-1}(k) \tag{B.35}$$

$$\frac{\partial \rho_i(k+1)}{\partial \beta_i} = \frac{T\lambda_{i-1}}{\Delta_i \lambda_i} \rho_{i-1}(k) v_{i-1}(k) \tag{B.36}$$

The derivative to all other state vector elements is equal to zero. A special case is the derivative of $\rho_1(k+1)$:

$$\frac{\partial \rho_1(k+1)}{\partial \rho_1} = 1 - \frac{T}{\Delta_1} v_1(k) \tag{B.37}$$

$$\frac{\partial \rho_1(k+1)}{\partial v_1} = -\frac{T}{\Delta_1} \rho_1(k) \tag{B.38}$$

$$\frac{\partial \rho_1(k+1)}{\partial q_0} = \frac{T}{\Delta_1 \lambda_1} \tag{B.39}$$

The speed state vector elements v_i

$$\begin{aligned} v_i(k+1) = v_i(k) &+ \frac{T}{\tau} \left[V(\rho_i(k)) - v_i(k) \right] + \frac{T}{\Delta_i} v_i(k) \left[v_{i-1}(k) - v_i(k) \right] \\ &- \frac{\nu T}{\tau \Delta_i} \frac{\left[\rho_{i+1}(k) - \rho_i(k) \right]}{\rho_i(k) + \kappa} - \frac{\delta T}{\Delta_i \lambda_i} \frac{r_i(k) v_i(k)}{\rho_i(k) + \kappa} + \xi_i^v(k) \end{aligned} \tag{B.40}$$

with

$$V(\rho) = v_f exp\left[-\frac{1}{a}\left(\frac{\rho}{\rho_{cr}}\right)^a\right] \tag{B.41}$$

The v-elements in $\mathbf{A}(k)$ lead to... Note that we still assume that there are no on- and off-ramps. If these do exist some extra parts should be added. These hold for $i = 2, \ldots, N-1$.

$$\frac{\partial v_i(k+1)}{\partial v_{i-1}} = \frac{T}{\Delta_i} v_i(k) \tag{B.42}$$

$$\frac{\partial v_i(k+1)}{\partial \rho_i} = \frac{T v_f}{\tau} \frac{exp\left[-\frac{1}{a}\left(\frac{\rho_i(k)}{\rho_{cr}}\right)^a\right]\left(\frac{\rho_i(k)}{\rho_{cr}}\right)^a}{\rho_i(k)} + \frac{\nu T}{\tau \Delta_i} \frac{[\rho_{i+1}(k) + \rho_i(k)]}{(\rho_i(k)+\kappa)^2} \tag{B.43}$$

$$\frac{\partial v_i(k+1)}{\partial v_i} = 1 - \frac{T}{\tau} - 2\frac{T}{\Delta_i} v_i(k) + \frac{T}{\Delta_i} v_{i-1}(k) \tag{B.44}$$

$$\frac{\partial v_i(k+1)}{\partial \rho_{i+1}} = -\frac{\nu T}{\tau \Delta_i} \frac{1}{\rho_i(k)+\kappa} \tag{B.45}$$

$$\frac{\partial v_i(k+1)}{\partial v_f} = \frac{T}{\tau} exp\left[-\frac{1}{a}\left(\frac{\rho_i(k)}{\rho_{cr}}\right)^a\right] \tag{B.46}$$

$$\frac{\partial v_i(k+1)}{\partial \rho_{cr}} = \frac{T v_f}{\tau} \frac{exp\left[-\frac{1}{a}\left(\frac{\rho_i(k)}{\rho_{cr}}\right)^a\right]\left(\frac{\rho_i(k)}{\rho_{cr}}\right)^a}{\rho_{cr}} \tag{B.47}$$

$$\frac{\partial v_i(k+1)}{\partial a} = \frac{T v_f}{\tau} \frac{\left(\frac{\rho_i(k)}{\rho_{cr}}\right)^a exp\left[-\frac{1}{a}\left(\frac{\rho_i(k)}{\rho_{cr}}\right)^a\right]\left[1 - alog\left(\frac{\rho_i(k)}{\rho_{cr}}\right)\right]}{a^2} \tag{B.48}$$

In the case of an on-ramp:

$$\frac{\partial v_i(k+1)}{\partial \rho_i} = \ldots + \frac{\delta T}{\Delta_i \lambda_i} \frac{r_i(k) v_i(k)}{(\rho_i(k)+\kappa)^2} \tag{B.49}$$

$$\frac{\partial v_i(k+1)}{\partial v_i} = \ldots - \frac{\delta T}{\Delta_i \lambda_i} \frac{r_i(k)}{\rho_i(k)+\kappa} \tag{B.50}$$

$$\frac{\partial v_i(k+1)}{\partial r_i} = -\frac{\delta T}{\Delta_i \lambda_i} \frac{v_i(k)}{\rho_i(k)+\kappa} \tag{B.51}$$

In this case, there are two special cases, namely $i = 1$ and $i = N$. For $i = 1$ and $i = N$, respectively the upstream speed $v_0(k)$ and the downstream density $\rho_{N+1}(k)$ have to be taken into account

$$\frac{\partial v_1(k+1)}{\partial v_0} = \frac{T}{\Delta_1} v_1(k) \tag{B.52}$$

$$\frac{\partial v_1(k+1)}{\partial v_1} = 1 - \frac{T}{\tau} - 2\frac{T}{\Delta_1} v_1(k) + \frac{T}{\Delta_i} v_0(k) \tag{B.53}$$

$$\frac{\partial v_N(k+1)}{\partial \rho_N} = \frac{T v_f}{\tau} \frac{exp\left[-\frac{1}{a}\left(\frac{\rho_N(k)}{\rho_{cr}}\right)^a\right]\left(\frac{\rho_N(k)}{\rho_{cr}}\right)^a}{\rho_N(k)} - \frac{\nu T}{\tau \Delta_N} \frac{[\rho_{N+1}(k) + \rho_N(k)]}{(\rho_N(k) + \kappa)^2} \tag{B.54}$$

$$\frac{\partial v_N(k+1)}{\partial \rho_{N+1}} = -\frac{\nu T}{\tau \Delta_N} \frac{1}{\rho_N(k) + \kappa} \tag{B.55}$$

The other (non-ramp) elements q_0, v_0, ρ_{N+1}, v_f, ρ_{cr} and a are all random walks. Therefore, their derivative (to their own element) is equal to 1 and the rest is 0.

B.2.2 Derivative of traffic model to the model error vector $\mathbf{\Gamma}(k)$

$$\mathbf{\Gamma}(k) = \frac{\partial \mathbf{f}}{\partial \xi}\left(\hat{\mathbf{x}}(k|k), \mathbf{0}\right) \tag{B.56}$$

The noise elements of the density and

$$\rho_i(k+1) = \frac{T}{\Delta_i \lambda_i}\left[(1 - \beta_i(k))\xi^q_{i-1} - \xi^q_i\right] \tag{B.57}$$

$$v_i(k+1) = \xi^v_i \tag{B.58}$$

Again we may make a distinction between the speed, density and random walk elements. For the speed elements

$$\frac{\partial v_i(k+1)}{\partial \xi^v_i} = 1 \tag{B.59}$$

$$\tag{B.60}$$

Similarly the derivatives for the random walk elements are equal to 1. Only the density elements differ (again for now no ramps are assumed)

$$\frac{\partial \rho_i(k+1)}{\partial \xi^q_{i-1}} = \frac{T}{\Delta_i \lambda_i} \tag{B.61}$$

$$\frac{\partial \rho_i(k+1)}{\partial \xi^q_i} = -\frac{T}{\Delta_i \lambda_i} \tag{B.62}$$

In the case of an off-ramp:

$$\frac{\partial \rho_i(k+1)}{\partial \xi_{i-1}^q} = \frac{T}{\Delta_i \lambda_i}(1-\beta_i(k)) \tag{B.63}$$

Special case is the density in the first cell

$$\frac{\partial \rho_1(k+1)}{\partial \xi_1^q} = -\frac{T}{\Delta_i \lambda_i} \tag{B.64}$$

$$\frac{\partial \rho_1(k+1)}{\partial \xi_i^{q_0}} = \frac{T}{\Delta_i \lambda_i} \tag{B.65}$$

B.2.3 Derivative of measurement equations to the state vector $\mathbf{C}(k+1)$

$$\mathbf{C}(k+1) = \frac{\partial \mathbf{g}}{\partial \mathbf{x}}\left(\hat{\mathbf{x}}(k+1|k), \mathbf{0}\right) \tag{B.66}$$

which is

$$\mathbf{C}(k+1) = \begin{bmatrix} \frac{\partial g_1}{\partial x_1}\left(\hat{\mathbf{x}}(k+1|k), \mathbf{0}\right) & \frac{\partial g_1}{\partial x_2}\left(\hat{\mathbf{x}}(k+1|k), \mathbf{0}\right) & \cdots & \frac{\partial g_1}{\partial x_n}\left(\hat{\mathbf{x}}(k+1|k), \mathbf{0}\right) \\ \vdots & \vdots & \vdots & \vdots \\ \frac{\partial g_m}{\partial x_1}\left(\hat{\mathbf{x}}(k+1|k), \mathbf{0}\right) & \cdots & \cdots & \frac{\partial g_m}{\partial x_n}\left(\hat{\mathbf{x}}(k+1|k), \mathbf{0}\right) \end{bmatrix}_{m \times n}. \tag{B.67}$$

Again the elements of matrix $\mathbf{C}(k+1)$ can be determine separately. In our application we may have a measurement estimation for either the flow q or the space-mean speed v (or u^S in other parts), which are given by m_i^q or m_i^v. For m_i^v, the only non-zero element is the derivative to v_i

$$\frac{\partial m_i^v(k)}{\partial v_i} = 1 \tag{B.68}$$

For m_i^v, the derivatives to ρ_i and v_i are the only non-zero elements

$$\frac{\partial m_i^q(k)}{\partial \rho_i} = v_i(k+1)\lambda_i \tag{B.69}$$

$$\frac{\partial m_i^q(k)}{\partial v_i} = \rho_i(k+1)\lambda_i \tag{B.70}$$

For m_i^r, the derivative to r_i is the only non-zero elements

$$\frac{\partial m_i^r(k)}{\partial r_i} = 1 \tag{B.71}$$

For m_i^s, the derivatives to β_i and v_{i-1} are the only non-zero elements

$$\frac{\partial m_i^s(k)}{\partial \rho_{i-1}} = \beta_i(k) v_{i-1}(k) \lambda_{i-1} \tag{B.72}$$

$$\frac{\partial m_i^s(k)}{\partial v_{i-1}} = \beta_i(k) \rho_{i-1}(k) \lambda_{i-1} \tag{B.73}$$

$$\frac{\partial m_i^s(k)}{\partial \beta_i} = \rho_{i-1}(k) v_{i-1}(k) \lambda_{i-1} \tag{B.74}$$

B.2.4 Derivative of measurement equations to the measurement error vector $\mathbf{\Sigma}(k+1)$

$$\mathbf{\Sigma}(k+1) = \frac{\partial \mathbf{g}}{\partial \eta} \left(\hat{\mathbf{x}}(k+1|k), \mathbf{0} \right) \tag{B.75}$$

This is an identity matrix. Note that $m_i^q(k)$ and $m_i^s(k)$ contain two noice factors, namely the state and measurement noise. This may have to be included in the determination of $\mathbf{R}(k)$.

Appendix C

Adaptive smoothing method

In this research the proposed methodology is compared to the Adaptive Smoothing Method (ASM) (Treiber & Helbing, 2002). The relevant equations and parameters settings are depicted in this appendix.

C.1 Methodology

A traffic state estimator based on the ASM is able to provide estimates for all combinations of location x and time t based on the surrounding observations. Yuan & Hoogendoorn (2014) provides the important ASM equation used in this research. With these equations the speed and flow may be estimated from the LDD. The estimated variable z_{ij} at location x_i and time t_j is given by:

$$z_{ij} = \alpha^{ASM} z_{ij}^{cong} + \left[1 - \alpha^{ASM}\right] z_{ij}^{free} \tag{C.1}$$

This function weighs z_{ij}^{free} and z_{ij}^{cong} using the factor α^{ASM}. The former estimate is the estimated variable assuming that (x_i, t_j) is in free-flow and the latter that it is in congestion. The factor α^{ASM} ways these factors based on the expected conditions. The free-flow and congested estimates, z_{ij}^{free} and z_{ij}^{cong}, are determined using the following equations. The ASM only considers observations within a certain range. This range is given by $x_I \in x_i \pm X_{max}, t_J \in t_j \pm T_{max}$.

$$z_{ij}^{free} = \frac{1}{\beta_{ij}^{free}} \sum_{x_I \in x_i \pm X_{max}, t_J \in t_j \pm T_{max}} \left[\phi(x_i - x_I, t_j - t_J, c_{free}) z_{IJ}\right] \tag{C.2}$$

$$z_{ij}^{cong} = \frac{1}{\beta_{ij}^{cong}} \sum_{x_I \in x_i \pm X_{max}, t_J \in t_j \pm T_{max}} \left[\phi(x_i - x_I, t_j - t_J, c_{cong}) z_{IJ}\right] \tag{C.3}$$

Where the normalizing terms β_{ij}^{free} and β_{ij}^{cong} are given by:

$$\beta_{ij}^{free} = \sum_{x_I \in x_i \pm X_{max}, t_J \in t_j \pm T_{max}} [\phi(x_i - x_I, t_j - t_J, c_{free})] \tag{C.4}$$

$$\beta_{ij}^{cong} = \sum_{x_I \in x_i \pm X_{max}, t_J \in t_j \pm T_{max}} [\phi(x_i - x_I, t_j - t_J, c_{cong})] \tag{C.5}$$

The importance of the individual observations is represented by the kernel ϕ, which is a negative exponential function:

$$\phi(dx, dt, c) = exp\left(-\left|\frac{dx}{\sigma}\right| - \left|\frac{dt - dx/c}{\tau}\right|\right) \tag{C.6}$$

This kernel bases the importance of the individual observation on its relative position in the time-space domain, dx and dt, and the wave speed c of the related traffic state, free-flow or congestion. The factor α^{ASM} may be calculated based on estimated free-flow and congested state speeds, namely v_{ij}^{free} and v_{ij}^{cong}. The minimum speed-estimate is $v_{min} = min\left(v_{ij}^{free}, v_{ij}^{cong}\right)$.

$$\alpha^{ASM} = \frac{1}{2}\left[1 + tanh\left(\frac{V_c^{ASM} - v_{min}}{dV^{ASM}}\right)\right] \tag{C.7}$$

For the parameters $V_c^{ASM} \pm dV^{ASM}$ the values 80 ± 10 km/h may be taken (Yuan & Hoogendoorn, 2014).

C.2 Involving FCD in ASM

Following Yuan & Hoogendoorn (2014) the FCD observations may be involved. The FCD will included in a similar way as the LDD. Figure C.1 shows observation of both data-types in the space-time domain. As it is expected that individual floating car observations are less representative than loop-detector observations (thus one-minute aggregates), a weighting factor for the difference between single LDD and FCD observations will be involved.

$$v_{ij}^{free} = \frac{1}{\beta_{ij}^{free}} \sum_{x_I \in x_i \pm X_{max}, t_J \in t_j \pm T_{max}, d \in [LDD, FCD]} \left[\phi^d(x_i - x_I, t_j - t_J, c_{free}) v_{IJ}^d\right] \tag{C.8}$$

$$v_{ij}^{cong} = \frac{1}{\beta_{ij}^{cong}} \sum_{x_I \in x_i \pm X_{max}, t_J \in t_j \pm T_{max}, d \in [LDD, FCD]} \left[\phi^d(x_i - x_I, t_j - t_J, c_{cong}) v_{IJ}^d\right] \tag{C.9}$$

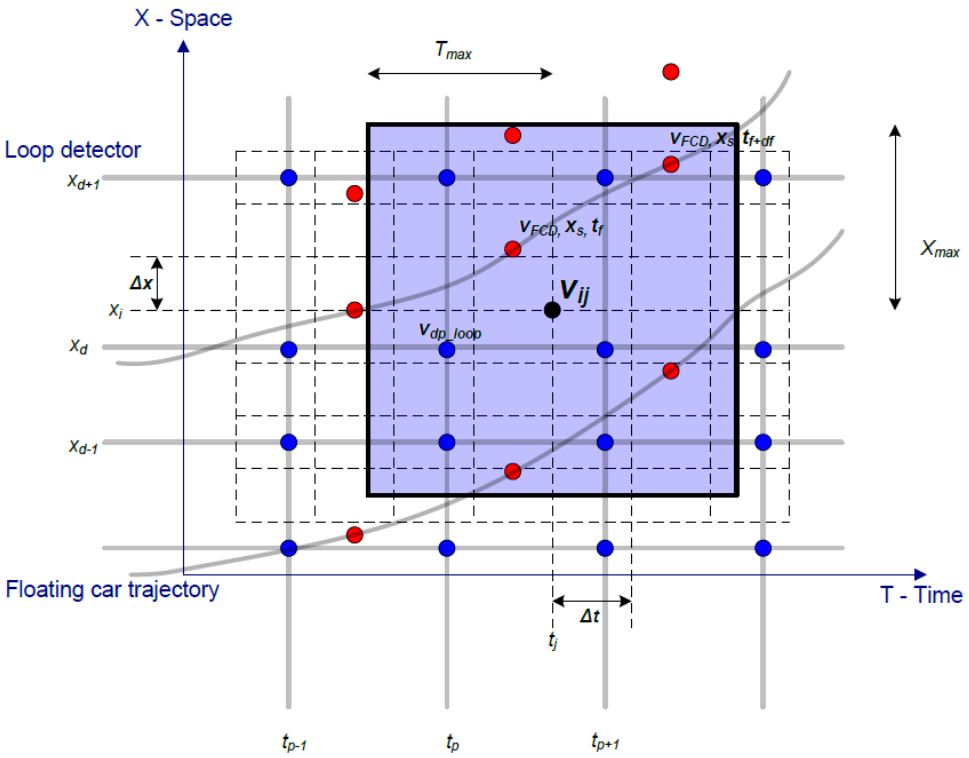


FIGURE C.1: Visualization of the observations within the time-space domain (Yuan & Hoogendoorn, 2014). The loop-detector and floating car observations are respectively denoted by the blue and red dots. The light blue box indicates which observations are taken into account to estimate u^S and q at {i,j}.

$$\phi(dx, dt, c) = f^d exp\left(-\left|\frac{dx}{\sigma}\right| - \left|\frac{dt - dx/c}{\tau}\right|\right) \tag{C.10}$$

where f^d represents a factor which weighs the importance of individual floating car observations with respect to the one-minute aggregated loop-detector observations. For LDD, a value of $f^{LDD} = 1$. To find the optimal f^{FCD} it was varied. Out of the options $[0.01, 0.05, 0.1, 0.2, 0.4]$, $f^{FCD} = 0.01$ yielded the best estimation performance and was thus selected.

C.3 Parameters

For the ASM-based estimates the same conditions in data-availability hold as for the proposed methodology-based estimates. It is assumed that the data is available until the minute for which the conditions are estimated. As the aggregated loop-detector observations are assigned to the middle of the one-minute aggregation period, the forward time-range is $30s$. For the backwards time-range a value of $120s$ is taken. The space-range is not restricted by the data-availability. A value of $4000m$ is taken, such that in low LDD-availability settings multiple loop-detector observations are available.

The other parameters are based on literature. Following Treiber & Helbing (2002), $\sigma = 0.6$ km and $\tau = 1.1$ min. These parameters are important for the weight assigned to individual observations based on the distance from the estimated point. Furthermore, the propagation velocities of perturbations in free-flow c_{free} and congested c_{cong} traffic are respectively chosen to be 80 and -15 km/h. Following Yuan & Hoogendoorn (2014) $V_c^{ASM} = 80$ km/h and $dV^{ASM} = 10$ km/h.

C.4 Evaluate influence of LDD and FCD

The space-mean speed u^S estimates may be based on both data-types. Both have a certain contribution to the combined estimate. We may get more insight in these contributions by evaluating the individual estimates and their weight in the combined estimate.

$$v_{ij}^{free} = \frac{1}{\beta_{ij,LDD}^{free} + \beta_{ij,FCD}^{free}} \sum_{x_I \in x_i \pm X_{max}, t_J \in t_j \pm T_{max}} \left[\sum_{d \in LDD,FCD} \phi^d(x_i - x_I, t_j - t_J, c_{free}) v_{IJ,d} \right] \tag{C.11}$$

Where the FCD and LDD estimates may be separated:

$$v_{ij,LDD}^{free} = \frac{1}{\beta_{ij,LDD}^{free}} \sum_{x_I \in x_i \pm X_{max}, t_J \in t_j \pm T_{max}} \left[\phi^{LDD}(x_i - x_I, t_j - t_J, c_{free}) z_{IJ,LDD} \right] \tag{C.12}$$

$$v_{ij,FCD}^{free} = \frac{1}{\beta_{ij,FCD}^{free}} \sum_{x_I \in x_i \pm X_{max}, t_J \in t_j \pm T_{max}} \left[\phi^{FCD}(x_i - x_I, t_j - t_J, c_{free}) z_{IJ,FCD} \right] \tag{C.13}$$

The combined (free-flow and congested) speed estimates are weighted using α^{ASM}. If the estimates are performed purely based on the LDD and FCD, this factor may differ from the combined α^{ASM}. However, to show contribution of the LDD-based and FCD-based estimate for the combined estimate the combined α^{ASM} will be used.

$$v_{ij,LDD} = \alpha^{ASM} v_{ij,LDD}^{cong} + \left[1 - \alpha^{ASM}\right] v_{ij,LDD}^{free} \tag{C.14}$$

$$v_{ij,FCD} = \alpha^{ASM} v_{ij,FCD}^{cong} + \left[1 - \alpha^{ASM}\right] v_{ij,FCD}^{free} \tag{C.15}$$

Additionally, the weights may be compared. Combined the weight equal to one: $w_{LDD} + w_{FCD} = 1$, thus $w_{FCD} = 1 - w_{LDD}$. Due to this relation it is enough to express the only one of the weights and derive the other. Therefore, we will only express the weight of the LDD. If $w_{LDD} = 1$, the estimate is

fully based on the LDD and if $w_{LDD} = 0$ the estimate is fully based on the FCD. The weight of the LDD in the estimate is given by:

$$w_{LDD} = \frac{\alpha^{ASM}\beta_{ij,LDD}^{cong} + \left[1 - \alpha^{ASM}\right]\beta_{ij,LDD}^{free}}{\alpha^{ASM}\left(\beta_{ij,LDD}^{cong} + \beta_{ij,FCD}^{cong}\right) + \left[1 - \alpha^{ASM}\right]\left(\beta_{ij,LDD}^{free} + \beta_{ij,FCD}^{free}\right)} \tag{C.16}$$

The effect of varying f^{FCD}, ceteris paribus, may be evaluated in more detail by making contour plots of w_{LDD} and the errors for LDD-based and FCD-based estimations. The error is expressed as $u - \hat{u}$, thus a positive error means that the variable is underestimated. Also the estimates of the LDD and FCD may be evaluated. This provides insight in the ability of the LDD and FCD based contributions to the estimations to describe the true conditions. Again as stated before these differ from the estimations which would purely be based on one of the data-types, due to the factor α^{ASM}.

Appendix D

Estimator inputs

The estimator inputs have to be defined. In this appendix the road lay-out characteristics, the loop-detector data availability settings and parameter inputs are presented.

D.1 Road lay-out characteristics

A FOSIM model of the A13-S is considered in this research. In this appendix the characteristics and discretization of this road are discussed. Table D.1 provides the length of each cell, the number of lanes and the location of the ramps.

Table D.1: Segmentation of the A13 freeway stretch

Cell	Length [m]	Number of lanes	On-ramp	Off-ramp
1	555	3		
2	520	3		
3	520	3		
4	520	3		
5	520	3		
6	520	3		
7	520	3		
8	520	3		
9	520	3		
10	520	3		
11	520	3		
12	520	3		
13	680	3	1	
14	620	3		1
15	670	3	2	
16	678	3		
17	500	3		2
18	500	3		
19	670	3	3	
20	660	3		3
21	665	3		
22	565	3	4	
23	770	3		4
24	540	3	5	
25	540	3		
26	540	3		
27	540	3		
28	540	3		
29	570	3		
30	640	3	6	5
31	860	3		
32	635	4		
33	510	2		6
34	515	2		

D.2 Loop-detector data availability settings

The LDD availability setting is defined as the loop-detectors which are available. For each loop-detector this results in a binary variable whether it is available or not. The three considered LDD availability settings are provided in Table D.2.

TABLE D.2: Available loop-detectors for the three LDD availability settings. An x denotes that the loop-detector is available.

Cell	1	2	3
0	x	x	x
1	x		
2	x	x	
3	x		
4	x	x	x
5	x		
6	x	x	
7	x		
8	x	x	x
9	x		
10	x	x	
11	x		
12	x	x	x
13	x	x	x
14	x	x	x
15	x		
16	x	x	x
17	x		
18	x	x	x
19	x	x	x
20	x		
21	x	x	x
22	x	x	x
23	x	x	x
24	x		
25	x	x	
26	x		
27	x	x	x
28	x		
29	x	x	x
30	x	x	
31	x		
32	x	x	x
33	x		
34	x	x	x

D.3 Selection of parameters of Q and R(k)

The selection of the parameters is based on a single data-availability setting. This is the LDD-availability setting 2 and for the FCD-density a penetration rate of 10 % is taken.

The parameters are selected consecutively. Firstly, the error variance of LDD-based u^S estimates, $E[(\eta_i^{u^S})^2]$, are selected. Secondly, the error variance of LDD-based q estimates, $E[(\eta_i^q)^2]$, are selected. In the third and fourth step, the model error variances are considered, namely respectively $E[(\xi_i^{u^S})^2]$ and $E[(\xi_i^q)^2]$. Finally, the remaining parameters are selected. The last step will be extensive, as it is expected that the influence on the estimator performance of these parameters is less than of the other parameters.

TABLE D.3: Estimation performance for different error variances of the LDD-based space-mean speed estimate. The other fixed error variance parameters are $E[(\xi_i^q(k))^2] = 250^2$, $E[(\xi_i^{u^S}(k))^2] = 25^2$, $E[(\eta_i^q(k))^2] = 250^2$.

	$E[(\eta_i^{u^S}(k))^2]$		
	10^2	25^2	50^2
MAE ρ [veh/km/lane]	3.53	18.45	286.25
MAE u^S $[km/h]$	3.97	7.98	25.65

The setting with a error variance of 10^2 $[(km/h)^2]$ has the best performance. Therefore, this LDD-speed error variance will be used. Next the LDD-based q model error variance, $E[(\eta_i^q)^2]$, is selected.

TABLE D.4: Estimation performance for different LDD-based flow error variances. The other fixed error variance parameters are $E[(\xi_i^q(k))^2] = 250^2$, $E[(\xi_i^{u^S}(k))^2] = 25^2$, $E[(\eta_i^{u^S}(k))^2] = 10^2$.

	$E[(\eta_i^q(k))^2]$		
	100^2	250^2	500^2
MAE ρ [veh/km/lane]	241.56	2.20	2.27
MAE u^S [km/h]	9.00	3.29	3.16

The setting with a error variance of 250^2 $[(veh/h)^2]$ has the best performance. This choice is made based on the ρ estimation performance. Therefore, this LDD-flow error variance will be used. Next the model-based u^S model error variance, $E[(\xi_i^{u^S})^2]$, is selected.

TABLE D.5: Estimator performance for different model-based space-mean speed error variances. The other fixed error variance parameters are $E[(\xi_i^q(k))^2] = 250^2$, $E[(\eta_i^q(k))^2] = 250^2$, $E[(\eta_i^{u^S}(k))^2] = 10^2$.

	$E[(\xi_i^{u^S})^2]$		
	10^2	25^2	50^2
MAE ρ [veh/km/lane]	714.71	2.17	2.20
MAE u^S [km/h]	13.60	3.32	3.48

The setting with a error variance of 25^2 $[(km/h)^2]$ has the best performance. Therefore, this model-based space-mean speed error variance will be used. Next the model-based q model error variance, $E[(\xi_i^q)^2]$, is selected.

TABLE D.6: Estimator performance for different model-based flow error variances. The other fixed error variance parameters are $E[(\xi_i^{u^S}(k))^2] = 25^2$, $E[(\eta_i^q(k))^2] = 250^2$, $E[(\eta_i^{u^S}(k))^2] = 10^2$.

	$E[(\xi_i^q)^2]$		
	100^2	250^2	500^2
MAE ρ [veh/km/lane]	2.27	2.30	2.36
MAE u^S [km/h]	3.43	3.35	3.25

Table D.6 shows small differences in the performance statistics. The density estimates seem to improve with decreasing model flow error variances, while the space-mean speed estimate accuracy decreases. This may be caused by an increase in accuracy of the flow at the expense of the space-mean speed estimate. As the density is considered as the most important variable, and the error differences are small, a model flow error variance of 100^2 $[(veh/h)^2]$ is selected. Next the influence of the ramp-related model-based error variances are evaluated.

TABLE D.7: Estimator performance for different model error variances of the space-mean speed prediction. The other fixed error variance parameters are $E[(\xi_i^q)^2] = 100^2$, $E[(\xi_i^{u^S}(k))^2] = 25^2$, $E[(\eta_i^q(k))^2] = 250^2$, $E[(\eta_i^{u^S}(k))^2] = 10^2$.

	$E[(\xi_i^r)^2]$ and $E[(\xi_i^\beta)^2]$	
	fixed	ramp-dependent
MAE ρ [veh/km/lane]	2.19	2.17
MAE u^S [km/h]	3.40	3.37

Table D.7 shows that the ramp-dependent model-based error variances lead to a better estimation performance. In this case the variance is dependent on the expected on-flows and exiting rates. In other words, if more vehicles are expected to use the ramp a higher error variance is assigned. Finally, it is investigated if a fixed or dynamic fundamental diagram is considered. A fixed fundamental diagram is achieved by setting the stationary flow equation, which describes the continuous fundamental diagram, error variances to zero.

TABLE D.8: Estimator performance for a fixed (zero) and dynamic (non-zero) fundamental diagram. The other fixed error variance parameters are $E[(\xi_i^q)^2] = 100^2$, $E[(\xi_i^{u^S}(k))^2] = 25^2$, $E[(\eta_i^q(k))^2] = 250^2$, $E[(\eta_i^{u^S}(k))^2] = 10^2$.

	Stationary flow equation error variances	
	non-zero	zero
MAE ρ [veh/km/lane]	2.14	2.15
MAE u^S [km/h]	3.37	3.37

Table D.8 only shows a small difference, but indicates that the variable fundamental diagram yields a better estimation performance. The small difference is caused by the high value assigned to τ which limits the influence of the fundamental diagram. The resulting parameter settings are depicted in Table D.9.

Table D.9: Assimilation technique parameters

Parameter	Value	
$E[(\eta^{u^S}_{i,LDD})^2]$	10^2	$(km/h)^2$
$E[(\eta^{q}_{i,LDD})^2]$	250^2	$(veh/h)^2$
$E[(\xi^{u^S}_{i,LDD})^2]$	25^2	$(km/h)^2$
$E[(\xi^{q}_{i,LDD})^2]$	100^2	$(veh/h)^2$
$E[(\xi^{r}_{i,LDD})^2]$	ramp-specific	-
$E[(\xi^{\beta}_{i,LDD})^2]$	ramp-specific	-
$E[(\xi^{u^S_0}_{i,LDD})^2]$	25^2	$(km/h)^2$
$E[(\xi^{q_0}_{i,LDD})^2]$	100^2	$(veh/h)^2$
$E[(\xi^{\rho_N}_{i,LDD})^2]$	1.5^2	$(veh/km)^2$
$E[(\xi^{v_f}_{i,LDD})^2]$	0.5^2	$(km/h)^2$
$E[(\xi^{\rho_{cr}}_{i,LDD})^2]$	0.1^2	$(veh/km)^2$
$E[(\xi^{a}_{i,LDD})^2]$	0.01^2	-

Appendix E

Solving errors in the MATLAB algorithm

When developing a complex methodology as proposed in this thesis and writing the proposed methodology-based estimator algorithm, errors may occur. In most cases, a small derivation or programming error leads to an unstable and incorrect estimator. To solve this problem, the errors should be found. For this purpose a procedure is designed, see Figure E.1. With this procedure multiple errors were found and solved, leading to the algorithm which was used.

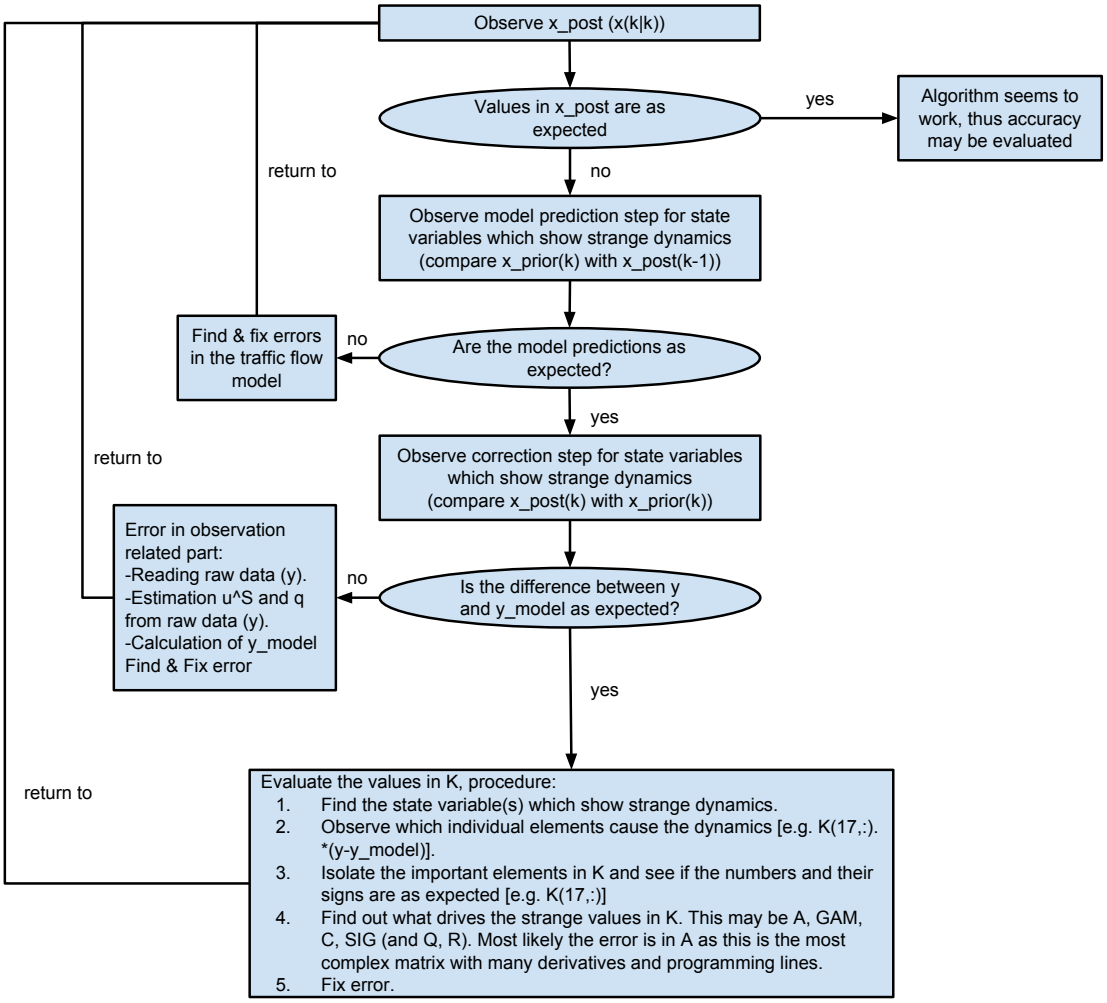


FIGURE E.1: Procedure to find and fix errors in the algorithm. This procedure proposed a number of steps which may be taken to systematically solve potential errors.