

Lagrangian Multi-Class Traffic State Estimation

Yufei Yuan

This thesis is a result from a project funded by
Delft University of Technology (TUDelft) and the Netherlands
Research School for Transport, Infrastructure and Logistics (TRAIL).



Cover illustration: Yufei Yuan and Jing Wei

Lagrangian Multi-Class Traffic State Estimation

Proefschrift

ter verkrijging van de graad van doctor
aan de Technische Universiteit Delft,
op gezag van de Rector Magnificus prof. ir. K.C.A.M. Luyben,
voorzitter van het College voor Promoties,
in het openbaar te verdedigen op dinsdag 19 maart 2013 om 10:00 uur
door

Yufei YUAN

Master of Science in Transport and Planning
geboren te Guilin, China

Dit proefschrift is goedgekeurd door de promotor:

Prof. dr. ir. S.P. Hoogendoorn

Copromotor: Dr. ir. J.W.C. van Lint

Samenstelling promotiecommissie :

Rector Magnificus,	voorzitter
Prof. dr. ir. S.P. Hoogendoorn,	Technische Universiteit Delft, promotor
Dr. ir. J.W.C. van Lint,	Technische Universiteit Delft, copromotor
Prof. dr. ir. C. Vuik,	Technische Universiteit Delft
Prof. ir. L.H. Immers,	Technische Universiteit Delft
Prof. dr. R.E. Wilson,	University of Bristol
Prof. dr. L. Leclercq,	École Nationale des Travaux Publics de l'État
Prof. dr. P.B. Mirchandani,	Arizona State University
Prof. dr. ir. B. van Arem,	Technische Universiteit Delft, reservelid

TRAIL Thesis Series no. T2013/5, the Netherlands Research School TRAIL

TRAIL

P.O. Box 5017

2600 GA Delft

The Netherlands

Phone: +31 (0) 15 278 6046

E-mail: info@rsTRAIL.nl

ISBN: 978-90-5584-162-2

Copyright © 2013 by Yufei Yuan

All rights reserved. No part of the material protected by this copyright notice may be reproduced or utilized in any form or by any means, electronic or mechanical, including photocopying, recording or by any information storage and retrieval system, without written permission from the author.

Printed in the Netherlands

路漫漫其修远兮,吾将上下而求索.

*The road ahead will be long and the climb will be steep,
but I will never terminate pursuit.*

Preface

Traffic and transport research is truly attractive to me. To explore the problem that relates closely to our life also motivates me. This is the reason why I dive into the field of traffic science. After my bachelor graduation in engineering mechanics in Shanghai China, I came to the Netherlands to start my master study at the department of Transport and Planning, Delft University of Technology in 2006. I succeeded in obtaining my degree within two years. With my ongoing curiosity in traffic, I decided to pursue a PhD degree in the same department. The four year PhD life is really an unforgettable and precious experience to me. It expands my view and horizon not only in the field of traffic science but also in the society and the world. Upon the completion of my PhD thesis, I would like to take the opportunity to thank all the people who helped me during my PhD study.

First of all, I would like to express my heartfelt thanks to my promotor, Serge Hoogenboom, for giving me sufficient guidance and confidence. I feel very grateful and proud of being one of your master and PhD students. Your research attitude has strongly influenced me and your ideas always inspire me. Thank you! I am also deeply indebted to my excellent daily supervisor and copromotor, Hans van Lint, for giving me kind and patient monitoring. Your critical comments and suggestions enabled me to make this research and my thesis better than I could have done by myself. When there were moments that I doubted about my research, you always supported me and gave me confidence. Thanks a lot!

I would like to express my thanks to my supervisor Jos Vrancken from the TBM faculty, for letting me join the C4C project and for all the useful advice and support. I feel rather grateful to my supervisor Eddie Wilson in England, for hosting me, for his kind supervision and collaboration when I exchanged at the University of Southampton as a guest researcher, and also for the final checking of my thesis. I also want to thank other committee members for their useful comments.

During my PhD research, I have the pleasure to work with my wonderful colleagues at the department. I would like to address my special thanks to Thomas, Femke and Olga. Thomas, thank you very much for all the inspiring discussion during the past four years and for the collaboration on projects, papers and Ping-Pong tournaments. Femke, thank you for the discussion and collaboration on papers. I learned a lot from you. Olga, thank you for always patiently answering lots of basic questions about the Dutch

language, and helping me with the translation of my summary and propositions. I also want to thank my other colleagues: Winnie, Victor, Adam and Chris, for their help and the chance to exchange ideas; Kees, Edwin, Peter, Nicole, Piet, Priscilla, Dehlaila, Charelle and Conchita, for all the technical and administrative support; Meng, Yaqing, Mo, Yubin, Gijs, Eric-sander, Tamara, Giselle, Kakpo, Bernat, Pavle, Daniel, Mario, Wouter, Mahtab, Mignon, Guus, Ramon and Raymond, for all the wonderful moments. Furthermore, my gratitude also goes to my (former) office mates: Tamara, Thomas, Olga, Fangfang, Lucas, Clarie, Feifei and Gerdien. I feel comfortable to work with you. I had a great time and enjoyed the multi-culture environment in our department.

My special thanks goes to Zhiwei for the friendship since the first year of undergraduate, Xiaoxiao for all the sharing and discussion, Xuying for all her help and support, Kai and Feng for the proof reading and all the valuable comments on my thesis, Yong and Jing for helping me visualise my research concept.

Finally, I would like to thank my family, who have always been supporting me. Special gratitude to my mother, who is consistently supporting, caring and listening to me. I dedicate this thesis to my mom, with love and thanks for her endless and unconditional love throughout my life.

Yufei Yuan, February 2013

*Dedicated to my mother
for her endless and unconditional love*

Contents

Preface	i
List of Figures	x
List of Tables	xii
Notation	xiii
1 Introduction	1
1.1 Background	2
1.2 Traffic state estimation	2
1.3 Dynamic traffic flow models	4
1.4 Data assimilation	5
1.5 Observation models and empirical data	6
1.6 Research contributions and relevance	7
1.6.1 Scientific contributions	7
1.6.2 Practical contributions	9
1.7 Outline of this thesis	10
2 The state-of-the-art in traffic state estimation	13
2.1 Introduction	14
2.2 A new classification framework for model-based traffic state estimation research	16
2.3 Choices in traffic process models	17
2.3.1 Eulerian formulated traffic process models	17
2.3.2 Lagrangian formulated traffic process models	19

2.4	Choices for incorporating observation models	20
2.5	Data-assimilation techniques	23
2.5.1	Overview of recursive assimilation techniques	23
2.5.2	Motivation for applying the EKF	24
2.6	Research direction and main challenges	26
2.7	Summary	26
3	Model-based mixed-class state estimation in Lagrangian coordinates	29
3.1	Introduction	30
3.2	Process models: Eulerian and Lagrangian formulations of mixed-class first-order traffic model	30
3.2.1	Mixed-class Eulerian formulated process model	30
3.2.2	Mixed-class Lagrangian formulated process model	31
3.3	Modelling network discontinuities in the Lagrangian formulation	34
3.3.1	Eulerian formulated node models	35
3.3.2	Lagrangian formulated node models	36
3.4	Observation models for mixed-class Lagrangian formulation	39
3.5	Mixed-class Lagrangian traffic state estimation based on the Extended Kalman Filter	43
3.6	Advantages of Lagrangian formulation for traffic state estimation	46
3.7	Summary and discussion	47
4	Model-based multi-class state estimation in Lagrangian coordinates	49
4.1	Introduction	50
4.2	Multi-class Lagrangian traffic flow models: continuum forms and different discretisation approaches	50
4.2.1	Eulerian formulated multi-class models	50
4.2.2	Lagrangian formulated multi-class models	51
	“Piggy-back” formulation	51
	“Multi-pipe” formulation	52
4.2.3	Discussion and choice	55
4.3	Multi-class node model for network discontinuities	57

4.4	Observation models for multi-class Lagrangian formulation	59
4.5	Multi-class Lagrangian traffic state estimation based on the Extended Kalman Filter	60
4.6	Summary and discussion	64
5	Case studies for Lagrangian traffic state estimation	65
5.1	Introduction	66
5.1.1	Experimental setup	66
5.1.2	Experimental objectives	66
5.2	Link-level validation of Lagrangian traffic state estimation	67
5.2.1	Data and test network	67
5.2.2	Experimental scenarios	68
5.2.3	Performance criteria	69
5.2.4	Results and discussion	70
	Quantitative results and discussion	70
	Discussion	71
5.2.5	Conclusion	71
5.3	Comparison between Lagrangian and Eulerian approaches at a link level	73
5.3.1	Data and test network	73
5.3.2	Experimental scenarios	74
5.3.3	Results and discussion	75
	Quantitative analyses	75
	Qualitative analyses	77
5.3.4	Conclusion	78
5.4	Comparison between Lagrangian and Eulerian approaches at a network level	78
5.4.1	Data and test network	80
5.4.2	Experimental scenarios	80
5.4.3	Results and discussion	81
	Quantitative analyses	81
	Qualitative analyses	82

5.4.4	Conclusion	84
5.5	Verification of multi-class Lagrangian traffic state estimation and comparison with its mixed-class formulation	84
5.5.1	Data and test network	84
5.5.2	Experimental scenarios	85
5.5.3	Results and discussion	86
	Quantitative analyses	86
	Qualitative results	89
5.5.4	Conclusion	89
5.6	Multi-class Lagrangian traffic state estimation in a real traffic network	92
5.6.1	Data and test network	92
5.6.2	Experimental scenarios	93
5.6.3	Results and discussion	94
5.6.4	Conclusion	96
5.7	Summary and discussion	99
6	Preparation and data pre-processing	101
6.1	Introduction	102
6.2	Speed-bias correction	104
6.2.1	Problem analysis and background	104
6.2.2	Overview of speed-bias correction algorithms	106
6.2.3	A new correction algorithm based on flow-density relations	108
6.2.4	Validation of the speed-bias correction algorithm	112
	Description of model and data	112
	Definition of scenarios	114
	Performance criteria	114
6.2.5	Results and discussion	115
	Qualitative impression	115
	Quantitative analysis of the normal correction	115
	Robustness study	119
	Qualitative analysis	119

6.2.6	Conclusions	119
6.3	Estimation of multi-class and multi-lane counts	121
6.3.1	Problem analysis	121
6.3.2	Methodology	123
6.3.3	Individual vehicle data environment	125
6.3.4	Aggregate data scenarios	125
6.3.5	Results	128
6.3.6	Error analysis	133
6.3.7	Conclusions	135
6.4	Summary	136
7	Conclusions and recommendations	139
7.1	Main findings and conclusions	140
7.2	Research implications	141
7.3	Future research directions	142
	Bibliography	144
	Appendices	155
A	Estimation of class-specific information from mixed-class data	157
A.1	Problem description	157
A.2	Solution procedure for two vehicle classes	158
A.3	Solution procedure for U vehicle classes	158
	Summary	159
	Samenvatting (Dutch summary)	161
	Summary (Chinese)	163
	TRAIL Thesis Series	165
	About the author	167

List of Figures

1.1	Schematic representation of an ideal model-based decision support system.	3
1.2	Outline of the thesis.	10
2.1	Schematic procedure of the data-assimilation method	15
3.1	Lagrangian fundamental diagrams.	33
3.2	Nodes with locally numbered links.	35
3.3	Vehicle and time discretisation approaches in this thesis.	38
3.4	Influence area of a detector (in congested state).	42
4.1	Vehicle discretisation at a time instant in the Piggy-back formulation (two-class case).	53
4.2	Vehicle discretisation at a time instant in the Multi-pipe formulation (two-class case).	54
4.3	On-ramp node modelling for a two-class case.	58
4.4	Multi-class Lagrangian fundamental relation (Smulders').	63
5.1	Illustration of a lane-drop case.	68
5.2	Comparison of speed contour plots.	72
5.3	Geometry of the enhanced carriageway showing the location of inductance loops.	74
5.4	Error measurement comparison between two methods for each of the 8 simulation runs of scenario 1.	76
5.5	Snapshots of a small region of the whole spatiotemporal speed map for scenario 1 based on both Eulerian and Lagrangian approaches.	79
5.6	Error comparison between two methods for each of the ten simulation runs using Eulerian sensing data.	82

5.7	Reference speed map with the related error maps of the speed estimates from both Eulerian and Lagrangian approaches.	83
5.8	Reference speed map with the related error maps of the speed estimates from both mixed-class and multi-class Lagrangian state estimations.	91
5.9	Google map picture of the chosen A15 freeway network.	92
5.10	A15 freeway section (eastbound) modelling in MATLAB.	93
5.11	Speed estimates from the multi-class Lagrangian traffic state estimation in four different cases.	95
5.12	Speed estimate comparison between case 1 and case 2.	96
5.13	Trajectory data collected from a helicopter and their application in traffic state estimation.	97
5.14	Class-specific speed estimates from the multi-class Lagrangian traffic state estimation, compared with raw speed loop data.	98
6.1	Simplified structure of a traffic control loop.	102
6.2	(a) Scatter plot of instantaneous speed variance versus time-mean speed; (b) Scatter plot of instantaneous variance versus “local density” and “time series” approximation.	107
6.3	Speed contour plot measured at the Dutch A13 freeway southbound between Delft North and Rotterdam Airport.	109
6.4	Schematic of correction principle.	109
6.5	Procedure of the speed-bias correction algorithm.	110
6.6	Illustration of a Dutch freeway A13.	113
6.7	Examples of speed-bias correction at a cross-section level.	116
6.8	Reference speed map from A13 FOSIM model.	120
6.9	Performance of polynomial-fit correction.	120
6.10	Comparison of time series for estimates and ground truth data in scenario 1 using the UK data set.	130
6.11	Comparison of time series for estimates and ground truth data in scenario 1 using the Dutch (NL) data set.	131

List of Tables

2.1	Classification framework for model-based traffic state estimation(TSE) research in terms of two mathematical formulations of process and observation models in two coordinate systems	17
2.2	Classification of previous studies based on the proposed framework for traffic state estimation research in terms of first two (X-Y) dimensions	22
2.3	Overview of the main characteristics of different recursive assimilation techniques and the relating applications	25
3.1	Extended Kalman Filtering algorithm (Pseudo code)	45
4.1	Comparison between two formulations of the multi-class Lagrangian model	57
4.2	Overview of observations and the related observation models in the Lagrangian formulation	61
5.1	Overview of studied cases	66
5.2	Objectives of case studies on Lagrangian TSE approaches	67
5.3	Experimental scenarios	69
5.4	<i>RMSE</i> errors of all the scenarios	70
5.5	<i>MAPE</i> errors of all the scenarios	71
5.6	State estimation results of seven scenarios with different loop and FCD resolutions	76
5.7	State estimation results based on two formulations (average over ten simulation runs)	81
5.8	Performance of mixed-class traffic state estimation with respect to different truck shares, using loop data	86
5.9	Performance comparison between mixed-class (a) and multi-class (b) traffic state estimation, in terms of loop observations	88

5.10	Performance comparison between mixed-class (a) and multi-class (b) traffic state estimation, in terms of FCD observations	88
5.11	Performance of multi-class traffic state estimation with noise in observations and biased model inputs, for three data scenarios	90
5.12	Experimental data scenarios on A15	94
5.13	Overview of the conclusions for all case studies	99
6.1	Error indicators for three correction methods of three test scenarios, in terms of effectiveness and robustness	117
6.2	<i>MPE/MAPE</i> errors in the normal-correction scenario, partitioned over five speed ranges	118
6.3	Matrix of the output variables q_{ij}	127
6.4	The numbers of input and output variables for each scenario	127
6.5	<i>RTCD</i> error measures for different scenarios using both British and Dutch inputs	129
6.6	Validation results for scenario 1 over several days using both the British and the Dutch IVD	132
6.7	Error indicators for scenario 1 in terms of different analysis intervals	134
6.8	Error indicators for scenario 1 in terms of different portions of data set	134
6.9	List of parameters and their standard deviations for scenario 1 based on the UK calibration set using bootstrapping (100 samples)	135

Notation

Coordinates

n	vehicle number	veh.
t	time	s
x	space (distance)	m

Variables

q	flow	veh./s
s	spacing	m/veh.
s_{tot}	(total) effective spacing	m/PCE
v	speed (velocity)	m/s
v_M	space-mean speed	m/s
v_H	harmonic mean speed	m/s
v_L	(local) arithmetic mean speed	m/s
η_u	passenger car equivalent (PCE) value of vehicle class u	-
k	density	veh./m
k_{tot}	(total) effective density	PCE/m
D	traffic demand	veh./s
S	traffic supply	veh./s

Parameters

h_t	minumum time headway (excluding vehicle length)	s
T	minumum time headway (including vehicle length)	s
h_x	minumum space headway (excluding vehicle length)	m
L	minumum space headway (including vehicle length)	m
s_{jam}	jam spacing	m/veh.
s_{cri}	critical spacing	m/veh.
v_{max}	maximum velocity	m/s
v_{free}	free-flow velocity	m/s
v_{cri}	critical velocity	m/s
k_{cri}	critical density	veh./m
k_{jam}	jam density	veh./m

Statistical symbols

$RMSE$	root mean squared error	-
MPE	mean percentage error	%
$MAPE$	mean absolute percentage error	%
SPE	standard deviation of percentage error	%

Chapter 1

Introduction

The first chapter of this thesis introduces the scope of this thesis and highlights the main contributions of this research. After briefly discussing advanced model-based decision support systems in Dynamic Traffic Management (DTM), it emphasizes one of the three tasks within the management decision loop: traffic state estimation. Then, the role and the requirements of state estimation systems in relation to DTM are explained. An introduction to the different components of model-based traffic state estimation is also presented. The scope and main focus of this research is thereafter discussed, after which we address the contributions and relevance of this research. Finally, the outline of the thesis is presented.

1.1 Background

Road traffic plays a major role in the daily life of human beings. It represents a significant percentage of people and goods transport services in most countries in the world, and consequently contributes to economic growth and social progression. However, it has also led to a number of undesirable side effects, such as accidents, pollution and congestion. For most commuters, congestion has become the rule rather than the exception. In general, congestion can be explained by a mismatch between supply (e.g., infrastructure capacities) and demand (e.g., traffic flows, travel activities). One of the potential solutions to alleviate congestion is to bridge this mismatch in road networks by developing and designing so-called Intelligent Transportation Systems (ITS). These systems constitute a collection of solutions to balance traffic supply and demand, directly and indirectly influencing the whole transportation network. In addition, traffic managers (such as governments, traffic management centres) play a leading role in Dynamic Traffic Management (DTM) by implementing such ITS systems effectively. One of the main purposes of DTM is to alleviate traffic jams. DTM requires real-time and reliable ITS systems to support its performance. DTM has been applied around the world successfully and has impacted society in a positive way. ([Transportation-Research-Board, 2000](#); [Rijkswaterstaat, 2003](#); [Louis et al., 2006](#))

1.2 Traffic state estimation

As a first step, the success of DTM depends on accurate, timely and reliable traffic monitoring. The surveillance of traffic in a network entails the measurement of typical network characteristics of traffic (e.g., flows, speeds), and the derivation of various network-related quantities, such as predicted travel times, lengths of queues, space-mean speed, and density. This procedure can be seen as a state estimation process, which is embedded in the network-wide control loop. This control loop can be further facilitated by advanced model-based decision support systems (DSSs). In the context of DTM, advanced model-based DSSs allow traffic managers to assess different traffic control and information provision scenarios in real time. Figure 1.1 schematically outlines such an ideal model-based DSS, which generally performs three closely intertwined tasks. These tasks are (i) traffic state estimation, in which data from various traffic sensors (loops, cameras, probe vehicle reports, etc.) and traffic flow models are used to reconstruct a network-wide picture of the traffic state (e.g., in terms of traffic densities and/or speeds). These can in turn be used as a basis for (ii) traffic state prediction and (iii) (the optimization of) traffic control measures (e.g., algorithms to compute settings of measures, such as speed-limit control, or ramp metering, etc.). The heterogeneity of traffic data from different sensors/detectors makes state estimation a complex and challenging task. As a foundation, reliable and accurate state estimation promotes the efficiency and safety of the whole traffic system. In this thesis, the focus is on the development for the first of these tasks, that is, traffic state estimation.

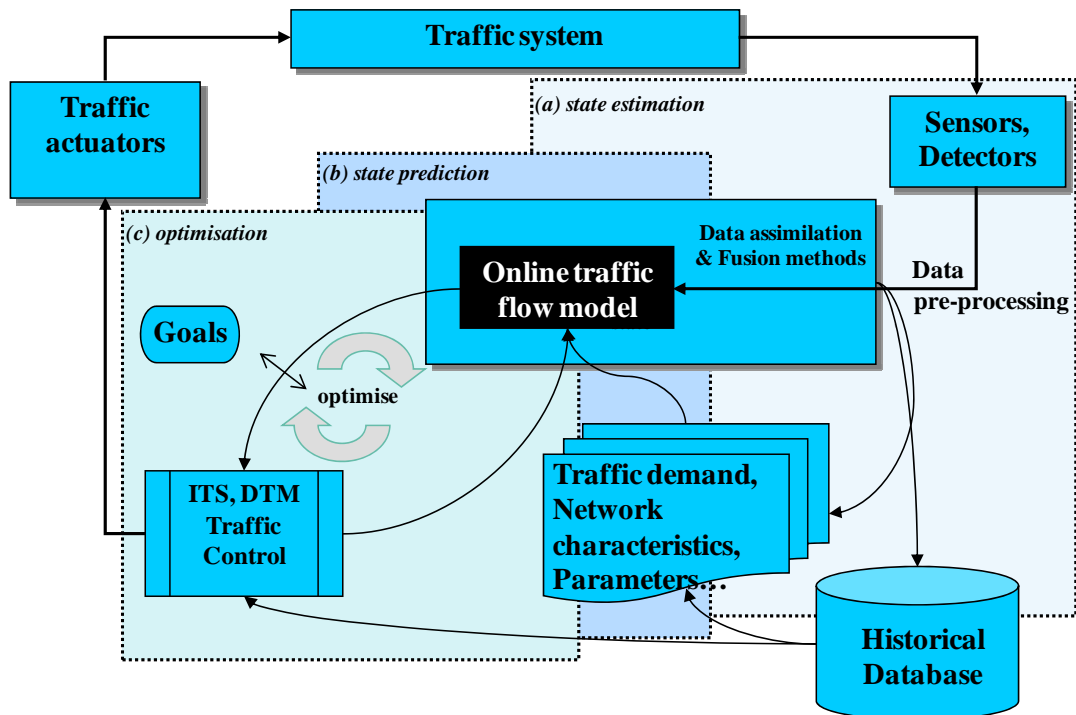


Figure 1.1: Schematic representation of an ideal model-based decision support system.

Therefore, the main question of this research boils down to: *how to provide reliable and accurate traffic state information in real time for DTM?*: specifically, how to develop efficient and accurate model-based traffic state estimation approaches for real-time DTM. This question will be investigated thereafter, where traffic flow models that describe traffic flow physics are incorporated.

Model-based traffic state estimation consists of three components: a dynamic traffic model, an observation model, and an assimilation technique. To compute and predict state variables (e.g., density k , speed v or spacing s), dynamic traffic flow models (see Section 1.3) are used. So-called observation models (e.g., the fundamental diagrams, see Section 1.5) are used to compute and predict, from these state variables, the expected observations from sensors. Finally, a data-assimilation technique (Section 1.4) is needed to estimate the most probable traffic state using both the model predictions and the actual sensor observations. Different choices with respect to these three elements have led to different state estimation approaches. In the following sections, a brief discussion of previous research and the available choices is presented. It is followed by a distinct provision of our research choice (main research scope), which offers a new perspective and opportunities for high-quality state estimation. Chapter 2 will further specify the design choices for traffic state estimation in both the literature and this thesis.

1.3 Dynamic traffic flow models

As the first component in state estimation, dynamic traffic flow models are formulated to describe various traffic phenomena. Existing traffic models serve to compute and predict the evolution of traffic state variables. These state variables consist of (but are not limited to) traffic density (number of vehicles per unit of road length), traffic flow (number of vehicles per unit of time passing a fixed point), traffic speed (space-mean speed of collections of individual vehicles), and vehicle spacing (average distance headway between two successive vehicles). Examples of macroscopic traffic flow models include first-order traffic models, such as the Lighthill-Whitham and Richards (LWR) model (Lighthill & Whitham, 1955; Richards, 1956) and the cell transmission type model (as a special case of the LWR model) (Daganzo, 1994, 1995a) (all referred to as LWR or kinematic wave models further below), and second-order traffic models, such as Payne-type (Payne, 1971) and METANET models (Papageorgiou et al., 1990). Apart from this classic traffic flow modelling, there are some other traffic models used in traffic state estimation, describing basic/empirical relations between traffic variables. For instance, the so-called “MARCOM” model used to estimate density is developed by Davis & Kang (1994), a two-level speed and flow linear model is presented by Cheng et al. (2006) to assimilate cell-phone data, and a linear model is used to estimate speed by Byon et al. (2010). Although these models are different in format, they all attempt to formulate the same traffic phenomena.

In this research, the choice of the process model is the first-order traffic flow model. This model is adequate to reproduce the fundamental phenomena observed in traffic (Daganzo, 1994, 1995b,c, 2002b; Lebacque, 1996; Newell, 1993; Van Wageningen-Kessels et al., 2011a); these are conservation of vehicles, traffic anisotropy, the onset and dissolution of congestion at bottlenecks and the fact that disturbances propagate over space and time in different directions as a function of the prevailing state (congested or not) (Lighthill & Whitham, 1955; Richards, 1956). Although a number of phenomena are not well or fully represented in the first-order model (Helbing, 2001; Kerner, 2009), most of which are related to the capacity drop and to traffic instability, there are still strong arguments as to why the LWR model is a valid choice for the purpose of state estimation. First of all, there is no undisputed alternative model that is able to reproduce these phenomena under all circumstances. Secondly, there is the principle of parsimony. The first-order model contains less parameters than more involved alternatives; it is a model that is mathematically tractable, that can be solved analytically, and that provides the analyst straightforward tools to switch between two different coordinate formulations - the usefulness of which will become clear in the subsequent discussion. Note that the concept and results in this research are not only limited to the first-order traffic flow model but also can be extended to more involved (high-order and/or other gas-kinetic-based) macroscopic models.

Traffic flow can be analysed with respect to three two-dimensional coordinate systems: space-vehicle number coordinates, space-time coordinates, and vehicle number-

time coordinates (Leclercq & Bécarie, 2012; Laval & Leclercq, 2013). The latter two are also known as Eulerian and Lagrangian coordinates, respectively. In this thesis, we restrict the discussion and comparison to the Eulerian and Lagrangian coordinate systems. Eulerian (space-time) coordinates are fixed in space; Lagrangian (vehicle number-time) coordinates move with the traffic. Commonly, both process and observation models are formulated in Eulerian coordinates. In such a formulation, traffic flow is described by state variables over consecutive spatially-fixed road segments. Recent studies by Leclercq et al. (2007); Van Wageningen-Kessels et al. (2009b, 2010a); Van Wageningen-Kessels (2013) show that the LWR model can be formulated and solved more efficiently and accurately in Lagrangian coordinates than in Eulerian coordinates. In such a new coordinate system, traffic flow is divided into vehicle platoons, by which state variables are formulated. The new traffic formulation and its simpler numerical scheme are supposed to yield benefits for state estimation. So far, none of previous research has focused on a Lagrangian form of state estimation. Therefore, we have been motivated to investigate this opportunity.

Another interesting and important perspective in traffic flow modelling is to consider driver and vehicle heterogeneity, which has received considerable attention recently in the research literature (Hoogendoorn & Bovy, 1999; Daganzo, 2002a; Wong & Wong, 2002; Ngoduy & Liu, 2007; Logghe & Immers, 2008; Van Lint et al., 2008b). As a simple example, the distinction can be made between the flows in different lanes (fast or slow vehicle lanes, dedicated lanes etc.) or between the flows from different origins to different destinations or between the flows in different vehicle-classes (trucks, buses, passenger-cars, and high-occupancy vehicles). By considering such heterogeneities in traffic modelling, not only are these models able to describe traffic flow more accurately, but also the control applications for such models can be made more elaborate. Very little previous research has implemented traffic heterogeneity into traffic state estimation. In this research, by considering one aspect of those heterogeneities, vehicle-user classification is addressed in the traffic formulation for state estimation.

To sum up, the main innovation of this work is to put forward an improved state estimation approach formulated in Lagrangian coordinates, providing vehicle-specific (also referred to as multi-class in this thesis) traffic state estimates.

1.4 Data assimilation

A farther component in state estimation is data assimilation, discussed in this section. Applications of data assimilation arise in many fields, such as aerospace, weather forecasting, hydrology, as well as traffic systems. It proceeds by analysis cycles. In each cycle, observations of the current available state of a system are combined with the predictions from a system (process) model to produce “the best” estimate of the current state of the system. Essentially, each cycle tries to balance the uncertainty in the

data and in the forecast. A variety of such data-assimilation techniques have been developed, such as: the *Kalman Filter* (KF) (Kalman, 1960), rooted in the state-space formulation of linear dynamic systems, that provides a recursive solution to the linear optimal filtering problem. However, most models in the real world are nonlinear, and KF can be extended through a linearisation procedure, resulting in an *Extended Kalman Filter* (EKF) (Jazwinsky, 1970). The basic idea of EKF is to linearise the state-space system model and apply the linear KF procedure. Contrary to the EKF, the *Unscented Kalman Filter* (UKF) (Julier & Uhlmann, 1997) does not require a first-order linearisation of the nonlinear system. Instead, it computes the Gaussian error variables by using a deterministic sampling approach. Similarly, the *Ensemble Kalman Filter* (EnKF) (Evensen, 2007) uses a Monte Carlo or ensemble integration method instead of a linearisation procedure. All these methods assume Gaussian error terms to represent the uncertainty in both model predictions and observations. However, the *Particle Filter* (PF) (Gordon et al., 1993; Doucet et al., 2001) relaxes the Gaussian assumption within a Monte Carlo framework. There are also many other data-assimilation techniques available, which are not listed here.

As an essential component in traffic state estimation, (recursive) data-assimilation methods aim to make an optimal estimate of the traffic system state at each time step, which start from the current estimate, predict the future state and then correct it based on new observations. Previous research on traffic state estimation adopts the Kalman Filter technique and/or its advanced variants. Some popular examples are given by Wang & Papageorgiou (2005); Van Lint et al. (2008a); Ngoduy (2008); Work et al. (2008); Herrera et al. (2010). In this thesis, to validate a Lagrangian and multiple-user perspective as well as to fulfill the real-time requirement, a relatively-simple and real-time applicable technique, the EKF, is selected for data assimilation. The detailed reasons for this choice are provided in Chapter 2.

As an answer to the main question in Section 1.2, this thesis concentrates on a model-based multi-class state estimation approach in Lagrangian coordinates, based on the EKF technique. The detailed specification and its research motivation will be explained in the following chapters.

1.5 Observation models and empirical data

In the state estimation procedure, observation models deal with the data collected from sensors. In the current road network, empirical (raw) traffic data are collected mainly from road-side traffic sensors (see the right-upper box in Figure 1.1). Local sensors on cross sections, such as inductive loops, radars and cameras, measure local traffic quantities, such as aggregated traffic counts and spot speeds. These types of devices are classified as Eulerian sensors and the related measurements as Eulerian sensing data, which are related to Eulerian coordinates. Increasingly, data from probe vehicles and mobile phones (Herrera et al., 2010) have also become available, providing posi-

tion and speed information of individual vehicles/travellers. These provide Lagrangian sensors and Lagrangian sensing data.

A Lagrangian formulation of traffic flow provides a natural set of observation equations to deal with such Lagrangian sensing data. The challenge however then lies in incorporating Eulerian-type data. Moreover, these various data sources in both Eulerian and Lagrangian types are typically different in formats and semantics. In a state estimation procedure, observation models are needed to predict and compute the expected traffic measurement data in different formats from system-state variables. Therefore, corresponding Lagrangian observation models that are used to analyse various types of data sources will be developed in this thesis.

There is another important aspect of this problem area, related to the (quality and usage of) empirical data. First of all, unreliable measurements and disturbances are an unavoidable part of the raw data from traffic sensors. For instance in the Netherlands, 5-10% on average of the available data from the dual loop system (named MoniCa data) are missing or otherwise deemed unreliable (Van Lint & Hoogendoorn, 2009). Although data-assimilation methods can balance the uncertainty/noise in the observation data and in the model forecast, observations with strongly-biased and unreliable features are meaningless and useless to efficient state estimation. Secondly, multi-class traffic state estimation requires class-specific observations as input, which are not directly available from most traffic sensors. Moreover, all the three components in traffic state estimation require predefined parameters, which can be derived from empirical data. Thus, pre-processing work (see the outgoing arrow from the “sensor” box in Figure 1.1), such as data cleaning and preparation, and model-parameter generation, should be included to achieve a high quality state estimation procedure. Two examples related to the first two problems are presented in this thesis. One aims to correct biased information inherited in dual-loop systems, and the other tries to infer more (class-specific) information than those from direct observations.

1.6 Research contributions and relevance

This thesis focuses on traffic state estimation research. New approaches for estimating traffic state information are developed, and application issues have also been considered. In the following discussion, the main contributions of this research are highlighted, and distinctions are made between purely scientific contributions and practical contributions.

1.6.1 Scientific contributions

Our main theoretical and methodological contributions are listed below. These are related to developing new approaches for traffic state estimation, synthesising the state-

of-the-art in traffic state estimation, and addressing new methods for pre-processing observation-data.

- *A Lagrangian formulation of the traffic state estimation problem*

This thesis (Chapter 3 and 4) presents new methods for state estimation, by taking a “Lagrangian” perspective. This means that the traffic system models used in state estimation are formulated in Lagrangian (moving observer) coordinates instead of the traditional Eulerian (spatially-fixed) coordinates. Under a Lagrangian traffic formulation, some problems in Eulerian coordinates (e.g., the mode-switching problem) are overcome, and the Lagrangian type of sensing data (such as probe vehicle data, cell-phone data) are naturally incorporated into state estimation. In addition, the corresponding observation models in Lagrangian coordinates dealing with both the Eulerian and Lagrangian sensing data are developed. To assemble a state estimator, a real-time applicable technique, EKF, is used for data assimilation. This study (Section 3.6, 5.3, and 5.4) reveals that the Lagrangian state estimator is significantly more accurate and offers computational (it is more efficient) and theoretical benefits over the Eulerian approach.

- *Development of a multi-class state estimation approach*

In this work, a multi-class traffic state estimator has been developed (Chapter 4). This is done by equipping a Lagrangian first-order traffic flow model with a distinction between different vehicle classes (such as cars, trucks, and buses). Experiments (Sections 5.5 and 5.6) show that such a multi-class Lagrangian state estimator, based on an EKF framework, succeeds in providing class-specific state estimates on traffic networks.

- *New insights into the existing continuum multi-class traffic flow model*

Two modelling and discretisation choices for Lagrangian multi-class traffic flow models are discussed and compared in Section 4.2, which are respectively the “Piggy-back” model and the “Multi-pipe” model. The process model used in our Lagrangian state estimator applies a “Piggy-back” formulation, due to its suitability for on-line traffic state estimation.

- *Generalisation of state estimation to a network-wide level*

This work (Sections 3.3, 4.3, 5.4, and 5.6) shows that Lagrangian state estimation is scalable to a network level by implementing node models to account for network discontinuities.

- *New insights into the characteristics of model-based state estimation*

During the development and analysis of our state estimation method, new insights are gained regarding the mechanism of each component in model-based state estimation. For instance, the reasons for the improvement in including multi-class features in state estimation, are identified; how the traffic formulation affects the performance of state estimation is addressed.

- *A new taxonomy of the state-of-the-art and the state-of-the-practice in traffic state estimation*

This thesis (Chapter 2) proposes a new perspective to classify state estimation research,

with respect to two different coordinate systems that allow different mathematical formulations. Along the line of this classification, it provides a comprehensive discussion on the state-of-the-art in traffic state estimation and its applications. A large number of previous studies are reviewed in terms of traffic dynamics models, observation models and the data-assimilation methods used in state estimation.

- *New methods for data pre-processing and estimation preparation*

Preparation is needed to apply the state estimation process in the real world. New techniques for data cleaning and pre-processing are developed to overcome the shortcomings of raw data (Chapter 6). A speed-bias correction algorithm is presented to deal with inaccurate (biased) aggregated-speed input. Multi-class and multi-lane flow estimation aims to infer additional information from existing loop data for its subsequent use.

1.6.2 Practical contributions

Four main practical contributions have been identified, along with their significance to society.

- *Improvement of traffic state estimation for real-time traffic network management*

In general, technologies and models for real-time traffic network management are formulated in the conventional Eulerian coordinate system. This research provides a new Lagrangian formulated multi-class state estimator for network applications. Due to the fact that it leads to faster computation and more accurate results, it can be implemented in a real-time context and thus promotes practical model-based decision support systems in traffic networks.

- *New framework to incorporate GPS-type data*

The Lagrangian formulation is based on the perspective of moving observers. Therefore, Lagrangian state estimation provides an ideal framework for the assimilation of data from those moving observers, such as mobile phone tracking data, GPS equipped probe vehicles, etc..

- *Implication for in-car state estimation applications*

Meanwhile, this state estimation method sheds some light on in-car localised information (state estimation) systems. Individual vehicles or vehicle platoons can be treated as independent state estimation units, and concepts in this thesis can be used in vehicle-wise cooperative systems.

- *Providing practitioners with smart tools to tackle the problems with empirical data*

The data-processing techniques that we develop can be used to solve several practical problems. The speed-bias correction overcomes the speed-bias problems in empirical data. A multi-linear regression approach is used to estimate multi-class and multi-lane counts from aggregate data formats. These algorithms provide practitioners with simple but effective and efficient tools to process empirical data.

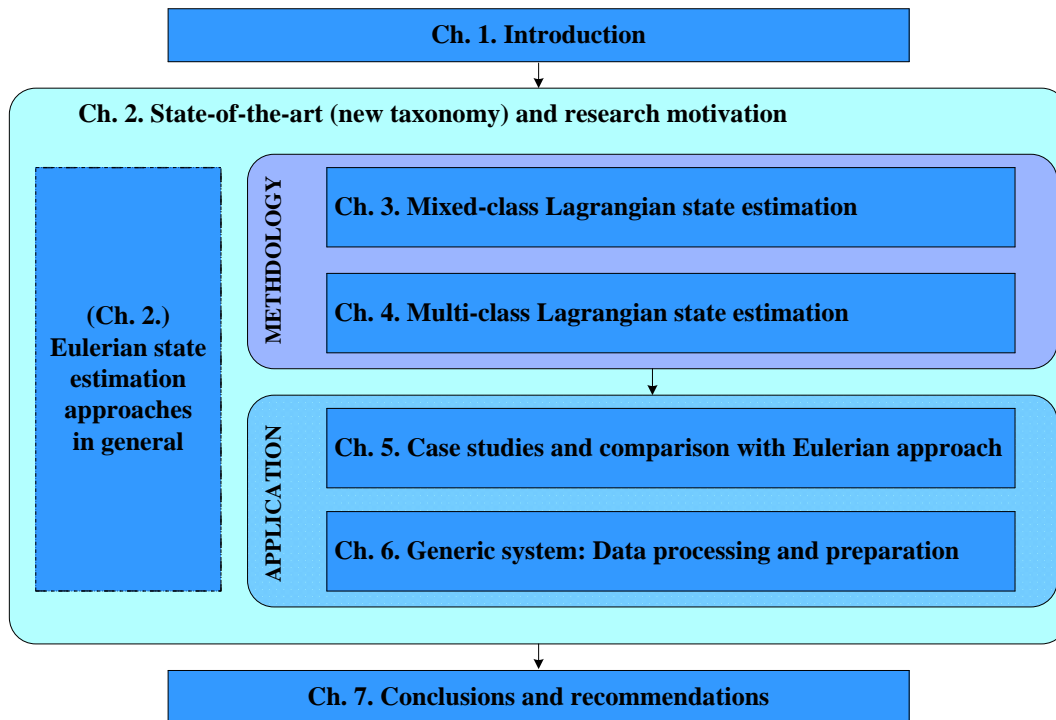


Figure 1.2: Outline of the thesis.

1.7 Outline of this thesis

The outline of this thesis is presented schematically in Figure 1.2 and discussed in more detail in this section. This thesis consists of seven chapters.

Chapter 2 first sets up a novel classification framework for traffic state estimation research with respect to two different (Eulerian and Lagrangian) observation coordinate systems. Under this framework, previous studies are classified into the related categories. Thereafter, research gaps and motivation are distinguished and proposed within the new taxonomy. The main methodological contribution, bridging these gaps, is elaborated further in Chapters 3 and 4. Furthermore, the applications related to the new method are addressed in Chapters 5 and 6. To make a distinction within this framework, the work that has previously been done relating to the Eulerian classification is presented as a comparable “dashed-block” to Chapters 3 to 6, as illustrated in Figure 1.2.

Chapter 3 presents the Lagrangian state estimation approach based on a mixed-class formulation. For the purposes of comparison, each of the three components in state estimation is elaborated for both the Lagrangian and Eulerian approaches. Their relative advantages and challenges are discussed. In addition, discontinuity modelling in Lagrangian coordinates is addressed for the extension of these techniques to traffic network modelling.

In Chapter 4, a multi-user perspective is considered in the Lagrangian state estimation approach. New applications and interpretations of the existing multi-class traffic flow

models are addressed for multi-class state estimation. The related observation models for class-specific data and node models, at a multi-class level, are developed.

In Chapter 5, the mixed-class Lagrangian state estimation model developed in Chapter 3 is tested in both synthetic and real-world data environments. In the former, FOSIM (Dijker, 2012), a microscopic simulation environment, is used to provide ground truth data to validate the Lagrangian approach at both the link level and the network level. In the latter case, the mixed-class Lagrangian approach is compared with its Eulerian counterpart based on empirical data taken from a British motorway. The multi-class Lagrangian state estimation, as presented in Chapter 4, is first verified with diverse class-specific data sources in the computational environment, and is then tested on a real freeway network (the Dutch A15) on the basis of diverse empirical data sources (aggregate loop data, individual vehicle data, trajectory data). Finally, the main findings are interpreted and summarised.

Chapter 6 deals with model applications in the real world. Before implementing a traffic state estimation procedure, preparations regarding raw data cleaning, model parameter and input generation are necessary. Methods and algorithms are developed to provide estimation inputs with accuracy. With this purpose, two examples of dealing with raw data are shown, which tackle the speed-bias problem and class-specific data inputs, respectively. These examples are based on two edited versions of published articles (Yuan et al., 2010, 2012. In Press).

The main conclusions are drawn in Chapter 7. Furthermore, the research implications for state estimation studies are highlighted, reviewing the main contributions of this research from both methodological and practical perspectives. We end the thesis by discussing a number of possible future research directions.

Chapter 2

The state-of-the-art in traffic state estimation

This chapter reviews the state-of-the-art in traffic state estimation. First of all, we establish a novel classification framework for model-based state estimation research with regards to the two different (Eulerian and Lagrangian) coordinate systems. This taxonomy allows the identification of potentially beneficial research angles. Both mixed-class and multi-class traffic descriptions/modelling are distinguished. A discussion of previous research efforts in this domain are then presented within this taxonomy. Different modelling choices are distinguished in terms of traffic process models, observation models and assimilation techniques used in model-based state estimation. The literature review addresses all the important aspects of the state estimation architecture, in order to make clear to the readers which design choices and trade-offs one needs to make in state estimation. Furthermore, gaps are identified in the proposed classification framework: whereupon, the motivation for a multi-class Lagrangian state estimation is provided accordingly.

2.1 Introduction

This thesis focuses on the estimation of traffic states, which aims to provide reliable and accurate traffic state information for real-time dynamic traffic management. The essence of traffic state estimation is to reproduce traffic conditions based on available traffic data. One class of available estimation methods does not make use of traffic flow dynamics, but relies on basic statistics and interpolation. These are referred to as data-driven methods. Another class of estimation methods relies on dynamic traffic flow models. These are referred to as model-based methods. The focus of this thesis is on the latter because it potentially provides better results than the former class.

First of all, the basic concept of model-based traffic state estimation is addressed. As mentioned in Chapter 1, model-based state estimation usually encompasses three components: (often nonlinear) dynamic traffic flow (or process) models, observation models and data-assimilation techniques. The first two components constitute the (macroscopic) traffic system models used in state estimation. These system models describe the underlying traffic dynamics and the relations between system states and observations. Based on system models, data assimilation methods (the third component) estimate the most probable traffic states. There are many data-assimilation techniques, ranging from simple techniques to more sophisticated algorithms. For instance, a simple Newtonian relaxation (nudging) method (Anthes, 1974) relaxes system models towards observations, meaning observation models are not required in performing data assimilation. The Kalman filtering method (Kalman, 1960) provides solutions to the optimal filtering problem: the best state estimates, in a recursive fashion. It is rooted in a (linear) system featuring process and observation models. As the test bed, this thesis chooses existing recursive data-assimilation techniques. This type of technique iteratively reuses one or more of the outputs as the input. This feedback typically results in either exponentially growing, decaying, or sinusoidal signal output components. In the field of transportation research, it indicates that the estimation errors by a recursive assimilation technique tend to get smaller and smaller over time.

Macroscopic (nonlinear) traffic system models can be generally cast in a discrete state-space form, which makes them suitable for recursive data-assimilation techniques. Here, it is assumed that traffic system models consist of both dynamic process models and observation models. Generally, these models can be formulated as follows:

$$\begin{aligned} \mathbf{z}_{\tau+1} &= f(\mathbf{z}_{\tau}, \mathbf{d}_{\tau}) + \mathbf{w}_{\tau} && \text{(process model)} \\ \mathbf{y}_{\tau} &= h(\mathbf{z}_{\tau}, \mathbf{d}_{\tau}) + \mathbf{u}_{\tau} && \text{(observation model)}. \end{aligned} \quad (2.1)$$

Here, the process model $f(\cdot)$ is a discrete equation (e.g., the conservation of vehicles equation) describing the evolution of the system state (e.g., vehicle density k , spacing s), and $h(\cdot)$ is a static relationship that relates observations to the system state (e.g., the fundamental diagrams relating average flow q and speed v to density k). Note that in equations (2.1), the subscript τ depicts discrete time instants. Henceforth, the state vector \mathbf{z}_{τ} typically denotes a vector of average densities (and/or speeds, spacing) on

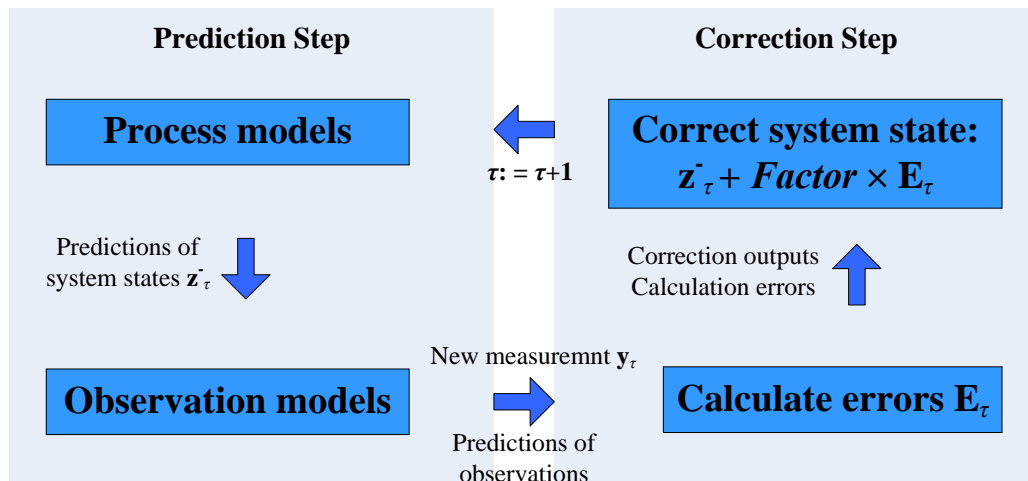


Figure 2.1: Schematic procedure of the prediction-correction data-assimilation method.

road segments (or vehicle platoons) over small time periods $[\tau, \tau + 1)$ or $(\tau - 1, \tau]$, \mathbf{y}_τ is a vector of observations (e.g., speeds, flows, vehicle spacing, headways, etc.) from fixed or vehicle-based sensors, and \mathbf{d}_τ is a vector of model inputs, including all disturbances and parameters (e.g., traffic demands, capacity constraints, link capacities, critical speeds and densities). The two white noise terms \mathbf{w}_τ and \mathbf{u}_τ in (2.1) represent errors in the process and observation models, respectively. The first term may stem from errors in the model input and parameters, and/or model mis-specification (e.g., many models necessarily simplify certain physical aspects and thus may cause errors). The error term in the observation model reflects the combined effect of modelling errors and observation errors (e.g., miscounts, equipment failure, etc.).

Based on the state-space model, recursive data-assimilation techniques aim to make an optimal estimate of the system state \mathbf{z}_τ given all observations \mathbf{y}_τ until the current time instant. To this end, these methods use an intuitive predictor-corrector structure, as shown in Figure 2.1. For each discrete time step, first a prediction of the system states \mathbf{z}_τ^- is made on the basis of the process model and the previous available estimate. Next in the correction step, this prior estimate \mathbf{z}_τ^- is corrected with an optimal weighting *Factor* (such as the Kalman gain matrix) proportional to the distance (errors \mathbf{E}_τ) between the available sensor data (\mathbf{y}_τ) and the predictions made by the observation model. The optimal weighting factor is determined in terms of minimising state estimation errors. This procedure iteratively provides state estimates at each time step.

In the following, we first present a new classification framework for model-based state estimation research. Previous research efforts are examined under the framework, which leads to the conclusion that there are several gaps. This motivation drives our main research direction.

2.2 A new classification framework for model-based traffic state estimation research

This section presents a novel classification framework for model-based traffic state estimation research, with respect to the embedded traffic process models (as the first “X” dimension) and traffic observation models (as the second “Y” dimension) in different coordinate systems, see Table 2.1. In this taxonomy table, the mixed-class and multi-class traffic descriptions are also identified, resulting in eight categories.

There are two selected coordinate systems as discussed in Chapter 1, in which traffic flow theories can be mathematically formulated: the Eulerian coordinate system, which is fixed in space; and the Lagrangian coordinate system, which moves with the traffic stream. Within the first dimension, we can, therefore, classify traffic state estimation research with respect to the system modelling in these two coordinate systems, leading to the Eulerian and the Lagrangian traffic state estimation approaches. The state estimators with Eulerian formulated traffic system (process) models are classified as Eulerian state estimators, whereas Lagrangian state estimators are embedded in Lagrangian formulated traffic system models.

Similarly, one can divide traffic observations into two main functional categories. Local traffic quantities such as aggregated traffic counts and spot speeds are observed at a fixed point in space (at a cross-section). This type of traffic measurement device, such as loop detectors, video sensors, and radar detectors, is classified as an Eulerian sensor and the related observations as Eulerian sensing (observation) data. One can also observe traffic flow characteristics moving along vehicle trajectories. These data are referred to as Lagrangian sensing (observation) data (Herrera et al., 2010). Lagrangian data can be obtained via GPS technology or any tracking devices providing position and velocity of individual vehicles. Moreover, vehicles equipped with distance sensors can even provide distance headways (spacing) between successive vehicles using infrared or radio technology. The two types of data and the relating observation models feature the second dimension in the new taxonomy.

Apart from the two dimensions, different data-assimilation techniques also form the basis for conventional classification approaches for traffic state estimation. For instance, Ou (2011) presented a classification with respect to this criterion. This thesis considers assimilation techniques as the third “Z” dimension in the new taxonomy. However, this dimension will not be visually presented in the “X-Y” dimension classification table. Instead, it will be discussed in Section 2.5, with a focus on “recursive” assimilation techniques. The new classification scheme is proposed with respect to the three components in traffic state estimation, therefore, it clarifies the design choices and trade-offs one needs to make and face in state estimation.

It is necessary to determine how previous studies fit into the taxonomy along the three classification dimensions. In the following three sections, we focus firstly on the distinction of traffic state estimation studies with respect to different traffic process mod-

Table 2.1: Classification framework for model-based traffic state estimation(TSE) research in terms of two mathematical formulations of process and observation models in two coordinate systems

			Y: Observations (Models)	
			Eulerian Sensing (Loop, camera, radar)	Lagrangian Sensing (GPS or cellphone)
Eulerian TSE	Traffic Process Models	Mixed-class Eulerian Traffic Dynamics Model		
		Multi-class Eulerian Traffic Dynamics Model		
Lag. TSE	X: Traffic Process Models	Mixed-class Lagrangian Traffic Dynamics Model		
		Multi-class Lagrangian Traffic Dynamics Model		

els, secondly with regards to different traffic data and observation models used in traffic state estimation, and finally in terms of assimilation techniques.

2.3 Choices in traffic process models

Traffic system models consist of process models and observation models. Traffic state variables in these models are traffic densities, speeds, flows and/or travel times, depending on different estimation purposes. Process models describe the evolution of these state variables. The choices of process models vary between different traffic flow theories (e.g., first-order or higher-order traffic flow theory). Most of the traffic process models (and the relating observation models) applied in traffic state estimation are formulated in Eulerian coordinates. Although diverse data-assimilation methods (the third dimension) have been used in previous studies, the traffic dynamics models used as process models are generally either Eulerian formulated mixed-class first-order or mixed-class second-order traffic flow models.

2.3.1 Eulerian formulated traffic process models

As one of the mixed-class first-order traffic flow models, the Lighthill-Whitham and Richards (LWR) model (Lighthill & Whitham, 1955; Richards, 1956) or the Cell Transmission Model (CTM, as a special case of the LWR model when the fundamental relation between flow and density is assumed to be triangular) (Daganzo, 1994, 1995a) have been widely used in this field. The LWR model is adequate to reproduce some

of the basic phenomena observed in traffic (Daganzo, 1994, 1995b,c, 2002b; Newell, 1993; Lebacque, 1996; Van Wageningen-Kessels et al., 2011a). These reproduced phenomena include conservation of vehicles, the onset and dissolution of congestion at bottlenecks, the fact that disturbances propagate over space and time in different directions as a function of the prevailing state (congested or not), and traffic anisotropy. Sun et al. (2003); Tampère & Immers (2007); Pueboobpaphan et al. (2007); Herrera et al. (2010); Van Hinsbergen et al. (2012) all used the LWR-type model for freeway traffic state estimation. Recently, a speed reformulation of the LWR model has been developed. This model describes the evolution of traffic speeds and it has been applied to better assimilate speed observations (Work et al., 2008; Chen et al., 2011; Coric et al., 2012). In these studies, a freeway is divided into spatially-fixed cells, in which traffic states are described by densities or speeds. The continuous LWR model is discretised and solved by numerical schemes. Since traffic characteristics in Eulerian coordinates might move either upstream or downstream depending on the prevailing traffic condition (mode/regime: congested or not), the mode needs to be identified. Mode identification can be done in several ways: 1) One can calculate the transition flows (fluxes) between cells by a minimum supply and demand principle (e.g., the widely applied Godunov scheme (Lebacque, 1996)), to automatically distinguish traffic regimes. 2) In (Sun et al., 2003), a switching mode model (SMM) was used with the assumption that every cell in one section only had the same mode, while the mode was determined by the comparison of the predicted density with critical density. 3) Tampère & Immers (2007) presented an implicit mode switching scheme for the CTM applied in traffic state estimation. Correct mode identification is important for accurate traffic state estimation.

There are also other first-order traffic flow models. For instance, a MARKov COmpartment Model (MARCOM) was developed by Davis & Kang (1994), adapted by Diet et al. (2010) to estimate arterial traffic densities. Similar to the CTM, the MARCOM model describes traffic flows within spatially-fixed compartments (cells), as a density-dependent birth and death process according to cell boundary flux transitions. When a vehicle makes a transition from the upstream compartment to the downstream one, a death occurs in the upstream compartment while correspondingly a birth happens in the downstream one. One advantage of this model is that more possible modes (as typical in an urban network) instead of the two modes (on freeways) can be identified. Clearly, when applying the Eulerian formulated mixed-class first-order traffic model in traffic state estimation, the mode-switching problem needs to be addressed and properly solved.

Second-order models include a second speed equation to account for the fact of gradual vehicle acceleration/deceleration, capacity drop, traffic hysteresis, oscillatory congested traffic, and so forth. They also allow to incorporate the observations of speeds and flows which are not directly related via fundamental diagrams (non-equilibrium states). With more elaborate descriptions of traffic flows, these models potentially improve the performance of traffic state estimation. Therefore, they have also been

widely used. For instance, a Payne model (Payne, 1971) was employed by Nanthawichit et al. (2003) to perform real time traffic state estimation. As one of Payne-type extensions, the METANET model (Papageorgiou et al., 1990) has been used in several state estimation or data-assimilation related studies (Wang & Papageorgiou, 2005; Hegyi et al., 2006; Mihaylova et al., 2012). Moreover, a second-order gas-kinetic-based traffic model (Treiber et al., 1999) was adopted by Ngoduy (2011) as the dynamic system model. Boel & Mihaylova (2006) developed an second-order extension of the CTM by adding probability distributions on the sending and receiving functions. This model was then implemented to perform data assimilations in (Mihaylova et al., 2007). Cheng et al. (2006) used a two-level second order model to perform speed and flow estimation. Due to more complex traffic modelling (more parameters) and highly nonlinear traffic dynamics in these second-order models compared to the first-order models, the related state estimation requires more sophisticated assimilation techniques. This aspect will be further elaborated in Section 2.5.

Apart from second-order extensions, some studies have additionally considered driver and vehicle heterogeneities in Eulerian formulated traffic flow models to improve the accuracy of the estimation. Different vehicle classes have different characteristics, such as maximum speeds, vehicle lengths, reaction times, minimum distance headways, and so forth. Multi-class models take into account this heterogeneity by distinguishing vehicle-user classes. For instance, Van Lint et al. (2008a) and Ngoduy (2008) have implemented multi-class first-order traffic flow models to estimate class-specific traffic states (densities) with a certain success. These models are based on the Eulerian formulated first-order traffic flow model under the assumption that the conservation law holds for each of the vehicle classes. Hoogendoorn (2001) has successfully applied a multi-class second-order traffic flow model (Hoogendoorn, 1999) for travel time estimation. The embedded process model distinguishes between different vehicle classes and its formulation is based on a gas-kinetic principle. Note that, since most of the multi-class and the second-(or higher-) order models are derived from the (Eulerian) mixed-class first-order traffic flow models, the applications with these models should also take into account mode-switching problems.

2.3.2 Lagrangian formulated traffic process models

Recent studies by Leclercq et al. (2007); Van Wageningen-Kessels (2013) show that the LWR model formulated in Lagrangian coordinates can be solved more efficiently and accurately than in Eulerian coordinates. In such an alternative coordinate system, traffic flow is divided into vehicle platoons, over which state variables (spacing) are described. The most favourable advantage is that traffic characteristics only move in one direction, independent of the prevailing traffic conditions. Mode identification is no longer required, and the “Godunov” scheme can in turn be simplified into an “upwind” scheme. The simplified scheme leads to “easy” numerical discretisation, and accurate simulation results. The new traffic formulation and its simpler numerical

scheme are supposed to yield benefits for state estimation. Hence, the Lagrangian formulations potentially promote accuracy of traffic state estimation, which slots into the main objective of this research.

However, few previous studies have focused on a Lagrangian form of traffic state estimation. Therefore, the characteristics of the Lagrangian traffic state estimation remain unclear to researchers. As a first attempt, we deploy a first-order traffic flow (LWR) model formulated in Lagrangian coordinates as the process model, to investigate the advantages and disadvantages of the Lagrangian traffic state estimation. Although a number of phenomena are not well or fully represented in the first-order model (Helbing, 2001; Kerner, 2009), most of which are related to the capacity drop and to traffic instability, there are still strong arguments as to why the LWR model is a valid choice for the purpose of state estimation. First of all, there is no undisputed alternative model that is able to reproduce the foregoing basic phenomena under all circumstances sufficiently. Secondly, there is the principle of parsimony: the first-order model contains less parameters than more involved alternatives (higher-order models); it is a model that is mathematically tractable, that can be analytically solved, and that provides the analyst straightforward tools to switch between two different coordinate formulations. However, note that the concept and results in this research are not only limited to the first-order traffic flow model but also can be extended to more involved (high-order and/or other gas-kinetic-based) macroscopic models.

Moreover, regarding traffic flow modelling in Lagrangian coordinates, we can also consider the driver and vehicle heterogeneity by applying a Lagrangian formulation of the multi-class first-order traffic flow model. By including heterogeneities in modelling, not only these traffic flow models are able to describe traffic flow more accurately (Bellomo & Dogbe, 2011), but also the control applications for such models can be made more elaborate (Schreiter, 2013). This type of multi-class state estimation and control is especially valuable in areas with high truck percentages.

2.4 Choices for incorporating observation models

Similar to process models, observation models can also be formulated in both coordinate systems, resulting in Eulerian and Lagrangian observation models. Traditionally traffic observations are collected by spatially-fixed Eulerian sensors, such as loop detectors, video sensors, and radar devices. They are classified as Eulerian sensing data. These data have been the dominant information sources in the field of transportation research for decades. Most of the studies discussed in the previous section apply only Eulerian sensing data based on Eulerian formulated observation models when performing state estimation. In recent years, wireless communication technologies, including GPS, cellular probe and bluetooth etc., have been increasingly used for ITS applications. These Lagrangian data have also been incorporated with Eulerian formulated observation models for traffic state estimation.

One of the first studies that succeeded in incorporating Lagrangian sensing data into the traffic state estimation was done by Nanthawichit et al. (2003), where by incorporating simulated Lagrangian data it turned out to improve estimations from the method that use only the Eulerian data. Usually, by additionally applying the Lagrangian sensing data, the performance of state estimation approaches can be substantially improved. For instance, Di et al. (2010) and Byon et al. (2010) successfully incorporated both loop and GPS data to estimate traffic states on urban arterials. Chu et al. (2005) also showed that better freeway travel time estimation can be achieved by using both Eulerian and Lagrangian data.

Increasing amounts of Lagrangian sensing data appear to be available for the transportation community. For example, traffic data from the Next Generation Simulation (NGSIM) project (FHWA, 2012) and the Mobile Millennium experiment (UCBerkeley, 2008), have been used as observation input for several CTM-based state estimation studies (Work et al., 2008; Herrera et al., 2010; Coric et al., 2012), yielding good results. These data are usually incorporated into an Eulerian formulated observation model, together with other available Eulerian data within an Eulerian state estimation framework. The main assumption in Eulerian formulated observation models is that the Lagrangian sensing data represent conditions in a spatial-temporal fixed “cell”. This might not be an appropriate approximation since Lagrangian sensors move with the traffic. An intuitive question remains: can the Lagrangian sensing data be better incorporated into a Lagrangian formulated framework? Leclercq et al. (2007); Tchakian & Verscheure (2011); Van Wageningen-Kessels (2013) recently present the Lagrangian first-order traffic formulations which can be applied for better Lagrangian sensing (GPS) data assimilation. Owing to the Lagrangian formulation, traffic flow models and observation models are described by the relation between vehicle spacing and speed. Assuming both the spatial location of the vehicle and the time instant of communicating its location and speed are available, these data are directly related to the spacing and speed (the states) of a vehicle platoon based on its spatial location at the same moment. Therefore, Lagrangian sensing data can be naturally incorporated into the Lagrangian formulated observation models without any further assumptions.

None of previous research has incorporated Lagrangian (and/or Eulerian) sensing data using Lagrangian observation models into Lagrangian formulated state estimation. Thus far, we can determine how previous studies fit into the proposed taxonomy along the first and the second dimensions. This leads to several research gaps, as indicated in Table 2.2. Due to the advantages of Lagrangian formulation in terms of both traffic simulations and incorporating traffic data, this thesis will perform a Lagrangian-formulated traffic state estimation research to fill these research gaps. In the next section, we investigate how the existing research accommodates the third classification dimension, in the context of traffic state estimation.

Table 2.2: Classification of previous studies based on the proposed framework for traffic state estimation research in terms of first two (X-Y) dimensions (1st and Higher indicate methods using the 1st- and the higher-order traffic flow models, respectively.)

		Y: Observations (Models)	
		Eulerian Sensing (Loop, camera, radar)	Lagrangian Sensing (GPS or cellphone)
X: Traffic Process Models	Mixed-class	1st: Sun et al. (2003); Chu et al. (2005); Tampère & Immers (2007); Pueboobpaphan et al. (2007); Work et al. (2008); Herrera et al. (2010); Di et al. (2010); Byon et al. (2010); Chen et al. (2011); Coric et al. (2012); Van Hinsbergen et al. (2012)	1st: Chu et al. (2005); Work et al. (2008); Herrera et al. (2010); Di et al. (2010); Byon et al. (2010); Coric et al. (2012)
	Eulerian Model	Higher: Nanthawichit et al. (2003); Wang & Papageorgiou (2005); Hegyi et al. (2006); Mihaylova et al. (2007); Ngoduy (2011); Mihaylova et al. (2012)	Higher: Nanthawichit et al. (2003); Cheng et al. (2006)
	Multi-class Eulerian Model	1st: Van Lint et al. (2008a); Ngoduy (2008) Higher: Hoogendoorn (2001)	
	Mixed-class Lagrangian Model	(this research)	(this research)
	Multi-class Lagrangian Model	(this research)	(this research)

2.5 Data-assimilation techniques

Previous studies based on different data-assimilation techniques (the “Z” dimension) are discussed and reviewed in this section. We restrict the discussion on the recursive assimilation techniques. In the following, we first briefly introduce these techniques and then review their applications in traffic state estimation.

2.5.1 Overview of recursive assimilation techniques

There are several well-known data-assimilation methods on the basis of the recursive structure in Figure 2.1. One of them is the Kalman Filter (KF) (Kalman, 1960; Haykin, 2001), rooted in a state-space formulation of linear dynamic systems, provides a recursive solution to the linear optimal filtering problem. However, most models in the real world are nonlinear, and there exist more advanced “relatives” to the KF to deal with that, such as the Extended Kalman Filter (EKF) (Jazwinsky, 1970), the Unscented Kalman Filter (UKF) (Julier & Uhlmann, 1997) and the Ensemble Kalman Filter (EnSKF) (Evensen, 2007). The fundamental idea of EKF is to linearise the state-space system model and apply the linear KF procedure. Contrary to the EKF, the UKF does not require a first-order linearisation of the nonlinear system. Instead, it computes the Gaussian error variables by using a deterministic sequential Monte Carlo sampling approach. Similarly, the EnSKF uses Monte Carlo or ensemble integration method instead of a linearisation procedure. Moreover, the Particle Filter (PF) (Gordon et al., 1993; Doucet et al., 2001) performs state estimation also within a Monte Carlo sampling framework.

The differences between the data-assimilation methods (EKF, UKF, EnSKF, and PF) lie in their assumptions related to the process and observation models and to the error (noise) terms. In the EKF method, it is assumed that both process and observation models are continuously differentiable functions, which can be locally (i.e., around the current state) approximated by a first-order Taylor-series expansion (linearisation), and in addition that the noise terms are Gaussian and independent over time. The benefit of these assumptions is that only the mean and error covariance of the traffic state are required to be calculated to approximate an optimal estimate of the system state using the classic KF equations. The UKF and EnSKF relax the first assumption that system models need to be differentiable, as it does not require local linearisation to calculate the posterior state distribution, whereas the PF additionally relaxes the Gaussian assumption. One can consider the UKF and EnSKF are affiliated to the PF family as special cases. As the derivative-free techniques, they do not require the first-order approximation regardless if the system model is differentiable. Therefore, they have been used in several state estimation studies combining with highly nonlinear second-order models (Hegyí et al., 2006; Mihaylova et al., 2007; Ngoduy, 2011; Mihaylova et al., 2012) and/or multi-class models (Ngoduy, 2008). The disadvantage of these more advanced filters is increased computational cost since they require a significant

number of process and observation model instances to run in parallel, whereas the EKF is a one-shot procedure. Table 2.3 overviews the main characteristics of different assimilation methods and the related existing applications. This table also presents the requirements (assumptions) of filtering techniques for process and observation models.

2.5.2 Motivation for applying the EKF

There are two classical arguments for using more involved and theoretically superior assimilation methods, instead of the EKF in traffic state estimation. These are related to the two main assumptions in the EKF: 1) that the first-order Taylor-series expansion is not a reasonable approximation for highly nonlinear process, and 2) that the Gaussian assumption for the noise terms in both process and observation models is invalid. Although realistic knowledge of the noise distributions is not utilised in the EKF, previous studies (Wang & Papageorgiou, 2005; Hegyi et al., 2006; Van Lint et al., 2008a) show that the Gaussian assumptions is not a bad choice. This first statement is justified particularly around capacity, where due to the upwind/downwind numerical scheme (mode-switching), the (Eulerian) traffic flow model (process model) indeed is highly nonlinear (Mihaylova et al., 2007; Ngoduy, 2008). Therefore, the first-order approximation for Eulerian traffic models might be invalid at transition states, and thus it might lead to a conflict with the requirements of the EKF. As a result, the EKF estimates around capacity point may quickly diverge from the true state since the linearisation may result in corrections with the “wrong” sign (i.e., the estimator may infer congested traffic while in reality traffic is flowing freely). It will in such a case drive the estimated state away from the true state rather than towards the true state. One solution to the problem is to better predict the mode (free-flowing or congested), for example, 1) by using a superior numerical scheme (e.g., Godunov scheme (Lebacque, 1996)), this alternative is sensitive to the critical density (speed/flow) values; 2) by using Lagrangian data (Herrera et al., 2010). However, as noted in (Herrera et al., 2010), “Since observations are sparse in time and space, however, the mode identification is a challenging task,” this makes traffic state estimation in Eulerian coordinates on the basis of the EKF a cumbersome exercise.

There are, nonetheless, also strong arguments in favour of the EKF approach for traffic state estimation. Firstly, it is computationally much more efficient than the UKF, EnKF and PF approaches and hence more suitable for real-time estimation in large traffic networks (Wang & Papageorgiou, 2005; Van Lint et al., 2008a), particularly with scalable cell-based methods (Van Hinsbergen et al., 2012). Secondly, although traffic propagation may be highly nonlinear around capacity, there are both theoretical and empirical evidences that in strictly free-flowing and strictly congested conditions the linear approximation is quite good on average (Daganzo, 1994; Newell, 1993; Treiber & Helbing, 2002; Van Lint & Hoogendoorn, 2009). A straightforward solution to the mode-switching problem is to apply the Lagrangian formulation of the traffic flow model. As discussed in Section 2.3, solving the Lagrangian kinematic wave model

Table 2.3: Overview of the main characteristics of different recursive assimilation techniques and the relating applications

		Z-dimension: (Recursive) Assimilation techniques				
		KF	EKF	UKF	EnKF	PF
Characteristics (Pros and Cons)		One-shot analytical procedure; easy to implement; only applicable for linear systems	One-shot analytical procedure; easy to implement; inaccurate linearisation for highly nonlinear problems	Deterministic sequential sampling; derivative free; computational cost scales with state dimension	Monte Carlo ensemble sampling; derivative free; heuristic choice of sample size	Sequential Monte Carlo sampling; theoretically superior for nonlinear estimation; large number of samples required
Requirements for process models (Assumptions in X-dimension)		Gaussian errors; linear	Gaussian errors; continuously differentiable	Gaussian errors	Gaussian errors	-
Requirements for observation models (Assumptions in Y-dimension)		Gaussian errors; linear	Gaussian errors; continuously differentiable	Gaussian errors	Gaussian errors	-
Existing applications		Nanthawichit et al. (2003); Sun et al. (2003); Chu et al. (2005); Herrera et al. (2010); Di et al. (2010); Byon et al. (2010)	Hoogendoorn (2001); Wang & Papageorgiou (2005); Hegyi et al. (2006); Tampère & Immers (2007); Van Lint et al. (2008a); Van Hinsbergen et al. (2012)	Hegyi et al. (2006); Pueboobpaphan et al. (2007); Ngoduy (2008, 2011)	Work et al. (2008); Coric et al. (2012)	Cheng et al. (2006); Mihaylova et al. (2007, 2012); Chen et al. (2011)

does not require mode switching, in a sense it “reduces” the nonlinearity level. The linearisation of the Lagrangian system model will always lead to EKF corrections with the “correct” sign. The Lagrangian formulated traffic flow model is continuously differentiable and thus it can be well represented by a linearised model, which entirely accommodates one of the EKF requirements for system models (see Table 2.3). The Lagrangian formulation can potentially improve the performance of an EKF-based state estimation. In this thesis, the EKF is selected to validate the Lagrangian state estimation method in a real-time context. Note that, the concept of Lagrangian formulation is not restricted to the EKF technique, but can furthermore apply to other data-assimilation techniques (e.g., UKF, EnsKF and PF).

2.6 Research direction and main challenges

In the previous sections, we have concluded that the new Lagrangian formulation is potentially suitable and beneficial for traffic state estimation, in terms of traffic simulations, incorporating Lagrangian sensing data, and the application of the data-assimilation (EKF) method. This thesis aims to investigate the properties of Lagrangian traffic state estimation, and thus to fill the four main blanks in the current traffic state estimation research (see, Table 2.2).

This thesis will develop a model-based first-order multi-class (and mixed-class) state estimation approach in Lagrangian coordinates, based on an EKF framework, where both Eulerian and Lagrangian sensing data can be incorporated. Note that, the idea of Lagrangian formulation is not restricted to first-order traffic flow model with the EKF technique, but can furthermore apply to other data-assimilation techniques combining with more involved macroscopic traffic flow models. The main challenges to this research are twofold: first of all, appropriate Lagrangian observation models need to be formulated to incorporate Eulerian sensing data into Lagrangian state estimation; and second of all, proper modelling network discontinuities is required, because the network essentially moves with respect to the coordinate system in the Lagrangian formulation. These issues will be elaborated in the rest of the thesis.

2.7 Summary

This thesis aims to provide reliable and accurate traffic state estimation for real-time dynamic traffic management. The performance of traffic state estimation can be improved in several aspects. As one aspect, people can develop more advanced filtering techniques to relax limiting methodological assumptions (e.g., linear system models, Gaussian-errors), in order to perform more sophisticated data assimilation. This thesis focuses on improving the foundation of traffic state estimation, by selecting more appropriate traffic system models: Lagrangian formulated traffic system models, while

maintaining those filtering assumptions (that is, Lagrangian first-order system models can be well represented by linear models). A new classification framework is proposed with respect to traffic process and observation models in two different coordinate systems, and with a “Z” dimension along assimilation techniques. With the taxonomy, we cannot only characterise available filtering approaches, but also identify potentially beneficial research angles. The identified missing gaps in the current research based on this taxonomy indicate that there are potentially more proper traffic system models formulated in Lagrangian coordinates to solve existing problems in state estimation for further improvement. Therefore, the multi-class Lagrangian formulated traffic state estimation becomes the core-value of this research and it will be elaborated in the remaining chapters.

Chapter 3

Model-based mixed-class state estimation in Lagrangian coordinates

Within the classification framework presented in Chapter 2, we develop a new approach of Lagrangian traffic state estimation. In this chapter, the approach is formulated at a mixed-class level. To apply this approach in real traffic networks, node models in Lagrangian coordinates are developed for network discontinuities. For the purpose of comparison, the corresponding Eulerian formulation is introduced in this chapter as well. Finally, these two formulations are compared in terms of theoretical and practical advantages. The extension of the Lagrangian formulation to a multi-class level will be given in the next chapter.

3.1 Introduction

The previous chapter presented a new classification framework for traffic state estimation studies, which is based on the two coordinate systems (Eulerian and Lagrangian). We have motivated why we need to develop a new Lagrangian formulation. Therefore, this chapter presents such a traffic state estimation approach in Lagrangian coordinates at a mixed-class level. Specifically, the three main components in traffic state estimation, namely dynamic traffic models, observation models and data assimilation techniques (the Extended Kalman Filter (EKF) in this thesis) are elaborated in both the Eulerian and Lagrangian formulations in the following sections. Similarly, the node models for network discontinuities in both formulations are also presented. In Section 3.6, the Eulerian and Lagrangian formulations are compared in terms of theoretical and practical advantages.

3.2 Process models: Eulerian and Lagrangian formulations of mixed-class first-order traffic model

First recall the nonlinear state-space form for macroscopic traffic system models presented in Chapter 2:

$$\begin{aligned} \mathbf{z}_{\tau+1} &= f(\mathbf{z}_{\tau}, \mathbf{d}_{\tau}) + \mathbf{w}_{\tau} && \text{(process model: } f(\cdot)) \\ \mathbf{y}_{\tau} &= h(\mathbf{z}_{\tau}, \mathbf{d}_{\tau}) + \mathbf{u}_{\tau} && \text{(observation model: } h(\cdot)). \end{aligned} \quad (3.1)$$

Here, \mathbf{z}_{τ} denotes system states, \mathbf{y}_{τ} denotes sensor observations, and \mathbf{d}_{τ} depicts model inputs. \mathbf{w}_{τ} and \mathbf{u}_{τ} are respectively the two white noise terms in the process and observation models.

3.2.1 Mixed-class Eulerian formulated process model

In most traffic applications, traffic system models (both $f(\cdot)$ and $h(\cdot)$) are formulated and discretised in Eulerian (space x , time t) coordinates. In this thesis, the first-order traffic flow model (also known as the Lighthill-Whitham and Richards (LWR) model (Lighthill & Whitham, 1955; Richards, 1956)) is employed as the process model to describe the evolution of traffic state variables. The conventional Eulerian formulation reads

$$\frac{\partial k}{\partial t} + \frac{\partial q}{\partial x} = 0, \quad \text{(Eulerian conservation of vehicles equation), (3.2)}$$

$$q = \begin{cases} Q(k) \\ kv \end{cases} \quad \text{with } v = V(k), \quad \text{(Eulerian fundamental diagrams: } Q \text{ and } V). \quad (3.3)$$

Here, k denotes vehicle density, and q denotes average flow (flux). Equation (3.2) does not distinguish different vehicle classes. Instead, it depicts the average performance

of traffic flows in that the change of traffic density over time should be equal to the change of traffic flow over space.

For simulation purposes (e.g., state estimation), the continuum model needs to be discretised. This yields the following state-space equations for a single cell i ($i = 1, 2, \dots$) of length Δx_i :

$$k_{\tau+1}^i = k_{\tau}^i + \frac{\Delta t}{\Delta x_i} (q_{\tau}^{i-1 \rightarrow i} - q_{\tau}^{i \rightarrow i+1}), \quad (3.4)$$

Here, τ depicts discrete time instants, and Δt denotes the period length. There are different (explicit) numerical schemes to calculate the numerical flux $q_{\tau}^{i \rightarrow i+1}$ between cells i and $i + 1$ in equation (3.4). The most widely used scheme is the Godunov (minimum supply (S) demand (D)) scheme (Lebacque, 1996), which reads

$$q_{\tau}^{i \rightarrow i+1} = \min(D_{\tau}^i, S_{\tau}^{i+1}) \quad (3.5)$$

with

$$D_{\tau}^i = \begin{cases} Q(k_{\tau}^i) & k_{\tau}^i < k_{C,\tau}^i \\ C_{\tau}^i & \text{otherwise} \end{cases} \quad (3.6)$$

$$S_{\tau}^{i+1} = \begin{cases} C_{\tau}^{i+1} & k_{\tau}^{i+1} < k_{C,\tau}^{i+1} \\ Q(k_{\tau}^{i+1}) & \text{otherwise.} \end{cases} \quad (3.7)$$

Here, $k_{C,\tau}^i$ denotes the critical density at which the flow reaches capacity C_{τ}^i in cell i .

Courant-Friedrichs-Lewy's (CFL) condition (Courant et al., 1967) defines the stability and convergence domain of the numerical method, which reads

$$\Delta x_i \geq \max_k |\partial(Q(k))/\partial k| \Delta t, \quad \text{for all } i. \quad (3.8)$$

This constraint ensures that the perturbations in the flow cannot travel faster than the free-flow speed.

3.2.2 Mixed-class Lagrangian formulated process model

In this thesis, we adopt a Lagrangian formulation of traffic flow models. The continuous representation of the mixed-class LWR model in Lagrangian (vehicle number n , time t) coordinates reads (Leclercq et al., 2007)

$$\frac{\partial s}{\partial t} + \frac{\partial v}{\partial n} = 0, \quad (\text{Lagrangian conservation equation}), \quad (3.9)$$

$$v = V^*(s), \quad (\text{Lagrangian fundamental diagram: } V^*). \quad (3.10)$$

Equation (3.9) states that the change in spacing s (which equals $1/k$ (m/veh.), i.e., the reciprocal of density) of a platoon of vehicles over time t is equal to the change in speed v over this platoon. This equation describes this undisputed principle. Suppose

that there are two vehicles, driving along the road. If the first vehicle drives faster than the second one, the spacing (or distance) between them would increase, and vice versa. For an intuitive explanation and the mathematical details of this Lagrangian formulation, we refer to (Van Wageningen-Kessels et al., 2009b, 2010a). Note that the vehicle number n decreases in the driving direction and this variable is not necessarily an integer. Analogous to the Eulerian case, the fundamental diagram V^* in Lagrangian coordinates (3.10) expresses speed v as a function of spacing s (Figure 3.1). For instance, the fundamental diagram proposed by Smulders (1989) in Lagrangian coordinates reads as follows:

$$v = V^*(s) = \begin{cases} v_{\text{free}} - s_{\text{cri}}(v_{\text{free}} - v_{\text{cri}})/s, & \text{if } s \geq s_{\text{cri}}, \\ v_{\text{cri}}(s - s_{\text{jam}})/(s_{\text{cri}} - s_{\text{jam}}), & \text{otherwise,} \end{cases} \quad (3.11)$$

with free-flow speed represented by v_{free} , critical speed by v_{cri} , critical spacing by s_{cri} , and jam spacing by s_{jam} , respectively. This relation is an invertible and piecewise function.

As an essential step in some of the data-assimilation methods (e.g., the EKF, which will be introduced in Section 3.5), it is required to conduct the first-order approximation (linearisation) of both the process and observation models. Therefore, the embedded fundamental relations need to be differentiable. However, there is a nondifferentiable point in piecewise functions (as shown in Figures 3.1(a) and 3.1(b)). This problem can be easily remedied by using a smooth approximation at the non-differentiable point of the fundamental diagram. Alternatively, a smooth v - s relation can be used in the model, for example, the fundamental diagram of Greenshields (1934):

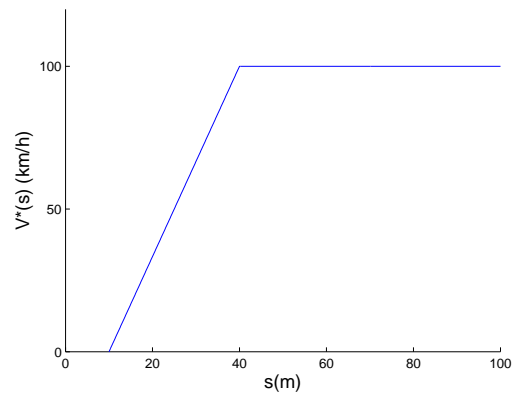
$$v = V^*(s) = v_{\text{free}}(1 - s_{\text{jam}}/s), \quad (3.12)$$

or a hyperbolic function (Bando et al., 1995) (Figure 3.1(c)):

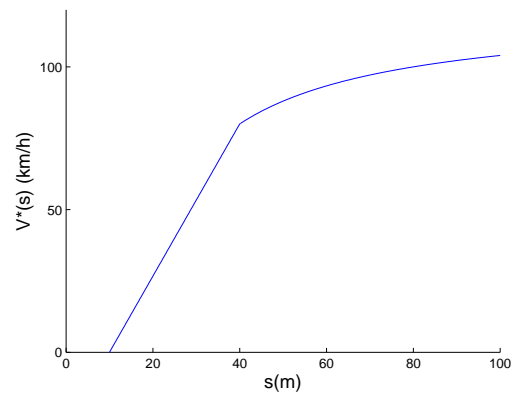
$$v = V^*(s) = A \tanh((s - B)/C - 2) + \tanh 2 \quad (3.13)$$

where A , B and C are the parameters which need to be calibrated based on specific road stretches. The Greenshields' relation deviates from the current reality on free-ways (unrealistic speed description in congested states, and unreasonable critical spacing/density), and thus it is not used in this thesis. The latter case is not used either, since the parameters in expression (3.13) are not directly related to physical meanings of traffic phenomena (e.g., free-flow speed or critical speed), which is difficult to estimate from empirical data. In the following, we will use a smooth version of Smulders' fundamental diagram as illustrated in Figure 3.1(b). The precise method to fix the nondifferentiable point will be given when calculating the Jacobian matrix in the EKF method in Section 3.5.

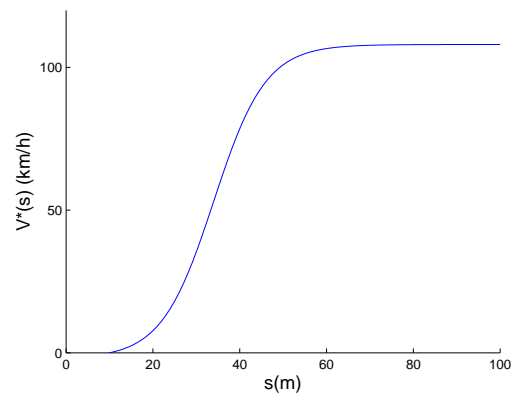
In Lagrangian coordinates, traffic characteristics only move upstream (in the direction of increasing n , opposite of the driving direction), which implies that vehicles only react to vehicles in front of them (traffic anisotropy). The (favorable) consequence is



(a) Daganzo



(b) Smulders



(c) Smoothing

Figure 3.1: Lagrangian fundamental diagrams for Daganzo (a) (Daganzo, 1994), Smulders (b) (Smulders, 1989), and Smoothing (c) (Bando et al., 1995). Note that all the fundamental relations in the Lagrangian form can be transformed into the equivalent Eulerian relations.

that the Godunov scheme in Lagrangian coordinates simplifies to an upwind scheme. The process model (the sending and receiving functions are smooth) can thus be easily linearised. For the convenience of state estimation, the process model (3.9) is discretised using an upwind and explicit time-stepping scheme, as shown in equation (3.14). More specifically, traffic flow on a freeway stretch is divided into vehicle platoons of size Δn with index i ($i \in N$), and time-steps τ are of size Δt , i.e.,

$$s_{\tau+1}^i = s_{\tau}^i - \frac{\Delta t}{\Delta n_i} (V^*(s_{\tau}^i) - V^*(s_{\tau}^{i-1})). \quad (3.14)$$

A platoon size of one would imply a microscopic simulation (on one-lane segments).

The CFL condition for the Lagrangian model becomes

$$\Delta n_i \geq \max_s |\partial(V^*(s))/\partial s| \Delta t, \quad \text{for all } i. \quad (3.15)$$

This constraint ensures that the variation in vehicle number at one location within one time step of size Δt cannot be larger than the size of the platoon itself.

The vectorised Lagrangian mixed-class formulation for freeway stretches then reads

$$\begin{cases} \mathbf{s}_{\tau+1} &= \mathbf{s}_{\tau} - \frac{\Delta t}{\Delta \mathbf{n}} (\mathbf{v}_{\tau} - \mathbf{v}_{\tau}^{\text{front}}), \\ \mathbf{v}_{\tau} &= V^*(\mathbf{s}_{\tau}), \quad \mathbf{v}_{\tau} = \mathbf{q}_{\tau} \cdot \mathbf{s}_{\tau}. \end{cases} \quad (3.16)$$

All boldface variables in equations (3.16) represent vectors of the related quantities (e.g., $\mathbf{s}_{\tau} = [\dots, s_{\tau}^i, \dots]^T$, $i \in N$), where $\mathbf{v}_{\tau}^{\text{front}}$ denotes the speed of related successive vehicle platoon, and \mathbf{q}_{τ} denotes the vectorised average flow. The above equations constitute the nonlinear state-space traffic system model in Lagrangian coordinates. There additionally exist two noise terms, respectively representing errors in the process and observation models. In the EKF technique, the noise terms are assumed to be Gaussian and independent over time. Therefore, independent zero-mean Gaussian noise terms $\boldsymbol{\sigma}_{\tau}$ and \mathbf{r}_{τ} need to be added in Lagrangian process and observation equations to suit the application of the EKF, yielding

$$\begin{aligned} \mathbf{s}_{\tau+1} &= f(\mathbf{s}_{\tau}, \mathbf{d}_{\tau}) + \boldsymbol{\sigma}_{\tau} && \text{(process equation)} \\ \mathbf{y}_{\tau} &= h(\mathbf{s}_{\tau}, \mathbf{d}_{\tau}) + \mathbf{r}_{\tau} && \text{(observation equation)}. \end{aligned} \quad (3.17)$$

3.3 Modelling network discontinuities in the Lagrangian formulation

Real-time traffic management and control require state estimators that work on networks. To this end, traffic state estimators need to be completed with boundary conditions and node models (network discontinuities). In Lagrangian coordinates, however, implementing boundary conditions and node models is not straightforward. This is because the coordinates are moving with the vehicles in this formulation, while the boundaries and nodes are fixed in space. As a result, in the Lagrangian formulation

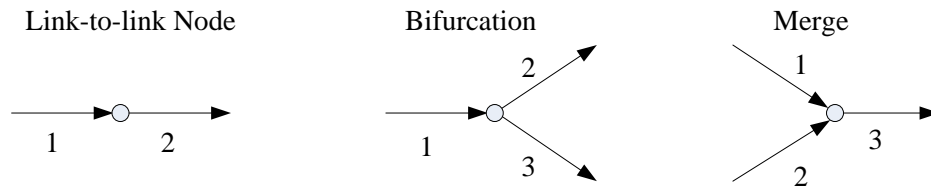


Figure 3.2: Nodes with locally numbered links.

the network (including discontinuities) essentially moves with respect to the coordinate system, which yields both modelling and computational difficulties. This section shows how to include such discontinuities, that is, sources (on-ramps, in-flows at entries) and sinks (off-ramps, out-flows at exits) in the Lagrangian formulation by means of a node model. This node model can be seen as an extension or a supplement to the foregoing traffic process model, which is used for the prediction step in the data-assimilation (e.g., the EKF) method (see Section 3.5).

3.3.1 Eulerian formulated node models

The essence of the LWR model is the conservation of vehicles equation. In case of on-ramps (merges) or off-ramps (bifurcations) on a freeway, node models are needed to deal with in-flows from sources and out-flows into sinks. In the network description used in our study, three node types are considered: link-to-link nodes, bifurcations and merges (see Figure 3.2). Note that, other more complex nodes, such as crossings, are not implemented yet. Here, we first recall traffic dynamics (fluxes) across these nodes in Eulerian coordinates (Van Lint et al., 2008b), then develop Lagrangian node models.

In case of a link-to-link node, it is a simple interface between two network links (links 1 and 2 in figure 3.2). The flux over the node is calculated by:

$$q^{1 \rightarrow 2} = \min(D^1, S^2). \quad (3.18)$$

This implementation at the link-to-link nodes is similar to the (Godunov) numerical scheme, which is used to calculate fluxes between two neighbouring cells, as shown in equations (3.5, 3.6 and 3.7). In many cases, a link-to-link node (also a bifurcation or merge node) describes a spatial or temporal discontinuity, such as a lane drop/expansion, curves and gradients, or a change in speed limits. Therefore, this kind of nodes also indicates a change in fundamental relations. As a result, the fundamental relations for network links can be formulated as space and time dependent. The demand D and supply S of each link can be determined via the related fundamental relations.

At bifurcation nodes, turn fractions γ for both outgoing links (i.e., links 2 and 3 in Figure 3.2) are defined, which depict the distribution of the total flow over the outgoing links. The demand on link 1 to links 2 and 3 can be determined using these turn

fractions:

$$\begin{aligned} D^{1 \rightarrow 2} &= \gamma D^1 \\ D^{1 \rightarrow 3} &= (1 - \gamma) D^1. \end{aligned} \quad (3.19)$$

Then we have the minimum supply-demand scheme for a bifurcation node:

$$\begin{aligned} q^{1 \rightarrow 2} &= \min(D^{1 \rightarrow 2}, S^2) \\ q^{1 \rightarrow 3} &= \min(D^{1 \rightarrow 3}, S^3). \end{aligned} \quad (3.20)$$

Whenever the supply at one (or both) of the outgoing links is not large enough, the corresponding excess vehicles will stay in the incoming link for the next time step.

At merge nodes, the supply of the outgoing link 3 is distributed proportional to the number of lanes L of each incoming link (also refer to Figure 3.2):

$$\begin{aligned} S_*^{1 \rightarrow 3} &= \frac{L_1}{L_1 + L_2} S^3 \\ S_*^{2 \rightarrow 3} &= \frac{L_2}{L_1 + L_2} S^3. \end{aligned} \quad (3.21)$$

Note that, this is just a specific and arbitrary choice. It is possible to consider realistic behaviours at nodes (e.g., priority and politeness) with more complicated modelling. Thereafter, if there is any supply left, it can be used by either incoming link:

$$\begin{aligned} S^{1 \rightarrow 3} &= S_*^{1 \rightarrow 3} + \max(0, S_*^{2 \rightarrow 3} - D^2) \\ S^{2 \rightarrow 3} &= S_*^{2 \rightarrow 3} + \max(0, S_*^{1 \rightarrow 3} - D^1). \end{aligned} \quad (3.22)$$

Similar to equation (3.20), we have the minimum supply-demand scheme for merges:

$$\begin{aligned} q^{1 \rightarrow 2} &= \min(D^1, S^{1 \rightarrow 3}) \\ q^{2 \rightarrow 3} &= \min(D^2, S^{2 \rightarrow 3}) \end{aligned} \quad (3.23)$$

Whenever the demand at one (or both) of the incoming links is larger than the assigned supply, the corresponding excess vehicles will stay in the related incoming link for the next time step.

3.3.2 Lagrangian formulated node models

To model sources and sinks in the Lagrangian formulation, Lagrangian node models are developed (see also in (Van Wageningen-Kessels et al., 2011c)). This contribution shows how sources (sinks) in Lagrangian coordinates effectively generate (remove) vehicle platoons into (from) the flow, and as a result change the spacing of platoons which pass the source (sink) location. This node model is based on the analytical expression for sources and sinks in the Lagrangian formulation of the LWR model (equation (3.24)), which is derived from its Eulerian counterpart:

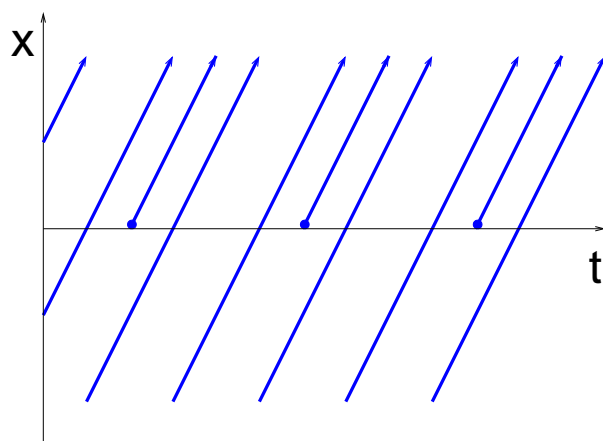
$$\frac{\partial s}{\partial t} + \frac{\partial v}{\partial n} = -s^2 \psi(x(n), t). \quad (3.24)$$

Here, the definitions of s , v , t and n are consistent with those in equations (3.9) and (3.10), the term $\psi(x(n), t)$ ($:= \alpha(x(n), t) - \beta(x(n), t)$) denotes the time and space dependent combined influence from source $\alpha(x(n), t)$ and sink $\beta(x(n), t)$, which is usually related to on-ramps or off-ramps. Note that, the source term α times the length of the influencing platoons at on-ramps results in inflows (vehicles per time unit). Similarly, from the sink term β , the related off-ramp outflows can be derived.

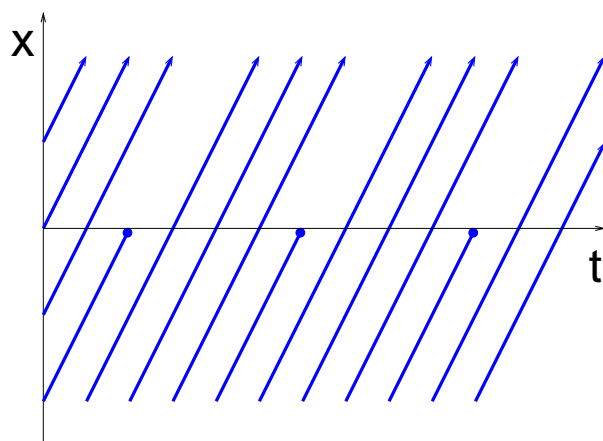
As discussed above, node models in Eulerian coordinates are based on the minimum demand-supply principle. The fundamental relations should also be formulated depending on spatiotemporal discontinuities. These main principles can be adapted for the node implementation in Lagrangian coordinates. However, we notice that the resulting fluxes from the Eulerian node models can be used in the Eulerian process models explicitly, whereas these fluxes cannot be directly incorporated into the Lagrangian process model (flow/flux quantities are implicit). The difficult part for the Lagrangian implementation mainly relates to the discretisation. When implementing sources and sinks into a Lagrangian simulation model, there are several options to choose from the vehicle discretisation and the time discretisation. In our application, whole vehicle groups (platoons) are only added to (or removed from) the flow at the beginning of a new time step. As an example in Figure 3.3(a), full vehicle groups are either continued on the main road, or added to the flow at the source location; similarly in Figure 3.3(b), full vehicle groups are either continued on the main road, or removed from the flow at the sink location. In terms of the time discretization, Figure 3.3(c) shows the new conditions applying from the beginning of the next time step.

More specifically, in the on-ramp (source) case, two consecutive platoons are considered at the corresponding time step, in which the upstream one covers the related on-ramp. This upstream platoon will be influenced by an entry platoon. The space of this platoon will be shared with the entry platoon. While the downstream one will not be influenced by the entry platoon, as vehicles (platoons) only react to vehicles (platoons) which are in front of them. A similar concept applies to the off-ramp (sink) case. When a platoon is leaving from the mainstream, the space of this platoon will be taken by its successive platoon. This ensures there is no vacuum space between two platoons while keeping the mainstream flow conserved.

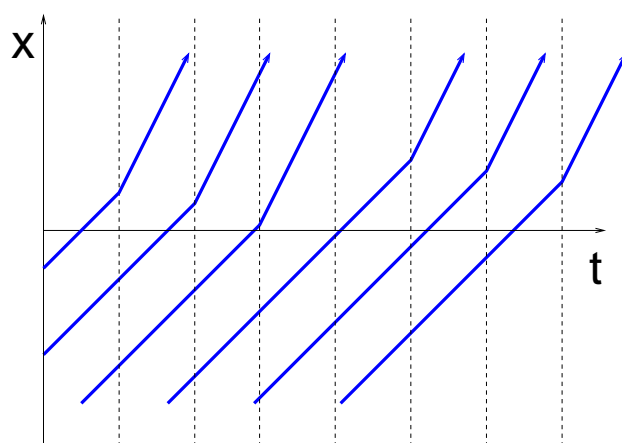
The presented node model is straightforward for implementations and leads to plausible and accurate simulation results (see (Van Wageningen-Kessels et al., 2011c)). Additionally, it is suitable for incorporating the current available observations at nodes (on-ramps and off-ramps). More importantly, it enables the extension to a network-wide traffic state estimation. More advanced modelling and its discretisation over nodes have been presented in (Van Wageningen-Kessels et al., 2011b), by considering more complex network discontinuity situations. For instance, in that contribution, priority ratios at mergers are introduced to define the flow distribution from incoming links. Thus, spillbacks onto the main road upstream of ramps and/or onto on-ramps can be modelled when congestion sets in. At bifurcations, turn fractions are defined for outgoing links, the assigned amount of traffic that cannot access to off-ramps will stay



(a) Vehicle discretisation for merges: full vehicle groups are added to the flow.



(b) Vehicle discretisation for bifurcations: full vehicle groups are removed from the flow.



(c) Time discretisation: new conditions apply at the beginning of the first time step after the vehicle group reached the node.

Figure 3.3: Vehicle and time discretisation approaches in this thesis (plotted lines denote vehicle group trajectories).

in the main road. In this thesis, on-ramp sources and off-ramp sinks are considered as a simple example of network discontinuities to test the concept of network-wide state estimation. In that, full priority is given to ramp flows.

The boundaries of traffic networks are mainly referred to as the origins and destinations. The Lagrangian boundary conditions are similar to the Eulerian case, which are based on the minimum supply-demand principle. At origin boundaries, the in-flows should be predefined. At out-flow boundaries, either a homogeneous Von Neumann condition or a Dirichlet condition can be applied (Van Wageningen-Kessels et al., 2010b). The former suggests that out-flow is always possible at exit boundary: no congestion forms at the exits. The latter is used when influence from downstream boundaries is modeled.

In our application, the in-flow from both upstream boundaries (origins) and on-ramps, and the out-flow to off-ramps (or turn fraction) need to be known or estimated; on the basis of which new platoons are added and out-flow platoons are removed. The node models effectively change the spacing of influencing platoons which pass the source or sink location, after that the spacing of all platoons are updated based on the process model (conservation law). Note that, the platoon size cannot be too large that covers two discontinuity locations (e.g., one on-ramp and one off-ramp). On the other hand, small platoon size will lead to a gain in accuracy for simulation and modelling discontinuities, but this also results in high computation time. One has to trade off between numerical accuracy and computational cost.

As a final remark, the proposed node models take effect mainly at the prediction step in recursive assimilation techniques (e.g., the EKF, see Section 3.5), whereas the correction procedure of the state estimation remains the same. Therefore, traffic characteristics (traffic states, errors) can be easily linearised at nodes.

3.4 Observation models for mixed-class Lagrangian formulation

Observation models relate observations, collected using traffic sensors (e.g., inductance loops, cameras or probe vehicles), to system state variables (i.e., vehicle spacing in the Lagrangian formulation). As presented in Chapter 2, two main (functional) data categories are identified, namely Eulerian sensing data and Lagrangian sensing data. The Eulerian fundamental diagrams in equation (3.3) provide a straightforward observation model for Eulerian sensing data, and also for Lagrangian sensing data (Herrera et al., 2010). In the former case, mixed-class flows and (harmonic) mean speeds measured by local detection equipments can be directly related to spatially-fixed discretised cells in Eulerian coordinates. In the latter, the Lagrangian data (e.g., floating car data (FCD)) is considered to be representative for the local traffic conditions (of corresponding discretised spatial cells) around the location and time of data transmission.

In Lagrangian coordinates, the Lagrangian fundamental diagram (see equation (3.10)) provides a natural observation model for Lagrangian sensing data (FCD). Assuming both the spatial location of the vehicle and the time instant of communicating its location and speed are available, these data are directly related to the spacing and speed of a vehicle platoon based on its spatial location at the same moment. This relaxes the assumption used in the Eulerian formulation that FCD represent conditions in a spatial-temporal fixed “cell”. If such a probe vehicle would also transmit the detected spacing (distance headway) between its successor and predecessor, then this piece of data can provide a direct measurement of the system state s . The observation equation, in that case, simplifies to an identity relation. The following equations are applied when trajectory-based data are available, where the former one is applied to incorporate spacing observations and the latter one is for speed observations:

$$\begin{aligned} \mathbf{y}_\tau &= \mathbf{s}_\tau^{\text{obs}} = \mathbf{s}_\tau + \mathbf{r}_\tau && \text{(spacing observation - linear equation)} \\ \mathbf{y}_\tau &= \mathbf{v}_\tau^{\text{obs}} = V^*(\mathbf{s}_\tau) + \mathbf{r}_\tau && \text{(speed observation).} \end{aligned} \quad (3.25)$$

Here, $\mathbf{s}_\tau^{\text{obs}}$ and $\mathbf{v}_\tau^{\text{obs}}$ respectively denote spacing and speed observations at a time instant τ . If both observations are available, they can be incorporated at the same time by using a combined observation equation, which reads

$$\mathbf{y}_\tau = \begin{bmatrix} \mathbf{s}_\tau^{\text{obs}} \\ \mathbf{v}_\tau^{\text{obs}} \end{bmatrix} = \begin{bmatrix} \mathbf{s}_\tau \\ V^*(\mathbf{s}_\tau) \end{bmatrix} + \mathbf{r}_\tau. \quad (3.26)$$

These (combined) observation equations are further used to calculate the Jacobian matrix with respect to the system state, and thus the weighting factor (see Section 3.5). With both the weighting factor and observations, correction can be performed.

However, incorporating Eulerian sensing data in the Lagrangian formulation is not straightforward. Consider a dual-loop detector installed at a certain cross-section. At the end of each measurement interval ΔT (e.g., 60 s), the detector will provide aggregated (mixed-class) flow q_{det} and speed v_{det} . The question is how to relate these spot speed or flow data to the observation of the state for a particular platoon. Intuitively, one might use flow to determine the number of vehicles (platoons) that pass the detector during one measurement interval. Then the observation speeds of those vehicles (platoons) are equal to the detected speed. This observation model seems to be plausible, assuming traffic conditions are stationary during measurement intervals. However, this relation can not be justified when shockwaves pass the detector. Instead, we will propose an alternative method, where only the spot speed observations are used.

According to the kinematic wave theory, in Eulerian coordinates, the traffic characteristics, e.g., flow, density and, speed, move along the road. The propagation speed of traffic information perturbations is equal to the derivative of the fundamental (q - k) diagram. There are two types of wave speeds, namely the congested wave speed v_{cong} and the free-flow wave speed v_{free} . In the case of a Daganzo fundamental diagram, the wave speeds are both constant. If the prevailing traffic condition around a detector is congested, traffic information measured by the detector will travel upstream with the

congested wave speed (e.g., see the slanted dashed line with a slope of v_{cong} in Figure 3.4). The same logic applies to the free-flow traffic state; the only difference is that the travel direction is downstream. The travel distance of traffic information can be easily calculated by the characteristic wave speed v_{chr} multiplied by the elapsed time ΔT , yielding

$$L_{\text{inf}} = v_{\text{chr}} \cdot \Delta T \quad (3.27)$$

where the characteristic wave speed v_{chr} is either equal to v_{free} or v_{cong} , depending on the prevailing traffic condition. The distance L_{inf} demarcates the “information influence” area of the detector (one example of the congestion case is shown in Figure 3.4(b)). Note that if there is an overlap between a congested influence area of a downstream detector and a free-flow influence area of an upstream detector, the congested influencing area “overrules” the free-flow part. The reason is that vehicles are constrained in their driving behaviour by the lowest prevailing speed.

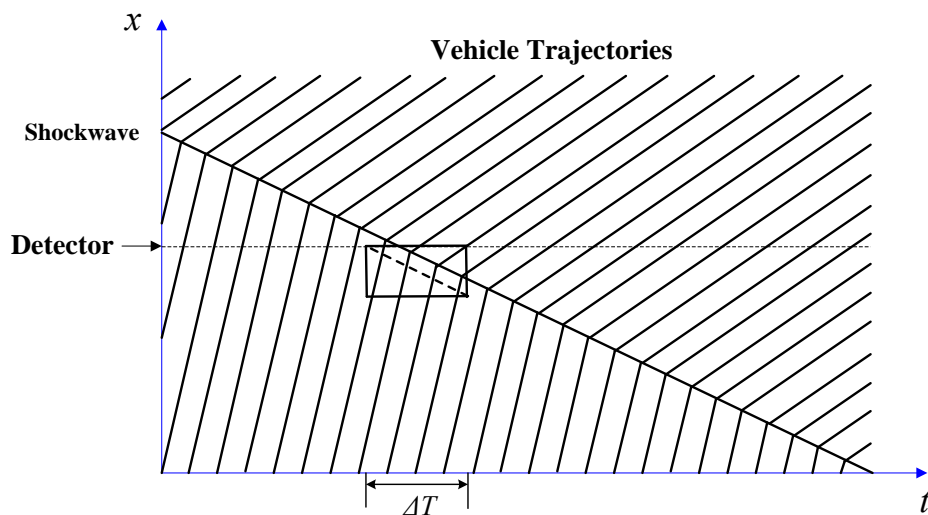
Based on the spatial positions (x^i) of all vehicle platoons and the influence areas of a loop detector, the platoons, of which the vehicle spacing needs to be corrected, can be located. The assumption is that the observed speeds are equal to the harmonic mean of each individual speed within the measurement interval ΔT . This yields

$$y_{\tau}^i = v_{\tau,H}^j, x_{\tau}^i \in L_{\tau,\text{inf}}^j \quad (3.28)$$

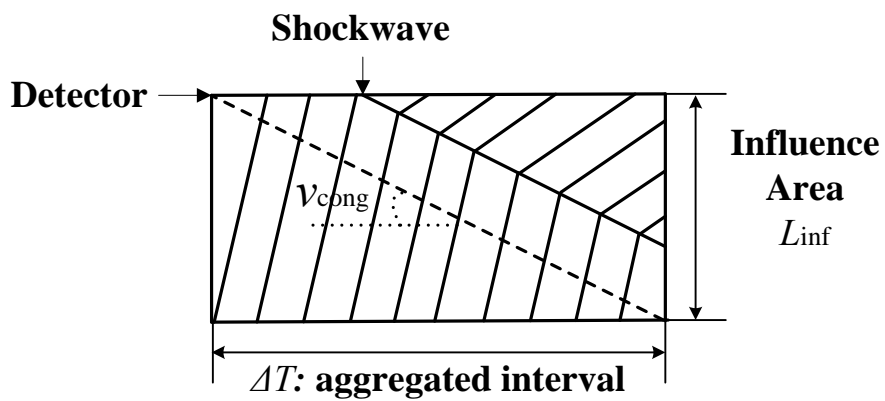
where y_{τ}^i denotes the (speed) observation of a platoon i at the τ th interval, and x_{τ}^i defines the position of this platoon. $v_{\tau,H}^j$ denotes the harmonic mean speed from a detector j at the same interval and j is the index of detectors.

In case of a homogeneous traffic situation as shown in Figure 3.4(a), where the shock-wave speed is equal to the congested wave speed, the equivalence of this approximation can be easily proven. According to Edie’s definition (Edie, 1965) of traffic variables for a space-time region, the speed at a cross-section (where the time period of space-time region is infinitesimally small) is the harmonic mean of all individual speeds within the interval. This harmonic mean speed is always equal to the space-mean speed within the influence area. In this observation model, there are two key components: the characteristic wave speeds, and the harmonic mean speed observations. Both the free-flow wave speed and the congested wave speed are treated as constant values in our application, since the linear approximation is reasonable (Daganzo, 1994; Newell, 1993). Alternatively, these two values can be dynamically estimated in real time via image processing techniques (Schreiter et al., 2010b) or even via the fundamental diagrams. When loop detectors provide arithmetic mean speeds instead of harmonic mean speeds, one has to firstly correct the raw detection speed, based on some speed-bias correction algorithms, for example, in (Soriguera & Robusté, 2011) or a method proposed in Chapter 6.

As a final remark, the assumption of homogeneity (same wave speeds) does not always hold, which implies this approximation of the observation speeds might contain errors.



(a) Typical shockwave propagation



(b) Zooming in on the rectangle

Figure 3.4: Influence area (L_{inf}) of a detector (in congested state). The detector is indicated by a horizontal dashed line. v_{cong} denotes the congested wave speed. The curved lines represent vehicle trajectories.

Nonetheless, the proposed observation model provides a straightforward way to relate spot speed data to the Lagrangian system state (spacing) via Lagrangian fundamental diagrams.

3.5 Mixed-class Lagrangian traffic state estimation based on the Extended Kalman Filter

For data assimilation, a relatively simple but online applicable technique, the Extended Kalman Filter (EKF) is chosen to validate the concept for a Lagrangian formulation of traffic state estimation. Note that, the Lagrangian approach can also be generalised to be incorporated into other data-assimilation methods, such as the foregoing Unscented Kalman Filtering (UKF) or the Particle Filtering (PF) technique.

The Extended Kalman Filter (EKF), which is rooted in the state-space formulation of nonlinear dynamic systems, provides a recursive solution to the optimal filtering problem. The basic idea of the EKF is to linearise the state-space traffic system model and apply the linear Kalman Filter procedure (Kalman, 1960). In a traffic system, suppose that an observation \mathbf{y}_τ has been made at a time instant τ , the EKF state estimator uses this observation to update the estimate of the unknown traffic state \mathbf{s}_τ , minimising the error covariance matrix \mathbf{P}_τ . On the basis of the state-space model (3.17), the mixed-class EKF state estimator in Lagrangian coordinates can be formulated. Note that the same (EKF) framework can be applied to the Eulerian formulation of the mixed-class traffic flow model, or any other analytical traffic models in a discrete state-space form (e.g., equation (3.1)).

Just as the standard Kalman filter, the EKF comprises two main steps: (1) a prediction step (for which the nonlinear process model is used), and (2) a correction step (for which the linearised system is used). Process models are used in the prediction step whereas both process and observation models are used in the correction step. According to the classification principle in Chapter 2, the state estimators using Eulerian formulated models (e.g., equations (3.2) and (3.3)) are classified as Eulerian state estimators. Whereas Lagrangian state estimators are embedded in Lagrangian formulated traffic system models (e.g., equations (3.9) and (3.10)).

In the prediction step, for each time step, the system state ($\hat{\mathbf{s}}_\tau^-$) is predicted with the process model. This predicted state serves as a prior state estimate, therefore, the best estimate before using any data is given by

$$\hat{\mathbf{s}}_\tau^- = f(\hat{\mathbf{s}}_{\tau-1}, \mathbf{d}_{\tau-1}). \quad (3.29)$$

Besides a prior for the mean, a prior for the error covariance is also computed by:

$$\hat{\mathbf{P}}_\tau^- = \mathbf{F}_{\tau,\tau-1} \hat{\mathbf{P}}_{\tau-1} \mathbf{F}_{\tau,\tau-1}^T + \mathbf{Q}_\tau \quad (3.30)$$

where \mathbf{Q}_τ represents the covariance matrix associated with the Gaussian noise term σ_τ in equation (3.17). The Jacobian matrix

$$\mathbf{F}_{\tau,\tau-1} = \left. \frac{\partial f}{\partial \mathbf{s}} \right|_{\mathbf{s}=\hat{\mathbf{s}}_{\tau-1}} \quad (3.31)$$

is the partial derivative of $f(\cdot)$ with respect to the system state \mathbf{s} , which is evaluated at $\hat{\mathbf{s}}_{\tau-1}$.

One of the disadvantages of implementing the EKF is, that the linearisation may not be possible when system (process) models are so complex and nondifferentiable. With the mixed-class traffic system models in both formulations, we can still apply the EKF for the on-line state estimation. Note that when a piecewise fundamental relation (e.g., equation (3.11)) is used instead of a smooth relation (e.g., equations (3.12) and (3.13)), the derivative with respect to the state of a particular vehicle platoon at the nondifferentiable point (capacity point) can also be provided by differentiation on an estimated smooth function through this point. This function connects two discontinuous branches and fulfills the (zero- and first-order) continuity condition. For instance, a third-order polynomial function ($v = V^*(s) = as^3 + bs^2 + cs + d$, with four deterministic parameters a, b, c and d) can be used. Essentially this smooth function provides a made-up value at the nondifferentiable point to make both the process and observation equations differentiable so that the EKF can be implemented. Although there is some approximation error at this point, the derivative values in the v - s relation maintain the “same” sign, which is an advantage over the similar application to a flow-density relation (with both positive and negative signs) in the Eulerian case.

In the correction step, the predictions of mean and covariance are corrected based on the observations obtained by traffic sensors. The Kalman gain determines the optimal weight put on both the model-predicted state and observation input, it is defined as

$$\mathbf{K}_\tau = \frac{\hat{\mathbf{P}}_\tau^- \mathbf{H}_\tau^T}{\mathbf{H}_\tau \hat{\mathbf{P}}_\tau^- \mathbf{H}_\tau^T + \mathbf{R}_\tau} \quad (3.32)$$

where \mathbf{R}_τ depicts the covariance matrix of the Gaussian noise term \mathbf{r}_τ in equation (3.17). The Jacobian matrix

$$\mathbf{H}_\tau = \left. \frac{\partial h}{\partial \mathbf{s}} \right|_{\mathbf{s}=\hat{\mathbf{s}}_\tau^-} \quad (3.33)$$

is the partial derivative of the observation model $h(\cdot)$ with respect to the system state \mathbf{s} , which is evaluated at $\hat{\mathbf{s}}_\tau^-$. The matrices \mathbf{Q} and \mathbf{R} reflect the uncertainty in the process and observation equation, and influence the weights that put on the model and observations, respectively. The state estimate $\hat{\mathbf{s}}_\tau$ and error covariance $\hat{\mathbf{P}}_\tau$ are updated as follows:

$$\hat{\mathbf{s}}_\tau = \hat{\mathbf{s}}_\tau^- + \mathbf{K}_\tau (\mathbf{y}_\tau - h(\hat{\mathbf{s}}_\tau^-)), \quad (3.34)$$

$$\hat{\mathbf{P}}_\tau = (\mathbf{I} - \mathbf{K}_\tau \mathbf{H}_\tau) \hat{\mathbf{P}}_\tau^-. \quad (3.35)$$

Table 3.1: Extended Kalman Filtering algorithm (Pseudo code)

<p>- Set $\tau = 0$, initialise \mathbf{s}_0, \mathbf{P}_0, \mathbf{Q}, \mathbf{R} and other model variables.</p> <p>for $\tau = 1 : T$ (for each time step)</p> <p>- Prediction step</p> <p>Calculate:</p> <p>a prior state estimate $\hat{\mathbf{s}}_{\tau}^{-}$: equation (3.29)</p> <p>a prior state estimate $\hat{\mathbf{P}}_{\tau}^{-}$: equation (3.30)</p> <p>- Correction step</p> <p>Calculate:</p> <p>Kalman gain \mathbf{K}_{τ}: equation (3.32)</p> <p>a posterior state estimate $\hat{\mathbf{s}}_{\tau}$: equation (3.34)</p> <p>a posterior state estimate $\hat{\mathbf{P}}_{\tau}$: equation (3.35)</p> <p>end</p>
--

Here, \mathbf{I} denotes the identity matrix. Equation (3.34) depicts the correction on the average spacing of each platoon based on the sensor observations. More information and details about the EKF can be found, for example, in (Haykin, 2001).

Table 3.1 schematically illustrates the entire procedure of the EKF model-based traffic state estimation. First, traffic system models are formulated in a state-space form. Then all the system states (\mathbf{s}_0), error covariance (\mathbf{P}_0), and predefine model variables (e.g., \mathbf{Q} and \mathbf{R}) are initialised. For each discrete time step, first a prediction of the system states ($\hat{\mathbf{s}}_{\tau}^{-}$) is made on the basis of the process model and the last available estimate ($\hat{\mathbf{s}}_{\tau-1}$). Meanwhile, a prior estimate of error covariance ($\hat{\mathbf{P}}_{\tau}^{-}$) is calculated. Next in the correction step, the prior estimate ($\hat{\mathbf{s}}_{\tau}^{-}$) is corrected with the so-called Kalman gain (\mathbf{K}_{τ}) proportional to the distance between the available sensor data (\mathbf{y}_{τ}) and the predictions made by the observation model ($h(\hat{\mathbf{s}}_{\tau}^{-})$). The prior error covariance ($\hat{\mathbf{P}}_{\tau}^{-}$), which is used in the \mathbf{K}_{τ} calculation, is again updated with this Kalman gain. This assimilation technique will converge since the error covariance (\mathbf{P}) tends to get smaller and smaller over time.

Note that when Eulerian sensing data (e.g., dual-loop detection) are incorporated as observations, a fully-recursive correction timing scheme is applied in the correction step of EKF. In this scheme, correction is performed in every prediction step (intermediate time step) within each measurement interval. The detailed method and its advantages can be found in (Schreiter et al., 2010a).

Due to the “non-mode-switching” (mode: congestion or free flow) numerical scheme (upwind scheme), the implementation of the EKF in Lagrangian coordinates is more straightforward than in the Eulerian case. In the latter case, the derivative in equation (3.31) (which is based on equation (3.33)) depends on the mode (congested or free flowing) of the considered cell and its up- and downstream neighboring cells, which

implies considering eight (2^3) different cases (two modes at three locations). In the Lagrangian case, the derivative (equation (3.31)) depends solely on the considered platoon and the one downstream of it (so four cases). Thus, the upwind scheme is computationally more efficient than the mode-switching Godunov scheme. Moreover, this numerical scheme for the Lagrangian model also improves the data-assimilation method, in that the linear approximation of the traffic system model near capacity is much better than in the Eulerian model. The derivative of the Eulerian ($q-k$) fundamental diagram shows a sudden sign change around capacity. Due to this mode-switching (between congestion and free flowing), the error in the Eulerian case may lead to EKF corrections with the “wrong” sign. Whereas in the Lagrangian case an error may still occur, but this error is guaranteed to pertain to the magnitude of the correction only.

The proposed state estimator can be applied to freeway networks. We first consider a typical freeway stretch without on/off ramps. In contrast to spatially-fixed road segments as discretised in Eulerian Cell Transmission Model (Daganzo, 1994), the traffic unit in Lagrangian applications is a platoon with a certain number of vehicles. The traffic flow on freeways is divided into “ platoons of vehicles” of size Δn . At every time interval Δt , the spacing of vehicles in each platoon is updated based on the Lagrangian conservation law. Then, the data-assimilation framework is applied using any available observations.

In the Lagrangian model, the coordinates (observers) travel along with the vehicles as they move through space. In a fixed-length road stretch, the platoons that are in the front of the whole vehicle platoon group will pass the downstream boundary of the road stretch (destination) after a certain time period. Hence, the states related to those leading platoons that are beyond the destination boundary will no longer be taken into account in the system state \mathbf{s}_τ . Furthermore, the gap between the upstream boundary and the tail of the platoon group needs to be filled by adding new vehicle platoons. On the basis of the (origin) inflow, new platoons can be added. At each time step of the system evolution, the system state vector \mathbf{s}_τ as well as the related error covariance matrix \mathbf{P}_τ need to be adjusted with respect to the freeway geometrical constraint.

3.6 Advantages of Lagrangian formulation for traffic state estimation

As already presented in (Leclercq et al., 2007; Van Wageningen-Kessels et al., 2009b, 2010a), the first-order traffic flow model can be formulated and solved more efficiently and accurately in Lagrangian coordinates than in Eulerian coordinates. The simplified “upwind” scheme leads to “easy” numerical discretisation, and accurate simulation results.

In the proposed state estimation approach, the first-order Lagrangian traffic flow model is used as the process model. Therefore, the advantages with numerical simulations remain in the state estimation process. The main challenge in the Lagrangian formulation

is to derive a slightly more complex observation model in case of Eulerian sensing data. Due to the “non-mode-switching” (upwind) numerical scheme, implementing the EKF in Lagrangian coordinates is more straightforward than in Eulerian coordinates. Moreover, this numerical scheme for the Lagrangian model also promotes the application of the data-assimilation (EKF) method, in that the linear approximation of the traffic system model near capacity is much better than in the Eulerian model. Due to mode-switching between congestion and free flowing, the error in the Eulerian case may lead to EKF corrections with the “wrong” sign. In contrast in the Lagrangian case, an error may still occur, but this error is guaranteed to pertain to the magnitude of the correction only (no sign change). Therefore, a Lagrangian formulation of an EKF-based traffic state estimation potentially delivers better estimation results. Chapter 5 will prove this concept by means of several experimental studies.

3.7 Summary and discussion

In this chapter, the mixed-class Lagrangian traffic state estimation model has been derived. Firstly, a state-space traffic flow model in Lagrangian coordinates is applied as the process model to describe the evolution of traffic states, where vehicle spacing is featured as the system state. Secondly, an improved (differentiable) Smulders’ fundamental relation is employed to describe the relations between traffic state variables, and applied as a natural observation model for incorporating Lagrangian sensing data. For Eulerian sensing data, a new observation model is developed additionally, based on the kinematic wave theory. As discussed in Chapter 2, to enable real-time filtering for online applications, the EKF is applied since it is more efficient than the other computational-expensive methods, such as UKF and PF. Meanwhile, it works well with non-mode-switching traffic systems. Hence, the EKF further combines the Lagrangian formulated traffic system model to perform state estimation, in a predictor-corrector structure. Note that, the advantage (benefit) of the Lagrangian approach can also apply to other data-assimilation methods. Furthermore, the newly developed node models enable the extension of the Lagrangian traffic state estimation to a network level. Theoretically, the Lagrangian formulation offers computational benefits for traffic state estimation over the Eulerian formulation. It potentially delivers more accurate estimation results.

In the next chapter, we will continue extending this approach to a multi-class level, whereupon the multi-class traffic system models, observation models, and network discontinuity node models will be discussed in detail. The model validation and application will be done in Chapter 5.

Chapter 4

Model-based multi-class state estimation in Lagrangian coordinates

In the previous chapter, a mixed-class state estimator was presented in terms of the three foregoing components. In that formulation, only one vehicle-user class was identified. In this chapter, the Lagrangian state estimator is extended with a multi-class distinction. Here, the methodology is firstly presented, and the model application and experimental study are further addressed in Chapter 5.

4.1 Introduction

In this chapter, we extend the mixed-class state estimator to the multi-class level, where the same modelling steps are followed as the mixed-class case, namely the process model, the observation models and the data-assimilation method. Firstly, two modelling and discretisation approaches for the multi-class Lagrangian traffic flow model are discussed. They are named as the “Piggy-back” model which is based on (only) one coordinate system, and the “Multi-pipe” model which is based on several independent coordinate systems for each of the vehicle classes (Section 4.2). The process model in the Lagrangian state estimator applies the “Piggy-back” formulation. To deal with network discontinuities in traffic state estimation, mixed-class node models in Lagrangian coordinates are generalised to a multi-class level based on the “Piggy-back” formulation (Section 4.3). Moreover, the Lagrangian observation models are developed (in Section 4.4) to incorporate different data sources, which facilitate a class-specific description of traffic flow. They are further used in the EKF-based state estimation (Section 4.5). Regarding the data assimilation, we discuss why the “Piggy-back” formulation is an appropriate choice.

4.2 Multi-class Lagrangian traffic flow models: continuum forms and different discretisation approaches

As the first component in the multi-class Lagrangian traffic state estimation, a Lagrangian formulation of the multi-class kinematic wave model is required as the process model. However, the formulation and the discretisation of the Lagrangian multi-class kinematic wave model can be done in different ways. This section focuses on two modelling and discretisation approaches of the Lagrangian multi-class first-order traffic flow model and discusses the related pros and cons.

4.2.1 Eulerian formulated multi-class models

Both mixed-class and multi-class traffic flow models can be formulated in either Eulerian coordinates or Lagrangian coordinates. We first recall the formulation of multi-class models in Eulerian coordinates (Van Lint et al., 2008b). It is based on the conservation of vehicle equation (3.2), stating that vehicles of all classes are conserved. The conservation equation describes that the change of traffic density k over time t should be equal to the change of traffic flow q over space x . This equation holds for each user class u . Additionally each user class has its class-specific fundamental diagram (Q_u or V_u), which relates the average class-specific flow (or speed) to the total (effective) density k_{tot} . The latter term is a weighted summation over all class-specific densities with respect to the passenger car equivalent (PCE: η) values. The complete model

formulation is given by

$$\frac{\partial k_u}{\partial t} + \frac{\partial q_u}{\partial x} = 0, \quad \forall u: \text{ class-specific conservation equation,} \quad (4.1)$$

$$q_u = Q_u(k_{\text{tot}}), \quad \forall u: \text{ class-specific fundamental relation } (q-k), \quad (4.2)$$

$$v_u = V_u(k_{\text{tot}}), \quad \forall u: \text{ class-specific fundamental relation } (v-k), \quad (4.3)$$

$$k_{\text{tot}} = \sum_u \eta_u k_u, \quad \text{total effective density.} \quad (4.4)$$

Here, the subscript u is a class-specific index. The class-specific PCE value η_u translates class-specific flows (e.g., trucks) into equivalent reference class flows (e.g., passenger cars). For example, a truck occupies more space than a passenger car and thus $\eta_{\text{truck}} > \eta_{\text{car}}$. These PCE values can be either constant or dynamically dependent on the total effective density. In the latter case, [Van Lint et al. \(2008b\)](#) propose that PCE can be dynamically calculated by

$$\eta_u = \frac{L_u + T_u V_u(k_{\text{tot}})}{L_{\text{car}} + T_{\text{car}} V_{\text{car}}(k_{\text{tot}})}. \quad (4.5)$$

Here, L_u denotes the class-specific gross stopping distance (average vehicle length), and T_u denotes class-specific reaction time (minimum time headway), which both increase for larger vehicle classes. The dynamic nature of the PCE value typically captures the ratio of the effective-occupied road lengths between one class- u vehicle and one passenger car, taking into account the prevailing traffic conditions (by speeds). [Van Lint et al. \(2008b\)](#) also propose a discretisation of this multi-class model in Eulerian coordinates. In fact, it is an adaptation of the minimum supply-demand method for multi-class models, including nodes.

4.2.2 Lagrangian formulated multi-class models

“Piggy-back” formulation

Now let us discuss two formulations in Lagrangian coordinates. For different applications, the formulations of the Lagrangian multi-class kinematic wave model can be addressed differently. In ([Van Wageningen-Kessels et al., 2010a](#)), a Lagrangian formulation of the multi-class kinematic wave model has been proposed. This multi-class derivation is partly based on the mixed-class Lagrangian conservation equation (3.9) for the reference class, and partly based on the conservation of vehicles law for the other classes. Each user class has its own fundamental relation (V_u^*). The related equations are defined as follows:

$$\frac{Ds_1}{Dt} + \frac{\partial v_1}{\partial n} = 0, \quad \text{conservation for class 1,} \quad (4.6)$$

$$\frac{Ds_u}{Dt} + \frac{s_u}{s_1} \frac{\partial v_u}{\partial n} + \frac{v_1 - v_u}{s_1} \frac{\partial s_u}{\partial n} = 0, \quad \text{conservation for other classes } u (\neq 1), \quad (4.7)$$

$$v_u = V_u^*(s_{\text{tot}}), \quad \forall u: \text{ class-specific fundamental relation,} \quad (4.8)$$

$$s_{\text{tot}} = \frac{1}{\sum_u \eta_u / s_u}, \quad \text{total (effective) vehicle spacing.} \quad (4.9)$$

Here, $\frac{D}{Dt} = \frac{\partial}{\partial t} + v_1 \frac{\partial}{\partial x}$ is the Lagrangian time derivative, which represents the total change as seen by an observer that is moving with traffic flows. The average vehicle spacing s is the reciprocal of the traffic density k . The class-specific equilibrium speed v_u is expressed as a function of the effective spacing s_{tot} , which is calculated based on the class-specific spacing s_u and PCE value η_u . The class-specific PCE value is defined similarly as the equation (4.5), taking the form:

$$\eta_u = \frac{L_u + T_u V_u^*(s_{\text{tot}})}{L_{\text{car}} + T_{\text{car}} V_{\text{car}}^*(s_{\text{tot}})}. \quad (4.10)$$

Note that vehicles are numbered in the opposite driving direction. In this formulation, n denotes the vehicle/platoon number of user-class 1, while the numbering for vehicles of other user-classes does not have actual meanings. Since this Lagrangian formulation is a continuum model, n can take any real value. For readability reasons, in the following we will refer to class 1 as passenger cars and assume that there is only one other class, namely trucks. The related analysis can be easily generalised to more classes.

For simulation applications, such as representing process models in traffic state estimations and other computer implementations of the model, the continuum model needs to be discretised and solved numerically. Based on the proposed multi-class formulation (equations (4.6) and (4.7)), the discretisation of vehicle number is based on user-class 1 only, since the spacing of all other classes is related to class 1 through equation (4.6). The vehicles from other user-classes are essentially travelling along with platoons of class 1. This formulation and the related discretisation are referred to as the ‘‘Piggy-back’’ method further below, since each platoon of class 1 of size Δn carries with it a certain number of other class vehicles while traversing over the network. Note that the numbers of other class vehicles within each defined platoon are not necessarily the same, depending on the relation of the spacing values between the other classes and the class 1. Figure 4.1 illustrates the vehicle discretisation in this formulation at a time instant. Since information in Lagrangian coordinates only travels in the opposite driving direction, a straightforward upwind numerical scheme is applied to solve the equations. For the time discretisation, both implicit and explicit time-stepping schemes can be used. The related comparison and discussion between these two methods have been presented in (Van Wageningen-Kessels et al., 2009a). Here, an explicit time-stepping scheme is used, because it is easy to calculate and thus suitable for online applications.

‘‘Multi-pipe’’ formulation

These discretisation choices lead to the following discretised equations. Equation (4.11) defines the state (vehicle spacing) evolution of user class 1 (cars), whereas equa-

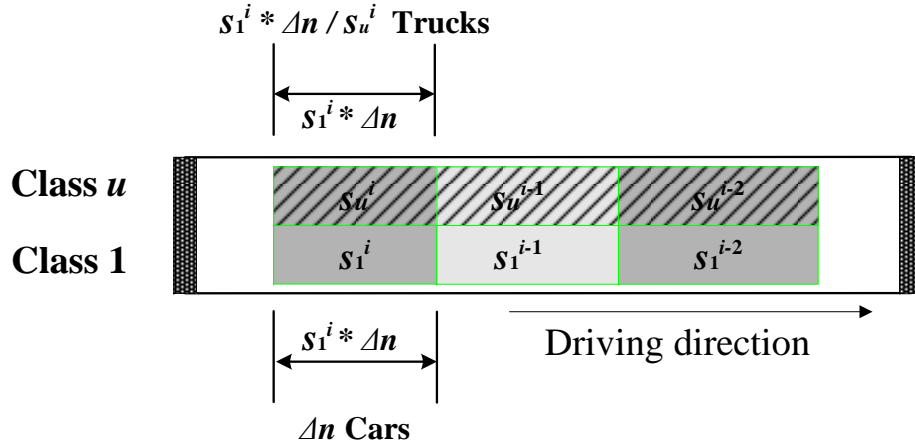


Figure 4.1: Vehicle discretisation at a time instant in the Piggy-back formulation (two-class case). The rectangular cell denotes vehicle platoon.

tion (4.12) calculates the states for other user classes (e.g., trucks) on the basis thereof:

$$s_{1,\tau+1}^i = s_{1,\tau}^i - \frac{\Delta t}{\Delta n} (v_{1,\tau}^i - v_{1,\tau}^{i-1}), \quad (4.11)$$

$$s_{u,\tau+1}^i = s_{u,\tau}^i - \frac{\Delta t}{\Delta n} \left[\frac{s_{u,\tau}^i}{s_{1,\tau}^i} (v_{u,\tau}^i - v_{u,\tau}^{i-1}) + \frac{v_{1,\tau}^i - v_{u,\tau}^i}{s_{1,\tau}^i} (s_{u,\tau}^i - s_{u,\tau}^{i-1}) \right], \quad u \neq 1. \quad (4.12)$$

Here, i denotes the vehicle platoon index (number) of user class 1, τ denotes the current time instant, and Δt denotes the period length.

The second alternative method to formulate the Lagrangian multi-class model can be directly derived from the mixed-class Lagrangian conservation equation which is based on the variation of platoon length (see equation (3.9)). This equation holds for each user class u . The concept for the resulting formulation is similar to the foregoing multi-class formulation in Eulerian coordinates, reading

$$\frac{Ds_u}{Dt} + \frac{\partial v_u}{\partial n_u} = 0, \quad \forall u: \text{ class-specific conservation equation}, \quad (4.13)$$

$$v_u = V_u^*(s_{\text{tot}}), \quad \forall u: \text{ class-specific fundamental relation}, \quad (4.14)$$

$$s_{\text{tot}} = \frac{1}{\sum_u \eta_u / s_u}, \quad \text{total (effective) vehicle spacing}. \quad (4.15)$$

Here, the Lagrangian time derivative takes the form $\frac{D}{Dt} = \frac{\partial}{\partial t} + v_u \frac{\partial}{\partial x}$. n_u denotes the vehicle number of user class u . The conservation equation (4.13) holds for each user class. This indicates the change of vehicle spacing over time should be equal to the change of vehicle speed over vehicle number in each class. Each user class has its own fundamental relation, and the total effective spacing is the link between different user classes. The class-specific equilibrium speed (equation (4.14)) is also determined by the effective spacing.

The main difference from the ‘‘Piggy-back’’ formulation is that an independent coordinate system is introduced for each of the user classes. This means that vehicles from

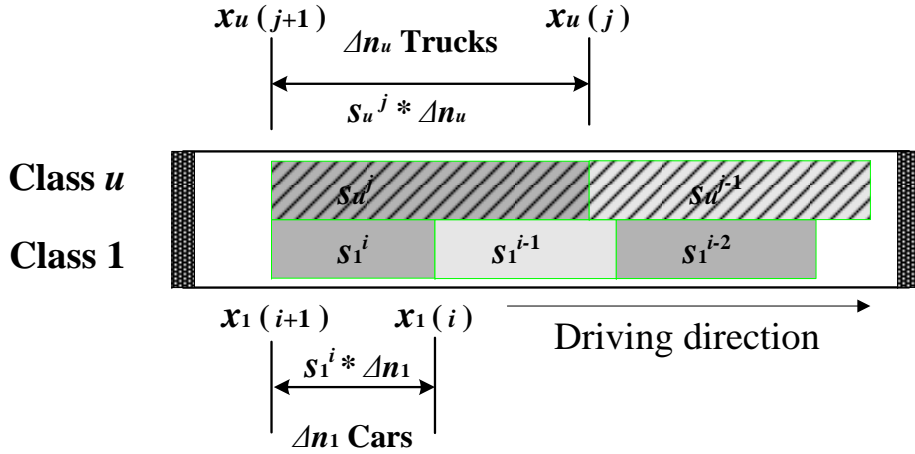


Figure 4.2: Vehicle discretisation at a time instant in the Multi-pipe formulation (two-class case). i and j denote the indices respectively for class 1 (passenger cars) and class u (trucks).

each class are clustered in platoons and numbered separately. Therefore, the platoon size for each class Δn_u can be chosen differently, but are still subject to the constraint of the CFL condition (Courant et al., 1967). According to the conservation law, the class-specific platoons only react to the consecutive same-class platoons via the equilibrium speed. This equilibrium speed, however, is as in the alternative formulation a function of the total (effective) vehicle spacing. This formulation is referred to as a “Multi-pipe” method further below. Figure 4.2 illustrates the related vehicle discretisation at a time instant.

For numerical solutions, the upwind scheme and the explicit time-stepping scheme can be applied to this formulation. This results in the following discrete equations:

$$s_{u,\tau+1}^k = s_{u,\tau}^k - \frac{\Delta t}{\Delta n_u} (v_{u,\tau}^k - v_{u,\tau}^{k-1}), \forall u, \quad (4.16)$$

$$v_{u,\tau}^k = V_u^*(s_{\text{tot},\tau}^k), \quad (4.17)$$

$$s_{\text{tot},\tau}^k = \frac{1}{\sum_u \eta_{u,\tau}^k / s_{u,\tau}^k}. \quad (4.18)$$

Here, k denotes the generalised vehicle platoon index (number).

The key to solving these equations lies in the fact that one needs the total effective spacing to compute equation (4.17). Since a platoon of class u may “span” over multiple (or partial) platoons of other classes, this requires an additional step in which the total effective vehicle spacing is calculated for all platoons of all user classes. Based on a two-class case (refer to Figure 4.2), the calculation is exemplified as follows. We consider a truck platoon j at a time instant τ . To simplify notation, the subscript for time (τ) is omitted. The equivalent spacing of class 1 (passenger cars) over this truck

platoon is given on the basis of the original definition (distance headway per vehicle):

$$s_1^j = \frac{\int_{x_u(j+1)}^{x_u(j)} s_1(n_1(x)) \cdot \Delta n_1(n_1(x)) dx}{\int_{x_u(j+1)}^{x_u(j)} \Delta n_1(n_1(x)) dx} = \frac{s_u^j \cdot \Delta n_u}{\int_{x_u(j+1)}^{x_u(j)} \Delta n_1(n_1(x)) dx}. \quad (4.19)$$

Here, the numerator of the right-hand side is the length of the truck platoon j , and the denominator calculates the number of cars over this truck platoon. $\Delta n_1(n_1(x))$ denotes the number of passenger cars as a function of locations $n_1(x)$. Accordingly the total effective spacing (s_{tot}^j) over this truck platoon j can be calculated based on s_u^j and s_1^j (in equation (4.19)) via equation (4.18). Similarly, one can calculate the equivalent spacing of any class u (s_u^i) and thus the total effective spacing (s_{tot}^i) over a passenger car platoon i .

4.2.3 Discussion and choice

So far, two modelling and (vehicle) discretisation choices of the Lagrangian multi-class first-order traffic flow model have been presented. The main difference between Lagrangian formulations and Eulerian formulations is that the numerical scheme for solving traffic flow model is simplified from a mode-switching scheme (e.g., the Godunov scheme) to an upwind scheme. As discussed in Section 3.6, there are several advantages in terms of numerical simulations, such as straightforward discretisation and accurate simulations. A simulation study in (Van Wageningen-Kessels et al., 2010a) using a Lagrangian multi-class formulation (the ‘‘Piggy-back’’ case) shows that the advantages with numerical simulations may also apply for the multi-class Lagrangian model.

In the Piggy-back method, only one vehicle coordinate system is used (the so-called reference class). On the basis of the reference class, the spacing and as a result, the equilibrium speed of all other classes can be derived. In the Multi-pipe method, vehicle coordinates are defined as many as the user classes. In this case, calculation of the equilibrium speed requires an additional (intermediate) expression (e.g., equation (4.19)), which calculates the effective total spacing and thus the equilibrium speed for all classes.

The main advantage of the Piggy-back method is the higher computational efficiency as compared to the Multi-pipe method. Due to the fact that only one coordinate system is used based on user-class 1, the computation for the evolution of multi-class system states is straightforward (equation (4.12)). The equilibrium speed based on the total effective spacing can be calculated directly, since by definition the spacing of all user classes is available on the same discretised platoons. Class 1 vehicles essentially carry portions of other vehicle classes on their backs (hence the name). These features are beneficial for real-time applications, for instance, traffic state estimation and prediction. In case of implementing an EKF technique, the Piggy-back formulation enables a relatively straightforward calculation of the Jacobian matrix (the derivative

of equations (4.11) and (4.12) with respect to each user class spacing). Therefore, it is possible to apply the EKF to a multi-class first-order traffic flow model in Lagrangian coordinates. This will be further clarified in Section 4.5. In contrast, computing such a Jacobian matrix in the Multi-pipe formulation is much more complex, due to the additional interaction term (equation (4.19)).

Nonetheless, there is a potential trade off between the ease of implementation and the simulation accuracy. Since the discretisation in the Piggy-back formulation is only based on class 1, problems may arise when the density of the first-class vehicle is small related to other classes. For example, we consider a traffic flow with 90% trucks and just 10% passenger cars (the reference class). In that case, a platoon of cars would carry a large number of trucks “on its back”, possibly much larger than the pre-defined platoon size Δn . The calculation for this number of other class vehicles is quite sensitive to the spacing values. As a result, large numerical errors may be introduced, and the total vehicle numbers of the other classes that are carried by the first-class platoons might change rapidly. In such cases, the Multi-pipe method might be more appropriate.

In contrast to the Piggy-back method, the Multi-pipe method introduces separate coordinates to different user-classes. The state evolution of each user-class is embedded in a relative independent coordinate system with its own discretisation. It is a more intuitive and understandable way to represent class-specific traffic. This formulation also results in a more efficient implementation of network discontinuities, class-specific flow calculations, class-specific control applications and even class-specific Origin-Destination (O-D) estimations. More specifically, for practical applications in real road networks, this Lagrangian multi-class model has to be completed with node models. In the Multi-pipe formulation, the mixed-class node model in Lagrangian coordinates (presented in Section 3.3.2) can be easily generalised (generally applied) to more classes. The main benefit lies in the fact that the in-flow of each user class is dealt with separately. We consider an extreme case, for example, an on-ramp where the truck percentage is very high. Intuitively in the Piggy-back model, these trucks can only enter the flow as soon as a class-1 platoon of size Δn is available. In the Multi-pipe model, trucks may enter as soon as Δn_u trucks are available. The Multi-pipe formulation hence provides a more natural model for class-specific event control, such as class-specific ramp metering and speed limit control, compared to the Piggy-back model. However, there is a disadvantage: the application of the Multi-pipe formulation in traffic state estimation is considerably more complex than the application of the Piggy-back formulations.

In sum, the different formulations of the (discretised) Lagrangian multi-class first-order traffic flow model have their own advantages and disadvantages. An overview is given in table 4.1. In the Lagrangian state estimator developed in this thesis, the process model applies a “Piggy-back” formulation, due to its suitability for on-line traffic state estimation. For network discontinuities, the mixed-class Lagrangian node models need to be extended and developed, to solve certain foregoing modelling difficulties with the “Piggy-back” formulation, as given in the next section.

Table 4.1: Comparison between two formulations of the multi-class Lagrangian model

	Piggy-back formulation	Multi-pipe formulation
Real-time applications, simulation efficiency (calculation of interconnected components)	+	-
Suitability for state estimation methods (e.g., multi-class EKF traffic state estimation)	+	-
Suitability for incorporating network discontinuities (e.g., node models, boundary conditions)	-	+
Class-specific control and O-D estimations	-	+

4.3 Multi-class node model for network discontinuities

To complete the traffic system modelling at a multi-class level, this section presents a multi-class node model in Lagrangian coordinates, which extends the mixed-class node models developed in Section 3.3. Essentially there are two aspects in node models:

1. to reformulate fundamental relations depending on spatiotemporal discontinuities at nodes (lane drop/extension, speed limits);
2. to generate (remove) vehicle platoons into (from) traffic flows, and as a result to change the spacing of platoons which pass the source (sink) location.

The first aspect can be easily adapted by making fundamental relations time and space dependent. Here, we focus on developing the second aspect. This multi-class node model is developed on the basis of and also applied to the Piggy-back formulation. In this formulation, only one vehicle coordinate system is used, in that class 1 (e.g., car class) is treated as the reference class. Therefore, solely for the reference class, the mixed-class node model can be easily adapted. The same discretisation choices are applied: whole (reference) vehicle groups (platoons) are only added to (or removed from) the flow at the beginning of a new time step.

In the Piggy-back formulation, the reference class vehicles carry portions of other vehicle classes on their backs. The amounts are determined by the corresponding spacing of other classes and the length of the platoon. To guarantee the conservation of vehicles, the multi-class node model updates the spacing of other vehicle classes based on in-flows or out-flows. We again consider a two vehicle-class Piggy-back formulation. In case of an on-ramp node, as soon as one car (class 1) platoon of size Δn_{car} (Δn) is available, the influencing car platoon and the related truck platoon at the node can be identified at the same moment (Figure 4.3). The space of this car platoon is shared with the entry platoon. Based on the in-flow information, the number of input trucks can also be computed, namely as $n_{\text{truck, inflow}}$. With $n_{\text{truck, inflow}}$ and the spacing

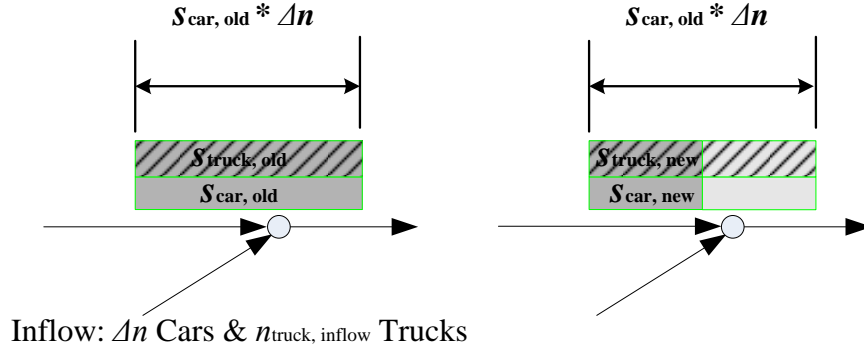


Figure 4.3: On-ramp node modelling for a two-class case.

(length) of the influencing platoon, we can further update the spacing of the influencing truck platoon as follows. We first calculate the original size of the influencing truck platoon by

$$n_{\text{truck,old}} = \frac{s_{\text{car,old}} \Delta n}{s_{\text{truck,old}}}. \quad (4.20)$$

Then, the updated spacing of the truck platoon is determined by the platoon length divided by the total number of trucks, reading

$$s_{\text{truck,new}} = \frac{s_{\text{car,old}} \Delta n}{n_{\text{truck,old}} + n_{\text{truck,inflow}}}. \quad (4.21)$$

Similar logic can be applied to the off-ramp case. The same as the mixed-class case, if a car platoon of size Δn is leaving from the mainstream, the space of this car platoon will be taken by its successive car platoon. The spacing of the influencing truck platoon (the successive one) will be also updated according to the number of out-flow trucks and the actual vehicle amounts in the influencing truck platoon.

However, there seems to be a limitation for the Piggy-back formulation when modelling nodes. For instance, when vehicle percentages for the non-reference classes are rather high in traffic flows, so that the reference class platoons would carry a large number of vehicles of other classes. This seems to indicate that at nodes, large number of other vehicles (larger than Δn) cannot enter or leave the flow until a reference vehicle platoon of size Δn is available. In this extreme case, we can update the spacing for each vehicle class according to a threshold value (n_{thrd}) separately. This value can be chosen as an identical value for all vehicle classes (equal to the discretisation size Δn) or set differently for each class. Once the in-flow or out-flow amount of a vehicle (non-reference) class exceeds this predefined value, the spacing of the related platoon will be updated by the related in-flow or out-flow information. For example, we consider an on-ramp case with two-class flow, in which truck demand is quite high. We can update the spacing of the truck platoon only based on the truck in-flow, given by

$$s_{\text{truck,new}} = \frac{s_{\text{car}} \Delta n}{n_{\text{truck,old}} + n_{\text{thrd}}}. \quad (4.22)$$

Regarding boundaries, the minimum supply-demand principle can also be applied. At origins, the class-specific in-flows should be given, which are used to determine input (initial) spacing of each class. If no multi-class information is available, two methods presented in Section 6.3 and Appendix A can be applied to estimate class-specific input based on the mixed-class information. At destinations, out flows are always possible. With the multi-class node model, we can perform a multi-class network-wide traffic state estimation. To be specific, in the prediction step of the data assimilation, the spacing of influencing platoons are first updated for each vehicle class via the node model, and then the spacing of all platoons are evaluated by the multi-class process model.

4.4 Observation models for multi-class Lagrangian formulation

For the full specification of the multi-class state estimation, the multi-class traffic systems need to be completed with multi-class observation models. First of all, class-specific observations are briefly discussed. Similar to the mixed-class case, Eulerian sensing data (observations) and Lagrangian sensing data (observations) can be identified. To be specific, we further distinguish both types of sensing data into two categorisations, namely event-based or aggregate data. Eulerian sensors (dual-loops, cameras) at fixed points can measure event-based individual vehicle passage information, including passage time, instantaneous speed, lane number and vehicle length. Based on the length information, vehicle classes can be distinguished. Moreover, local spacing observations can be inferred under an assumption of the local homogeneous condition, in which speeds of vehicles are constant in a short time period. This information can be directly transmitted to the end-user, or first averaged locally then transmitted with aggregate values. In principle, the end-user can obtain either class-specific individual vehicle data or aggregate speeds and flows. In practice, although most of Eulerian sensors (e.g., most of the MoniCa data systems in the Netherlands ¹) only provide mixed-class aggregate information owing to historical technical constrains, new monitoring systems allow for remote collection of individual passages or directly provide class-specific aggregate information (such as in parts of Italy and Germany, new MoniCa data systems). To demonstrate the concepts in this thesis, class-specific information is considered to be available as observations. Similarly, the Lagrangian sensors (probe vehicles, aerial cameras, helicopter videos) can also provide either class-specific individual probe information or class-specific space-mean speed and distance headway (spacing or density) over certain spatial distances (platoons).

In Lagrangian coordinates, the class-specific Lagrangian fundamental relation (4.8) provides a natural observation model for Lagrangian observations. It relates speed observations of different vehicle classes to the system state variable: spacing (total

¹A Dutch loop data collection system that provides 1-min aggregate speed and flow information.

effective spacing). If spacing observations are available, then we can also obtain an identity observation relation for the class-specific state variables (class-specific spacing). The following equations are used when Lagrangian sensing data (spacing and speed observations) are incorporated:

$$\begin{aligned} \mathbf{y}_{u,\tau} &= \mathbf{s}_{u,\tau}^{\text{obs}} = \mathbf{s}_{u,\tau} + \mathbf{r}_\tau && \text{(spacing observation)} \\ \mathbf{y}_{u,\tau} &= \mathbf{v}_{u,\tau}^{\text{obs}} = V^*(\mathbf{s}_{\text{tot},\tau}(\mathbf{s}_{1,\tau}, \dots, \mathbf{s}_{u,\tau})) + \mathbf{r}_\tau && \text{(speed observation)}. \end{aligned} \quad (4.23)$$

Here, $\mathbf{s}_{u,\tau}^{\text{obs}}$ and $\mathbf{v}_{u,\tau}^{\text{obs}}$ respectively denote class-specific spacing and speed observations at a time instant τ . Analogous to the mixed-class case, to incorporate both types of observations at the same time, a combined observation equation can be used, which reads

$$\mathbf{y}_{u,\tau} = \begin{bmatrix} \mathbf{s}_{u,\tau}^{\text{obs}} \\ \mathbf{v}_{u,\tau}^{\text{obs}} \end{bmatrix} = \begin{bmatrix} \mathbf{s}_{u,\tau} \\ V^*(\mathbf{s}_{\text{tot},\tau}) \end{bmatrix} + \mathbf{r}_\tau. \quad (4.24)$$

Similarly, multi-class observation models are used in the correction step of data-assimilation methods.

For class-specific Eulerian observations, we extend the previously developed observation model for the mixed-class Eulerian data (Section 3.4). This model is based on the kinematic wave theory. Given class-specific spot speeds, the speed observation for a specific vehicle (either car or truck) platoon within the influence area of a loop detector is determined by the related harmonic (space) mean speed obtained from this loop. Then, the multi-class Lagrangian fundamental relation applies to relate system states to spot observations. Note that, the mixed-class aggregate speeds and flows from dual-loop monitoring systems cannot be used as class-specific inputs directly. Therefore, an estimation procedure is developed to extract class-specific (density, speed, spacing) information out from the aggregate data. This method requires local density and traffic composition information. The details are illustrated in Appendix A. Table 4.2 provides an overview of both mixed-class and multi-class observations and the corresponding observation models in the Lagrangian formulations.

4.5 Multi-class Lagrangian traffic state estimation based on the Extended Kalman Filter

In this section, we discuss the possibility and the modelling of Extended Kalman Filter (EKF)-based traffic state estimation using the multi-class Lagrangian formulation.

For data assimilation, the EKF is applied. The detailed algorithm has been presented in Section 3.5. The difference is that the EKF is now rooted in the multi-class Lagrangian traffic system models. [Ngoduy \(2008\)](#) argued that the EKF method might not be a proper option for highly nonlinear traffic models, such as multi-class traffic flow models. This is justified for the Eulerian multi-class traffic flow model (refer to

Table 4.2: Overview of observations and the related observation models in the Lagrangian formulation

	Eulerian sensing data (loops, cameras)		Lagrangian sensing data (probe vehicles, cell-phone information, aerial cameras)	
	Event-based (Individual passages) v_i, l_i, t_i, x_i	Aggregate (MoniCa, NDW ^a ; over ΔT) $q_{\text{det}}, v_{\text{det}}, q_{u,\text{det}}, v_{u,\text{det}}$ ($k = q/v$)	Event-based (Individual probes) v_j, s_j, x_j $v_{u,j}, s_{u,j}, x_{u,j}$	Aggregate (over platoons Δn) $v_M, v_{u,M}$ $s_M, s_{u,M}$
mixed-	$y_i = s_i = v_i \cdot (t_i - t_{i-1})$ (Under assumption) ^b $y_{\Delta T} = \begin{cases} q_{\Delta T} = \sum_{\Delta T} i, \\ v_L = \sum_{\Delta T} v_i / \sum_{\Delta T} i, \\ v_H = \sum_{\Delta T} i / \sum_{\Delta T} (1/v_i) \end{cases}$	$y_{\Delta T} = \begin{cases} q_{\text{det}}, \\ v_{\text{det}} (v_L \text{ or } v_H) \end{cases}$	$y_{\Delta n} = \begin{cases} v_M = \sum_{\Delta n} v_j / \Delta n, \\ s_M = \sum_{\Delta n} s_j / \Delta n \end{cases}$	
multi-	$y_i = s_{u,i}$ (Under assumption) $y_{\Delta T} = v_{u,H}$	$y_{\Delta T} = \begin{cases} s_{u,\Delta T}^{\text{Est}}, (\text{Appx A}) \\ v_{u,\text{det}} (v_{u,L} \text{ or } v_{u,H}) \end{cases}$	$y_j = \begin{cases} v_{u,j}, \\ s_{u,j} \end{cases} \quad (\text{via } x_{u,j})$	$y_{\Delta n} = \begin{cases} v_{u,M} = \sum_{\Delta n} v_{u,j} / \Delta n, \\ s_{u,M} = \sum_{\Delta n} s_{u,j} / \Delta n \end{cases}$

Here, i denotes the index of sample observations from Eulerian sensors; j denotes the index of sample observations from Lagrangian sensors. u denotes different vehicle classes. ΔT denotes aggregate (measurement) interval. Δn denotes size of platoons. The subscripts L, H and M respectively denotes three different ways of averaging: local arithmetic average, harmonic average and average over space.

^aNDW: National Data Warehouse.

^bAssumption of the local homogeneous condition: speeds of vehicles are constant in a short time period.

(Van Lint et al., 2008b)) or the foregoing Multi-pipe formulation of the multi-class Lagrangian flow model. In that, the linearisation of the system models are complex and not possible, owing to the additional interaction terms with the flux calculations in the Eulerian multi-class formulation, and with the calculations of equivalent spacing in the Multi-pipe model.

However, there are no such complex interactions in the Piggy-back formulation of the multi-class traffic flow model. This enables a relatively straightforward derivation analysis, namely the calculation of the Jacobian matrix (3.31) in the EKF. Considering a two vehicle-class Piggy-back formulation, car class (class 1) and truck class (class 2) are identified, respectively. The system state variable \mathbf{s} consists of a vector of the states of the two classes, namely $\mathbf{s} = [\mathbf{s}_1, \mathbf{s}_2]^T$ ($\mathbf{s}_1 = [\dots, s_1^i, \dots]^T$, $\mathbf{s}_2 = [\dots, s_2^i, \dots]^T$, $i \in N$, i denotes the vehicle platoon index (number) of user class 1). The evolution of the system states for the two classes are described by discretised equations (4.11) and (4.12). In the following, we show how the discretisation of the process model can be used in the derivation analysis of the EKF. We reorganise the discretisations of the two-class Piggy-back formulation, yielding

$$s_{1,\tau+1}^i = f_{1,\tau}^i = s_{1,\tau}^i - \frac{\Delta t}{\Delta n} (v_{1,\tau}^i - v_{1,\tau}^{i-1}), \quad (4.25)$$

$$s_{2,\tau+1}^i = f_{2,\tau}^i = s_{2,\tau}^i - \frac{\Delta t}{\Delta n} \left[\frac{v_{1,\tau}^i (s_{2,\tau}^i - s_{2,\tau}^{i-1}) - s_{2,\tau}^i v_{2,\tau}^{i-1} + s_{2,\tau}^{i-1} v_{2,\tau}^i}{s_{1,\tau}^i} \right]. \quad (4.26)$$

Here, f_1 and f_2 denote the process equations of the two classes. The calculation the Jacobian matrix in this case is equivalent to calculate the derivative of equations (4.25) and (4.26) with respect to the spacing of each user class (s_1 and s_2). We notice that in these equations the class-specific speeds v_1 and v_2 are as functions of total effective spacing s_{tot} , which is further calculated based on the class-specific spacing s_1 and s_2 and the PCE value η_1 ($:= 1$) and η_2 . During each time (calculation) step, the PCE values are considered as constant based on the previous time step (do not update intermediately, since η_u also depends on v_1 and v_2). In this example, we only have η_2 . Moreover, Smulders' fundamental relations in Lagrangian coordinates are used, with two free-flow speeds for each of the two classes (see Figure 4.4 and equation (3.11)). The value at the nondifferential point in the Smulders' fundamental relation can also be made up by a smooth function through this point (the same method presented in Section 3.5). By doing this, the system model is theoretically reasonable to be applied for the EKF technique. Then, the chain rules of derivatives can be applied to the

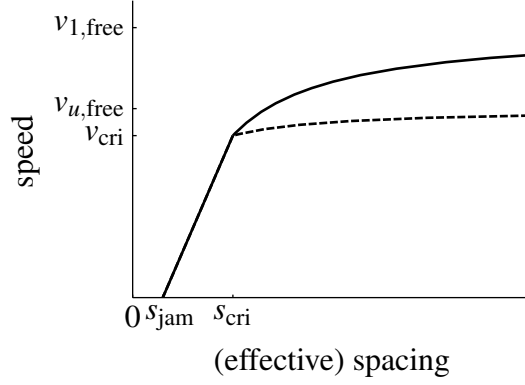


Figure 4.4: Multi-class Lagrangian fundamental relation (Smulders').

discretisations of the process model, yielding the following Jacobian matrix:

$$\mathbf{F} = \frac{\partial \mathbf{f}}{\partial \mathbf{s}} = \begin{bmatrix} \frac{\partial f_1^1}{\partial s_1^1} & 0 & 0 & \frac{\partial f_1^1}{\partial s_2^1} & 0 & 0 \\ \ddots & \ddots & 0 & \ddots & \ddots & 0 \\ 0 & \frac{\partial f_1^N}{\partial s_1^{N-1}} & \frac{\partial f_1^N}{\partial s_1^N} & 0 & \frac{\partial f_1^N}{\partial s_2^{N-1}} & \frac{\partial f_1^N}{\partial s_2^N} \\ \frac{\partial f_2^1}{\partial s_1^1} & 0 & 0 & \frac{\partial f_2^1}{\partial s_2^1} & 0 & 0 \\ \ddots & \ddots & 0 & \ddots & \ddots & 0 \\ 0 & \frac{\partial f_2^N}{\partial s_1^{N-1}} & \frac{\partial f_2^N}{\partial s_1^N} & 0 & \frac{\partial f_2^N}{\partial s_2^{N-1}} & \frac{\partial f_2^N}{\partial s_2^N} \end{bmatrix} \quad (4.27)$$

Here, we will not present each of the elements in the Jacobian matrix \mathbf{F} but the derivatives of the class-specific speeds with respect to the class-specific spacing, which are also used in the calculation of the Jacobian matrix \mathbf{H} (3.33). To simplify, neither time index τ nor platoon index i is included, the derivatives read

$$\frac{\partial v_1}{\partial s_1} = \begin{cases} \frac{s_{cri}(v_{1,free} - v_{1,cri})}{s_1^2}, & \text{free-flow} \\ \frac{v_{1,cri}}{(s_{cri} - s_{jam})} \frac{s_2^2}{(s_1 \eta_2 + s_2)^2}, & \text{congestion} \end{cases} \quad (4.28)$$

$$\frac{\partial v_1}{\partial s_2} = \begin{cases} \frac{s_{cri}(v_{1,free} - v_{1,cri})}{s_2^2} \eta_2, & \text{free-flow} \\ \frac{v_{1,cri}}{(s_{cri} - s_{jam})} \frac{\eta_2 s_1^2}{(s_1 \eta_2 + s_2)^2}, & \text{congestion} \end{cases} \quad (4.29)$$

$$\frac{\partial v_2}{\partial s_1} = \begin{cases} \frac{s_{cri}(v_{2,free} - v_{2,cri})}{s_1^2}, & \text{free-flow} \\ \frac{v_{2,cri}}{(s_{cri} - s_{jam})} \frac{s_2^2}{(s_1 \eta_2 + s_2)^2}, & \text{congestion} \end{cases} \quad (4.30)$$

$$\frac{\partial v_2}{\partial s_2} = \begin{cases} \frac{s_{cri}(v_{2,free} - v_{2,cri})}{s_2^2} \eta_2, & \text{free-flow} \\ \frac{v_{2,cri}}{(s_{cri} - s_{jam})} \frac{\eta_2 s_1^2}{(s_1 \eta_2 + s_2)^2}, & \text{congestion} \end{cases} \quad (4.31)$$

Here, $v_{1,cri} = v_{2,cri} = v_{cri}$. The related analysis can be easily generalised to more classes. These explicit expressions enable the computing of Jacobian matrices (linearisation of the process model), and thus the standard procedure of the EKF can be applied to the Piggy-back formulation. Within the EKF framework, at each calculation time step, the spacing of each vehicle class is updated (predicted) based on the process model. Then,

the prior estimate spacing is corrected with any available observations and the Kalman gain. Here, the Kalman gain is updated with the two Jacobian matrices \mathbf{F} and \mathbf{H} .

To sum up, based on the Piggy-back formulation of the multi-class traffic flow model and the related multi-class observation models, an EKF-based multi-class traffic state estimation approach is developed. As discussed above, the advantages with numerical simulations also remain for the multi-class Lagrangian formulation. This Lagrangian formulation renders both process and observation models more linear, and it is more suitable for the linear approximation and thus for the application with the EKF. We will further test its quality in the next chapter.

4.6 Summary and discussion

This chapter has extended the Lagrangian traffic state estimation method to a multi-class level. We apply the “Piggy-back” formulation of the multi-class Lagrangian traffic flow model as the process model. The observation models in Lagrangian coordinates have been developed for multi-class sensing data, and also summarised for the mixed-class case. The resulting multi-class Lagrangian traffic system model is more appropriate in the EKF-based framework compared with its Eulerian counterpart, which enables online multi-class traffic state estimation. Meanwhile, the multi-class node model is developed to extend the traffic system model and thus leads to network-wide multi-class traffic state estimation. In the following chapter, the model validation and application will be elaborated, where both multi-class and mixed-class traffic state estimators will be tested with synthetic and real-world data at link or network levels.

Chapter 5

Lagrangian traffic state estimation for freeway networks and case studies

In Chapter 3 and Chapter 4, we derived the mixed-class and the multi-class Lagrangian state estimation methods, respectively. In this chapter, we validate and verify the proposed methods, whereupon both of these two-level approaches are analysed and tested in synthetic (Sections 5.2, 5.4 and 5.5) and real-world (Sections 5.3 and 5.6) data environments. The Lagrangian state estimation method is also compared with the Eulerian counterpart in Sections 5.3 and 5.4. The experimental results are discussed thereafter.

Table 5.1: Overview of studied cases

Section No.	5.2	5.3	5.4	5.5	5.6
Lagrangian mixed-class	✓	✓	✓	✓	
Lagrangian multi-class				✓	✓
Eulerian comparison		✓	✓		
Link level	✓	✓		✓	
Network level			✓		✓
Synthetic data	✓		✓	✓	
Empirical data		✓			✓

5.1 Introduction

The mixed-class and multi-class Lagrangian traffic state estimation (TSE) methods were presented in Chapter 3 and Chapter 4, respectively. Therein, the three main components, system models, observation models and the data-assimilation method were elaborated in the Lagrangian traffic state estimation. For applications on networks, the corresponding node models were developed for the estimation methods. Chapter 3 also provided a qualitative comparison between the Lagrangian traffic state estimation and its Eulerian counterpart. This chapter further studies this Lagrangian approach in a quantitative manner.

5.1.1 Experimental setup

In this chapter, we perform several case studies to further explore the Lagrangian state estimation method by using both synthetic and realistic traffic data. With that, the Lagrangian state estimation method is validated and tested against different types (sources) of data and networks, and compared with the traditional Eulerian method. Table 5.1 presents all the studied cases in this chapter, and also specifies different aspects/dimensions considered in each of the cases.

5.1.2 Experimental objectives

For each case study, we define a specific experimental objective (hypothesis) as indicated in table 5.2. The first case study (Section 5.2) aims to validate the mixed-class Lagrangian state estimation on a small traffic stretch, incorporating both Eulerian and Lagrangian sensing data. We further hypothesize that the (EKF-based) state estimation works better with the Lagrangian formulation than the Eulerian formulation. To test this, the second one (Section 5.3) compares the Lagrangian TSE method with an Eulerian counterpart on a real motorway stretch. This comparison is additionally extended to a network level in the third experiment (Section 5.4). In the fourth experiment (Section 5.5), we intend to verify (validate) the multi-class Lagrangian estimator

Table 5.2: Objectives of case studies on Lagrangian TSE approaches

Lagrangian Model -	Validation (Verification)	Comparison Eulerian	Comparison Mixed-class	Application Integration
Mixed-class	5.2	5.3, 5.4		
Multi-class	5.5		5.5	5.6

via a synthetic simulation, and expect that better performance of the multi-class formulation over the mixed-class formulation can be achieved in terms of incorporating class-specific traffic data and estimation accuracy. Finally (Section 5.6), the application of the multi-class Lagrangian state estimation will be done in a real traffic network, incorporating all the available sources of empirical observations at a network level.

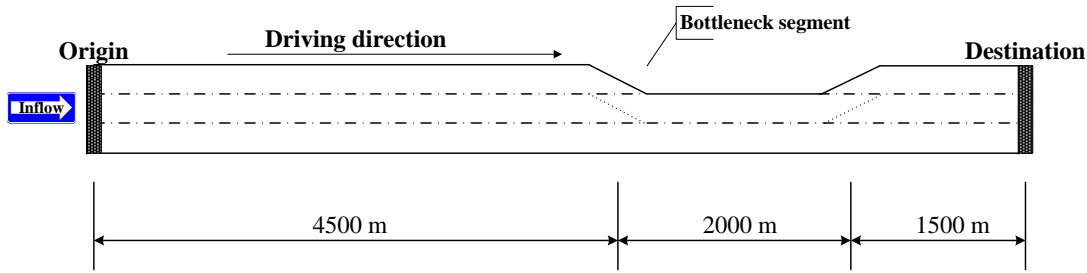
5.2 Link-level validation of Lagrangian traffic state estimation

The first experimental study aims to validate the proposed mixed-class Lagrangian state estimator, incorporating both Lagrangian and Eulerian sensing data (Yuan et al., 2011a). A freeway stretch with a lane-drop (without any on-ramps or off-ramps) is simulated in a microscopic simulation tool, which provides the ground-truth data. The detailed configurations of the simulation are given in the following.

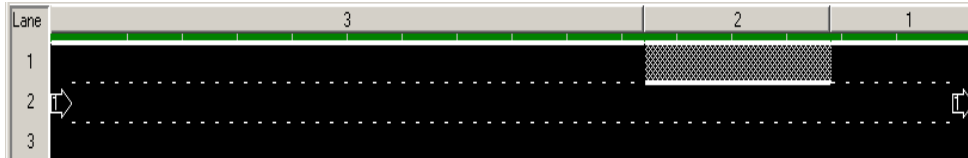
5.2.1 Data and test network

For validation of the traffic state estimator, any type of (synthetic or realistic) data can be used. The microscopic simulation model “FOSIM” (Dijker, 2012) is applied to provide synthetic traffic data. This model is developed at the Delft University of Technology, specially designed for the detailed analysis of discontinuities in freeway networks. All the parameters in terms of driving behaviour have been calibrated and validated based on data from Dutch freeways. In FOSIM, any type of traffic data (in both the Eulerian and Lagrangian sense) can be simulated based on the information of all the individual vehicles. Moreover, the ground-truth data are available in FOSIM. The state estimator was implemented in MATLAB 7.13.0, in which further analyses and evaluations were conducted.

A freeway stretch as shown in Figure 5.1(a) is considered for the simulation study. It is a road stretch of 8 km in length with three lanes and a single entrance and exit. There is a bottleneck segment with 2 lanes in the middle, which divides the whole stretch into three parts with lengths of 4500 m, 2000 m and 1500 m, respectively. This lane-drop model in FOSIM is shown in Figure 5.1(b). In this network, dual-loop detectors are located at every 500 m, which is comparable to the current monitoring situation on



(a) Freeway stretch with a lane drop.



(b) FOSIM model of the lane-drop case.

Figure 5.1: Illustration of a lane-drop case.

Dutch freeways, collecting 1-min (harmonic¹) averaged speeds and aggregate flows. A certain number of vehicles in FOSIM, depending on the chosen penetration rate, are treated as the floating cars. The floating car data (FCD) are obtained from the speed and location information of these individual vehicles at a fixed time interval. In this case, the ground-truth data (space-mean speeds, space density based on vehicle trajectories) are regarded as reference, which are derived from FOSIM over equidistant spatiotemporal regions of size 100 (m) x 60 (s).

5.2.2 Experimental scenarios

All the simulations are based on the same set of predefined parameters. The number of vehicles in each platoon (Δn) is chosen uniformly as 10 vehicles per platoon. The model-prediction time interval Δt is 1 s, and the measurement (updating) interval ΔT is 60 s for loop data and 1 s for FCD (this choice is too ideal but reasonable for model validation purpose). The total simulation period is 2 h (7200 s). The fundamental relation presented by Smulders (1989) is used, since it can describe the basic traffic characteristics in FOSIM without any complexity in the model itself (this fundamental relation can be easily transformed between Lagrangian coordinates and Eulerian coordinates via the relation of spacing and density). The related parameters, free-flow speed, critical speed, critical spacing and jam spacing, are estimated with FOSIM data as 115 km/h, 85 km/h, 41 m/veh./lane and 10 m/veh./lane, respectively. Note that, if one vehicle platoon covers a road stretch with different lanes (e.g., lane-drop area), then the fundamental parameters of the corresponding platoon will be weighted averaged, based on the length of the occupied segments.

¹In reality, Dutch dual-loop detectors provide arithmetic averaged speed, while detectors in FOSIM provide harmonic averaged speed to overcome the speed-bias problem. Please refer to Section 6.2

Table 5.3: Experimental scenarios

No.	Loop ($\Delta T = 60$ s)	FCD ($\Delta T = 1$ s)	Noise parameter (\mathbf{R})
Scn.	Spacing	Penetration rate	High to low reliability
12 (4 x 3)	500m, 1000m, 1500m, 2000m	-	$2^2, 5^2, 20^2$
12 (4 x 3)	-	0.5%, 1%, 2%, 3%	$2^2, 5^2, 20^2$

In this section, we also focus on testing the properties of the new approach in terms of the observation aspects, including the choice of observations (namely the choice of observation models), and the choice of noise parameters (used in the error covariance matrix). In the former, speed observations are used as input, both from dual-loop detectors of different spacings, and from floating cars of different penetration rates. As discussed in Section 3.5, the error covariance matrix \mathbf{Q} in the EKF reflects the uncertainty in process equations, which is, however, difficult to identify. In this study, the matrix \mathbf{Q} is set as a diagonal matrix with constant values of $5^2(\text{m/veh.})^2$ along the diagonal and zeros on off-diagonal elements. Likewise, the measurement error covariance matrix \mathbf{R} is related to the uncertainty in observations and fundamental diagrams (observation models). In an EKF process, \mathbf{R} reflects the learning rate from observations ($\mathbf{R} \propto \mathbf{1}/\text{Learning rate}$). Given a fixed process covariance \mathbf{Q} , if the uncertainty in the observation equation is small, then \mathbf{R} tends to be small, and thus the calculated gain \mathbf{K} is large. That means the learning rate from observations is high, and vice versa. In general, an appropriate combination of \mathbf{Q} and \mathbf{R} is quite important for providing good estimates. Given a fixed \mathbf{Q} , we will study different (constant) diagonal values of \mathbf{R} with respect to different reliability levels of observations. Alternatively, one can make the matrix \mathbf{R} adaptive to the real output errors. The latter means that the diagonal value in matrix \mathbf{R} is updated during the filtering process, see for example [Van Hinsbergen et al. \(2010\)](#). Table 5.3 overviews the simulation scenarios considered.

5.2.3 Performance criteria

To assess different scenarios, performance criteria are defined. The estimates from the state estimator are mainly vehicle spacing s and speed v of each platoon (or density k and speed v of each cell). First, this information is transformed into system states (q, k, v) with the format of the reference data. Then, the estimates (u - denotes system states) are compared with the ground truth (reference \hat{u}) data in terms of root mean square error ($RMSE$) and one relative error indicator, namely mean absolute percentage error ($MAPE$), which gives a combined indication of the relative error and the standard

Table 5.4: *RMSE* errors of all the scenarios (As the error covariance matrix \mathbf{Q} is fixed with the diagonal value of 5^2 , there are three choices of the diagonal values of \mathbf{R} : \mathbf{R} high reliability with 2^2 ; Equal to \mathbf{Q} with 5^2 ; \mathbf{R} low reliability with 20^2 .)

<i>RMSE</i>	Loop							
	500m		1000m		1500m		2000m	
	k	v	k	v	k	v	k	v
\mathbf{R} high reliability	22.2	12.9	25.2	13.2	25.7	14.4	27.9	14.5
Equal to \mathbf{Q}	23.2	13.8	30.0	16.6	31.4	17.4	32.8	18.0
\mathbf{R} low reliability	44.0	23.0	45.7	23.4	45.1	23.2	45.3	23.3

<i>RMSE</i>	FCD							
	0.5%		1%		2%		3%	
	k	v	k	v	k	v	k	v
\mathbf{R} high reliability	26.1	14.8	23.6	13.1	21.8	13.0	20.4	14.2
Equal to \mathbf{Q}	27.7	15.7	22.7	13.4	22.1	11.9	20.4	11.3
\mathbf{R} low reliability	44.5	22.9	42.3	22.0	42.5	22.1	39.6	20.9

deviation of the percentage error. These indices are defined as:

$$RMSE = \sqrt{\frac{\sum (u - \hat{u})^2}{NN}}, \quad (5.1)$$

$$MAPE = \frac{1}{NN} \sum \frac{|u - \hat{u}|}{\hat{u}}, \quad (5.2)$$

Here, u ($:= u(t)$) denotes either raw-input data or corrected data, \hat{u} ($:= \hat{u}(t)$) depicts the reference data. They are both as the functions of time (t). NN denotes the size of the data set. The definitions of *RMSE* and *MAPE* are as the same as those in Section 6.2.4, and they are used throughout the remaining chapter.

5.2.4 Results and discussion

Quantitative results and discussion

Tables 5.4 and 5.5 respectively list the *RMSE* and *MAPE* results of the different detection combinations with respect to the choice of the observation error covariance matrix \mathbf{R} . Clearly, with increasing loop observations, the improvement on speed and density estimation in terms of error performance is remarkable. Likewise, increasing FCD penetration rate also improves the accuracy for speed and density estimation.

Amongst the same observation-data scenarios with different noise levels, better speed and density estimates are achieved with the “high reliability” setting of observation noise parameters (see the boldface values in the tables). The reason is that, in our examples, reliable observation inputs (speed observations) are directly obtained from

Table 5.5: *MAPE* errors of all the scenarios

<i>MAPE</i> (%)	Loop							
	500m		1000m		1500m		2000m	
	<i>k</i>	<i>v</i>	<i>k</i>	<i>v</i>	<i>k</i>	<i>v</i>	<i>k</i>	<i>v</i>
R high reliability	13.7	13.9	15.9	15.4	15.4	16.5	16.8	16.8
Equal to Q	14.0	16.4	15.6	20.3	16.2	21.2	16.6	22.5
R low reliability	19.4	34.5	19.7	35.6	19.8	34.9	19.8	35.0
	FCD							
	0.5%		1%		2%		3%	
	<i>k</i>	<i>v</i>	<i>k</i>	<i>v</i>	<i>k</i>	<i>v</i>	<i>k</i>	<i>v</i>
R high reliability	16.9	17.2	14.6	14.5	13.5	14.7	12.1	14.7
Equal to Q	15.9	18.3	14.5	14.8	14.5	13.5	13.9	12.6
R low reliability	19.6	34.2	18.9	31.8	18.8	32.3	18.0	29.3

FOSIM. Therefore, observations should account for more weight over the model prediction to get accurate estimates. The choice of the noise in observation models is highly related to the quality of the observations. Appropriately tuning of this noise parameter will lead to good results.

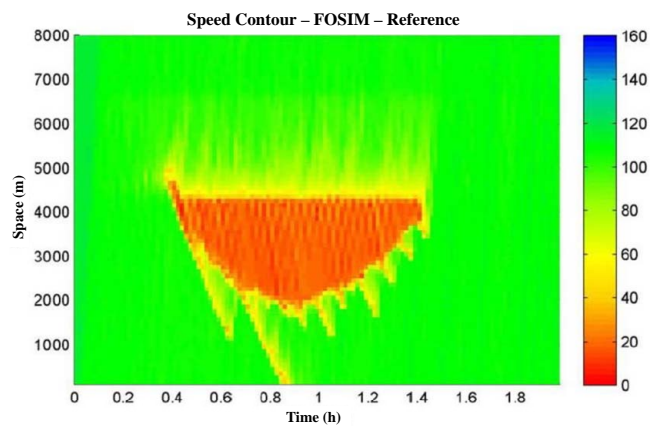
It can also be noticed that the relative errors (*MAPE*) for both speed and density estimations are mostly around 15%. This indicates that there still exist phase (systematic) errors between the reference FOSIM data and the state estimation; these errors are mainly from the congested states. The current results can be improved, for instance, by alternative choices of fundamental diagrams rather than the current choice with a linear congested branch (the Smulders' fundamental relations). Further study is needed by using the empirical data from real traffic networks.

Qualitative results and discussion

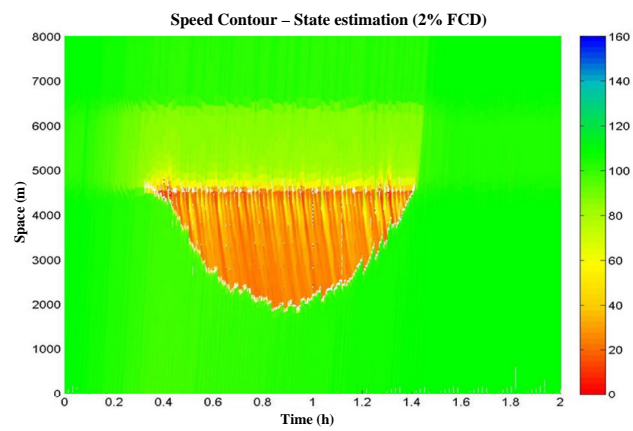
Due to the “upwind” numerical solution scheme, implementing the EKF in the Lagrangian case is far more straightforward than in the Eulerian case, which was already discussed in Section 3.5. As shown in Figure 5.2, the speed map (b) derived from the Lagrangian state estimator is comparable to the reference (ground truth) speed map (a). The starting time and dissolving time of congestion from estimation are similar to the reference plot. The typical traffic pattern (shockwaves) is also present in the estimation.

5.2.5 Conclusion

The mixed-class Lagrangian traffic state estimation method has been validated by a simulation study. Several properties of the Lagrangian state estimator in terms of the



(a) FOSIM reference case.



(b) State estimation result (2% FCD case).

Figure 5.2: Comparison of speed contour plots.

observation aspects have been studied. The experimental results demonstrate that both the Eulerian sensing data and the Lagrangian sensing data can be incorporated with the Lagrangian state estimation, while providing adequate results.

5.3 Comparison between Lagrangian and Eulerian approaches at a link level

The previous section has validated the mixed-class Lagrangian state estimation method at a link level. In this section, an empirical study is performed to compare the proposed Lagrangian state estimator with its Eulerian counterpart, based on the data from a British motorway (Yuan et al., 2012). First, the data environment and pre-processing are presented, followed by the description of simulation scenarios. Then, the result interpretation and discussion are provided.

5.3.1 Data and test network

The data used in this study were obtained from the Active Traffic Management section of the M42 motorway near Birmingham in the UK (Highway-Agency, 2011). The motorway is equipped with dynamic speed control systems and features hard shoulder running in peak hours, expanding the width of the carriageway from three lanes to four lanes in each direction. Moreover, this section has an unprecedented coverage of inductance loop detectors, with a nominal spacing of 100m. During 2008/09, 16 consecutive detectors on the Northbound carriageway (shown in Figure 5.3) were enhanced so that, amongst other improvements, the full Individual Vehicle Data of all vehicles driving through the 1-mile section were recorded (Wilson, 2011). The individual vehicle data include the passage time, speed, lane number and length of each vehicle as it passes each of the sites. With this high resolution, one can track most individual vehicles through the section in most traffic conditions and thus in effect reconstruct their trajectories (Wilson, 2008). However, the reconstructed trajectory data constitute only vehicle positions and speeds at fixed points (every 100 m). Therefore, certain methods are needed to interpolate the estimated vehicle trajectories in-between two detectors, based on which the FCD data samples are generated. For instance, the piecewise constant, linear (PLSB), and quadratic speed-based trajectory methods and the filtered inverse trajectory method could be used for this purpose (Van Lint, 2010). Clearly, interpolation (by any method) may induce errors, but due to the resolution of the data and the small distances between detectors, these errors will be very small. In our experiment, the PLSB method is used to complete and smooth the vehicle trajectories.

As a result of data pre-processing, “full” individual vehicle trajectories were estimated. The ground truth data (space-mean speeds, densities) were then derived from this trajectory dataset in Eulerian form over semi-equidistant spatiotemporal regions of size

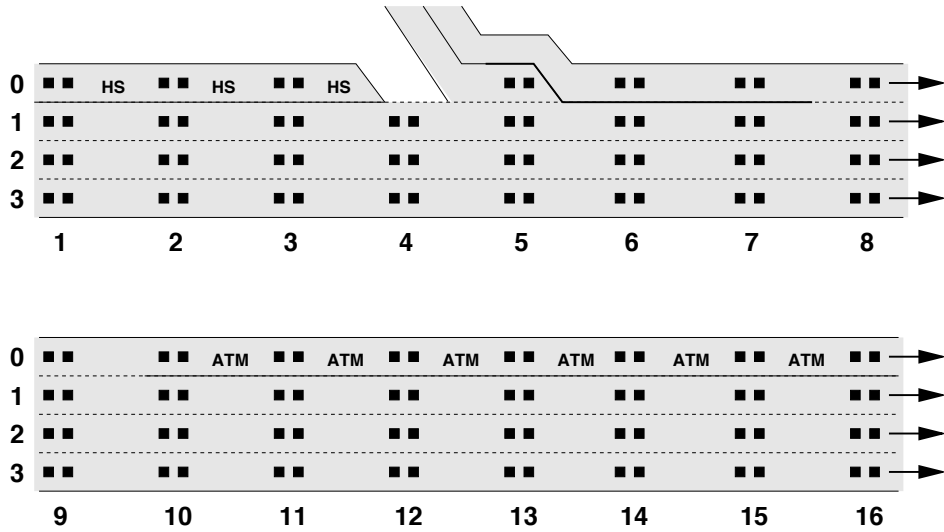


Figure 5.3: Geometry of the enhanced carriageway showing the location of inductance loops. HS denotes the hard shoulder (emergency lane) and Active Traffic Management (ATM) denotes hard shoulder activated during peak hours.

100 (m) x 60 (s), which were used as reference. The space-mean speeds over cells were calculated by the harmonic average of individual speed samples. The reference densities were calculated over spatiotemporal cells using trajectory data, under Edie's definition (Edie, 1965).

The study area is a straight stretch without on- and off-ramps between the ten most downstream detectors (between detectors 7 and 16 shown in Figure 5.3) along the enhanced carriageway (a section of approximately 1km in length). It contains an in-flow and an out-flow boundary. A homogeneous Von Neumann out-flow condition is applied in the process model. That means downstream congestion is not modeled in the prediction step, but it can be reproduced in the correction step. The in-flow boundary starts from the most upstream loop detector; thus the in-flow is known.

5.3.2 Experimental scenarios

The process models for both Eulerian and Lagrangian formulations are based on the discretised mixed-class first-order traffic flow model, given in equations (3.4) and (3.14), respectively. As an observation model, Smulders' fundamental relation (3.11) is used, since it is thus convenient to estimate the parameters from empirical data. Note that, the Eulerian form of the observation model can be easily derived from equation (3.11) via the relation between spacing s and density k . The related parameters are calibrated from the detector datasets and yield for v_{free} , v_{cri} , s_{cri} , s_{jam} , respectively, 115 km/h, 80 km/h, 51 m/veh/lane and 10 m/veh/lane. To obtain statistically reliable results, eight simulation runs for each scenario have been performed based on the empirical data of eight days of October in 2008. The simulation time periods for each run are 2 h, which are taken either from the morning peak or the afternoon peak. Within

these periods, the ATM lane is used, and a selection of representative traffic patterns (stop-and-go waves) occurs. Note that at the boundaries of stop-and-go waves, traffic states switch frequently between two traffic conditions (free flow and congestion). This indicates that the mode-switching problems occur frequently in these areas.

In this study, both the Eulerian (loop) measurements and the Lagrangian (floating car) observations are considered as input into the observation model of EKF. Loop detectors provide harmonic mean speed based on individual vehicle data aggregated over 60 s. The reconstructed FCDs provide instantaneous speed and location information reporting at a certain time frequency. Three different spacings of loop detection as well as four different penetration rates and reporting frequencies of FCD are studied independently.

To compare, an Eulerian formulation of the LWR state space model using the same fundamental relations is implemented in the same EKF framework. For fair comparison in terms of the same level of numerical stability, the sizes of discretized unit in two cases are determined based on the critical condition of the CFL condition. A time step $\Delta t = 1$ second is used; thus, the platoon size Δn and the cell size Δx are 2.2 (veh.) and 32 (m), respectively. Since the Gaussian noise parameters in \mathbf{Q} and \mathbf{R} are difficult to calibrate with the model and the surveyed road stretch, several combinations of values are analysed. The noise matrices \mathbf{Q} and \mathbf{R} are assumed to be diagonal with fixed diagonal values. As the same (speed) observations are used, the measurement error covariance matrices \mathbf{R} are fixed for both the Eulerian estimation and the Lagrangian estimation. Distinction has been made with respect to the process model noise \mathbf{Q} , since the filter gain (equation (3.32)) depends on the ratio of these two error parameters. In this study, different combinations of error noise matrices \mathbf{Q} with a fixed \mathbf{R} are tested based on a Monte-Carlo simulation. For the sake of brevity, the results of the combination ($\mathbf{Q} : 10^2$, $\mathbf{R} : 2^2$) leading to the highest performance in both Eulerian and Lagrangian state estimators are presented.

5.3.3 Results and discussion

Quantitative analyses

Table 5.6 summarises the considered scenarios and shows the averaged *RMSE* errors and their relative improvement compared to the Eulerian case for each scenario. Figure 5.4 illustrates the comparison of each run between the two state estimators in terms of the density and speed *RMSE* errors in scenario 1. The most important observation is that in all cases the Lagrangian state estimator outperforms its Eulerian counterpart by up to 24% for density and 75% for speed. In both the Lagrangian and Eulerian models, denser loop spacing yields improvements on speed and density estimation. Likewise, the more Lagrangian sensing data (observations) (from 5%/20s to 20%/5s) that are available, the better accuracy that can be achieved in the estimates.

Table 5.6: State estimation results of seven scenarios with different loop detector spacing and different penetration rates and reporting frequencies of FCD

Scenario	Data type		Eulerian state estimation		Lagrangian state estimation			
			$RMSE_k$ (veh/km)	$RMSE_v$ (km/h)	$RMSE_k$ (veh/km)	POI_k (%)	$RMSE_v$ (km/h)	POI_k (%)
1	Loop	200m	14.4	4.6	12.8	11.2	1.2	73.5
2		400m	15.6	5.3	14.0	10.7	2.4	55.3
3		800m	17.1	6.4	15.4	9.9	3.8	40.6
4	FCD	5%/20s	28.1	17.1	23.5	16.4	10.6	37.7
5		5%/5s	24.4	13.7	18.4	24.5	6.8	50.3
6		20%/20s	21.7	11.6	20.0	7.6	7.9	31.3
7		20%/5s	18.6	8.6	16.5	11.3	6.1	29.6

Note that: $RMSE$ error averaged over 8 Monte Carlo simulation runs and its related Percentage Of Improvement (POI) when compared with the Eulerian case. The POI index is calculated as $100 * (RMSE_{Euler} - RMSE_{Lag}) / RMSE_{Euler}$. When using Lagrangian sensing data, 5 different samplings (random seeds) with the same penetration rate and reporting frequency are simulated for each day.

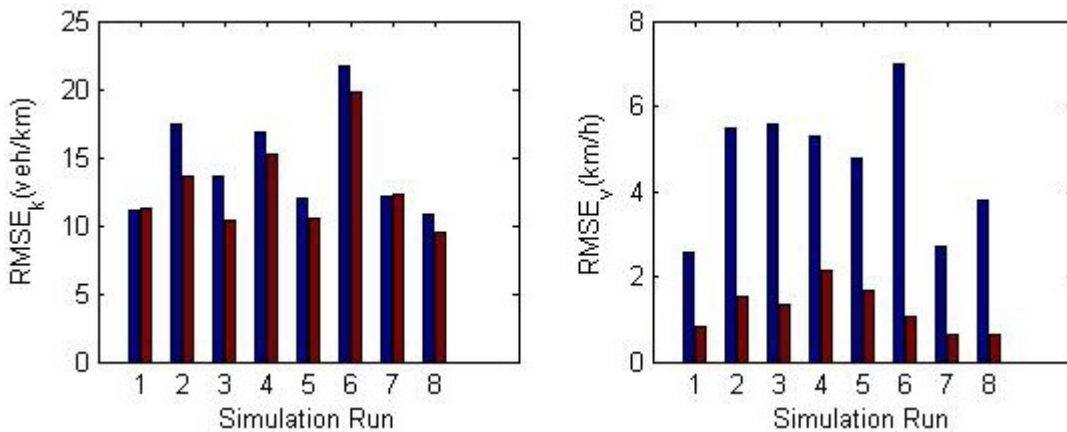


Figure 5.4: Error measurement comparison between two methods for each of the 8 simulation runs of scenario 1. Blue: Eulerian state estimation. Red: Lagrangian state estimation.

Note that the large relative improvement in speed estimation in the Lagrangian formulation can be partially explained by the fact that for very low speeds (errors), the relative difference between the RMSE results become large. Even though the error for the Eulerian state estimate in the absolute sense is also small (see, for example, the results in table 5.6 for scenario 1, where the relative improvement of 73.5% represents an absolute improvement of 3.4 km/h). On the other hand, in the case where the estimated speeds are also used for estimating travel times (proportional to $1/\text{speed}$), the absolute errors will also be large (Van Lint, 2010).

Nonetheless, the large improvement in speed estimation provides evidence that the Lagrangian formulation indeed improves the state estimation considerably. Although the Eulerian and Lagrangian traffic system models in the two coordinate systems describe the same traffic situation exactly, the numerical approximation of the Lagrangian model is more accurate than its Eulerian counterpart. Moreover, as discussed earlier, the “non-mode-switching” discretization scheme for the Lagrangian model also improves the data-assimilation method, since the linear approximation of the traffic system model near capacity is much better than in the Eulerian model. As a result, more accurate results are obtained in the Lagrangian formulation.

Qualitative analyses and discussion of the results

Note that a considerable error for the density estimates still exists. A reason for this might be owing to the inaccuracy of the observation models used. In the estimation procedure, speed observations are used to derive the “observed” system states (either density or spacing) from the estimated fundamental relations. In the ideal case, the derived state from speed should perfectly match the ground truth state. This is, however, not the case in practice. It is well known that in congestion state, a given density can be achieved at different speeds, indicating a set-valued congested branch in the fundamental diagrams (Cassidy, 1998). Therefore, a better interpretation of the fundamental relations between traffic variables can improve estimates of system states.

The simulation also shows that the Eulerian method mis-estimates the densities and the speeds in the first few cells. The explanation here relates to the way that the flow observations at the in-flow boundary are used. The in-flow is used to determine the incoming flux for the first cell in the Eulerian prediction model. The flow values are calculated by summing the counts during measurement intervals which can introduce errors. Moreover, an inaccurate fundamental relation (wrong capacity) might result in a biased flux calculation. Owing to the numerical scheme in Eulerian coordinates, the accuracy of the evolution of system states near the upstream boundary is sensitive to the in-flow. In contrast, in the Lagrangian model, the spacing of platoons is affected by what happens downstream, and the in-flow is used only to determine the number of vehicle platoons which need to be added to the flow. Therefore, in the Lagrangian estimator, the influence from the in-flow to the evolution of system states is limited

to just the first platoons entering the system. This yields another advantage of the Lagrangian state estimator over its Eulerian counterpart.

Finally, Figure 5.5 shows the simulation results based on the Eulerian and the Lagrangian state estimation with the Eulerian sensing data (scenario 1). In both simulations, at $t = 2100\text{s}$, a shock (low speed) is observed by a detector located at 670m. In the Lagrangian estimation, the shock propagation is clearly visible in the result, whereas the shock diffuses quickly further upstream in the Eulerian case. Moreover, the edges of vehicle platoons clearly distinguish the shock boundary, whereas a step-wise boundary (in a resolution determined by the grid size) is present in the Eulerian simulation. Figure 5.5 illustrates that indeed the numerical method in Lagrangian formulation causes less numerical diffusion and more accurate results than the Eulerian method. Note that more details about this issue of numerical diffusion can be found in (Van Wageningen-Kessels et al., 2010a).

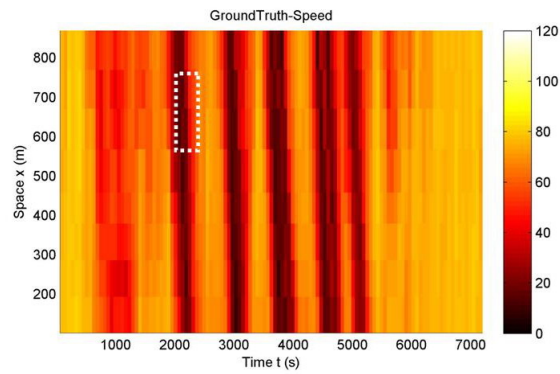
5.3.4 Conclusion

In this section, we have demonstrated that (mixed-class) Lagrangian state estimation outperforms Eulerian state estimation at a link level. First of all, the Lagrangian formulation enables more accurate and efficient simulation. Second, the Lagrangian formulation of the traffic model also leads to better data assimilation. It is important to note that the Eulerian and Lagrangian formulations are two ways to implement exactly the same underlying (LWR) theory and that in both cases exactly the same data from the same traffic situation were used in our experiments. The large increase in estimation accuracy hence stems entirely from a more suitable formulation of the same problem. Additionally, the simulation study with the empirical data shows that both the Eulerian sensing data and the Lagrangian sensing data are well incorporated with the Lagrangian state estimator and that the estimation provides better results than the Eulerian state estimates.

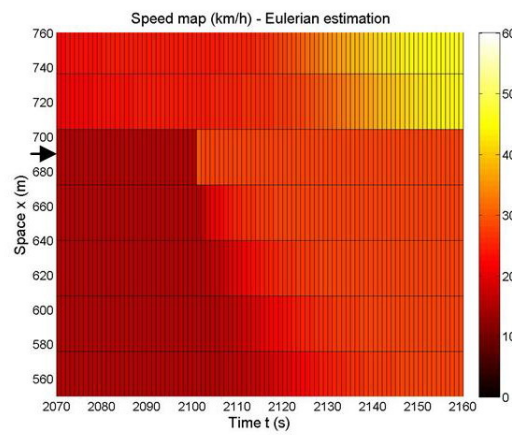
The Lagrangian formulation of state estimation enables more accurate application of the EKF method, owing to the solution to the mode-switching problem. This solution renders the estimation problem less nonlinear which results in more accurate results. This idea is not restricted to the EKF method, but can furthermore apply to other data-assimilation techniques (e.g., EnsKF, UKF and PF), for which we expect similar improvements as those shown therein.

5.4 Comparison between Lagrangian and Eulerian approaches at a network level

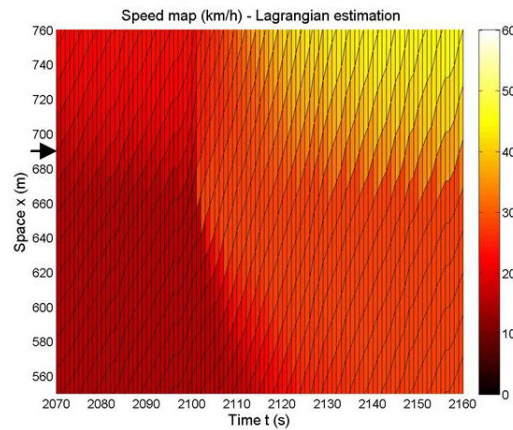
The previous section has demonstrated that the mixed-class Lagrangian state estimation outperforms the related Eulerian method at a link level. However, real-time traffic



(a) Reference spatiotemporal speed map



(b) Eulerian traffic state estimation



(c) Lagrangian traffic state estimation

Figure 5.5: Snapshots of a small region of the whole spatiotemporal speed map for scenario 1 based on the Eulerian estimation (b) and the Lagrangian estimation (c) (Data set: 7:30 am. - 9:30 am. 02-10-2008). The curved lines in (c) indicate trajectories of the vehicle groups. The upper plot (a) indicates the range (white dashed box) of the snapshots in the whole spatiotemporal reference speed map. The speeds are indicated as colors in $x-t$ plane (plot (a) scales to 0-120km/h, (b)&(c) scale to 0-60km/h).

management and control require state estimators that work on networks. By implementing the node models presented in Section 3.3, both Lagrangian and the Eulerian state estimators can be extended to a network level. In this section, a simulation study is conducted to further compare the performance of these two extensions (Yuan et al., 2011b).

5.4.1 Data and test network

In this section, the experiment and the comparison were also based on the synthetic data from the microscopic traffic simulator, FOSIM. As discussed previously in section 5.2.1, we can emulate all types of existing data formats in FOSIM, such as 1-min (harmonic) mean speeds and flows from dual-loop detectors, floating car data (FCD) with selected penetration rates and reporting frequencies. Similar to previously, ground-truth (reference) data derived from FOSIM were in the format of equidistant spatiotemporal regions of size 100 (m) x 60 (s). The simulations were conducted in the MATLAB environment.

For demonstration purpose, a typical highway stretch of about 7 km in length with a single on-ramp and an off-ramp was simulated within a period of 1 h. There was a bottleneck in the middle (starting from the on-ramp). FOSIM allows users to predefine demand patterns, generating vehicles at origins of freeways. In our simulation, traffic demand during each simulation period varied, so that a traffic jam emerged and dissolved at the bottleneck.

5.4.2 Experimental scenarios

For fair comparison in terms of the same level of numerical stability, the sizes of discretised unit in two formulations are determined based on the critical condition of the CFL conditions (equality applies in equations (3.8) and (3.15)). A time step $\Delta t = 1$ s is used, thus the platoon size Δn and the cell size Δx are 0.8 (veh.) and 32 (m), respectively. Since the Gaussian noise parameters in \mathbf{Q} and \mathbf{R} are difficult to calibrate with the model and the surveyed road stretch, several combinations of values are analysed. For the sake of brevity, the results of the \mathbf{Q} - \mathbf{R} combination ($\mathbf{Q} : 5^2, \mathbf{R} : 2^2$) leading to the highest performance in both Eulerian and Lagrangian state estimators are presented.

Both the Eulerian (loop) observations and the Lagrangian (floating car data - FCD) observations are inputted into the observation models of the EKF. This indicates that two types of data scenarios are studied independently. One uses speed observations from loop detectors, and there are three spatial resolutions investigated, namely 500 m, 800 m and 1000 m. The other one is with the FCD (speed) data, and the penetration rates and reporting frequencies are set respectively as 5%/30s, 5%/10s and 10%/10s. Although the choice for FCD might be too ideal compared to a realistic scenario, it is justified for the model demonstration purpose. In each data scenario, ten simulation

Table 5.7: State estimation results based on two formulations (average over ten simulation runs)

Scenario	Data type	Eulerian state estimation		Lagrangian state estimation			
		$RMSE_k$ (veh/km)	$RMSE_v$ (km/h)	$RMSE_k$ (veh/km)	POI_k (%)	$RMSE_v$ (km/h)	POI_k (%)
1	500 m	11.3	7.8	10.6	6.1	6.8	12.5
2	Loop 800 m	12.0	9.9	11.9	0.5	9.3	5.7
3	1000m	12.8	11.9	12.4	3.0	11.3	5.3
4	5%/30s	20.4	24.0	19.4	4.9	21.4	10.8
5	FCD 5%/10s	20.4	23.8	19.0	6.8	20.4	14.2
6	10%/10s	19.9	22.4	17.2	13.5	17.9	20.0

runs with diverse traffic patterns (different random seeds or different traffic demands) have been performed. The comparison results are shown in the following.

5.4.3 Results and discussion

Quantitative analyses

Table 5.7 shows the RMSE errors between the estimation results from the two formulations and the reference data (averaged on the basis of ten runs), together with the relative improvement (Percentage Of Improvement: POI) by the Lagrangian state estimation compared to the Eulerian case. Figure 5.6 illustrates the comparison of each run between the two formulations in terms of the density and speed RMSE errors in scenario 1 (500m apart Eulerian sensing). As Figure 5.6 indicates, in general, the Lagrangian approach provides more accurate estimates than the Eulerian method. When Lagrangian sensing data are used, the improvement of accuracy by the Lagrangian method is respectively up to 13% for density and 20% for speed estimates, compared with the Eulerian method (see scenario 6). Note that, in both Lagrangian and Eulerian approaches, the more Eulerian and/or Lagrangian observations (higher data resolutions) that are available, the more accurate estimates that can be achieved.

With all the positive POI values (see the boldface values in table 5.7), the experiment results show that the Lagrangian state estimation is systematically more accurate than the Eulerian method. This improvement stems from a more suitable formulation of the same problem. The numerical approach of the Lagrangian model is more accurate than its Eulerian counterpart. The “non-model-switching” discretization scheme for the Lagrangian model also improves the data-assimilation method. As a result, more accurate results are obtained.

Furthermore, the simulation study proves that the proposed Lagrangian state estimator combining with node model can provide sufficiently good traffic state estimates over

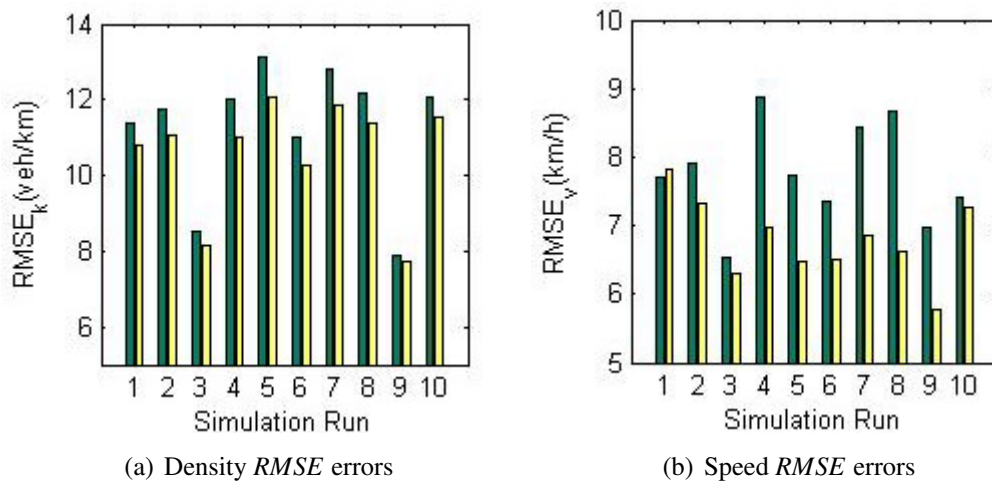


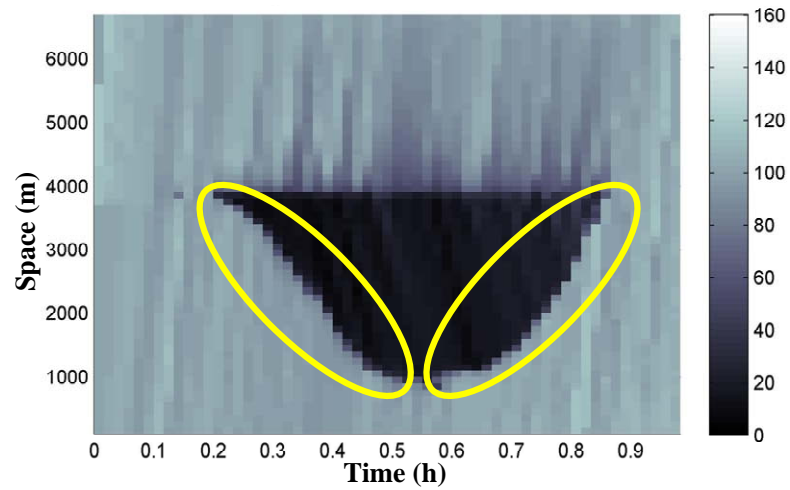
Figure 5.6: Error comparison between two methods for each of the ten simulation runs using Eulerian sensing data (500 m apart). Green: Eulerian state estimation. Yellow: Lagrangian state estimation.

a typical on-ramp and off-ramp freeway stretch. In the Lagrangian formulation, the node model takes effect mainly at the prediction step by effectively changing system states (spacing) over nodes, whereas the correction procedure of the state estimation remains the same. Therefore, traffic characteristics (traffic states, errors) can be easily linearised at nodes. This implies that large scale state estimation of traffic networks on the basis of the Lagrangian formulation becomes possible; the advantages of the Lagrangian formulation over its Eulerian counterpart remain at a network level.

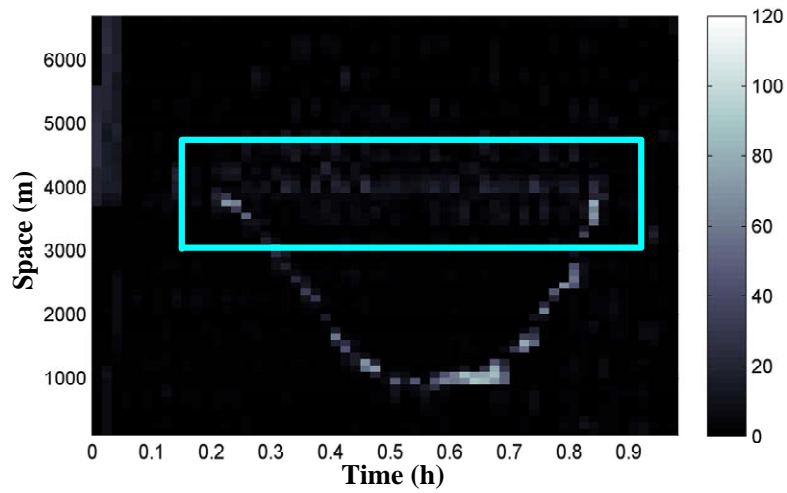
Qualitative analyses

Figure 5.7 shows a graphical representation of the simulation results. The reference speed map from one simulation run is presented, followed by the related error comparison maps from the two methods. The first observation is that the node model in Lagrangian coordinates does induce slightly larger errors than the one in Eulerian coordinates, which is visible by the (faint but visible) grey horizontal line (see the highlighted block) in subfigure 5.7(c) directly downstream of the bottleneck. These errors are barely present in the Eulerian case, see subfigure 5.7(b).

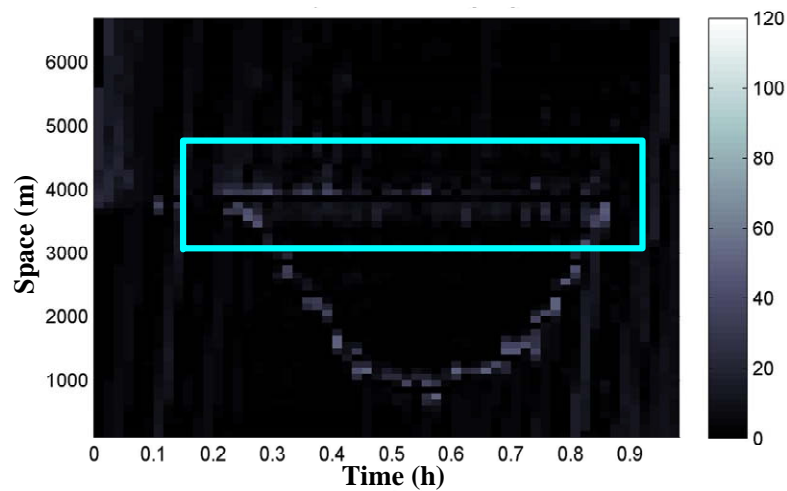
However, the largest errors occur at the upstream boundaries of congestion (see the highlighted circles in subfigure 5.7(a)), where traffic states switch frequently between freely flowing and congestion. Clearly, here the Eulerian method makes larger errors than the Lagrangian method. This can be attributed to the mode-switching problem as discussed earlier, in that the linearisation of Eulerian traffic model around those congestion-or-not boundaries leads to errors of “wrong” sign. These errors may also be partially due to larger numerical diffusion in the spatial dimension. In the Lagrangian estimation method, the errors at these upstream boundaries are much smaller than the Eulerian case.



(a) Reference speed map



(b) Eulerian speed error map (Compared with the reference)



(c) Lagrangian speed error map (Compared with the reference)

Figure 5.7: Reference speed map with the related error maps of the speed estimates from both Eulerian and Lagrangian approaches, taken from one simulation run for scenario 1 (in km/h).

5.4.4 Conclusion

This section has shown that, by incorporating node models, the Lagrangian traffic state estimator can be extended for generic freeway stretches/networks state estimation. The performance in the test case on the basis of synthetic data was consistently better than its Eulerian counterpart (recall Section 5.3). These results may be due to improvements in both the prediction step as well as the correction step of the EKF method. In accordance to earlier studies, the Lagrangian formulation enables more accurate and efficient simulation of freeway traffic and thus leads to more accurate predictions. Due to the non-mode-switching numerical solution, the Lagrangian method essentially is a more appropriate choice for the application in the EKF. In that linearisation of the process model around capacity is a much better approximation in Lagrangian coordinates than in Eulerian coordinates. In the latter case, it may lead to sign errors, and in the former the errors pertain to the magnitude of the corrections only. In our results, this led to improved performance, particularly at the upstream boundaries of congestion. The gain in accuracy at the upstream boundaries coincided with a slight (smaller) loss of accuracy at the downstream boundary (i.e., at the bottleneck itself). This implies that the current implementation of the node model may need improvements.

5.5 Verification of multi-class Lagrangian traffic state estimation and comparison with its mixed-class formulation

The previous sections focused on testing and demonstrating the advantages of the Lagrangian traffic state estimation at a mixed-class level. The next two sections concentrate on the multi-class level Lagrangian state estimation. To simplify, only two vehicle classes, namely car and truck classes, are identified from the traffic flow. At first, this section verifies the multi-class Lagrangian traffic state estimation. First of all to verify the proposed estimator with diverse types of observations, a simulation study is conducted based on the synthetic data from MATLAB-based simulations. In this study, we further explore the characteristics of the multi-class Lagrangian traffic state estimation and its advantages over the mixed-class counterpart.

5.5.1 Data and test network

For verification purpose, the experiments were based on the synthetic data generated by the multi-class (two distinguished vehicle classes: cars and trucks) Lagrangian process model in the MATLAB environment. The logic is to let the multi-class Lagrangian traffic flow model, which is also used as the process model in the state estimator, provide ground-truth data out from the so-called reference simulation runs. The related

parameters used in fundamental diagrams, namely free-flow speed for car class $v_{\text{car free}}$, free-flow speed for truck class $v_{\text{truck free}}$, critical speed v_{cri} , critical spacing s_{cri} and jam spacing s_{jam} , were set as, 120 km/h, 85 km/h, 80 km/h, 40 m/PCE/lane and 7 m/PCE/lane, respectively. The platoon size Δn in these reference runs was 1 vehicle, which implied a comparable microscopic simulation. The time step Δt was 0.2 s, which allowed reference runs generated vehicle trajectories with a high resolution. Based on the information of all the “individual vehicles”, we can emulate any type of observation data and define different experimental scenarios.

An initial condition (Riemann) problem was simulated in reference runs. At the initial instant, there was a temporary bottleneck (e.g., temporary road blockage) in the middle of a two-lane freeway stretch of 5 km in length. This bottleneck generated congestion, which was bounded by a backward forming and a backward recovery wave. The overall simulation period of this Riemann problem was 600 s, which allowed that the congestion emerged and eventually dissolved. Note that, in practice, data samples from loops are averaged over 60 s, for instance in the Netherlands. However, the time span of the simulation period (600 s) allowed only 10 observation instants. Therefore, the measurement interval ΔT for loops was adjusted to 10 s to enable more observations. To obtain statistically-reliable results, ten reference runs with diverse truck shares varying from 5% to 15% were conducted.

5.5.2 Experimental scenarios

For fair comparison between the mixed-class and multi-class Lagrangian state estimators, the sizes for both time and vehicle discretisations were set as the same in the two cases. The time step Δt and the platoon size Δn were chosen as 1 (s) and 3 (veh.), respectively. The Gaussian noise parameters were also set as the same for the two formulations.

Both the Eulerian (loop) observations and the Lagrangian (floating car data - FCD) observations were used. These two types of data scenarios were studied independently. The first data scenario used speed observations from loop detectors, there were three spatial resolutions investigated, namely 300 m, 500 m and 800 m. It provided harmonic (space) mean of individual speed samples over the interval ΔT of 10 s (as explained above to enable more observations). For multi-class state estimation, we further calculated class-specific mean speeds as observation input. The other one used FCD data. To simplify matters, the reporting frequency (ΔT) was chosen uniformly as 5 s. The penetration rates varied from 2% to 20%. It provided instantaneous speeds and locations of class-specific probe vehicles at polling intervals. We additionally assumed that the class-specific distance headways (spacing) were available from probe vehicles. This can be used in the multi-class Lagrangian state estimation as a direct observation of the system state (s).

In addition, we performed a sensitivity analysis with respect to observation inputs and model inputs. More specifically, the performance of the multi-class state estimation

Table 5.8: Performance of mixed-class traffic state estimation with respect to different truck shares, using loop data (300m)

Case	Truck share	$RMSE_k$	$MAPE_k$	$RMSE_v$	$MAPE_v$
1	5%	2.9	3.4	2.8	1.3
2	7%	4.2	5.5	3.8	1.7
3	8%	4.6	6.3	4.0	1.8
4	9%	4.6	6.5	3.7	1.6
5	10%	5.3	7.4	4.4	2.1
6	11%	5.4	7.5	4.2	2.3
7	12%	5.7	8.1	4.4	2.7
8	13%	5.6	8.3	3.9	2.4
9	14%	6.5	9.3	4.5	2.7
10	15%	7.0	9.7	4.7	2.8

in terms of different types of errors (mainly two: noise or bias) in the input data (\mathbf{Y} and \mathbf{d} , refer to Figure 6.1) was tested with three types of scenarios. First, random errors (Gaussian noise) were added into observation inputs. Second, truck shares, as one of the model parameters, were combined with structural errors (biases). Third, fundamental relations were added with structural errors, by defining biased parameters (critical/jam spacing).

5.5.3 Results and discussion

Quantitative analyses

The first experiment is to show why we benefit from a multi-class state estimation approach instead of a mixed-class one. We therefore test the performance of a mixed-class traffic state estimator with respect to different traffic compositions from the ten reference runs. As illustrated in table 5.8, as the truck share increases, the level of estimation performance decreases monotonously. In this experiment, the fundamental diagrams are set the same as the reference runs in all cases. As a result, the change in the traffic composition (truck share) is not captured by the mixed-class traffic state estimation model. This explains the unsatisfied performance. In fact, traffic-mix information can be adapted into the fundamental relations. However, there is no straightforward way to directly connect the information of truck share to the parameters of fundamental diagrams. Instead, it requires complex calibration procedure. Under this circumstance, if class-specific data are available, such as truck shares, multi-class observations, this information can be directly and more appropriately incorporated into the multi-class traffic state estimator. Therefore, in the following experiments, the multi-class traffic state estimator will utilise all the available class-specific information to improve its performance.

Tables 5.9 and 5.10 present the comparison between the mixed-class state estimation and the multi-class state estimation using both Eulerian and Lagrangian observations. As the first important result, the multi-class Lagrangian traffic state estimation has been verified within this experiment. In that the estimator succeeds to employ the observation data generated by the multi-class system (process) model, to reconstruct the traffic conditions, which are similar to the ground-truth provided by the same system model. In both mixed-class and multi-class cases, the more observation data that are available, the better accuracy that can be achieved in the estimates. Obviously, the multi-class traffic state estimation systematically provides more accurate results than the mixed-class method, with most positive values of *POI*. This result relies on a better utilisation of the class-specific information. When the Eulerian sensing data are used, we notice that the estimation errors by the multi-class model in the scenario of the sparse loop observations (800 m) are slightly bigger than the mixed-class case. This indicates that for multi-class state estimation, sufficient class-specific inputs should be provided to ensure its performance, as the current model also entitles more detailed outputs compared with the mixed-class case. In the case of using the Lagrangian sensing (FCD) data in the multi-class model, the performance with speed observations is generally better than spacing observations. This is reasonable since here all the simulations apply the same noise parameters, whereas the noise levels of observations are different. Spacing observations usually contain sampling errors (similar to 1-min loop-flow detection) and these errors are even larger in the free-flow state, while speeds are more stable observations than spacing. So it is possible that alternative choices for the noise parameters of spacing observations could deliver better results.

Table 5.11 presents the performance of the multi-class Lagrangian traffic state estimation with different types of input errors. This is a sensitivity study with respect to different data inputs. The results from the so-called “normal” scenarios (using the default parameters from one reference run: case 15 % truck share) are used as a benchmark. This study indicates that state estimation with noisy observations (both from loops and FCDs) can still provide adequately good estimates (compared with the normal case). The influence of noisy observations can be eliminated in the filtering procedure by appropriately defining error levels for observations. Observations with high error levels will be put a small weight when calculations are processed.

If the input of truck shares is biased, this results in a biased class-specific inflow from the origin. The speed and especially the spacing of the truck class are generated differently (inaccurately) from the reference run (or the normal scenario). Nonetheless, this influence can be compensated by the correction with (speed) observations from loops and probe vehicles. The results shows that state estimation using speed observations provides comparable results as the normal case, under the conditions of wrong truck shares. However, the data scenarios using spacing observations (from probe vehicles) is rather limited compared to the other two. Somehow the correction with spacing observations fails to correct the strong-bias in truck spacing.

The performance of the state estimation is unsatisfied with biased parameters in fun-

Table 5.9: Performance comparison between mixed-class (a) and multi-class (b) traffic state estimation, in terms of loop observations. Each value is averaged over ten simulation runs (ten reference cases). *POI* denotes percentage of improvement.

Observations		Loop ($\Delta T = 10$ s)		
		300m	500m	800m
		v	v	v
a	$RMSE_k$ (veh/km)	5.2	6.0	6.1
b		2.9	4.5	9.0
	<i>POI</i> (%)	44.3	24.0	-
a	$MAPE_k$ (%)	7.2	7.3	7.3
b		2.1	3.3	4.7
	<i>POI</i> (%)	70.8	54.8	-
a	$RMSE_v$ (km/h)	4.1	5.5	6.2
b		2.9	4.1	8.3
	<i>POI</i> (%)	28.1	24.7	-
a	$MAPE_v$ (%)	2.1	3.4	4.6
b		1.4	2.5	6.5
	<i>POI</i> (%)	36.5	25.4	-

Table 5.10: Performance comparison between mixed-class (a) and multi-class (b) traffic state estimation, in terms of FCD observations. Each value is averaged over ten simulation runs. *POI* denotes percentage of improvement.

		FCD ($\Delta T = 5$ s)							
		2%		5%		10%		20%	
		v	s	v	s	v	s	v	s
a	$RMSE_k$	4.6		4.4		4.1		4.0	
b		3.6	5.4	3.3	2.5	2.3	2.4	2.2	1.8
	<i>POI</i> (%)	20.8	-	25.9	43.1	44.7	42.5	46.0	54.0
a	$MAPE_k$	7.0		7.0		7.0		6.8	
b		2.5	4.4	1.8	2.8	1.2	2.2	0.9	1.6
	<i>POI</i> (%)	64.9	-	73.6	60.0	83.0	68.7	86.7	76.3
a	$RMSE_v$	3.3		2.6		2.1		1.7	
b		2.7	4.5	2.7	2.1	1.6	2.0	1.5	1.6
	<i>POI</i> (%)	16.4	-	-	20.9	23.4	2.7	6.9	2.8
a	$MAPE_v$	2.4		1.7		1.3		1.0	
b		1.6	3.3	1.2	1.9	0.7	1.6	0.6	1.1
	<i>POI</i> (%)	30.7	-	31.5	-	42.9	-	40.6	-

damental relations. It is because fundamental ($q-k$ or $s-v$) relations are used for both the process model and the observation model, and in both the prediction step and the correction step. Thus, the performance of the state estimation is quite sensitive to the quality of these relations. Data-assimilation methods can correct random errors (Gaussian white noise in this case) but fail to correct biased (observation and model) inputs. They require unbiased models together with unbiased parameters to ensure satisfied performance. Therefore, a data pre-processing step is needed to clean out biases and structural errors for estimation inputs. This issue will be discussed in the next chapter.

Qualitative results

Figure 5.8 illustrates the performance of the two formulations in terms of the speed estimation, from one test run. The loop detection is used in these two traffic state estimators. Neglecting the errors generated in the warming-up period of state estimation (the dark area at several initial steps), estimation errors mainly occurs at the upstream boundary of the congestion (see subfigures 5.8(b) and 5.8(c)). Clearly, here the multi-class method makes smaller errors than the mixed-class method. This is due to a better description of traffic flow at this transition (free flow to congestion) area. In that multi-class method distinguishes two vehicle classes of different flow characteristics (e.g., free-flow speed), and thus better capture the change of traffic states (speed in this figure).

The mixed-class traffic state estimator further introduces small (faint but visible) errors near the upstream origin (in subfigure 5.8(b)). This is because that the mixed-class state estimator fails to describe the state of inflow with diverse truck demands.

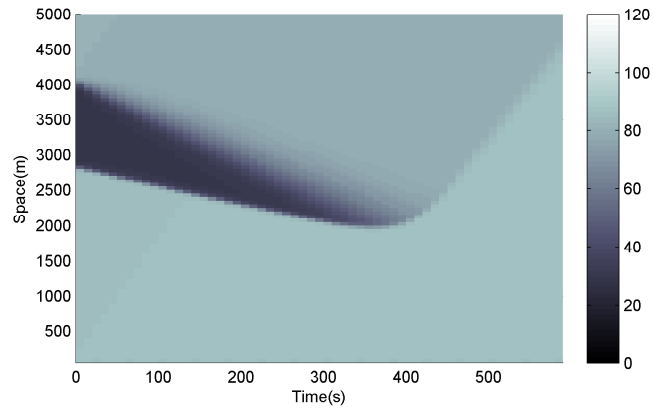
5.5.4 Conclusion

In sum, this section has verified the multi-class Lagrangian traffic state estimation method with a synthetic data environment. The method makes proper use of diverse types of class-specific data to provide more accurate state estimation results, compared with the mixed-class method. The resulting class-specific traffic state estimates would benefit class-specific control applications, such as multi-class ramp metering and speed limit control (Deo et al., 2009), multi-class route guidance (Schreiter et al., 2012).

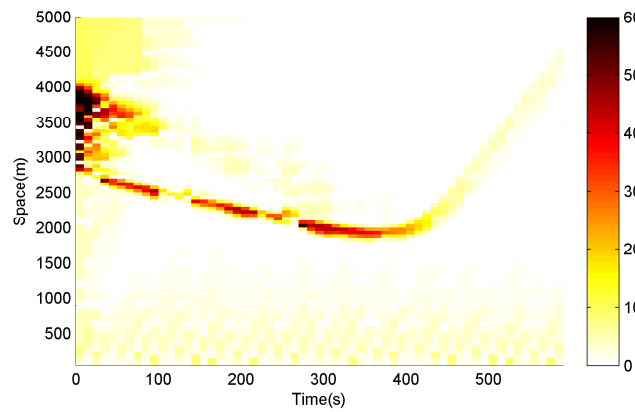
The sensitivity study has also suggested that traffic state estimators can handle with random errors but fail to correct structural errors and biases in estimation input data. This thus justifies that a preparation step in traffic state estimation is required. A reliable data pre-processing can potentially improve the performance of traffic state estimation. This will be elaborated in the following chapter.

Table 5.11: Performance of multi-class traffic state estimation with noise in observations and biased model inputs, for three data scenarios. Normal case: 15% truck share, $s_{\text{cri}} = 40$ m/PCE/lane, $s_{\text{jam}} = 7$ m/PCE/lane. Noise 1~3: noise power 1~3 dBW. Truck 10 %: 10 % truck share. Truck 5 %: 5 % truck share. Diff. FD. 1: $s_{\text{cri}} = 38$ m/PCE/lane, $s_{\text{jam}} = 10$ m/PCE/lane. Diff. FD. 2: $s_{\text{cri}} = 42$ m/PCE/lane, $s_{\text{jam}} = 10$ m/PCE/lane.

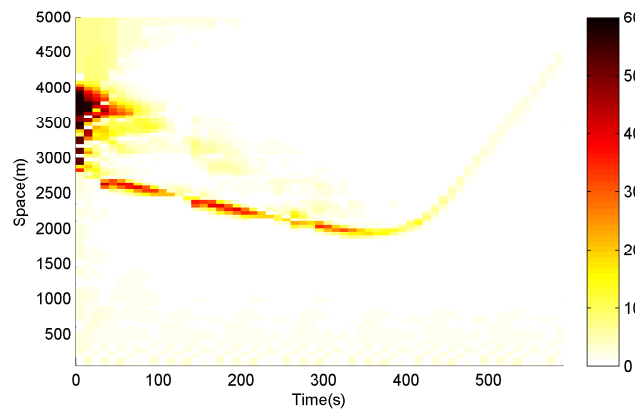
R-Q	Case	$RMSE_k$	$MAPE_k$	$RMSE_v$	$MAPE_v$
Observations: Class-specific speed - Loop 300m					
	Normal	3.5	2.1	3.3	1.6
	Noise 1	3.2	2.3	2.9	1.9
	Noise 2	3.1	2.2	2.9	1.8
	Noise 3	3.2	2.3	2.9	1.9
	Truck 10%	3.3	2.8	3.1	1.9
	Truck 5%	4.2	3.7	4.1	2.4
	Diff. FD. 1	6.4	5.2	5.1	3.6
	Diff. FD. 2	5.8	5.1	3.5	2.5
Observations: Class-specific speed - FCD 5s/20%					
	Normal	2.0	1.2	1.7	0.9
	Noise 1	1.9	2.0	1.6	1.4
	Noise 2	2.1	1.9	1.8	1.5
	Noise 3	2.1	1.8	1.7	1.4
	Truck 10%	1.8	1.2	1.7	1.0
	Truck 5%	1.9	1.3	1.8	1.1
	Diff. FD. 1	3.4	4.6	1.8	1.0
	Diff. FD. 2	4.8	4.2	1.8	1.1
Observations: Class-specific spacing - FCD 5s/20%					
	Normal	2.3	1.6	1.8	1.3
	Noise 1	2.5	2.0	2.0	1.6
	Noise 2	2.2	1.9	1.8	1.5
	Noise 3	2.8	2.1	2.0	1.7
	Truck 10%	5.5	7.2	4.3	5.0
	Truck 5%	5.1	6.7	5.0	5.7
	Diff. FD. 1	2.3	2.0	2.8	3.4
	Diff. FD. 2	3.4	4.2	7.1	9.2



(a) Reference spatiotemporal speed map



(b) Speed error map from mixed-class Lagrangian state estimation (Compared with the reference)



(c) Speed error map from multi-class Lagrangian state estimation (Compared with the reference)

Figure 5.8: Reference speed map with the related error maps of the speed estimates from both mixed-class and multi-class Lagrangian state estimations (using 300m loop detection, truck share: 10%).

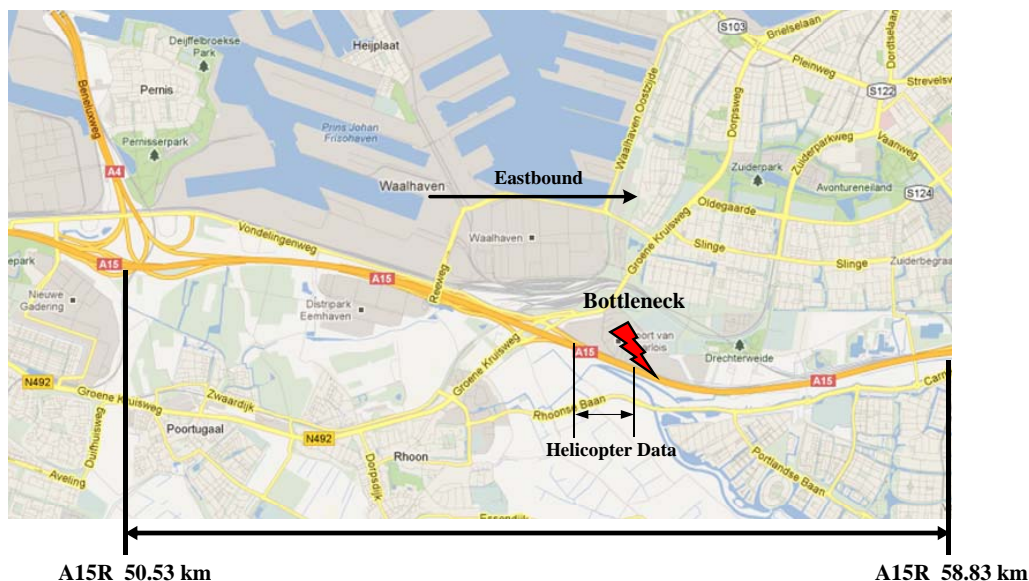


Figure 5.9: Google map picture of the chosen A15 freeway network.

5.6 Multi-class Lagrangian traffic state estimation in a real traffic network

The previous section has verified the multi-class Lagrangian traffic state estimator with a synthetic Riemann problem. It showed that different types of observations can be incorporated into the estimation process independently. In this section, we integrate different observation data sources into the multi-class traffic state estimation by means of the related observation models, and extend our approach to a network level by the node models presented in Section 4.3. This case study is based on a realistic traffic network, a part of the A15 freeway network in the Netherlands.

5.6.1 Data and test network

The data used in this experiment were obtained from a part of the east-bound carriageway of the Dutch freeway A15. The studied time period covered the afternoon peak hours, from 14:00 to 20:00. The chosen road segment is about 8300 m in length, with the milepost A15R-50.53km to A15R-58.83km as illustrated by Figure 5.9. Most of the carriageways in this segment contains three vehicle lanes. There are high truck percentages (higher than 20% sometimes) on the A15, since it is an important freight-transport corridor connecting the harbour city of Rotterdam to the hinterland.

On the A15, dual-loop detectors are located at about every 500 m, providing 1-min aggregate flow and speed profiles. The fundamental parameters were calibrated from these data sets. During October of 2011, several selected detectors, with the mileposts A15R-51.83km, A15R-52.48km and A15R-55.12km, additionally collected individual vehicle data (IVD) so that class-specific data (speeds, vehicle lengths) were available.

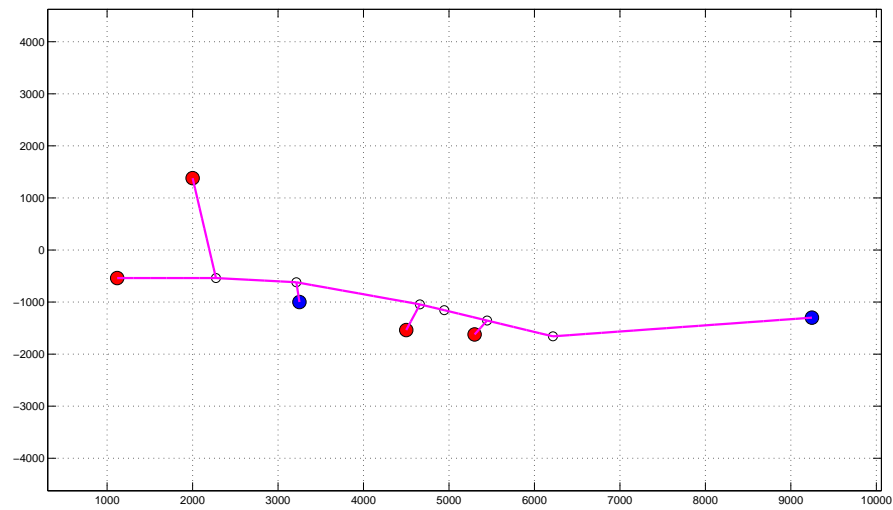


Figure 5.10: A15 freeway section (eastbound) modelling in MATLAB.

On 25th October 2011, a helicopter was used to collect vehicle trajectory data around the bottleneck segment (about 500 m in length, highlighted in Figure 5.9) within the study section during the afternoon peak (starting around 16:15). All the individual vehicle trajectory data (locations, reporting time stamps, vehicle lengths) can be extracted from video footage via image processing techniques (Knoppers et al., 2012). On the basis of these data, we can further interpolate vehicle speeds and vehicle classes, which can be used as detailed observation inputs (also refer to Figure 5.13). In addition, based on the data from the helicopter, we can calibrate some class-specific model parameters, such as driver reaction time and minimum distance headway (e.g., T_u and L_u in equation (4.10)). In principle, the ground truth data are available for the segment covered by the helicopter camera; however, no ground truth information is available for the rest of the road stretch.

This freeway segment contains four on-ramps and one off-ramp, one in-flow and one out-flow boundary, as modelled in MATLAB (see Figure 5.10). The in-flow boundary starts from a loop detector, where the in-flow can be determined. The multi-class node models are used to deal with the network discontinuity.

5.6.2 Experimental scenarios

The purpose of this experiment is to show that the multi-class traffic state estimation can incorporate multiple observation data sources at a network level, and explore its related advantages. Since there are no ground truth data available on the studied A15 section (only for a small segment), this experiment addresses qualitative research.

The time step Δt and the platoon size Δn in the simulation were chosen as 2 (s) and 7 (veh.), respectively. This discretisation choice was assumed to provide sufficient

Table 5.12: Experimental data scenarios on A15

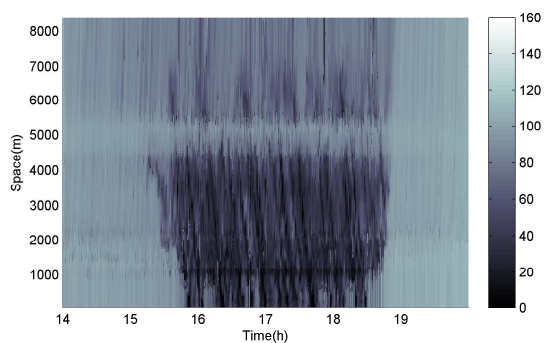
No. Sce.	Loop speed	IVD speed (class-specific)	Trajectories (class-specific)
	Resolution	Three loops	Data usage
6 (3 x 1 x 2)	500m, 1000m, 1500m	All included	Included or not

accuracy for this network application. All three types of observations from the chosen freeway segment (on 25th October 2011), namely aggregate loop data, IVD at selected locations and helicopter trajectory data, were input as observations. Several spatial resolutions of aggregate loops were investigated, namely around 500 m, 1000 m, and 1500 m apart. Table 5.12 overviews all the scenarios in this study. However, aggregate speeds and flows from dual-loop detectors cannot be directly used as class-specific inputs. Therefore, local truck share was assumed for each of these loops to estimate class-specific speed (and spacing). The estimation procedure is illustrated in Appendix A. Generally, spacing observations are also not directly available from those spatially-fixed sensors but can be inferred under an assumption of the local homogeneous condition (speeds of vehicles are constant in a short time period). To demonstrate concepts, we only used speed observations from loops, IVD and trajectory data, and distinguished only two vehicle classes, namely cars and trucks.

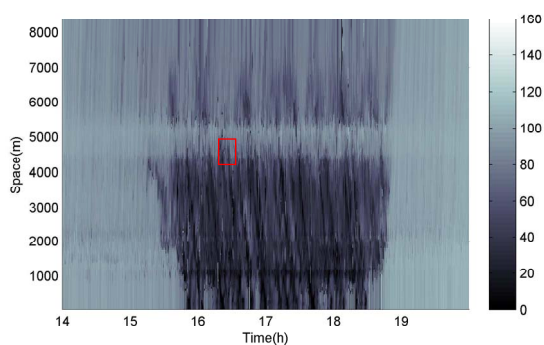
5.6.3 Results and discussion

The experiment represents a qualitative analysis. As shown in Figure 5.11 with all the speed contour plots, all three data sources are successfully incorporated into the multi-class state estimator, providing adequate estimation results. Traffic network discontinuities are sufficiently modelled and simulated. One on-ramp (at 1155 m) and one off-ramp (at 2095 m) are clearly illustrated by two horizontal regions at the locations around 1000 m and 2000 m, respectively. The area between 4000 m and 5000 m is related to a weaving section of four lanes (lane expansion) on the A15. In cases 1 and 2, stop-and-go waves are clearly visible in the congested area. As the detection resolution decreases in cases 3 and 4, the performance is limited compared with the 500 m detection scenario. Nonetheless, the basic traffic patterns and network discontinuities are reflected from the results.

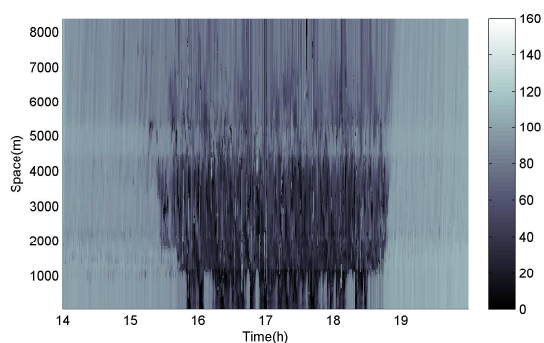
The only difference between case 1 and case 2 is that case 2 additionally incorporates vehicle trajectory data as observation inputs. Comparing Figure 5.11(a) with Figure 5.11(b), it is difficult to differentiate the differences between the two cases. Therefore, we further illustrate the difference in speed estimates between these two scenarios in Figure 5.12. We notice that the estimate difference at certain spatiotemporal cells is even over 100%, and this difference is starting the moment the trajectory data are available as observations. Figure 5.13(a) presents the spatiotemporal span of the collected



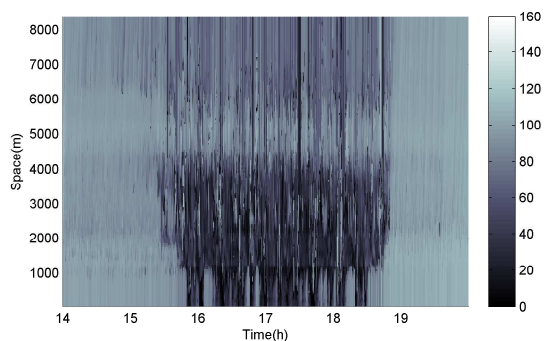
(a) Case 1: 500m loop and IVD data



(b) Case 2: 500m loop, IVD and trajectory data



(c) Case 3: 1000m loop, IVD and trajectory data



(d) Case 4: 1500m loop, IVD and trajectory data

Figure 5.11: Speed estimates from the multi-class Lagrangian traffic state estimation in four different cases. Note that the red rectangle box in (b) indicates the region of the zooming-in snapshot in Figure 5.13(a).

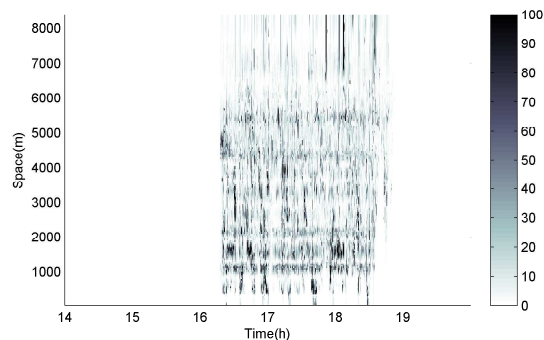


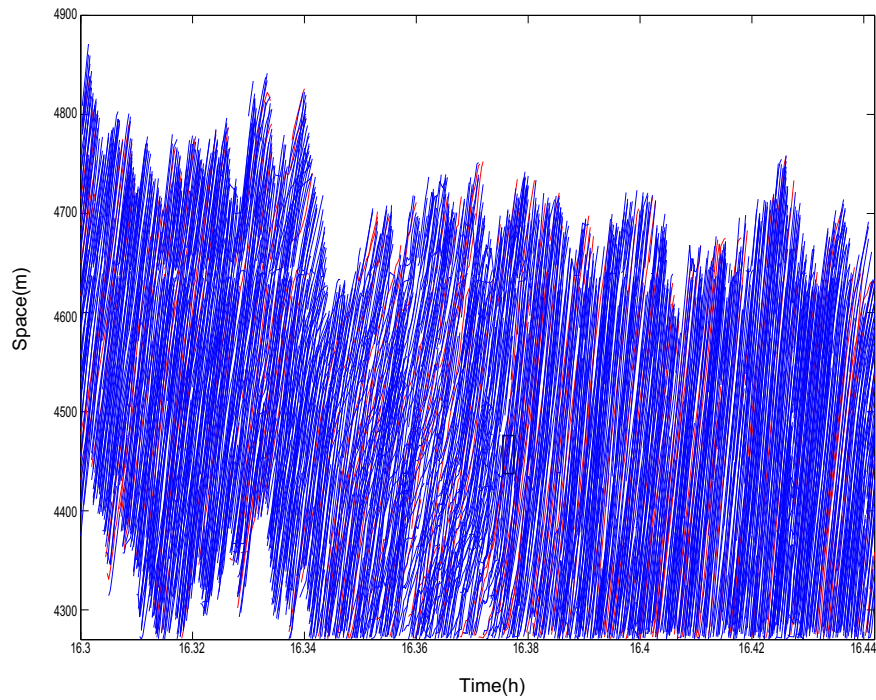
Figure 5.12: Speed estimate comparison between case 1 and case 2 (based on case 2, in %).

trajectory data corresponding to a small region within the whole spatiotemporal map (see Figure 5.11(b)). We further zoom in to a snapshot as shown in Figure 5.13(b). Obviously, class-specific trajectories are distinguished by different colours. These individual trajectories can be directly related to a vehicle (car or truck) platoon (discretised unit in the estimation model). For instance, the space-mean speed observation for a vehicle platoon (indicated by the black dashed box in Figure 5.13(b)) is determined by averaging all the speed samples within its range. As indicated, trajectory data generally provide more detailed information at the collection area compared to loop data. Hence, the estimates in case 2 (also in other cases using trajectory data) are expected to benefit from it.

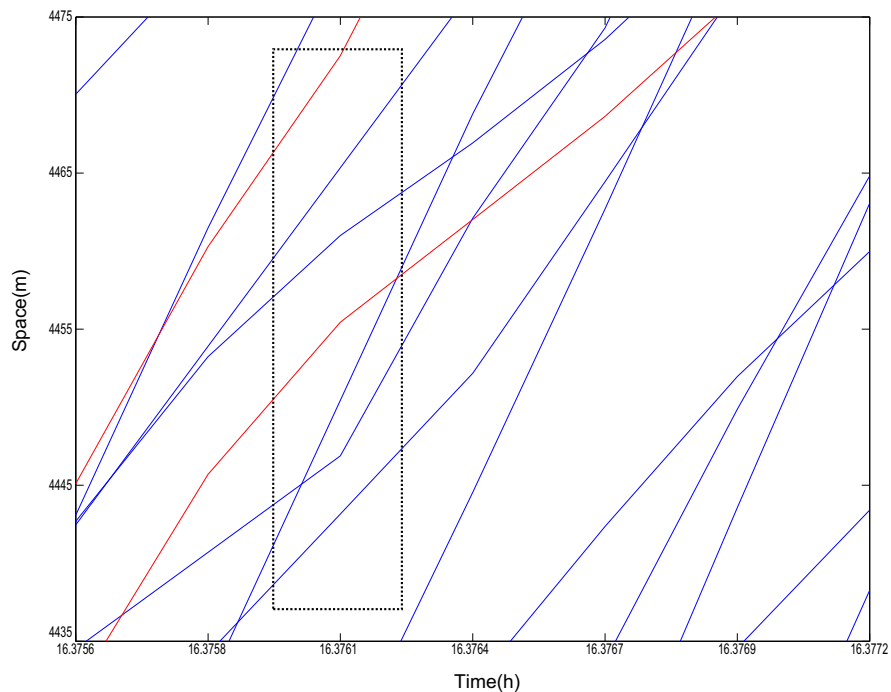
One of the main advantages of multi-class state estimation is to provide class-specific traffic state estimates, such as speeds, densities, and flows. Figure 5.14 shows that speed estimates for both car and truck classes are available from simulations. The speed patterns of the two classes are similar to each other in congestion, whereas in the free-flow state truck speeds are bounded by a maximum value of about 85 km/h. Moreover, traffic compositions over space and time can be accordingly calculated from class-specific densities. As a result, people might monitor the share in total flows of each vehicle class over time at specific locations. Compared to the mixed-class raw-speed observations (Figure 5.14(a)), multi-class estimation offers more useful information, which is important for class-specific traffic control and management.

5.6.4 Conclusion

This section has indicated that the multi-class Lagrangian state estimation approach provides class-specific state estimates in a real freeway network successfully, using all types of traffic observations (in both Eulerian and Lagrangian forms, in both mixed-class and multi-class patterns). By implementing multi-class node models, the state estimator is extended to model network discontinuities and provides adequately good results. It provides the foundation of information for a wide range of class-specific applications for both researchers and practitioners alike.

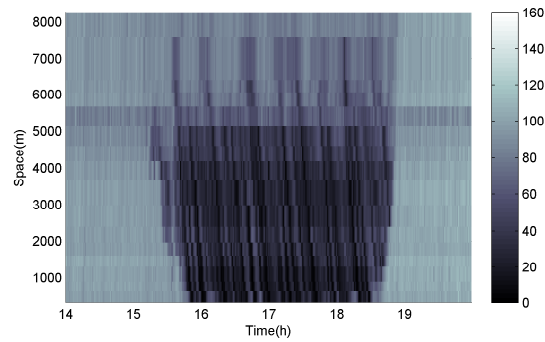


(a) Trajectory data collected from the zooming-in area in Figure 5.11(b). Blue curved lines stand for trajectories of car class, whereas red curves lines stand for trajectories of truck class.

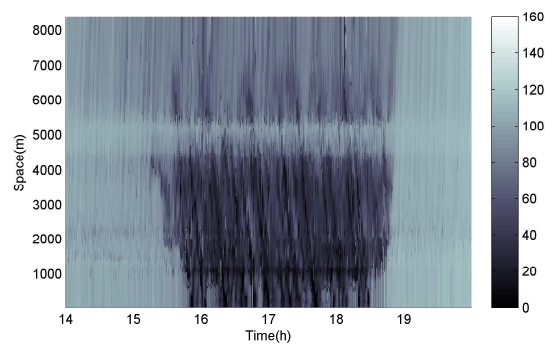


(b) Zooming-in of the dark-blue rectangle area in (a): Example of incorporating trajectory (speed) data into traffic state estimation. The black dashed box indicates one of the discretised units: vehicle platoon.

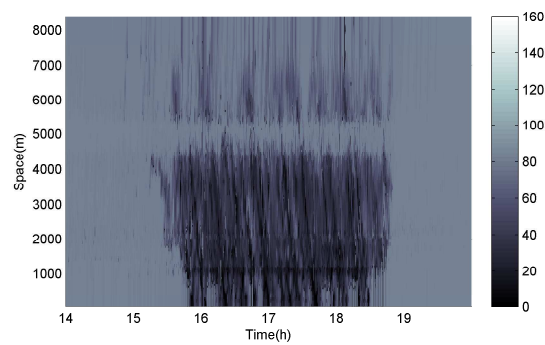
Figure 5.13: Trajectory data collected from a helicopter and their application in traffic state estimation.



(a) (Mixed-class) Speed observations from loops



(b) Speed estimates for car class (case 2)



(c) Speed estimates for truck class (case 2)

Figure 5.14: Class-specific speed estimates from the multi-class Lagrangian traffic state estimation, compared with raw speed loop data.

Table 5.13: Overview of the conclusions for all case studies

Section No.	All the hypotheses can be accepted
5.2	Validating the mixed-class Lagrangian TSE model Well incorporating both Lagrangian and Eulerian sensing data
5.3	Lagrangian formulation of the EKF-based TSE outperforms the Eulerian formulation at a link level
5.4	The advantages of the Lagrangian formulation succeed to extend to a network level
5.5	Verifying the multi-class Lagrangian TSE model Class-specific approach outperforms the mixed-class approach
5.6	Multi-class Lagrangian TSE is applicable in realistic network-wide systems, incorporating all available empirical data sources

5.7 Summary and discussion

In this chapter, we have validated both the mixed-class and the multi-class formulations of traffic state estimation with an EKF technique. It has been shown that both formulations have certain advantages via experimental studies. These experiments cover diverse aspects and test all the hypothesis presented in Section 5.1.2. The hypothesis for each case study can be accepted according to the simulation results. Table 5.13 summarises all the conclusions drawn from each of the case studies. To sum up, both the mixed-class and the multi-class formulation of traffic state estimation with an EKF technique are validated. It illustrates how the multi-class Lagrangian traffic flow model can also be iteratively used for such a fast data-assimilation technique (EKF). With the newly-developed observation models, both Eulerian and the Lagrangian sensing data are well incorporated with the Lagrangian state estimator. Meanwhile, Lagrangian node models sufficiently extend the Lagrangian state estimator to a network level. The experimental results demonstrates that the Lagrangian formulation of traffic state estimation, provides more accurate state estimates than its Eulerian counterpart. This is due to improvements in both the prediction step and the correction step of the data-assimilation method. In the final experiment, the multi-class Lagrangian traffic state estimation has been successfully implemented in a real traffic network, incorporating all available empirical data sources. This offers another efficient and reliable opportunity for real-time traffic state estimation in the traffic management context.

Chapter 6

Preparation for generic Lagrangian multi-class state estimation systems

Previous chapters presented the main methodology and the related validation study for Lagrangian multi-class traffic state estimation. This chapter deals with one of the application aspects in the practical world, which is data pre-processing related to the data inputs used in state estimation. State estimation requires model parameters and observation samples to generate outputs. In general, every traffic simulation model consists of three components: namely model parameters and input, models/algorithms themselves and output. As a consequence of its basic structure, each simulation model will process the most nonsensical of input data and produce nonsensical output. In this chapter, we highlight the importance of the input data from sensors (Section 6.1), presenting two exemplified methods (in Sections 6.2 and 6.3) respectively to process and exploit the raw data for the further state estimation procedure.

This chapter is an edited version of: Yuan, Y., J.W.C. Van Lint, T. Schreiter, S.P. Hoogendoorn, J.L.M. Vrancken (2010) Automatic speed-bias correction with flow-density relationships, in: *Proceedings of the 2010 International Conference on Networking, Sensing and Control (ICNSC)*, Chicago, pp. 1-7 (Yuan et al., 2010), and Yuan, Y., R.E. Wilson, J.W.C. Van Lint, S.P. Hoogendoorn (2012) Estimation of multi-class and multi-lane counts from aggregate loop detector data, *Transportation Research Record*. In press (Yuan et al., 2012. In Press).

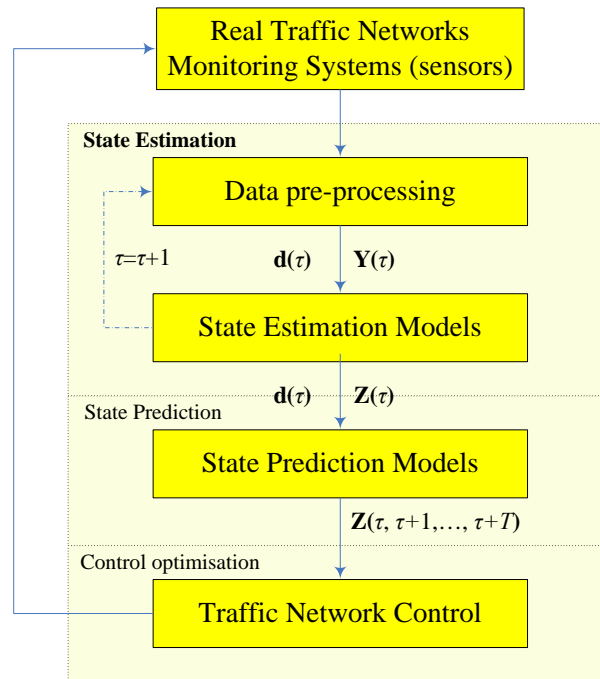


Figure 6.1: Simplified structure of a traffic control loop. Here, τ denotes the index of the individual time interval (current time instant); T denotes the prediction time interval; \mathbf{d} and \mathbf{Y} depict the model input (constraints and parameters) and the observation input, respectively; \mathbf{Z} is the model output.

6.1 Introduction

This chapter studies the data-processing and preparation problems for traffic state estimation. As a simplified version of Figure 1.1, Figure 6.1 illustrates a conceptual structure of a traffic control loop, where model-based state estimation is highlighted. Model-based traffic state estimation is an iterative process, in which at each time instant it requires data from monitoring systems (sensors) to be inputted into state estimation models, to generate estimation output \mathbf{Z} for further state prediction and traffic control. The intermediate step, data pre-processing, serves the purpose of data preparation. Here, two types of input data are distinguished: (i) observation-related data \mathbf{Y} which are used in observation models, and (ii) model-related inputs \mathbf{d} which are constraints and parameters used in both system and observation models as well as data assimilation methods. For successful applications of Lagrangian multi-class traffic state estimation, bias-free and class-specific input data are required. However, these input data are either unreliable or unavailable (directly) from traffic sensors in the practical world.

More specifically, in many countries, traffic network monitoring is largely based on stationary loop detectors (as Eulerian sensors). Unreliable measurements and disturbances are part of the raw data from these sensors. These raw data typically exhibit both random errors (noise) and structural errors (bias). As stated in Section 5.5.4,

although data assimilation techniques are suitable to correct the first category of errors (random errors), they typically fail to correct structural errors. For example, the extended Kalman filter (discussed in chapters 3 and 4) can balance the uncertainty/random noise in the observation data and in the model forecast, by respectively defining error levels (e.g., \mathbf{R} and \mathbf{Q}) for observations and system models. However, observation data (and/or model inputs) with strong bias will lead to inaccurate state estimates. A natural question is that how to get rid of the structural errors in observation data.

Moreover, traffic sensors (dual loops) generally provide only aggregate speed and flow profiles. Certain input data for model-based traffic state estimation are not directly available. These data consist of but not limited to parameters of fundamental diagrams, traffic compositions, traffic demand, boundary conditions, traffic turn fractions, and noise levels. Previous studies have offered lots of opportunities to estimate model-related inputs from sensors. For instance, parameters in fundamental diagrams can be estimated from loop detectors via image processing techniques (Schreiter et al., 2010b) and basic statistics (Cassidy, 1998; Kockelman, 1998; Dervisoglu et al., 2009). Noise parameters (error levels) used in data assimilation methods can also be derived from observation data based on statistical analyses (Wang & Papageorgiou, 2005; Van Hinsbergen et al., 2010). However, to implement a multi-class state estimation (presented in chapter 4), it additionally requires both model parameters and observations with a class-specific distinction. Hence, inputs such as class-specific inflows and observations, truck share, multi-class parameters in fundamental relations, should be obtained in the pre-processing step. Essentially, data pre-processing provides the state estimation step with reliable observations and prior information on parameters. It bridges the gap between (multi-class) traffic state estimation models and empirical data from the real world.

This chapter presents two new approaches for data pre-processing, which are essential to apply the first-order multi-class traffic state estimation in the practical world. The first is a speed-bias correction algorithm as in Section 6.2, which typically corrects the second category of errors (structural errors) in observation-related inputs (\mathbf{Y}). It corrects biased speeds inherited from dual-loop detectors, based on notions from the first-order traffic flow theory and empirical flow-density relationships. The second develops a procedure for estimating multi-lane multi-class counts from a variety of standard aggregate loop data formats from around the world, as in Section 6.3. It provides multi-class state estimation with both types of input data (\mathbf{d} and \mathbf{Y}). In that the estimates at a specific location can be transformed as truck share, class-specific inflows and observations (flows). The estimate rules can be even generalised to another location of similar traffic and infrastructure patterns to provide class-specific input. The estimation procedure involves the inference of multi-linear regression laws that relate multi-lane multi-class data to standard aggregate formats. The regression laws themselves then need to be trained with small samples of individual vehicle data on a site-by-site basis.

6.2 Speed-bias correction

This section presents a speed-bias correction algorithm. After background introduction, we will briefly review the previously proposed methods to eliminate this bias and derive the new method. Next we will present the setup of a simulated experiment on the basis of which the new method with two variants are evaluated and discussed. This section will close with conclusions and recommendations for further research.

6.2.1 Problem analysis and background

Traffic data collected by traffic sensors are critically important for many applications, both for online purposes (traffic information, management and control) and for offline use (travel time estimation, model calibration and validation, incident/bottleneck detection and ex-post policy evaluation). Traffic speed is one traffic variable that can be measured by many different sensors, both locally (with e.g., inductance loops and radar technology) as well as over short or longer distances (e.g., automated vehicle identification systems or floating car data). In particular, knowledge of the speed profile and/or the flow profile is useful to determine bottleneck locations, physical extent of queues, travel times, turn fraction, fundamental relations, and many other parameters used for traffic control systems. This also serves as the essential input for model-based traffic state estimation and prediction.

In many countries, such as the US, the UK, the Netherlands, France, Germany, Spain, and Italy etc., the traffic network monitoring is largely based on the stationary loop detectors. Single loops (mainly implemented in the US and in the urban environment) provide flow and occupancy information, which in turn can be used to estimate speeds (using average vehicle length). However, these estimates are quite noisy due to unobserved variables (e.g., vehicle length, traffic density) and measurement errors. Many researchers have achieved better estimates of speed from single loop detectors (Dailey, 1999; Coifman, 2001; Jain & Coifman, 2005; Coifman & Kim, 2009). Additional to flow and occupancy, dual loops also collect (average) speeds of vehicles. The Dutch freeway network in the Netherlands, for example, is monitored by dual loops which are located at about every 500m on the western part of the Dutch freeway network. 5-10% on average of the available data from this dual loop system (named MoniCa data) are missing or otherwise deemed unreliable (Van Lint & Hoogendoorn, 2009). Similar traffic monitoring systems, which are equipped with inductance loops, can be found in England (Highway-Agency, 2012) and in Germany (Schönhof & Helbing, 2007).

Locally measured and averaged speeds are not necessarily representative for spatial properties of traffic streams. This is due to two reasons. First of all, a local average speed at some cross section x_i over a time period ΔT is equal to the space-mean speed v_M on a road section m (including x_i) over the time period ΔT only, in case traffic conditions are homogeneous and stationary over the section m during ΔT (which does

not mean that all vehicles drive with the same speed). Secondly, the former holds only in case speeds of passing vehicles are averaged harmonically, and not arithmetically with $v_L = \sum v_i/N$, where N depicts the number of observations in the time period ΔT . The latter average (often referred to as time-mean speed) is biased due to an over-representation of faster vehicles in a time sample. In contrast, the harmonic mean speed

$$v_H = \frac{N}{\sum_{i=1}^N \frac{1}{v_i}} \quad (6.1)$$

essentially represents the reciprocal value of average “slowness” $1/v_i$, i.e., the average time per unit space each vehicle spends while passing the detector. Saving for the error due to the assumption of homogeneous and stationary traffic, this average is an unbiased estimator of the space-mean speed. The equivalence can be easily proven by Edie’s definition of traffic variables (see (Edie, 1965), also e.g., (Leutzbach, 1987)). The relationship between space-mean speed (v_M) and local arithmetic time-mean speed (v_L) can be analytically expressed by the following equation (Wardrop, 1952; Hoogendoorn, 2008):

$$v_L = v_M + \frac{\sigma_M^2}{v_M} \quad (6.2)$$

where σ_M^2 denotes instantaneous space speed variance.

Firstly, the implication of equation (6.2) is that time-mean speed is always equal or larger than space-mean speed. This is a structural difference proportional to instantaneous speed variance. Secondly, the effect of arithmetic time averaging cannot be “undone” afterwards using equation (6.2). There is no straightforward (analytical) formula to derive space-mean speeds from time-mean speeds directly, since the bias depends on a second unknown quantity (instantaneous space speed variance). Many empirical studies have revealed that the bias (σ_M^2/v_M) is significant and may result in speed overestimation of up to 25% in congested conditions, and even larger errors in for example travel times or densities derived/estimated from these speeds (Wardrop, 1952; Stipdonk et al., 2008; Stipdonk & Postema, 2009; Lindveld & Thijs, 1999; Van Lint & Van der Zijpp, 2003). A typical result from using arithmetic time averaged speeds in estimating densities (via $k = q/v_L$) is that the estimated traffic states $\{q, k\}$ (correct flow with biased density) do not lie on a nice semi-linear line in the congested branch of the flow-density diagram. Instead, they scatter widely below this line, resulting in a “P-shape” distortion of the estimated fundamental (q - k) diagram (Stipdonk et al., 2008), where the congested phase line looks like a convex curve as the upper part of the letter “P”. This effect has also been analytically proven in (Stipdonk & Postema, 2009). Additionally, the biased speeds may cause underestimates of route travel times (Lindveld & Thijs, 1999; Van Lint & Van der Zijpp, 2003) and inaccurate estimation of traffic states (e.g., densities).

Clearly, in case monitoring systems (such as the Dutch MoniCa system, the British and the German highway monitoring systems) collect time-averaged speeds, there is

a need for algorithms and tools, which are able to correct the inherent bias. The next subsection will first overview the approaches proposed in literature. Naturally following, a new algorithm is presented, based on the flow-density diagram and notions from the traffic flow theory.

6.2.2 Overview of speed-bias correction algorithms

Based on equation (6.2), there are two ways of correcting the bias. One could consider the bias term $B = \sigma_M^2/v_M$ as a whole entity, that is

$$v_M = v_L - B \quad (6.3)$$

where B could be a constant or some functions of quantities that are available (e.g., $B = \mathbf{B}(q, k, v, \dots)$). One can also solve v_M from equation (6.2), which gives us

$$v_M = \frac{1}{2}(v_L + \sqrt{v_L^2 - 4\sigma_M^2}), \quad \sigma_M \leq \frac{1}{2}v_L. \quad (6.4)$$

If the instantaneous space speed variance (σ_M^2) is known or estimated, the space-mean speed can be deduced accordingly. In the following, we will recall the previous research within these two categories.

The simplest method to correct speed-bias is to describe the space-mean speed v_M as a function of the time-mean speed v_L . Essentially with equation (6.4), we can consider the speed variance (standard deviation) is equal to a certain percentage of the time-mean speed, that is $\sigma_M = \beta v_L$, with β in the order of 0.05-0.3 (5-30%, according to (Lindveld & Thijs, 1999)). Lindveld & Thijs (1999) also illustrated that β is not a constant, but rather a function of traffic flow and speed. This relationship is derived using empirical data, in which the ratio of standard deviation over mean speed increases with decreasing mean speed.

Van Lint (2004) has proposed two other methods for estimating instantaneous speed variance in equation (6.4) on the basis of empirical individual data collected along the A9 freeway in the Netherlands in October 1994. Figure 6.2(a) illustrates these data and shows that there appear two distinct regimes, thus these data can be fitted by a bi-linear function (the solid line). For low (time-mean) speeds (i.e., in congestion), speed variance appears constant (although noisy) with a mean in the order of $5^2(\text{m/s})^2$. Above 80 km/h (i.e., in free-flow conditions), speed variance seems to steeply increase with time-mean speeds. It is because in congestion vehicles are constrained in their choices of speeds by another vehicle. The predominant cause of speed variation over space will be (collective) acceleration and deceleration e.g., due to passing stop-and-go waves. Under free-flow conditions, speed variance will logically increase due to (i) traffic heterogeneity (trucks versus person cars) and (ii) the fact that samples (number of vehicles per space and time unit) will become increasingly smaller with decreasing density (and higher speeds). Figure 6.2(b) shows a second scatter plot of instantaneous speed vari-

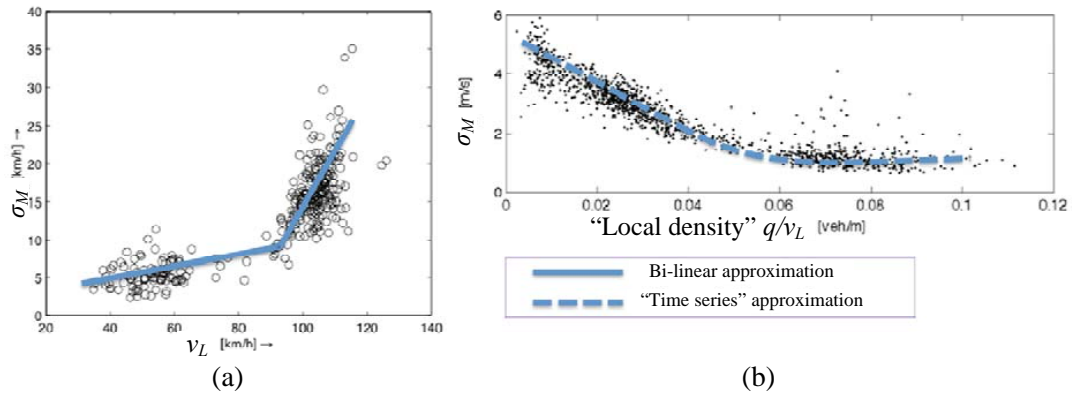


Figure 6.2: (a) Scatter plot of instantaneous speed variance versus time-mean speed, based on individual data collected at a single detector along the A9, 25/10/1994, and the bi-linear approximation between the two quantities; (b) Scatter plot of instantaneous variance versus “local density” ($k_L = q/v_L$) and “time series” approximation, both adapted from (Van Lint, 2004).

ance and “local density” k_L , which is flow divided by time-mean speed (q/v_L). The second approach uses an approximation of this relationship, illustrated by the dashed line in Figure 6.2(b). This method exploits the time-series relations between the instantaneous speed variance and the variance of time-mean speed (referred to as SVE-TS (Speed Variance Estimation - Time Series) further below). Note that this procedure contains three parameters (P is the size of consecutive time windows, k_L^{cri} denotes the critical local density, and γ is a scaling parameter), which need to be tuned with respect to the site-specific characteristics. More details can be found in (Van Lint, 2004).

Recently, Soriguera & Robusté (2011) have also developed a method to estimate speed variance based on the probabilistic theory. However, the method requires the input of speed stratification (vehicle counts over defined speed ranges), which is commonly available from Spanish traffic monitoring systems but not around the world. Therefore, the application of this method is rather limited.

The methods outlined above essentially correct speed bias based on local relationships found in the available detector data. Alternatively, one could estimate speed-bias also on spatiotemporal relationships found in the data, using for example the fundamental diagram, which relates average flow (a local quantity) to average density (a spatial quantity). In (Jain & Coifman, 2005), for example speed estimates from single loops are validated using flow-occupancy relationship. Coifman (2002) also exploits basic traffic flow theory and spatiotemporal propagation characteristics of traffic patterns to estimate link travel time using local dual-loop data. Although these methods do not apply to the problem of estimating the bias in time-mean speed directly, we are motivated for an alternative method to solve the problem.

6.2.3 A new correction algorithm based on flow-density relations

According to the kinematic wave theory (Lighthill & Whitham, 1955; Richards, 1956), perturbations (low speeds, high densities) propagate through a traffic stream at a speed equal to df/dk , with $q = f(k)$ depicting an equilibrium relationship (the fundamental diagram) between average flow and density. This characteristic speed is typically positive in free flow and negative in congested conditions. According to some authors (Newell, 1993; Daganzo, 1994; Kerner & Rehborn, 1997; Windover & Cassidy, 2001; Treiber & Helbing, 2002; Van Lint & Hoogendoorn, 2009), a simple still reasonable approximation is to assume only two main characteristic speeds, one for congested traffic and one for free-flowing traffic respectively. This results in a triangular flow-density relationship (see Figure 6.4), which reads

$$q = f(k) = \begin{cases} v_{\text{free}} \cdot k, & k \leq k_{\text{cri}} \\ q_{\text{cap}} + v_{\text{cong}} \cdot (k - k_{\text{cri}}), & \text{otherwise} \end{cases} \quad (6.5)$$

where q_{cap} is the capacity flow, and v_{free} and v_{cong} respectively depicts the characteristic propagation speed in free-flow and congested conditions. Note that v_{cong} is often parameterized with $v_{\text{cong}} = -q_{\text{cap}}/(k_{\text{jam}} - k_{\text{cri}})$, where k_{jam} and k_{cri} depict the jam density and critical density respectively.

Figure 6.3 presents a typical speed contour taken from a Dutch freeway, in which those approximate constant characteristic propagation speeds can be identified. In congestion (low speeds), for example, perturbations in a traffic stream move upstream with remarkably constant speeds, illustrated by the thick dark stop-and-go waves (low speed areas) in Figure 6.3, which propagate upstream over the entire freeway stretch of 8 km. Note that inside the speed waves themselves, the individual vehicle speeds are not uniform.

Using the same detector data, one would expect these phenomena to translate into traffic states on a straight line in the q - k diagram. As already mentioned above and in (Stipdonk et al., 2008; Stipdonk & Postema, 2009), the straight or semi-straight line for congested traffic is not observed when we use “local density” $k_L = q/v_L$ as a proxy for true (but unobserved) density k . Instead, in that case, we see a “P-shape” distortion, attributed to the non-zero bias term (σ_M^2/v_M in equation (6.2)). We can, however, with the assumption of a (approximately) straight congested branch of the fundamental diagram, estimate and correct this error in density, and as a result correct the bias in speed.

To this end, a few more conditions need to be met. First of all, if we assume that in congestion the true traffic state $\{q, k\}$ lies on a straight line with slope equal to the propagation speed v_{cong} , it is required to estimate this parameter. For instance, the estimation can be based on spatiotemporal plots (Figure 6.3) via e.g., image processing techniques (Schreiter et al., 2010b). Second, we need to assume that the measured flows are unbiased, in which case correcting $k_L (= q/v_L)$ boils down to estimating the error in v_L , which of course equals to σ_M^2/v_M . Figure 6.4 illustrates the basic correction principle.

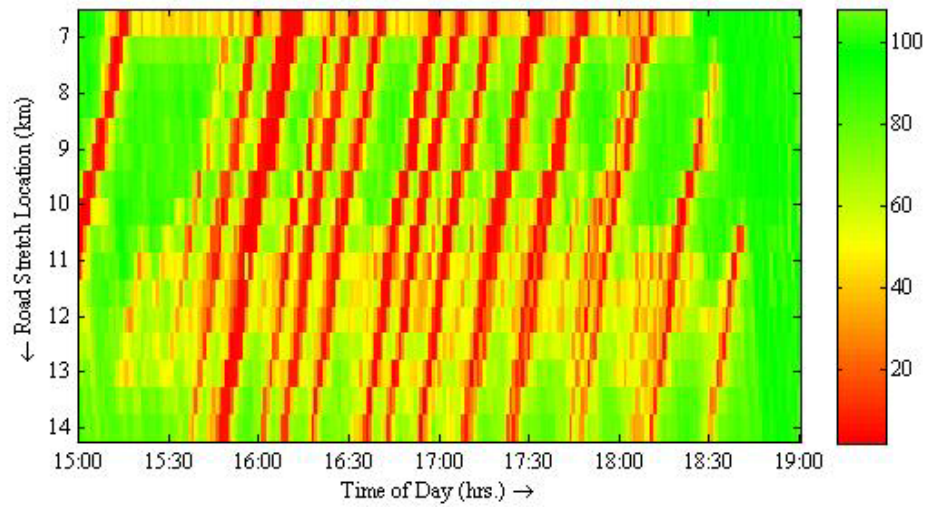


Figure 6.3: Speed contour plot measured at the Dutch A13 freeway southbound between Delft North (km6.5) and Rotterdam Airport (km14.5) on Monday 11/6/2009 from 3:00 PM to 7:00 PM. Colors correspond to one-minute arithmetically averaged speeds.

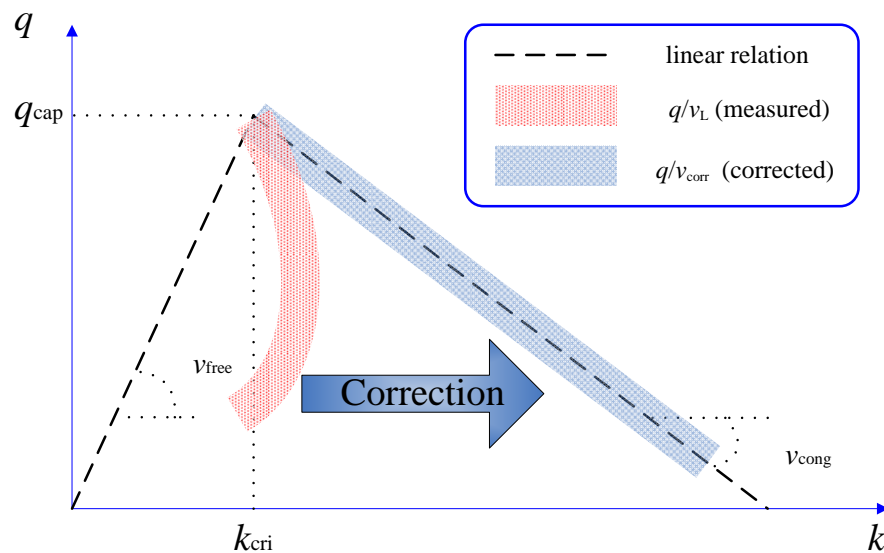


Figure 6.4: Schematic of correction principle.

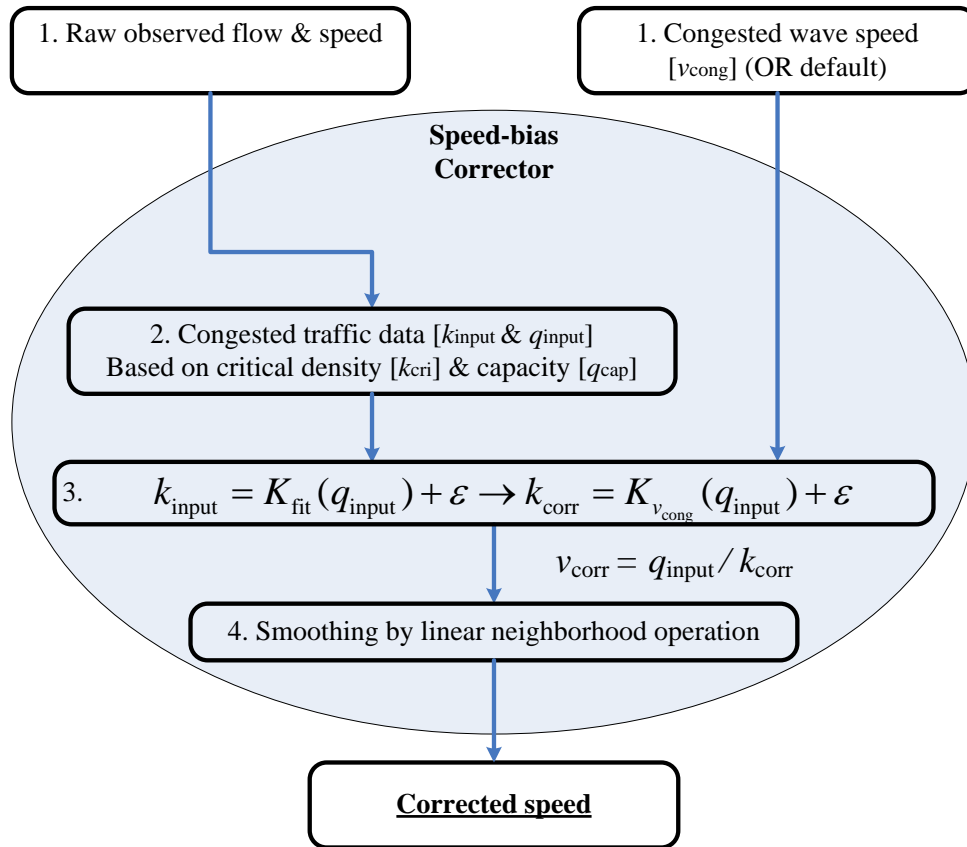


Figure 6.5: Procedure of the speed-bias correction algorithm.

The essence of this method lies in considering the bias term as a whole entity (B , refer to equation (6.3)) to correct local time-mean speed, based on notions from the traffic flow theory and traffic propagation characteristics. The detailed working principle is described as below, followed by a schematic procedure shown in Figure 6.5.

- Data input (The first step presented in Figure 6.5)

The input data are obtained from the monitoring system in the pattern of spatiotemporal ($x-t$) matrix. They are named as the raw observed flow q_{input} and speed v_L . Meanwhile, the propagation characteristic speed v_{cong} needs to be determined as the slope of the congested flow line used in this algorithm. This parameter can be estimated based on the historical data (e.g., speed contour plot) or set as default (e.g., -18 km/h as used in many applications).

- Data selection (The second step in Figure 6.5)

Based on local speed (v_L) and flow (q_{input}) data, the density data are generated (via $k_{\text{input}} = q_{\text{input}}/v_L$). As the focus is on the congested traffic, the free-flow state and the congested state need to be distinguished. In macroscopic traffic flow models, critical density on a road segment discriminates between the free-flow and congested flow conditions. Here, the critical density value (k_{cri}) and/or the capacity value (q_{cap}), which relate to the peak of a $q-k$ diagram, serve the same purpose. k_{cri} and q_{cap} values can be

easily estimated from the input data. The data sets in terms of flow and density, which meet the condition $k_{\text{input}} > k_{\text{cri}}$, are categorised as congested state data.

Additionally, we need to ensure that the data selected (from congested conditions) represent stable traffic states, that is, traffic states which are likely to lie on the (linear) congested branch of the fundamental diagram. This implies we cannot select detector data located downstream (“in”) a bottleneck. At those locations where congestion is resolving, one typically notices that flows are far below capacity with large spacing (low density) due to accelerating vehicles.

- Density-wise correction (The third step in Figure 6.5)

This step is the core of the correction algorithm. The mapping of this operation is expressed by:

$$k_{\text{input}} = K_{\text{fit}}(q_{\text{input}}) + \varepsilon \rightarrow k_{\text{corr}} = K_{v_{\text{cong}}}(q_{\text{input}}) + \varepsilon \quad (6.6)$$

where ε denotes the residual deviation of each density (k_{input}) scatter point away from the fitted curve (K_{fit}) of the congested data, k_{corr} denotes the corrected density values, $K_{v_{\text{cong}}}$ denotes the targeted congested branch.

Given the peak point (fit-origin: k_{cri} and q_{cap}) and congested traffic scatters (k_{input} and q_{input}), fit function (K_{fit}) applies. In q - k space, fit function would fit the congested data, regarding the peak point (k_{cri} and q_{cap}) as the point of origin. Density values (k_{input}) are then expressed as a function K_{fit} of the input flow values (q_{input}) with residual (ε). We then shift these scatters in congested phase part of the q - k diagram from the fitted curve K_{fit} to the targeted congested characteristic line ($K_{v_{\text{cong}}}$).

The fitted curve can be either linear or polynomial. This leads to two variations of the correction algorithm. The linear fit function is defined as:

$$K_{\text{fit}}(q_{\text{input}}) = a \cdot q_{\text{input}} + b, \quad (6.7)$$

with two fit coefficients (a and b). Intuitively, the polynomial fitting could describe the feature of the “P-shape” distortion presented in (Stipdonk et al., 2008; Stipdonk & Postema, 2009) better than the linear one. Here, a quadratic polynomial form of fit function is considered, written as:

$$K_{\text{fit}}(q_{\text{input}}) = c \cdot q_{\text{input}}^2 + d \cdot q_{\text{input}} + e, \quad (6.8)$$

with three fit coefficients (c , d and e). These two variations are referred to as Linear-fit and Polynomial-fit respectively further below.

For the targeted congested characteristic line, if only one degree of freedom for the congested state is considered, the $K_{v_{\text{cong}}}$ is linear featuring by congested propagation speed (v_{cong}), expressed by

$$K_{v_{\text{cong}}}(q_{\text{input}}) = \frac{q_{\text{input}} - q_{\text{cap}}}{v_{\text{cong}}} + k_{\text{cri}}. \quad (6.9)$$

Note that the more degrees of freedom in the targeted characteristic line that are taken into account to capture congestion dynamics, the more accurate correction that will be achieved.

Based on corrected density and the invariant flow values, the corrected speeds are generated directly (via $v_{\text{corr}} = q_{\text{input}}/k_{\text{corr}}$). In fact, the correction presented here is another form of equation (6.3):

$$v_M = v_L - \frac{q_{\text{input}}(k_{\text{corr}} - k_{\text{input}})}{k_{\text{corr}} \cdot k_{\text{input}}} \quad (6.10)$$

where the bias term B is expressed as a function of the corrected density k_{corr} .

- Post-processing (The last step in Figure 6.5)

In some cases, samples with high-speed (e.g. larger than 120 km/h) are identified as erroneous. Moreover, in the speed space, the corrected speed values (v_{corr}) of congested state might contain discontinuity with respect to those in free-flow state, so a linear smoothing can be used here. The smoothing is a neighborhood operation which has already proposed in (Treiber & Helbing, 2002; Van Lint & Hoogendoorn, 2009). The operation for each point is conducted within a small spatiotemporal area based on the values of its neighbouring points to overcome the discontinuity problem. Finally, the corrected speed data are obtained.

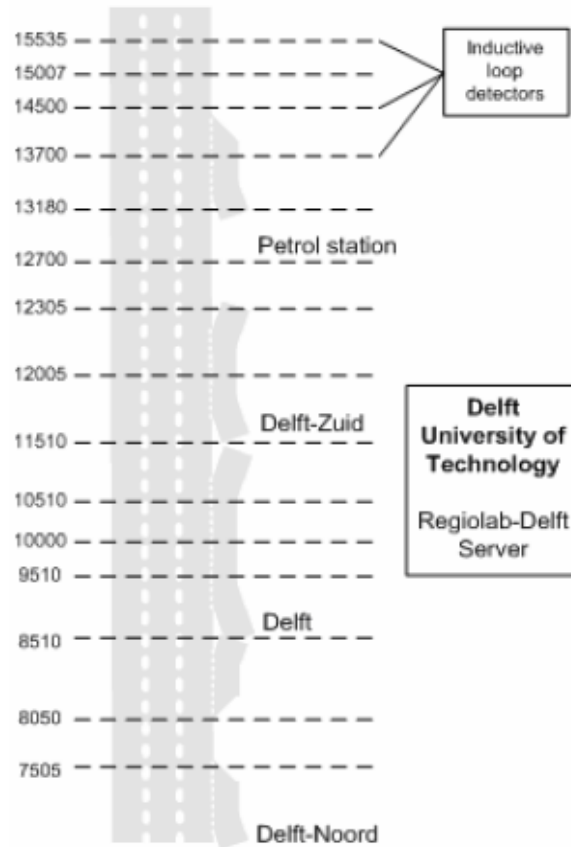
So far the correction procedure has been presented. The next section presents a microscopic simulation study to assess this algorithm.

6.2.4 Validation of the speed-bias correction algorithm

To test and validate the proposed algorithm, a simulation study is performed. The microscopic simulation model ‘‘FOSIM’’ (Dijker, 2012) is used for this purpose, since it provides both (synthetic) detector data as well as ground-truth data (space mean speed). This model is developed at the Delft University of Technology, specially designed for the detailed analysis of discontinuities in freeway networks. It has been calibrated and validated for the Dutch freeways in terms of driving behaviours. From FOSIM, traffic data from any type of detectors can be simulated and obtained. Data processing and analyses are further conducted in MATLAB (7.5.0).

Description of model and data

As mentioned above, the current monitoring system on Dutch motorways consists of dual inductance loops located about every 500 meters, collecting 1-min average time-mean speeds and 1-min aggregate flows. To assess the proposed algorithm, a part of the Dutch freeway A13 (Figure 6.6(a)) is modeled in FOSIM, as shown in Figure 6.6(b). It is a road stretch of 11.5 km in length from Delft North to Rotterdam Airport. The



(a) Dutch Freeway A13 southbound (from Delft North to Rotterdam Airport).



(b) FOSIM model of A13.

Figure 6.6: Illustration of a Dutch freeway A13.

simulation period is during afternoon peak hours from 14:00 to 20:00 (six hours). The traffic volume is defined to match typical pattern of a weekday afternoon, with 10% truck fraction. In FOSIM, we can simulate MoniCa data (time-mean speeds, aggregate flows), within semi-equidistant spatiotemporal regions of size 500(m) x 60(s). The ground truth data (space-mean speeds, that is, harmonic mean speed v_H in equation (6.1)) are derived from FOSIM over equidistant spatiotemporal regions of size 100(m) x 30(s). The data are treated as reference. The raw input for the correction algorithm is then pre-processed to match the same spatiotemporal grid (100(m) x 30(s)) as the ground truth data.

Definition of scenarios

Based on the speed contours from FOSIM via image processing techniques (Schreiter et al., 2010b), a constant value of -21 km/h is estimated as the congested propagation speed v_{cong} throughout the rest of the simulation. That indicates the processed data are corrected to approach a straight (linear) congested phase line featuring this value. Two variations of the new proposed method are evaluated, namely the Linear-fit correction and the Polynomial-fit correction. To test performance, the proposed methods are compared with one of the speed-variance estimation methods, namely the “time series” (SVE-TS) method discussed above. To simplify matters, we will use the default parameter settings proposed in (Van Lint, 2004) (namely $P = 30$, $k_L^{\text{cri}} = 0.02$ veh/m/lane and $\gamma = 0.24$). These parameters are also calibrated from a Dutch freeway (A9), which are assumed to adequately process the simulated data from FOSIM. These three methods are cross-compared in a so-called normal-correction scenario.

Additionally, the robustness of the proposed algorithm is tested with two other scenarios, namely the noisy-input and the parameter-variation scenarios. In the former, the input data are combined with some white Gaussian noise to emulate the data condition from traffic sensors. In the latter, the (only-one) parameter in correction v_{cong} varies within a small range (from -18km/h to -24km/h).

Performance criteria

To assess different scenarios, the performance criteria are selected. The raw-input / corrected speed data are compared with the ground truth (reference) speed data in terms of root mean squared error (*RMSE*) and three relative error indicators. They are mean percentage error (*MPE*) which reflects structural bias, mean absolute percentage error (*MAPE*) which gives a combined indication of the relative error, and standard deviation of the percentage error (*SPE*) which is an index for the variability around the *MPE*. Since the bias / error mostly occurs at low speeds (congestion), these relative error indicators are more informative than the *RMSE* error, which is calculated by the absolute speed values that are relatively low in congestion. Nonetheless, *RMSE* error can still provide an overview of performance on the whole data set. These indices are defined as follows:

$$RMSE = \sqrt{\frac{\sum (u - \hat{u})^2}{NN}}, \quad (6.11)$$

$$MPE = \frac{1}{NN} \sum \frac{u - \hat{u}}{\hat{u}}, \quad (6.12)$$

$$MAPE = \frac{1}{NN} \sum \frac{|u - \hat{u}|}{\hat{u}}, \quad (6.13)$$

$$SPE = \text{std}\left(\frac{u - \hat{u}}{\hat{u}}\right). \quad (6.14)$$

Here, u ($:= u(t)$) denotes either raw-input data or corrected data, \hat{u} ($:= \hat{u}(t)$) is the reference data. They are both as the functions of time (t). NN denotes the size of the

data set. Note that, these error indices can be generalised to any comparison between two data sets.

6.2.5 Results and discussion

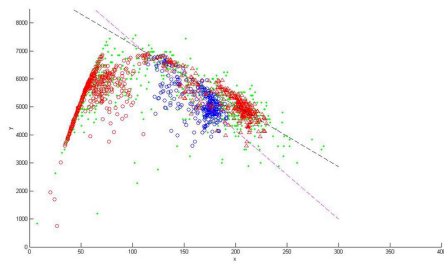
Qualitative impression

The proposed correction algorithm is tested by the synthetic data from FOSIM simulation. Figure 6.7 provides a first impression of the correction effect at a cross-section level. Specifically, subfigures 6.7(a)-(d) give a qualitative impression of how the proposed algorithm works. The targeted (raw) traffic data move rightward at the congested branch of q - k diagram. As a result, the corrected data approach closely to the ground truth data, scattering around the congested-phase line. As discussed above, the fit function can either be linear or polynomial, as illustrated in the subfigures (a)&(b) and (c)&(d) respectively. Obviously, the polynomial-fit curve captures most of the feature of the biased data, which is the foregoing “P-shape” distortion of phase diagram. This implies good performance of correction (this complies with the following quantitative results). Subfigures 6.7(e) and 6.7(f) present two fundamental diagrams that are calculated from typical downstream-bottleneck areas in FOSIM A13 model. The congested states do not exist herein, only the states of queue discharging are presented as the right (blue) part of circle area. This justifies the choice of non-correction.

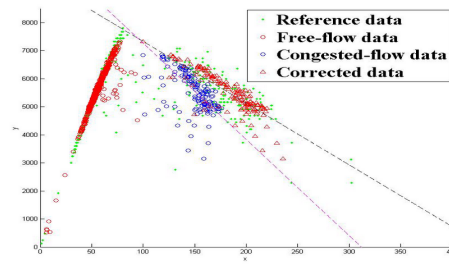
Quantitative analysis of the normal correction

The performance criteria for different test scenarios / correction methods are listed in table 6.1. The results are categorised by three test scenarios, namely the normal-correction scenario, the noisy-input scenario and the parameter-variation scenario (the latter two used in the robustness study). The MPE error is regarded as the most important indicator, which reflects the structural bias. Therefore, good performance leads to a low value of the MPE error, which indicates low structural error (speed bias).

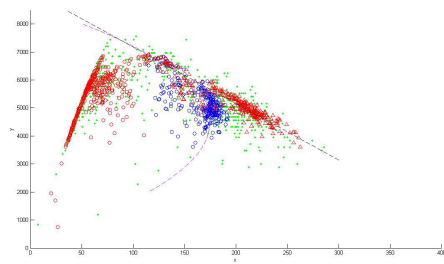
In the normal-correction scenario, the results are provided for the three correction methods, namely speed variance estimation method based on time-series of local mean speed (SVE-TS), Linear-fit correction and Polynomial-fit correction. The polynomial-fitting outperforms the other two methods, with an absolute improvement of 2% on MPE and 1% on $MAPE$, compared to the raw input data, and a relative improvement of 47% on MPE , which is quite substantial. When calculating the MPE error of a correction, the negative components can cancel out with the positive ones, while this is not the case for $MAPE$. So the higher improvement on MPE than $MAPE$ demonstrates that, the corrected data by polynomial fitting scatter around the ground truth data (or the reference congested phase line in q - k space) as expected. The linear fitting achieves a similar level of performance to the polynomial correction. The improvement by the



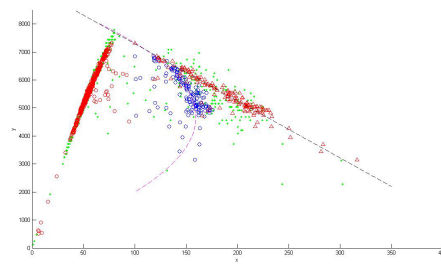
(a) km 2.5 in the FOSIM model



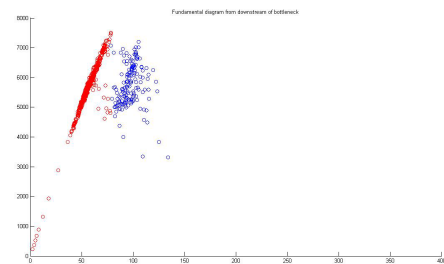
(b) km 5.8 in the FOSIM model



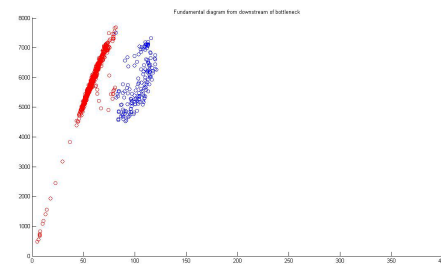
(c) km 2.5



(d) km 5.8



(e) km 6.5



(f) km 11

Figure 6.7: Examples of speed-bias correction at a cross-section level: (a)&(b) Linear-fit, (c)&(d) Polynomial-fit, (e)&(f) Exception for correction algorithm(downstream bottleneck). (The two dashed lines at the congested region in each plot are the fitted curves respectively for raw data and corrected data).

Table 6.1: Error indicators for three correction methods of three test scenarios, in terms of effectiveness and robustness (All error values accurate to two decimal places.)

Error indicators		<i>MPE(%)</i>	<i>MAPE(%)</i>	<i>SPE(%)</i>	<i>RMSE(km/h)</i>
Normal-correction scenario					
Raw Input		4.95	7.56	34.78	5.81
SVE-TS Method		4.56	7.59	34.39	6.35
Linear-fit		3.26	6.72	33.79	5.60
Polynomial-fit		2.64	6.60	33.27	5.64
Noisy-input scenario (robustness)					
Raw Input		5.04	8.48	35.51	6.19
SVE-TS Method		3.70	8.35	34.70	6.56
Linear-fit		4.08	8.34	35.30	6.18
Polynomial-fit		2.66	8.09	35.84	6.13
Parameter-variation scenario (robustness)					
	-18	2.10	7.30	33.72	5.84
Propagation	-19	2.33	7.07	33.56	5.78
speed (km/h)	-20	2.37	6.71	33.20	5.68
	-22	2.82	6.56	33.14	5.66
(Only Polynomial-fit)	-23	3.15	6.67	33.33	5.68
	-24	3.30	6.59	33.36	5.62

Table 6.2: *MPE/MAPE* errors in the normal-correction scenario, partitioned over five speed (in km/h) ranges

		Full speed range	0-20	20-40	40-60	60-80	>80
<i>MPE</i>	Raw Input	4.95	90.16	13.79	19.75	8.86	0.11
	SVE-TS	4.56	83.94	13.20	19.13	8.46	-0.17
	Linear-fit	3.26	75.02	6.40	14.45	7.99	0.11
	Poly.-fit	2.64	42.85	2.93	13.78	8.39	0.11
<i>MAPE</i>	Raw Input	7.56	90.20	15.30	23.94	16.97	2.07
	SVE-TS	7.59	83.99	14.85	23.60	16.90	2.29
	Linear-fit	6.72	75.20	11.30	21.65	17.08	2.07
	Poly.-fit	6.60	47.61	10.76	21.70	17.42	2.07

SVE-TS method on *MPE* is rather limited compared to the proposed correction methods. This method increases the *RMSE* error by 8% as a non-preferable negative effect. The similar observations are also presented in the noisy-input scenario.

To further analyse the correction effect on the raw-input data (at congested state), raw data samples from the normal-correction scenario are subdivided into five speed ranges. The correction performance in terms of the *MPE* and *MAPE* errors is shown in table 6.2. First, it is numerically demonstrated that the biases / errors are mainly derived from the congested traffic since the large error values are related to low speed regions. Additionally, the large values for *SPE* presented in table 6.1 can be accounted by the noticeable fluctuation in the *MPE* (and/or *MAPE*) over the different speed ranges. In “fitting” correction methods, the decrease on errors (*MPE* / *MAPE*) is mainly contributed by correcting on the congested-region data (speed values lower than 60km/h). For the SVE-TS method, the correction effects distribute over different speed ranges.

In the proposed methods, the amplitude of improvement is proportional to the direction of speed decreasing (or density increasing as shown in Figure 6.4). Compared to the linear-fit, the major advantage by the polynomial-fit on decreasing errors is reflected at the low speed ranges (significantly at speed ranges of 0-20km/h and 20-40km/h). The explanation is that traffic data at low speed region, which relates to the lower part of congested branch in *q-k* diagram (see Figure 6.4), can be fit more closely by polynomial curves, thus the correction is better.

More specifically, in the speed range of 0-20km/h, the *MPE* / *MAPE* error of the raw input data is rather high (more than 90%). It is remarkable that the relative improvement on *MPE* / *MAPE* by the polynomial-fit is still greater than 50% (42.85 versus 90.16). In the range of 20-40 km/h, the polynomial-fit even leads to a higher relative improvement of more than 79% (2.93 versus 13.79). This is because most of input data samples are in this speed range. However, the improvement by the new methods for the speed range of 60-80 km/h is marginal. This can be explained as follows: at high speed

ranges (related to the upper part of congested branch in Figure 6.4), the difference between the fitted curve and the targeted corrected line is marginal. In that the chosen targeted correction line (as expressed by equation (6.9)) considers only one degree of freedom. If more degrees of freedom of parameters are considered, the performance would be better.

So far we can firmly conclude that the proposed correction algorithm is effective and practical on reducing speed bias. The polynomial-fit method performs better than its linear-fit counterpart and the time-series methods.

Robustness study

The results of two test scenarios for robustness study are also presented in table 6.1. First, if the input data from traffic sensors contain noise, the errors of these data increase accordingly. However, the correction algorithm is still effective, resulting in relatively low *MPE* and *MAPE* errors. The improvement on *MPE* by the polynomial-fit in this case is comparable to that in the normal-correction case. Second, as in the new proposed algorithm, only one degree of freedom, the congested propagation speed (v_{cong}) is targeted. When varying this parameter within a small range, the correction still performs well within an acceptable error range. In reality, the propagation wave speed varies slightly depending on the accepted safe time clearance, the average vehicle length, traffic composition and weather conditions (Windover & Cassidy, 2001; Treiber & Helbing, 2002; Kerner, 2004). Therefore, the influence of this parameter would be marginal. This algorithm is concluded to be quite robust.

Qualitative analysis

To qualitatively assess the polynomial-fit correction, detailed corrected speed profiles are studied. Figure 6.9 shows the comparison of this correction variant, with respect to the reference (ground truth) data and the raw-input data at three chosen time instants, which are indicated with three arrows in the whole (reference) speed map of the A13 model (see Figure 6.8). The correction (indicated by small red circles) aims to resemble the reference data (indicated by black stars). Complying with the previous observation (in table 6.2), there is little difference among the three speed profiles at high speed regions (above 80 km/h). The correction algorithm works mainly for the data set at low speed regions (below 60km/h). As highlighted by the lower part of blue circles, the corrected speeds can represent the ground truth, while the raw-input (arithmetic mean) speeds overestimate the real speeds.

6.2.6 Conclusions

A new effective and robust method has been proposed for correcting speed biases caused by the common practice of arithmetic time-averaging. The simulation results

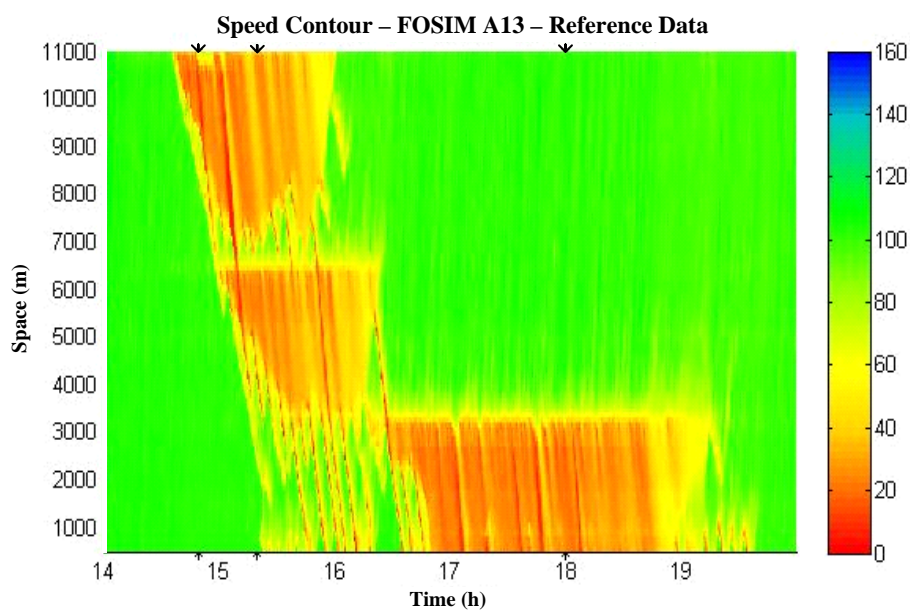


Figure 6.8: Reference speed map from A13 FOSIM model. The arrows indicate the chosen time instants for the speed profiles in Figure 6.9.

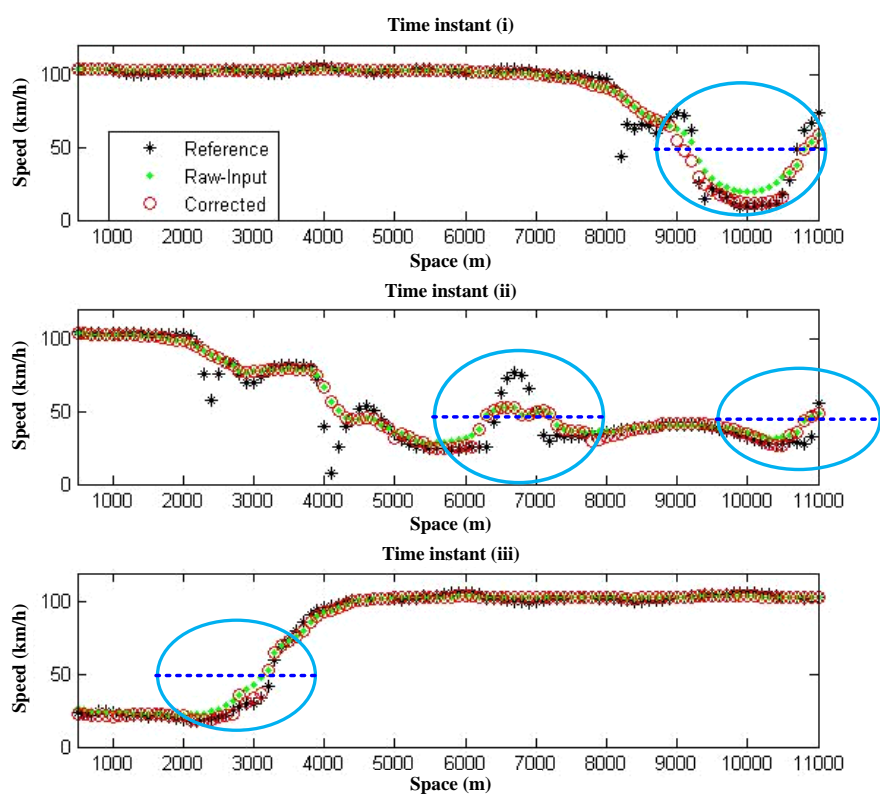


Figure 6.9: Performance of polynomial-fit correction: speed vs. space plots at three chosen time instants. The reference measurements are the target that the correction aims to resemble.

show that this algorithm succeeds to correct speed biases with relatively low errors and performs better than the previous proposed method (speed-variance estimation). The relative improvement reaches almost 50%. The robustness of this algorithm with regards to noisy traffic data and parameter variation is verified in the study.

The concept of the speed correction is straightforward and easy to implement. The computational speed of the correction algorithm is fast in MATLAB environment. All these advantages are beneficial for real-time operation. If harmonic mean speeds can be collected directly from monitoring systems at no cost, then the proposed algorithm would have no purpose. So the potential value in this algorithm is that it provides practitioners with an effective and efficient tool to overcome the speed-bias problem in the empirical data.

The algorithm is developed for correcting speed data from dual loops, and it can also be extended and applied to correct unreliable densities from single loops. The density profiles can be inferred from the occupancy information with the assumption of average vehicle length. However, this assumption is sensitive to low traffic demand (flow) states (at both free-flow end and congestion end refer to $q-k$ diagram), since the variance in average vehicle length increases as flow decreases (Coifman, 2001). This implies the density estimation at both two ends of $q-k$ diagram are unreliable or inaccurate. To correct noisy densities, similar rationale can be applied. One could correct density values to approach the congested and/or the free flow phase line(s) while keeping flow values invariant in the flow-density space.

There are some directions for further research. In the correction step, the targeted characteristic curve for speed correction can be better described by adding more degrees of freedom, besides the congestion propagation speed. Based on this proposed method, the insight to improve travel time estimation can be further addressed.

6.3 Estimation of multi-class and multi-lane counts

In the previous section, the proposed algorithm aims to improve the quality of the first type of input \mathbf{Y} (see Figure 6.1). The method presented in this section additionally deals with the second type of input \mathbf{d} , which attempts to infer class-specific data for multi-class traffic state estimation.

6.3.1 Problem analysis

Driver and vehicle heterogeneity is a commonly accepted fact and has received considerable attention recently (Daganzo, 2002a; Van Lint et al., 2008b) in the research literature. As a simple example, the distinction can be made between the flows in different lanes (fast or slow vehicle lanes, dedicated lanes etc.) or between the flows in different vehicle-classes (trucks, buses, passenger-cars, and high-occupancy vehicles).

This information is of great importance for realistic multi-class multi-lane vehicle generation in microscopic traffic simulation tools (e.g., VISSIM (PTV, 2010), AIMSUN (Barceló et al., 2004) etc.), on-line lane-specific capacity and Person-Car-Equivalents (PCE) estimation (Webster & Elefteriadou, 1999), infrastructural design/planning and safety analysis (Transportation-Research-Board, 2000) and their consequent use in ITS applications, such as Dynamic Traffic Management (DTM) (Louis et al., 2006). Reducing congestion on motorways is one of the main objectives of DTM, for instance, by using lane management. This can be done by reserving the use of some lanes to specific categories of vehicles (e.g., a dedicated truck lane), prohibition of truck overtaking, or by increasing the number of lanes, for example by hard shoulder running in peak hours. In order to properly implement and evaluate these control interventions, multi-lane and multi-class models and data are necessary. This also justifies the presentation of a multi-class traffic state estimation method in this thesis.

The provision of multi-class and multi-lane data relies on the characteristics of traffic monitoring systems, which in many countries are based on stationary loop detectors. These are of two main types, namely single loops and double loops. Single loops (implemented on US freeways and world-wide in urban networks) provide flow and occupancy information, but speed measurements are not directly available, although they may be indirectly estimated (Coifman, 2001). In addition to flow and occupancy, double loops can measure vehicles' speeds and lengths. Although the loops capture individual vehicle data, these are usually aggregated over a given time interval ΔT (usually 1-5 minutes) because of historical technical constraints on the communications hardware. In standard operation, individual vehicle data (IVD) is usually not available to the end-user, although some new detector systems allow for remote collection of IVD.

The question of how to study realistic multi-class and multi-lane behaviours boils down to how we can get class-specific and lane-specific information from the data in standard traffic monitoring systems. The basic idea is to estimate counts of vehicles disaggregated by lane and by vehicle class, over a given analysis interval T (usually $T > \Delta T$), using only aggregate data available from standard inductance loop implementations. In some countries, such as in parts of Italy and Germany, the dual-loop system already provides a complete disaggregation of flow by lanes and (at least two) vehicle classes. However, in most other countries, the standard aggregate loop data reports only speed, occupancy and counts by lane, but not by class.

We let q_{ij} denote the vehicle count in lane i and class j over a given analysis interval, which are the variables to be estimated. Here, $i = 1, 2, \dots, N_L$, and $j = 1, 2, \dots, N_C$, where N_L and N_C denote the numbers of lanes and classes, respectively. In all of the data that we consider, vehicle counts are available by lane, and we let q_i^L denote the observed count of vehicles in lane i . (Here, the superscript L denotes aggregation by lane.) It follows the constraints:

$$q_i^L = \sum_{j=1}^{N_C} q_{ij}, \text{ for each } i = 1, 2, \dots, N_L. \quad (6.15)$$

The main objective is to infer more count variables than those that are observed directly. The idea is to construct estimation rules which exploit equation (6.15) in combination with the other observed aggregate quantities (such as speed and occupancy). The underlying assumption is that the aggregate data might contain class-specific and/or lane-specific information, which can be used for the inference. This can be justified by empirical observations. The empirical flow and density values generally scatter around the equilibrium line of fundamental (q - k) diagrams. This indicates that a given flow value can be achieved by slightly different densities (occupancy rates), because of different traffic compositions. More specifically, an increase in density or occupancy without a corresponding increase in flow tends to suggest an increase in the proportion of long vehicles (e.g., trucks and buses), and vice versa. There are certain correlations between the class-specific (lane-specific) information and the aggregate data. The estimation rules can be trained to capture these plausible correlations, using individual vehicle data as “ground truth”, since then all of the (typically) observed variables and those to be estimated are available exactly. A natural question is to see how site-specific this technique is, and how much individual vehicle data do we need to collect when calibrating the estimation laws for a new site.

In our estimation rules, we need to take into account the full details provided in aggregate loop data, and this varies subtly from country to country. For example, in the UK, the aggregate data also reports counts disaggregated by vehicle class (but not jointly disaggregated by lane and class) besides the lane-specific speed, occupancy and count. In Spain, speed stratification information, giving the counts of vehicles separated by different speed thresholds, is also available. In contrast, single loops from most countries provide only aggregate count and occupancy by lane, yet still some progress is possible with our estimation technique.

The next section introduces the basic mathematical structure of the multi-linear regression methodology, including the error measure to evaluate its performance. Then, we describe the “ground truth” individual vehicle data to calibrate the regression, and the standard aggregate-data scenarios which are representative of typical loop implementations around the world. Finally, initial results and error analysis of the algorithm are provided.

6.3.2 Methodology

This section describes the structure of the rules for estimating multi-lane multi-class counts from standard aggregate loop data. Our study is a regression analysis from inductive statistics, where we make inferences for the correlation between lane/class disaggregated counts as output variables and standard aggregate data as input variables, see e.g., (Freedman, 2005). It is well known that the regression method can provide the forecasts/estimates and it is easy to apply. In general, the estimation rules take the form:

$$\mathbf{y}_\tau \approx \hat{\mathbf{y}}_\tau = \mathbf{f}(\mathbf{X}_\tau, \beta) \quad (6.16)$$

where τ is the index of the individual analysis time interval; \mathbf{y}_τ is the vector of N_{out} (exact) output variables and $\hat{\mathbf{y}}_\tau$ denotes the related estimates; \mathbf{f} denotes the model (fit) functions; \mathbf{X}_τ is the vector of N_{in} input variables; and β is a vector of a N_{para} parameters. We have experimented with various forms of nonlinear model functions but found that they improved the fit only very marginally over linear functions of the form:

$$\mathbf{f}(\mathbf{X}_\tau, \beta) = \mathbf{A}\mathbf{X}_\tau + \mathbf{b} \quad (6.17)$$

which we use throughout the remainder of this study. Here, \mathbf{A} is an $N_{\text{out}} \times N_{\text{in}}$ matrix, \mathbf{b} is a vector with N_{out} elements, and β is a vector of $N_{\text{out}}(N_{\text{in}} + 1)$ elements, listing all of the elements of \mathbf{A} and \mathbf{b} .

The fitting problem then concerns the optimal selection of the parameters β for a training set consisting of N_T pairs of the form $(\mathbf{X}_\tau, \mathbf{y}_\tau)$, which is constructed from individual vehicle data (see Sections 6.3.4 and 6.3.3). The method applied is least squares multi-linear regression analysis, where we aim to minimise the sum E of the squared errors/residuals between the estimates and the actual values in the data set, that is:

$$\arg \min_{\beta} E = \arg \min_{\beta} \sum_{\tau=1}^{N_T} \|\mathbf{r}_\tau\|_2^2, \text{ with } \mathbf{r}_\tau = \mathbf{y}_\tau - \hat{\mathbf{y}}_\tau. \quad (6.18)$$

Here, \mathbf{r}_τ is an N_{out} -element vector of residuals between the estimates and the actual values (ground truth) of the output variables at the τ^{th} time interval. This is a quadratic program which can be solved with standard computational techniques.

To assess the performance of the estimation, we use the Relative Total Count Difference (*RTCD*) which is the sum of absolute differences between the estimates and the ground truth divided by the total counts of the surveyed data set, given by:

$$RTCD = \frac{\sum_{\tau=1}^{N_T} \|\mathbf{r}_\tau\|_1}{\text{TotalFlow}}. \quad (6.19)$$

Here, the term ‘‘TotalFlow’’ is introduced to normalise the error measures between data sets with different total traffic demands. Note that the *RTCD* error is not exactly what is minimised in the regression analysis, but provides a more intuitive measure of the fitting performance. Later we will present this error measure both for calibration sets (where β is chosen to minimise E), and for evaluation sets (where β is held fixed at values determined by independent calibration sets).

Finally, in practice, the computations are simpler and have fewer degrees of freedom than the fully general structure that we have provided here. In particular, it may seem that there are $N_{\text{out}} = N_L N_C$ output variables, which are the multi-lane multi-class counts for a given analysis interval. However, constraints such as equation (6.15) imply that these outputs are linearly dependent. As a result, the entries of β are linearly dependent, and there are fewer than $N_{\text{out}}(N_{\text{in}} + 1)$ parameters to estimate. The size N_{in} of the input variable vectors depends on the precise format of the aggregate loop data under consideration, and which of its entries are used in the estimation. These details are mapped out in Section 6.3.4.

6.3.3 Individual vehicle data environment

Individual vehicle data (IVD) from the UK and the Netherlands are used to provide the “ground truth” in our study. IVD data typically include passage time (τ_κ), speed (v_κ), lane number (L_κ) and length (l_κ) of each vehicle ($\kappa = 1, 2, \dots$) as it passes each detector. Note that a vehicle (κ) can generally be classified according to its length: we let \hat{l}_j denote the upper length limit of class j ($= 1, 2, \dots, N_C$) so that $\hat{l}_{j-1} < l_\kappa < \hat{l}_j$ implies that vehicle κ belongs to class j , where $\hat{l}_0 := 0$.

This IVD can be rolled-up into any aggregate-data format that we choose: to emulate either existing data formats (see Section 6.3.4) or new ones that include the complete disaggregation by lane and class. Consequently it can be used to learn and test the relationship between multi-lane multi-class counts and existing aggregate formats.

The British IVD come from the *Active Traffic Management* (ATM) section of the motorway M42 near Birmingham in the UK (Highway-Agency, 2011). This section has an unprecedented coverage of inductive loop detectors, with a nominal spacing of 100m. During 2008/09, 16 consecutive detectors on the North-bound carriageway were enhanced so that, amongst other improvements, the full IVD of all vehicles driving through the one-mile section was recorded (Wilson, 2008). The motorway is equipped with dynamic speed control systems and features hard shoulder running in peak hours, expanding the width of the carriageway from three lanes to four lanes in each direction. Furthermore, the enhanced section includes an on-ramp. However, to simplify matters in this initial study, we consider only periods where the dynamic control systems are turned off (since they have a complex effect on lane utilisation) and we use the loop detectors which are furthest downstream from the ramp, in order to minimise its influence.

The Dutch IVD come from a group of consecutive loop detectors (about 500 meters apart) on the east-bound carriageway of freeway A15, which connects Rotterdam to the German border. The individual vehicle data are available for the period 8-17 April 2006, basically the same format as the British data, including the passage time, speed, lane number and length of each vehicle. In contrast to the British data, there are neither control interventions nor peak lanes implemented on the Dutch study section on the given period, so there is no need to discard data to simplify the analysis. The detector located with milepost A15-88.02km is used, at which point the carriageway consists of two lanes.

6.3.4 Aggregate data scenarios

As we have discussed, the standard inductance loop implementations in many countries (the US, and most of the countries in Europe, e.g. the UK, the Netherlands, Spain, France, Germany, Italy etc.) generally provide aggregate flow (q_i^L) and occupancy (Occ_i^L) by lane, in addition to aggregate speed (v_i^L) for dual-loop systems, over a given measurement interval ΔT .

However, the details of each implementation are slightly different and some provide extra data in addition to that listed above. If possible, these extra data should be used in the estimation procedure, i.e., they should be listed as additional elements in the input variables \mathbf{X}_τ . As examples of what might be achieved, we consider three representative international data scenarios. Scenario 1 models UK MIDAS (Motorway Incident Detection and Alert System) data, where counts are disaggregated by lane and by class, but not jointly by lane and class. Scenario 2 is the basic US (single loop) situation where only flow and occupancy are available. Scenario 3 is a model of Spanish data, where flow is disaggregated by speed thresholds.

Note that our approach is to emulate each of the given scenarios in the analysis that follows. This means that we need not use real aggregate data from the UK, US, Spain etc. Rather, we use British and Dutch individual vehicle data as described in Section 6.3.3, and create our own aggregates from it, according to the rules of the various scenarios. Of course, our approach is thus limited in that it cannot evaluate differences in lane utilisation behaviour between the US, Spain etc., because the underlying data is British/Dutch. Such study remains for future work.

The details of the three scenarios are as follows.

- Scenario 1: a model of UK MIDAS data, in which, in addition to the standard dual-loop by-lane data given above, counts are disaggregated into four vehicle classes, where class 1 corresponds roughly to passenger cars, class 4 to heavy goods vehicles, and classes 2,3 represent vehicles of intermediate lengths. Therefore, similar to lane counts q_i^L , we let q_j^C denote counts disaggregated by class j , where the superscript C denotes aggregation by class. Analogous to equation (6.15), we thus have

$$q_j^C = \sum_{i=1}^{N_L} q_{ij}, \text{ for each } i = 1, 2, \dots, N_C. \quad (6.20)$$

Whereas true UK MIDAS data have $N_C = 4$, to simplify matters here, we will group together all vehicles that are longer than typical passenger cars (of 5.0 meters), referred to as the “lorry” class, and work with $N_C = 2$. This choice will apply throughout this study for all the other scenarios too. In scenario 1, we have not used the speed and occupancy data in the estimation procedure, so each input vector \mathbf{X}_τ , in full, consists of $N_{\text{in}} = N_L + N_C$ entries, namely q_i^L ($i = 1, 2, \dots, N_L$) and q_j^C ($j = 1, 2, \dots, N_C$), and the output vector $\hat{\mathbf{y}}_\tau$ consists of $N_L N_C$ entries. However, as we have discussed, constraints (6.15) and (6.20) imply that the component flows are linearly dependent, and it thus follows that we may work with $N'_{\text{in}} = N_L + N_C - 1$ and $N'_{\text{out}} = N_L N_C - N_L - N_C + 1$, noting that the $N_L + N_C$ constraints (6.15) and (6.20) when taken together have rank degeneracy of one. See Table 6.3.

- Scenario 2: Single loop data format - i.e., a US-type situation, where only occupancy and flow are available, disaggregated by lane. Each input vector \mathbf{X}_τ consists of $N_{\text{in}} = 2N_L$ entries, and these may not be reduced because only constraints (6.15) apply, which in themselves are linearly independent. In full, each output vector $\hat{\mathbf{y}}_\tau$ has $N_{\text{out}} = N_L N_C$

Table 6.3: Matrix of the output variables q_{ij} . Equation (6.15) implies constraints on the row-sums, and equation (6.20) (for scenario 1) implies constraints on the column-sums

Count Matrix	Class $N_C \rightarrow$			counts by lane
Lane	q_{11}	...	q_{1N_C}	q_i^L
N_L	\vdots	q_{ij}	\vdots	
\downarrow	$q_{N_L 1}$...	$q_{N_L N_C}$	
counts by class	q_j^C			

Table 6.4: The numbers of input and output variables for each of the data scenarios. N'_{para} denotes the minimal number of parameters that need to be estimated so that all other quantities can be estimated from them

	Data Source	N_{out}	N'_{out}	N_{in}	N'_{in}	N_{para}	$N'_{\text{para}} := N'_{\text{out}}(N'_{\text{in}} + 1)$
Scenario 1	UK	6	2	5	4	36	10
	NL	4	1	4	3	20	10
Scenario 2	UK	6	3	6	6	42	21
	NL	4	2	4	4	20	10
Scenario 3	UK	6	3	5	5	36	18
	NL	4	2	4	4	20	10

entries, but one may exploit their linear dependence, from equation (6.15), to work with a reduced set of $N'_{\text{out}} = N_L N_C - N_L$ linearly independent output quantities.

- Scenario 3: a model of Spanish data formats - as in addition to the usual by-lane data, a speed stratification is provided. Speed stratification means that a count q_{v^*} of vehicles travelling at a speed lower than a threshold v^* is provided. Usually, the counts for two distinct speed thresholds are given, namely 50km/h and 100km/h (Soriguera & Robusté, 2011). In this case, each input vector \mathbf{X}_τ consists of $N_L + 2$ entries (the count for each lane and below each speed threshold), since in common with the other scenarios, we have chosen not to use the speed data - or indeed, the occupancy data (like scenario 1, but unlike scenario 2). The output vector $\hat{\mathbf{y}}_\tau$ has $N_{\text{out}} = N_L N_C$ entries, which as for scenario 2, may be reduced to $N'_{\text{out}} = N_L N_C - N_L$ linearly independent output quantities.

All this information concerning the numbers of input and output variables for each of the scenarios is summarised in Table 6.3 and Table 6.4. Note that, the selection of independent output quantities does not influence the results of the fitting procedures, because the error components of all the output variables have been incorporated in equation (6.18).

Example: consider the case of a two-lane carriageway with two vehicle classes. There are four unknown elements in the output “count matrix”. However, in scenario 1 (UK-like), there are four constraints, two relating to lane counts, and two relating to class counts. However, the rank of the constraints is three (four minus the rank degeneracy

of one), implying that there is only one degree of freedom (four degrees of freedom minus three linearly independent constraints) in the output variables. It is thus possible to write all output quantities in terms of a single output quantity, e.g. q_{11} . So how many (scalar) parameters need to be fitted? Using the matrix notation introduced in Section 6.3.2, we have $q_{11} = b_1 + A_{11}X_1 + A_{12}X_2 + A_{13}X_3 + A_{14}X_4$, with the usual indexing notation for the elements of \mathbf{A} and \mathbf{b} . We use $X_1 = q_1^L, X_2 = q_2^L, X_3 = q_1^C, X_4 = q_2^C$. However, because $X_1 + X_2 = X_3 + X_4$, there are only three independent input quantities. We might thus simplify matters by using the new input variables q_{TOT} , φ , and ψ , with $X_1 = q_{TOT}/2 + \varphi/2, X_2 = q_{TOT}/2 - \varphi/2, X_3 = q_{TOT}/2 + \psi/2$ and $X_4 = q_{TOT}/2 - \psi/2$. It follows that there are only four independent parameters to be determined, which are $b_1, \zeta = (A_{11} + A_{12} + A_{13} + A_{14}), \eta = (A_{11} - A_{12})$ and $\theta = (A_{13} - A_{14})$. The other cases may be worked out in a similar fashion, but with much more complicated algebra that we do not present here.

6.3.5 Results

To illustrate the entire procedure, we firstly take individual vehicle data (described in Section 6.3.3); we then roll it up into standard aggregate formats, that we call scenarios (see Section 6.3.4); furthermore, we also roll it up into an aggregation which describes the complete breakdown of flow by class and lane. All of these aggregates have time resolution ΔT (=1 minute). We then aggregate the data up into larger “analysis intervals”, which are of length T . In this section we choose $T = 10$ minutes. (An analysis of the effect of different choices of T follows in Section 6.3.6. In sum, if T is too small, our results become swamped by statistical sampling error, whereas if T is too large, the applicability of our technique for online traffic estimation becomes rather limited.) We then use the aggregates over T to perform the multi-linear regression procedure described in Section 6.3.2, which relates multi-class multi-lane counts to the standard aggregate variables. For all the scenarios, the calibration data sets are taken from the weekdays 2nd Oct 2008 in the UK and 10th Apr 2006 in the Netherlands (NL), respectively. For validation, other days with similar traffic patterns are tested.

Example continued. Let us follow through the procedure described above for the set-up that was introduced at the end of Section 6.3.4, namely for the Dutch (2-lane) individual vehicle data and the UK-like scenario aggregates. The regression determines the following optimal values for the parameters in the estimation rule for q_{11} (the count of cars in lane 1): $b_1 = 0.614, \zeta := (A_{11} + A_{12} + A_{13} + A_{14}) = 0.008, \eta := (A_{11} - A_{12}) = 0.996, \theta := (A_{13} - A_{14}) = 0.991$. Here, ζ describes the marginal effect of the total flow on q_{11} - which we can see is insignificant. The interpretation is that if the total flow is increased in equal proportion across lanes and vehicle classes, we would expect most of the extra cars to drive in lane 2. The parameter η describes the marginal effect on q_{11} with respect to the lane 1 flow minus the lane 2 flow, and the parameter θ describes the marginal effect with respect to the car flow minus the “lorry” flow. We can see their effects are significant: for example, the interpretation for η is

Table 6.5: *RTCD* error measures for different scenarios using both British and Dutch inputs

		Calibration	Validation
		UK: Thu. 2nd Oct 2008 NL: Mon. 10th Apr 2006	UK: Tue. 7th Oct 2008 NL: Wed. 12th Apr 2006
Scenario 1	UK	6.44	5.79
	NL	0.95	1.20
Scenario 2	UK	6.71	6.87
	NL	3.69	3.68
Scenario 3	UK	9.67	9.43
	NL	8.03	7.85

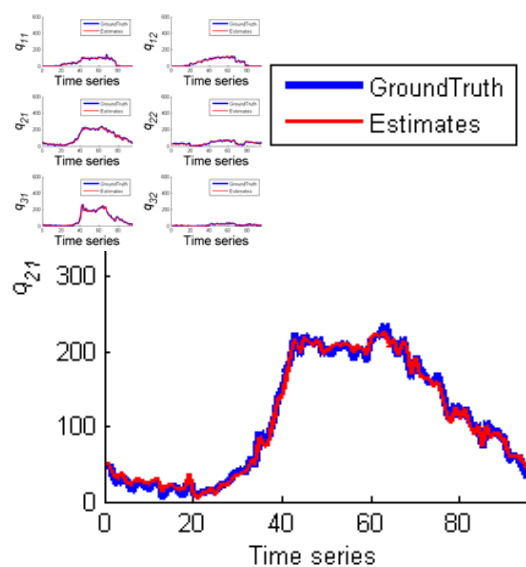
that if lane 1 flow alone is increased with no increase in the total flow, then half of the extra flow in lane 1 is cars. Regression results (parameters) from other cases can be interpreted in a similar way.

So how well does the estimation procedure work throughout a given days data? See figures 6.10 and 6.11, which respectively present time series of estimates and the ground truth data from scenario 1 (UK-like) using the UK and Dutch data sets. In the case of calibration (Subfigure 6.10(a) for UK-IVD and subfigure 6.11(a) for NL-IVD), all the estimated curves lie almost exactly on top of the ground truth data (see the example zoom in the subfigure 6.10(a)). This indicates the estimation rule works well on both the British and the Dutch input data sets that it was calibrated with, in terms of all the output quantities. Moreover, the calibrated estimation rules are validated by good predictions on independent data sets corresponding to different days of the week (see subfigures 6.10(b) and 6.11(b)).

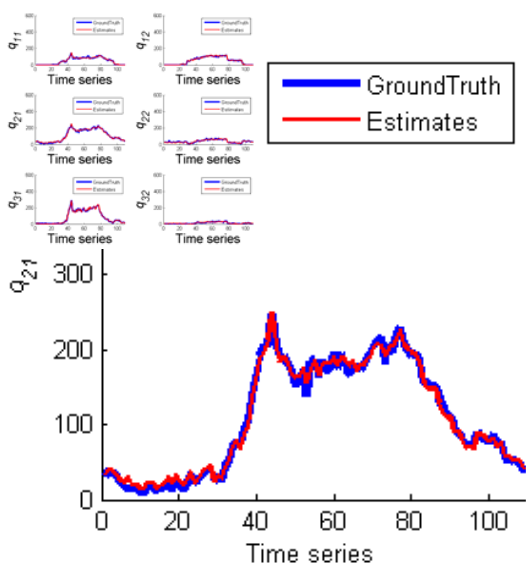
Subfigures 6.10(c) and 6.10(d) show the *RTCD* errors over all the individual output quantities, which are at the same level. The slightly larger *RTCD* error on q_{22} and q_{32} (“lorry” class on the fast-vehicle lanes 2 and 3) can be explained by the fact that for relatively low lorry flow the relative count error becomes large.

Let us now analyse the quantitative performance of the method, for each of the aggregate data scenarios, and for each of the input sets. See Table 6.5. The error indicators for the calibration sets and corresponding validation sets are of the same order, indicating that the method is able to estimate the flow decomposition when trained on relatively small amounts of ground truth data.

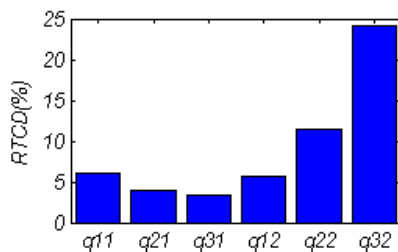
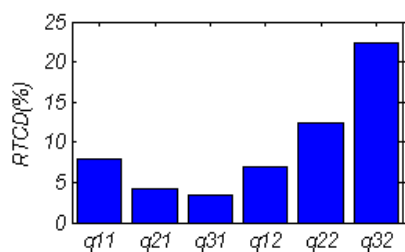
The estimation with the UK-like scenario outperforms other scenarios in terms of both the calibration and the validation, to a quite remarkable degree when using the Dutch input data. (Since counts by class are provided in scenario 1, the performance is expected to be good.) The estimation with the single-loop scenario 2 surprisingly achieves a similar level of accuracy to scenario 1. Although class counts are not available in scenario 2, it seems the occupancy information can partially reflect the class



(a) UK - Calibration set (Thu. 2nd Oct. 2008)

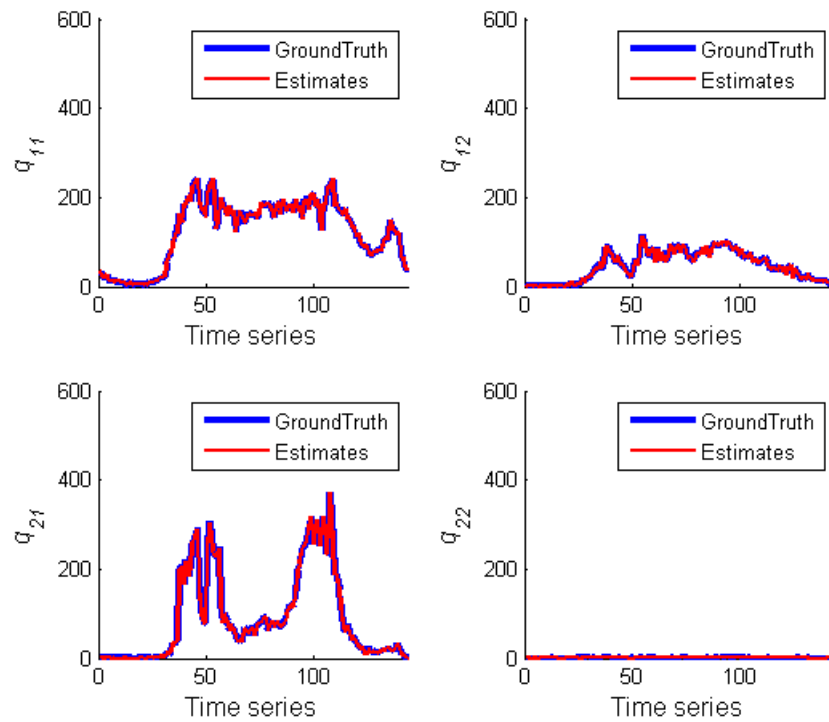


(b) UK - Validation set (Tue. 7th Oct. 2008)

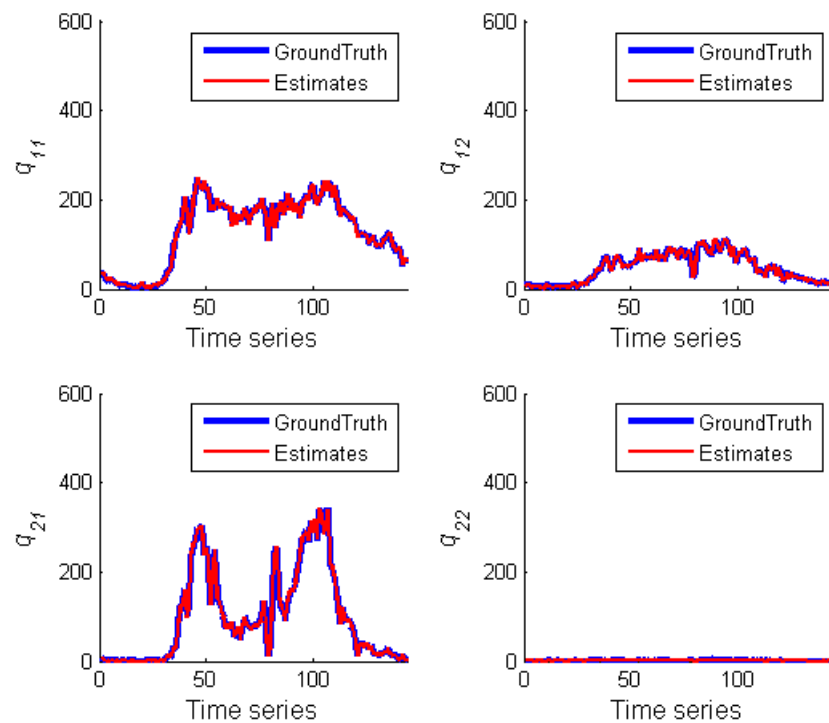


(c) Individual *RTCD* errors in Calibration (d) Individual *RTCD* errors in Validation

Figure 6.10: Comparison of time series for estimates and ground truth data in scenario 1 using the UK data set.



(a) NL - Calibration set (Mon. 10th Apr. 2006)



(b) NL - Validation set (Wed. 12th Apr. 2006)

Figure 6.11: Comparison of time series for estimates and ground truth data in scenario 1 using the Dutch (NL) data set.

Table 6.6: Validation results for scenario 1 over several days using both the British and the Dutch IVD (Note that: Weekend sets are marked with grey shading.)

Scenario 1	Calibration	Validation							
		1	2	3	4	5	6	7	8
UK	6.44	5.07	5.91	5.79	6.29	5.93	6.56	12.08	9.19
NL	0.95	1.05	1.20	1.10	1.11	0.67	0.86	1.02	0.79

information, since a high occupancy might imply a high flow of lorries. The fitting in the speed-stratification scenario 3 is rather limited compared to scenarios 1 and 2. The standard speed stratification uses two speed thresholds, namely 50 and 100 km/h, whereas the typical free flow speed of lorries is 80-90 km/h. Hence, it is possible to assume that a different set of thresholds may help classify the flow better. Note that the performance of scenario 3 is quite sensitive to speed limit controls, because these typically result in a very small speed variance, so that all vehicles are counted in the same speed bin, irrespective of their classes.

Compared to the British data set, two lanes rather than three lanes of carriageway are distinguished in the Dutch data set. This implies that there is less complexity in the lane utilisation behaviour. For instance, the lorry flow drives mostly in lane 1 (the right lane) at our Dutch site (see the right half of the subfigures 6.11(a) and 6.11(b)) rather than distributed over lanes as in the British case. Intuitively, it seems that the estimation errors would decrease accordingly and this is confirmed by the results of each scenario using the two data sources.

Table 6.6 investigates the validation question in more details. We focus on scenario 1 calibrated on a single day of data for each of the UK and Dutch data sets, and test against eight different validation days. The errors for all the test cases are of about the same level (except for two UK weekend sets), which implies that the calibrated estimation rule from one day can be applied to other days (weekdays). Similar results can be obtained for the other scenarios. Note that, the British data sets for validation (from 1 to 8) are respectively from Wed.1st Oct, Mon.6th Oct, Tue.7th Oct, Wed.8th Oct, Thu.9th Oct, Fri.10th Oct, **Sat.4th Oct and Sun.5th Oct** (2008). The Dutch data sets for validation are respectively from Tue.11th Apr, Wed.12th Apr, Thu.13th Apr, Fri.14th Apr, Mon.17th Apr, **Sun.9th Apr, Sat.15th Apr and Sun.16th Apr** (2006).

These validation sets are taken from both weekdays and weekends, the main difference between which is that the proportion of lorry (class 2) flow tend to be significantly lower at weekends. According to our preliminary calculations, the proportion of lorry flow drops from about 30% at weekdays to about 10% at weekends in the British IVD, while in the Dutch case, this value changes from around 20% to lower than 5%. This will have a direct influence on lane utilisation. As a result, the estimation rule calibrated from a weekday may not be applied to the weekend without further investigation and analysis. To illustrate this, note that the UK validation sets 7 and 8, which correspond to the weekend, show a significant increase in error. In contrast,

the weekday calibrated estimation rule works well on any day of the week (including weekends) in the Dutch data. This result is a major surprise: it is almost certainly due to the less complicated lane utilisation in the Dutch data, and the tendency there for most of the lorries to drive in lane 1.

6.3.6 Error analysis

We now give a preliminary error analysis of the estimation method that we have developed. There are two main questions:

- Q1: What is the effect of varying the analysis interval T ?
- Q2: How large do the training sets need to be? For a given size of training set, what are the error bars on the regression parameters?

To simplify the presentation, we illustrate our study by using the (UK-like) scenario 1 and training sets based on the UK IVD set from Thu. 2nd Oct 2008. In particular, we avoid using the Dutch IVD because the lane utilisation appears too simple for this data to yield an error analysis that is representative of more general situations.

Let us first consider Q1. Up to now, we have worked with an analysis interval of $T=10$ minutes. Table 6.7 shows how the *RTCD* error (for the calibration set) changes as we vary T . To clarify, when we use e.g. $T=30$ minutes, there are only one third as many data pairs to feed into the regression analysis as there are with $T=10$ minutes. Note that the normalisation of the *RTCD* error by the total flow implies that there is no inherent bias in the error measure due to the number of analysis intervals over which it is computed.

As we might expect, the *RTCD* error decreases as T increases, owing to the effect of sampling error. To clarify, even if the traffic conditions were entirely homogeneous over the calibration set (and they are not), then in any one interval we would expect random variations in the proportions of the multi-lane multi-class counts, which the estimator cannot reproduce. This effect is most severe for short analysis intervals, e.g. for $T = \Delta T = 1$ minute. According to classical statistical theory, the sampling error should decay like the reciprocal of the square root of the length of the sampling interval. However, our errors decay a little more slowly, because as we increase T , we also mix together different sorts of traffic conditions which in itself introduces error. In practice, there will always be a play-off between the sampling error and the time resolution, and the optimal choice of T will depend on the application in question.

Let us now address Q2. We expect that the details of lane utilisation will vary subtly from site to site, even within the same country. Thus a limitation of our method is that for each new site where it is applied, individual vehicle data must be collected in order that the estimator can be calibrated. In practice, this means the installation of extra equipment at the roadside (so-called engineers' terminals) to capture individual vehicle data at source. Alternatively, we may use other techniques, such as the analysis

Table 6.7: Error indicators for scenario 1 in terms of different analysis intervals

	T (minute)	$RTCD$ (%)
Scenario 1	1	15.29
UK-Calibration	5	8.03
Thu. 2nd Oct 2008	10	6.51
	20	5.44
	30	4.94

Table 6.8: Error indicators for scenario 1 in terms of different portions of data set

	% of data set (random samples)	$RTCD$ (%)
Scenario 1	5	8.33
UK-Calibration	10	7.87
Thu. 2nd Oct 2008	25	6.71
$T = 10$	50	6.63
	75	6.60
	100	6.51

of CCTV¹ video pictures, to generate the ground truth data set. For whichever method is used, it involves considerable expenses and inconveniences. (If it did not, then there would be no purpose to the present study - because multi-lane multi-class counts would then be easily accessible in an operational context, without any of the estimation procedures that we have developed here.)

The upshot is that there is a strong practical interest in deriving an acceptable estimator with the bare minimum of calibration data. The results derived so far each use one full day of data. We now attempt to calibrate the estimator by using subsets of one full day of data, see Table 6.8, where we report the $RTCD$ when the estimator is then evaluated over the whole of that day's data. Because the flow characteristics vary with time of day, we have selected the analysis intervals at random so as to give a representative spread of lane utilisation behaviours. In that sense, the results presented here are slightly optimistic, because in practice one may only wish to collect data for a short contiguous period during the day, in which case the performance of the estimator which is thus calibrated may be very poor at other times of the day. However, the headline result is remarkable - the error is almost unchanged when we use one quarter of the original data, and acceptable results may be derived with just 5% - that is, with only one hour of individual vehicle data.

The second part of Q2 concerns the formal identification of error bars for the estimated parameters from any given (fixed-size) calibration set. To answer this question, we return to calibrating the parameters with one day of ground-truth data, and with the

¹CCTV: Closed-Circuit TeleVision.

Table 6.9: List of parameters and their standard deviations (in proportion to its mean) for scenario 1 based on the UK calibration set using bootstrapping (100 samples) ($T=10$, fitting for counts q_{11} , q_{22})

	Constant terms		Linear coefficients							
	b_1	b_2	A_{11}	A_{12}	A_{13}	A_{14}	A_{21}	A_{22}	A_{23}	A_{24}
mean	1.593	4.246	0.152	-0.316	-0.371	0.461	0.115	0.540	0.408	-0.404
std.	0.488	0.557	0.034	0.031	0.055	0.051	0.032	0.034	0.055	0.052
(%)	30.6	13.1	22.2	9.8	14.8	11.1	27.8	6.2	13.4	12.8

analysis interval fixed at $T=10$ minutes. To analyse the statistical distributions of the regression parameters, we should ideally have access to many days of individual vehicle data when the traffic conditions were similar, and perform the calibration procedure for each. However, because individual vehicle data are at a premium, this procedure will usually not be possible.

The technique we thus apply is the bootstrap method (Wu, 1986) to a single day of data which has been aggregated with $T=10$ minutes. In this method we randomly sample the real data (with replacement) to generate many emulated days' data, each of which should be broadly characteristic of the original. The regression procedure may then be carried out for each of the emulated days in turn. Thus the statistical properties of the regression parameters may be computed. Results are presented in Table 6.9 for a bootstrap giving a population of 100 emulated days. The standard deviation for each parameter is in relatively small proportion to its mean, thus generally lending support to the stability of the technique when calibrated over a full day of data. A more thorough analysis with formal confidence intervals etc. remains for future work.

6.3.7 Conclusions

The study in this section has proposed a new method to estimate multi-class and multi-lane counts from standard formats of aggregate loop detector data. We considered three formats as typical representatives of those captured around the world, namely (i) a UK-style format where counts are captured by class and by lane, but not jointly by class and lane; (ii) a US-style single-loop format in which only occupancy and counts are available (by lane); and (iii) a Spanish-style format in which counts are given by lane, and also according to a speed stratification - meaning that counts of vehicles in different speed bins are provided. In all cases, a simple linear regression method was able to estimate with some success the multi-lane multi-class flows, provided the analysis interval was taken sufficiently long (e.g. $T=10$ minutes) to reduce the effect of sampling error.

Not surprisingly, the method works best on the UK-style data, because this contains the most information. However, the minimal US-style data also yield good estimates of the

multi-lane, multi-class counts - probably because an increase in occupancy without a corresponding increase in flow provides information which tends to suggest an increase in the proportion of long vehicles. The method works least well on the Spanish-style data, although it is possible that alternative choices for the speed thresholds could deliver better results - this analysis remains for future work. Note that more complicated fits than those presented here are possible - for example, the UK-style calculation could incorporate speed and occupancy as input variables - in which case it would presumably deliver even more refined results. Similarly, although our initial analyses indicated that there was little benefit in incorporating nonlinear terms in the model functions, it is possible that they deliver improvements for more complicated regressions than those presented here. Throughout this study our approach has demonstrated the potential power of this very simple regression technique, rather than to exhaustively compute all the different ways in which it can possibly be applied.

A significant drawback of the method is that it requires the acquisition of representative samples of initial vehicle data (IVD) for each new site in which it will be operated. Obviously, if such IVD were continuously available at no cost, then our method would have no purpose - because the multi-lane multi-class counts could be computed exactly from the IVD in an ongoing operational fashion. So the potential added value in our method is that quantities which would normally need to be computed from IVD can instead be computed from standard aggregate loop data, albeit IVD must be collected for a short start-up period in order to calibrate the regression parameters. Our initial error analysis indicates that very short periods of IVD collection - perhaps as little as one hour - would be adequate for calibration, providing they were to contain a representative mix of the various flow regimes that will be observed at that site.

Unfortunately, one hour will not usually span a representative mix of flow regimes. In particular, the proportion of trucks is often significantly different at nights and at weekends, and so at the very least there must be calibration data for a range of different time periods. Indeed, for optimal accuracy, each site should probably have a different regression law depending on the day of week / time of day, and possibly on other aspects of the macroscopic traffic conditions. For instance, if speed controls / lane management are in operation, lane utilisation will usually be different, and a different regression law may be required to capture this behaviour.

Detailed studies of (i) how best the site-dependent calibration may be performed; (ii) of how lane utilisation varies from country to country; and (iii) of how these methods may be incorporated in dynamic traffic management, remain for future work.

6.4 Summary

This chapter has proposed two new methods to process data from traffic sensors in the real world, and thus can enhance their usefulness for multi-class traffic state estimation. The experimental results show that both methods are effective, efficient and robust.

They have a wide range of potential applications for both researchers and practitioners alike.

All these data pre-processing methods serve the purpose of accurate and reliable traffic state estimation. After the data pre-processing step, speed bias can be eliminated from observations, class-specific inputs can be obtained for multi-class state estimation. This will certainly benefit generic Lagrangian multi-class traffic state estimation systems in reality.

Chapter 7

Conclusions and recommendations

This last chapter synthesises the main findings of the thesis. Furthermore, the research implications for traffic state estimation studies are highlighted in terms of both scientific originality and societal relevance. The chapter ends by discussing a number of possible future research directions.

7.1 Main findings and conclusions

This thesis has developed a Lagrangian multi-class traffic state estimation method, which offers both computational and theoretical advantages over the conventional Eulerian method, as well as providing timely, accurate and reliable class-specific traffic information for Dynamic Traffic Management (DTM) at a network level. The data pre-processing methods that we have developed improve both model and observation inputs, and thus can additionally benefit real-world traffic state estimation.

In this study, the Lagrangian approach has been generalised to both mixed-class and multi-class levels. Three components in traffic state estimation are specified: 1) The Lagrangian mixed-class first-order traffic flow model is applied as the process model to describe the evolution of traffic states, where vehicle spacing is used as the system state. For the multi-class case, two modelling and discretisation choices are identified, namely the “Multi-pipe” and “Piggy-back” formulations. Their related pros and cons are discussed. The latter one is applied as the process model in the estimation framework that follows. 2) An improved differentiable Lagrangian fundamental relation is used to define the relations between traffic states and Lagrangian sensing observations. In addition, a new observation model for incorporating Eulerian sensing data is developed based on notions from first-order traffic flow theory. These models are applied for both mixed-class and multi-class methods. 3) As a real-time applicable method, the EKF is adopted for data assimilation. Because it is more efficient than other computational methods, it can be applied with non-mode-switching mixed-class and multi-class traffic systems. Moreover, the newly developed node models enable the extension of the Lagrangian traffic state estimation to a network level. A series of experimental studies based on both synthetic and real-world data have been performed to test and validate the proposed methods. Both Eulerian and Lagrangian sensing data are incorporated within the Lagrangian state estimation.

Firstly, experimental studies have validated both the mixed-class and the multi-class traffic state estimation approaches. Secondly, since more accurate estimation results can be achieved in the experiments, the Lagrangian mixed-class traffic state estimation has been seen to outperform its Eulerian counterpart. It also offers both theoretical and computational benefits. This is due to improvements in both the prediction step and the correction step of the data-assimilation method: (i) the Lagrangian formulation enables more accurate and efficient simulation of freeway traffic and thus leads to more accurate predictions; (ii) due to the non-mode-switching numerical solution, the Lagrangian method is a more appropriate choice for the application in the EKF. Accordingly linearisation of the process model around capacity is a much better approximation in Lagrangian coordinates than in Eulerian coordinates. In the latter case, it may lead to sign errors, whereas in the former the errors pertain to the magnitude of the corrections only. Therefore, this leads to improved performance particularly at the boundaries of congestion (state transitions). Thirdly, the multi-class approach makes proper use of diverse types of class-specific observation data, and it thus improves the

performance of state estimation compared with the mixed-class approach. The experiment that we presented on a real traffic network (the A15) demonstrates that multi-class Lagrangian state estimation can provide accurate and reliable class-specific traffic information at a network level. This offers another efficient and reliable opportunity for real-time traffic state estimation in traffic management.

This thesis additionally addresses the importance of data pre-processing and preparation for the application of multi-class traffic state estimation in the real world. Two methods have been developed. The first one is a speed-bias correction algorithm. It corrects biased speeds inherited from dual-loop detectors, based on the first-order traffic flow theory and empirical flow-density relationships. The other one develops a procedure for estimating multi-class and multi-lane traffic counts based on a variety of standard aggregate loop data formats. It relies on the notions from inductive statistics. We can conclude from the experimental results that both methods are quite effective, efficient and robust.

7.2 Research implications

The Lagrangian multi-class formulations for traffic state estimation research are scientifically innovative and provide practical solutions for problems which cannot be properly addressed by existing methods. This thesis has demonstrated that the proposed Lagrangian formulation is suitable and beneficial for traffic state estimation, in terms of traffic simulations, incorporating Lagrangian data, and the application of the data-assimilation (EKF) method. Therefore, it suggests the future employment and investigation of the Lagrangian formulation for traffic state estimation, as a new research direction. The multi-class state estimator can make proper use of class-specific observations (e.g., multi-class loop or trajectory data) to estimate class-specific traffic states. Therefore, the newly developed Lagrangian multi-class traffic state estimator can be applied to promote a model-based decision support system for real-time traffic network management.

This study has presented a novel classification framework for model-based state estimation research with regards to different coordinate systems (mathematical formulations). This taxonomy allows the identification of potentially beneficial research angles. The identified research gaps based on the new taxonomy have been filled and investigated by the current study. However, this study only applies the first-order traffic flow theory for traffic modelling and the EKF for data assimilation. This implies that other traffic flow modelling and data-assimilation approaches can be further studied in the framework of Lagrangian formulation, with advantages expected along the same lines as shown in this thesis.

Traditionally, the performance of traffic state estimation can be improved by using more advanced and sophisticated assimilation methods. This research has shown that

its performance can alternatively be enhanced by choosing more appropriate and suitable traffic system models (which are used in the data-assimilation framework). First of all, the Lagrangian formulation of the first-order traffic flow model does not require mode switching. This characteristic provides more accurate simulation results and straightforward numerical discretisations, which is beneficial for state estimation. Second, the Lagrangian formulation provides an ideal framework to incorporate the data from probe vehicles and mobile phones. Therefore, these Lagrangian sensing data can be used for assimilation to improve accuracy.

Meanwhile, the Lagrangian state estimation method sheds some light on in-car localised information (state estimation) systems. The concept in this thesis can be used in vehicle-to-vehicle cooperative systems. Individual vehicles or vehicle platoons can be treated as independent state estimation units, locally estimating the states of their adjacent moving units or transmitting the states of themselves.

This thesis has also developed two new methods to pre-process data from traffic sensors, with the purpose of data cleaning and data mining. These methods enable the application of Lagrangian multi-class traffic state estimation approach in the real world. They also have a wide range of potential applications for both researchers and practitioners alike.

7.3 Future research directions

In the final section we provide some possible directions for future research. First of all, the concept of Lagrangian formulation is not restricted to the first-order traffic flow model with the EKF technique, but can be further applied to other data-assimilation techniques (e.g., UKF, EnsKF, and PF) combining with other types of macroscopic traffic flow models (e.g., second- or even higher-order traffic flow models). For instance, the Lagrangian formulation can be applied to a multi-class second-order traffic flow model within a particle filtering framework, for which we expect similar computational and theoretical improvements as those shown in this study.

Regarding multi-class traffic flow modelling and discretisation in Lagrangian coordinates, this thesis presents two alternatives: 1) the “Piggy-back” formulation, where only one vehicle coordinate system is used for the reference class; 2) the “Multi-pipe” formulation, where separate coordinates for different vehicle classes are introduced. The “Piggy-back” formulation is used in this study since it is suitable for the EKF method. However, there are limitations regarding class-specific control and modelling network discontinuities. To further improve state estimation for class-specific traffic management, a “multi-pipe” model could be employed within a different data-assimilation framework. Meanwhile, the additional computational effort should be taken into account.

The proposed state estimation method in this study applies a global approach for the EKF. This means the assimilation procedure operates on a global matrix containing

all system state variables in the entire network. When the state dimension increases with the network scale, the efficiency of this method will be limited by the required computational effort. This problem can be solved by using a localised EKF (L-EKF) technique (Van Hinsbergen et al., 2012). In the L-EKF, corrections are only performed to the states of cells that have considerable error covariance with the observation cells, indicating that operation with a global state matrix is no longer required. A Lagrangian formulated state estimation based on the L-EKF remains for future work.

In this thesis, the concept of network-wide state estimation has been validated based on a relatively simple example considering on-ramp sources and off-ramp sinks. More complex situations related to network discontinuities should be taken into account into traffic state estimation research, such as spillbacks onto the main road upstream of ramps and/or onto on-ramps. Further work is needed to improve the current implementation of the node models, and to test and demonstrate the Lagrangian estimator on more complex and realistic traffic networks.

Bibliography

- Anthes, R. A. (1974) Data assimilation and initialization of hurricane prediction models, *Journal of the Atmospheric Sciences*, 31(3), pp. 702–719.
- Bando, M., K. Hasebe, A. Nakayama, A. Shibata, Y. Sugiyama (1995) Dynamical model of traffic congestion and numerical simulation, *Physical Review*, 51(2), pp. 1035–1042.
- Barceló, J., E. Codina, J. Casas, J. L. Ferrer, D. García (2004) Microscopic traffic simulation: A tool for the design, analysis and evaluation of intelligent transport systems, *Journal of Basic Engineering*, 41, pp. 173–203.
- Bellomo, N., C. Dogbe (2011) On the modeling of traffic and crowds: A survey of models, speculations, and perspectives, *SIAM Review*, 53, pp. 409–463.
- Boel, R., L. Mihaylova (2006) A compositional stochastic model for real time freeway traffic simulation, *Transportation Research Part B: Methodological*, 40(4), pp. 319 – 334.
- Byon, Y.-J., A. Shalaby, B. Abdulhai, S. El-Tantawy (2010) Traffic data fusion using scaat kalman filters, in: *Proceedings of the Transportation Research Board 89th Annual meeting*, Transportation Research Board, Washintong, D.C.
- Cassidy, M. (1998) Bivariate relations in nearly stationary highway traffic, *Transportation Research Part B: Methodological*, 32(1), pp. 49–59.
- Chen, H., H. Rakha, S. Sadek (2011) Real-time freeway traffic state prediction: A particle filter approach, in: *Proceedings of the 2011 IEEE Conference on Intelligent Transportation Systems (ITSC)*, Washington, D.C., pp. 626–631.
- Cheng, P., Z. Qiu, B. Ran (2006) Traffic estimation based on particle filtering with stochastic state reconstruction using mobile network data, in: *Proceedings of the Transportation Research Board 85th Annual Meeting*, Washington D.C.
- Chu, L., S. Oh, W. Recker (2005) Adaptive kalman filter based freeway travel time estimation, in: *Proceedings of the Transportation Research Board 84th Annual Meeting*, Washington D.C.

- Coifman, B. (2001) Improved velocity estimation using single loop detectors, *Transportation Research: Part A*, 35(10), pp. 863–880.
- Coifman, B. (2002) Estimating travel time and vehicle trajectories on freeways using dual loop detectors, *Transportation Research: Part A*, 36(4), pp. 351–364.
- Coifman, B., S. Kim (2009) Speed estimation and length based vehicle classification from freeway single-loop detectors, *Transportation Research: Part C*, 17(10), p. 349C364.
- Coric, V., N. Djuric, S. Vucetic (2012) Traffic state estimation from aggregated measurements using signal reconstruction techniques, in: *Proceedings of the Transportation Research Board 91th Annual Meeting*, Washington D.C., pp. 1–15.
- Courant, R., K. Friedrichs, H. Levy (1967) On the partial difference equations of mathematical physics, *IBM Journal on Research Development*, 11(2), pp. 215–234.
- Daganzo, C. F. (1994) The cell transmission model: A dynamic representation of highway traffic consistent with the hydrodynamic theory, *Transportation Research Part B: Methodological*, 28(4), pp. 269–287.
- Daganzo, C. F. (1995a) The cell transmission model, part ii: Network traffic, *Transportation Research Part B: Methodological*, 29(2), pp. 79–93.
- Daganzo, C. F. (1995b) A finite difference approximation of the kinematic wave model of traffic flow, *Transportation Research Part B: Methodological*, 29(4), pp. 261–276.
- Daganzo, C. F. (1995c) Requiem for second-order fluid approximations of traffic flow, *Transportation Research Part B: Methodological*, 29(4), pp. 277–286.
- Daganzo, C. F. (2002a) A behavioral theory of multi-lane traffic flow. part i: Long homogeneous freeway sections, *Transportation Research Part B: Methodological*, 36, pp. 131–158.
- Daganzo, C. F. (2002b) A behavioral theory of multi-lane traffic flow. part ii: Merges and the onset of congestion, *Transportation Research Part B: Methodological*, 36, pp. 159–169.
- Dailey, D. (1999) A statistical algorithm for estimating speed from single loop volume and occupancy measurements, *Transportation Research: Part B*, 33(5), pp. 313–322.
- Davis, G. A., J.-G. Kang (1994) Estimating destination-specific traffic densities on urban freeways for advanced traffic management, *Transportation Research Record*, 1457, pp. 143–148.
- Deo, P., B. De Schutter, A. Hegyi (2009) Model predictive control for multi-class traffic flows, in: *Proceedings of the 12th IFAC Symposium on Transportation Systems*, Redondo Beach, California, pp. 25–30.

- Dervisoglu, G., G. Gomes, J. Kwon, A. Muralidharan, P. Varaiya, R. Horowitz (2009) Automatic calibration of the fundamental diagram and empirical observations on capacity, in: *Proceedings of the Transportation Research Board 88th Annual Meeting*, Washington, D.C., pp. 1–15.
- Di, X., H. Liu, G. Davis (2010) A hybrid extended kalman filtering approach for traffic density estimation along signalized arterials using gps data, in: *Proceedings of the Transportation Research Board 89th Annual Meeting*, Washington D.C.
- Dijker, T. (2012) Fosim (freeway operations simulation), URL <http://www.fosim.nl>.
- Doucet, A., N. Freitas, N. Gordon (2001) *Sequential Monte Carlo methods in practice*, Springer, New York.
- Eddie, L. (1965) Discussion of traffic stream measurements and definitions, in: *Proceedings of the 2nd International Symposium on Theory of Traffic Flow*, Paris, France, pp. 139–154.
- Evensen, G. (2007) *Data Assimilation: The Ensemble Kalman Filter*, Springer-Verlag, Berlin Heidelberg.
- FHWA (2012) Next generation simulation, URL <http://ngsim-community.com>.
- Freedman, D. A. (2005) *Statistical Models: Theory and Practice*, Cambridge University Press, Cambridge.
- Gordon, N., D. Salmond, A. Smith (1993) A novel approach to nonlinear/non-gaussian bayesian state estimation, *IEEE Proceedings on Radar and Signal Processing*, 40, pp. 107–113.
- Greenshields, B. (1934) A study of traffic capacity, *Highway Research Board Proceedings*, 14, pp. 468–477.
- Haykin, S. (2001) *Kalman Filtering and Neural Networks*, John Wiley & Sons, Inc., New York.
- Hegy, A., D. Girimonte, R. Babuska, B. De Schutter (2006) A comparison of filter configurations for freeway traffic state estimation, in: *Proceedings of Intelligent Transportation Systems Conference, ITSC '06. IEEE*, Toronto, Canada, pp. 1029–1034.
- Helbing, D. (2001) Traffic and related self-driven many-particle systems, *Reviews of Modern Physics*, 73(4), pp. 1067–1141.
- Herrera, J. C., D. B. Work, R. Herring, X. Ban, Q. Jacobson, A. M. Bayen (2010) Evaluation of traffic data obtained via gps-enabled mobile phones: The mobile century field experiment, *Transportation Research Part C: Emerging Technologies*, 18(4), pp. 568–583.

- Highway-Agency (2011) URL <http://www.highways.gov.uk/knowledge/1334.aspx>, accessed July 2011.
- Highway-Agency (2012) Live traffic information covering england's motorways and major a-roads, URL <http://www.trafficengland.com>.
- Hoogendoorn, S. P. (1999) *Multiclass Continuum Modelling of Multilane Traffic Flow*, Ph.d. dissertation, TRAIL Thesis series. Delft University of Technology.
- Hoogendoorn, S. P. (2001) Model-based multiclass travel time estimation, in: *Proceedings of the 9th World Conference on Transport Research*, Seoul, Korea.
- Hoogendoorn, S. P. (2008) *Traffic Flow Theory and Simulation*, Delft University of Technology, Delft.
- Hoogendoorn, S. P., P. H. L. Bovy (1999) Multiclass macroscopic traffic flow modelling: a multilane generalisation using gas-kinetic theory, in: A., C., ed., *Proceedings of the 14th International Symposium on Transportation and Traffic Theory*, Jerusalem, Israel, pp. 27–50.
- Jain, M., B. Coifman (2005) Improved speed estimates from freeway traffic detectors, *ASCE Journal of Transportation Engineering*, 131(7), pp. 483–495.
- Jazwinsky, A. (1970) *Stochastic processes and filtering theory*, Academic Press, New York.
- Julier, S. J., J. K. Uhlmann (1997) A new extension of the kalman filter to nonlinear systems, in: *Proceedings of AeroSense: The 11th Int. Symp. on Aerospace/Defence Sensing, Simulation and Controls*, pp. 182–193.
- Kalman, R. (1960) A new approach to linear filtering and prediction problems, *Journal of Basic Engineering*, 82(1), pp. 35–45.
- Kerner, B. (2004) *The Physics of Traffic*, Springer-Verlag, Berlin, Germany.
- Kerner, B. (2009) *Introduction to Modern Traffic Flow Theory and Control*, Springer-Verlag, Berlin.
- Kerner, B., H. Rehborn (1997) Experimental properties of phase transitions in traffic flow, *Physical Review Letters*, 79(20).
- Knoppers, P., J. W. C. Van Lint, S. P. Hoogendoorn (2012) Automatic stabilization of aerial traffic images, in: *Proceedings of the Transportation Research Board 91st Annual Meeting*, Washington, D.C., pp. 1–13.
- Kockelman, K. (1998) Changes in flow-density relationship due to environmental, vehicle, and driver characteristics, *Transportation Research Record*, 1644, pp. 47–56.

- Laval, J. A., L. Leclercq (2013) Hamilton-jacobi partial differential equation and three representations of traffic flow, in: *Proceedings of the Transportation Research Board 92nd Annual Meeting*, Washington D.C., pp. 1–21.
- Lebacque, J. P. (1996) The gudunov scheme and what it means for first order traffic flow models, in: Lesort, J., ed., *Proceedings of the 13th International Symposium on Transportation and Traffic Theory*, Lyon, France, pp. 647–677.
- Leclercq, L., C. Bécarie (2012) A meso lighthill-whitham and richards model designed for network applications, in: *Proceedings of the Transportation Research Board 91th Annual Meeting*, Washington D.C., pp. 1–15.
- Leclercq, L., J. A. Laval, E. Chevallier (2007) The lagrangian coordinates and what it means for first order traffic flow models, in: Allsop, R., M. Bell, B. Heydecker, eds., *Proceedings of the 17th International Symposium on Transportation and Traffic Theory*, Elsevier, London, U.K., pp. 735–753.
- Leutzbach, W. (1987) *Introduction to the Theory of Traffic Flow*, Springer-Verlag, Berlin.
- Lighthill, M., G. Whitham (1955) On kinematic waves ii: A theory of traffic flow on long crowded roads, *Proceedings of Royal Society*, 229A(1178), pp. 317–345.
- Lindveld, C., R. Thijs (1999) On-line travel time estimation using inductive loop data: The effect of instrumentation peculiarities on travel time estimation quality, in: *Proceedings of The 6th ITS World Congress*, Toronto, Canada.
- Logghe, S., L. Immers (2008) Multi-class kinematic wave theory of traffic flow, *Transportation Research Part B: Methodological*, 42(6), pp. 523–541.
- Louis, G., P. Neudorff, E. Jeffrey, P. Randall, P. Robert Reiss, P. Robert Gordon (2006) Freeway management and operations handbook, Tech. rep., FHWA, U.S. Department of Transportation.
- Mihaylova, L., R. Boel, A. Hegyi (2007) Freeway traffic estimation within recursive bayesian framework, *Automatica*, 43(2), pp. 290–300.
- Mihaylova, L., A. Hegyi, A. Gning, R. Boel (2012) Parallelized particle and gaussian sum particle filters for large-scale freeway traffic systems, *Intelligent Transportation Systems, IEEE Transactions on*, 13(1), pp. 36–48.
- Nanthawichit, C., T. Nakatsuji, H. Suzuki (2003) Application of probe-vehicle data for real-time traffic-state estimation and short-term travel-time prediction on a freeway, *Transportation Research Record*, 1855, pp. 49–59.
- Newell, G. (1993) A simplified theory of kinematic waves in highway traffic, part i: General theory, *Transportation Research Part B: Methodological*, 27B(4), pp. 281–287.

- Ngoduy, D. (2008) Applicable filtering framework for online multiclass freeway network estimation., *Physica A: Statistical Mechanics and its Applications*, 387(2/3), pp. 599–616.
- Ngoduy, D. (2011) Low rank unscented kalman filter for freeway traffic estimation problems, in: *Proceedings of the Transportation Research Board 90th Annual Meeting*, Transportation Research Board, Washington D.C.
- Ngoduy, D., R. Liu (2007) Multiclass first-order simulation model to explain non-linear traffic phenomena, *Physica A: Statistical Mechanics and its Applications*, 385(2), pp. 667–682.
- Ou, Q. (2011) *Fusing Heterogeneous Traffic Data: Parsimonious Approaches Using Data-Data Consistency*, Ph.d. dissertation, TRAIL Thesis series. Delft University of Technology.
- Papageorgiou, M., J. M. Blosseville, H. Hadj-Salem (1990) Modelling and real-time control of traffic flow on the southern part of boulevard peripherique in paris: Part i: Modelling, *Transportation Research Part A: General*, 24(5), pp. 345–359.
- Payne, H. (1971) Models of freeway traffic and control, *Simulation Councils Proceedings Series: Mathematical Models of Public Systems*, 1(1), pp. 51–61.
- PTV (2010) Vissim user manual, Tech. rep., PTV.
- Pueboobpaphan, R., T. Nakatsuji, H. Suzuki (2007) Unscented kalman filter-based real-time traffic state estimation, in: *Proceedings of the Transportation Research Board 86th Annual Meeting*, Washington D.C., pp. 1–19.
- Richards, P. (1956) Shock waves on the highway, *Operations Research*, 4(1), pp. 42–51.
- Rijkswaterstaat (2003) Sustainable traffic management handbook, Tech. rep., Dutch Ministry of Transport, Public Works and Water Management.
- Schönhof, M., D. Helbing (2007) Empirical features of congested traffic states and their implications for traffic modeling, *Transportation Science*, 41(2), pp. 135–166.
- Schreiter, T. (2013) *Vehicle-class specific control of freeway traffic*, Ph.d. dissertation, TRAIL Thesis series. Delft University of Technology.
- Schreiter, T., R. L. Landman, J. W. C. Van Lint, A. Hegyi, S. P. Hoogendoorn (2012) Vehicle-class specific route-guidance of freeway traffic by model-predictive control, in: *Proceedings of the Transportation Research Board 91st Annual Meeting*, Washington, D.C., pp. 1–18.
- Schreiter, T., C. P. I. J. Van Hinsbergen, F. S. Zuurbier, J. W. C. Van Lint, S. P. Hoogendoorn (2010a) Data-model synchronization in extended kalman filters for accurate

- online traffic state estimation, in: *Proceeding of the Summer Meeting of the Transportation Research Board*, Annecy, France, pp. 1–17.
- Schreiter, T., J. W. C. Van Lint, Y. Yuan, S. P. Hoogendoorn (2010b) Propagation wave speed estimation of freeway traffic with image processing tools, in: *Proceedings of the Transportation Research Board 89th Annual Meeting*, Washington D.C., pp. 1–17.
- Smulders, S. A. (1989) *Control of Freeway Traffic Flow*, Ph.d. dissertation, TU Twente.
- Soriguera, F., F. Robusté (2011) Estimation of traffic stream space mean speed from time aggregations of double loop detector data, *Transportation Research Part C: Emerging Technologies*, 19(1), pp. 115–129.
- Stipdonk, H., M. Postema (2009) On the congested motorway traffic paradox.
- Stipdonk, H., J. Van Toorenburg, M. Postema (2008) Phase diagram distortion from traffic parameter averaging, in: *Proceedings of European Transport Conference*, pp. 1–13.
- Sun, X., L. Munoz, R. Horowitz (2003) Highway traffic state estimation using improved mixture kalman filters for effective ramp metering control, in: *Proceedings of the 42nd IEEE Conference on Decision and Control*, Maui, pp. 6333–6338.
- Tampère, C., B. Immers (2007) An extended kalman filter application for traffic state estimation using ctm with implicit mode switching and dynamic parameters, in: *Proceedings of the 2007 IEEE Intelligent Transportation Systems Conference*, IEEE, Seattle, WA, USA, pp. 209–216.
- Tchraikian, T. T., O. Verscheure (2011) A lagrangian state-space representation of a macroscopic traffic flow model, in: *Proceedings of the 2011 IEEE Conference on Intelligent Transportation Systems (ITSC)*, Washington, D.C., pp. 632–637.
- Transportation-Research-Board (2000) Hcm2000: Highway capacity manual, Tech. rep., TRB, National Research Council.
- Treiber, M., D. Helbing (2002) Reconstructing the spatio-temporal traffic dynamics from stationary detector data, *Cooperative Transportation Dynamics*, 1, pp. 3.1–3.24.
- Treiber, M., A. Hennecke, D. Helbing (1999) Derivation, properties, and simulation of a gas-kinetic-based, nonlocal traffic model, *Phys. Rev. E*, 59, pp. 239–253.
- UCBerkeley (2008) Mobile millennium: A community-based traffic information system, URL <http://traffic.berkeley.edu>.
- Van Hinsbergen, C. P. I. J., T. Schreiter, J. W. C. Van Lint, S. P. Hoogendoorn, H. J. Van Zuylen (2010) Online estimation of kalman filter parameters for traffic state estimation, in: *Proceedings of TRISTAN VII*, Norway.

- Van Hinsbergen, C. P. I. J., T. Schreiter, F. S. Zuurbier, J. W. C. van Lint, H. J. van Zuylen (2012) Localized extended kalman filter for scalable real-time traffic state estimation, *IEEE Transactions on Intelligent Transportation Systems*, 13(1), pp. 385–394.
- Van Lint, J. W. C. (2004) *Reliable Travel Time Prediction for Freeways*, Ph.d. dissertation, Delft University of Technology.
- Van Lint, J. W. C. (2010) Empirical evaluation of new robust travel time estimation algorithms, *Transportation Research Record*, 2160, pp. 50–59.
- Van Lint, J. W. C., S. P. Hoogendoorn (2009) A robust and efficient method for fusing heterogeneous data from traffic sensors on freeways, *Computer-Aided Civil and Infrastructure Engineering*, 25(8), pp. 596–612.
- Van Lint, J. W. C., S. P. Hoogendoorn, A. Hegyi (2008a) Dual ekf state and parameter estimation in multi-class first-order traffic flow models, in: *Preprints of the 17th IFAC World Congress*, IFAC, Seoul, Korea, pp. 14078–14083.
- Van Lint, J. W. C., S. P. Hoogendoorn, M. Schreuder (2008b) Fastlane - a new multiclass first-order traffic model, *Transportation Research Record: Journal of the Transportation Research Board*, 2088(1), pp. 177–187.
- Van Lint, J. W. C., N. J. Van der Zijpp (2003) Improving a travel time estimation algorithm by using dual loop detectors, *Transportation Research Record*, 1855, pp. 41–48.
- Van Wageningen-Kessels, F. L. M. (2011) Initialization of fastlane, Tech. rep., TU Delft.
- Van Wageningen-Kessels, F. L. M. (2013) *Multi-class continuum traffic flow models: Analysis and simulation methods*, Ph.d. dissertation, TRAIL Thesis series. Delft University of Technology.
- Van Wageningen-Kessels, F. L. M., J. W. C. Van Lint, S. P. Hoogendoorn, C. Vuik (2009a) Implicit time stepping schemes applied to the kinematic wave model in lagrangian coordinates, in: *Traffic and Granular Flow 2009*, Shanghai, China.
- Van Wageningen-Kessels, F. L. M., J. W. C. Van Lint, S. P. Hoogendoorn, C. Vuik (2009b) Multiple user classes in the kinematic wave model in lagrangian coordinates, in: *Proceedings of the Conference on Traffic and Granular Flow 2009*, Shanghai, China, pp. 1–7.
- Van Wageningen-Kessels, F. L. M., J. W. C. Van Lint, S. P. Hoogendoorn, C. Vuik (2010a) Lagrangian formulation of multiclass kinematic wave model, *Transportation Research Record: Journal of the Transportation Research Board*, 2188(1), pp. 29–36.

- Van Wageningen-Kessels, F. L. M., B. van't Hof, S. P. Hoogendoorn, K. Vuik, J. W. C. van Lint (2011a) Anisotropy in generic multi-class traffic flow models, *Transportmetrica*, Online, pp. 1–22.
- Van Wageningen-Kessels, F. L. M., Y. Yuan, S. P. Hoogendoorn, J. W. C. Van Lint, C. Vuik (2011b) Discontinuities in the lagrangian formulation of the kinematic wave model, *Transportation Research Part C: Emerging Technologies*, Online.
- Van Wageningen-Kessels, F. L. M., Y. Yuan, J. W. C. Van Lint, S. P. Hoogendoorn, C. Vuik (2010b) Discontinuities in the lagrangian formulation of the kinematic wave model., in: *Proceeding of the Summer Meeting of the Transportation Research Board*, Annecy, France.
- Van Wageningen-Kessels, F. L. M., Y. Yuan, J. W. C. Van Lint, S. P. Hoogendoorn, C. Vuik (2011c) Sinks and sources in lagrangian coordinates: derivation, interpretation and simulation results, in: *Proceedings of the Transportation Research Board 90th Annual Meeting*, Washington D.C.
- Wang, Y., M. Papageorgiou (2005) Real-time freeway traffic state estimation based on extended kalman filter: A general approach, *Transportation Research Part B: Methodological*, 39(2), pp. 141–167.
- Wardrop, J. G. (1952) Some theoretical aspects of road traffic research, in: *Proceedings of the Institute of Civil Engineers, Part II*, pp. 325–378.
- Webster, N., L. Elefteriadou (1999) A simulation study of truck passenger car equivalents (pce) on basic freeway sections, *Transportation Research Part B: Methodological*, 33(5), pp. 323–336.
- Wilson, R. E. (2008) From inductance loops to vehicle trajectories, in: *Proceedings of Symposium on the Fundamental Diagram: 75 Years (Greenshields 75 Symposium)*, Woods Hole, MA, pp. 134–143.
- Wilson, R. E. (2011) URL <http://www.enm.bris.ac.uk/trafficdata>, accessed July 2011.
- Windover, J., M. J. Cassidy (2001) Some observed details of freeway traffic evolution, *Transportation Research Part A: Policy and Practice*, 35, pp. 881–894.
- Wong, G. C. K., S. C. Wong (2002) A multi-class traffic flow model - an extension of lwr model with heterogeneous drivers, *Transportation Research Part A: Policy and Practice*, 36(9), pp. 827–841.
- Work, D., O.-P. Tossavainen, S. Blandin, A. Bayen, T. Iwuchukwu, K. Tracton (2008) An ensemble kalman filtering approach to highway traffic estimation using gps enabled mobile devices, in: *Proceedings of the 47th IEEE Conference on Decision and Control*, Cancun, Mexico, pp. 2141–2147.

- Wu, C. F. J. (1986) Jackknife, bootstrap and other resampling methods in regression analysis, *Annals of Statistics*, 14, pp. 1261–1295.
- Yuan, Y., J. W. C. van Lint, S. P. Hoogendoorn, J. L. M. Vrancken, T. Schreiter (2011a) Freeway traffic state estimation using extended kalman filter for first-order traffic model in lagrangian coordinates, in: *Proceedings of the 2011 IEEE International Conference on Networking, Sensing and Control (ICNSC)*, Delft, The Netherlands, pp. 121–126.
- Yuan, Y., J. W. C. Van Lint, T. Schreiter, S. P. Hoogendoorn, J. L. M. Vrancken (2010) Automatic speed-bias correction with flow-density relationships, in: *Proceedings of the 2010 International Conference on Networking, Sensing and Control (ICNSC)*, Chicago, pp. 1–7.
- Yuan, Y., J. W. C. Van Lint, F. L. M. Van Wageningen-Kessels, S. P. Hoogendoorn, J. L. M. Vrancken (2011b) Lagrangian traffic state estimation for freeway networks, in: *Proceedings of MT-ITS 2011*, Leuven, Belgium, pp. 1–4.
- Yuan, Y., J. W. C. Van Lint, R. E. Wilson, F. L. M. Van Wageningen-Kessels, S. P. Hoogendoorn (2012) Real-time lagrangian traffic state estimator for freeways, *IEEE Transactions on Intelligent Transportation Systems*, 13(1), pp. 59–70.
- Yuan, Y., R. E. Wilson, J. W. C. Van Lint, S. P. Hoogendoorn (2012. In Press) Estimation of multi-class and multi-lane counts from aggregate loop detector data, *Transportation Research Record*.

Appendices

Appendix A

Estimation of class-specific information from mixed-class data

In this section, a method is developed to extract class-specific (density, speed, spacing) information out from the mixed-class aggregate data. The class-specific outputs can be further used in multi-class traffic state estimators (as inflow inputs, observations). In the following, the methodology is presented without giving the proving procedure. Details about this method can be found in (Van Wageningen-Kessels, 2011).

A.1 Problem description

Initial data for this algorithm consist of total mixed-class densities (in veh./m) and flow shares per class (e.g., in cars per hour/vehicles per hour). From this data, the class-specific densities (in veh./m), speeds (in m/s) and spacing (in m/veh.) have to be derived. We will first do this assuming only two vehicle classes. Later the method will be extended for the general case with u user classes. In the following we will use these symbols:

k_{mix}	total mixed-class density	veh./m
k_{tot}	effective density	PCE/m
k_u	density of class u , $\forall u$	veh./m
s_u	spacing of class u ($s_u = 1/k_u$)	m/veh.
v_u	speed of class u	m/s
q_u	flow of class u	veh./s
$V_u(k_{\text{tot}})$	fundamental diagram for class u	m/s
p_u	flow share of class u	[]
η_u	PCE value of class u	[]
$E_u(v_1, v_u)$	PCE function of class u (function of speeds)	[]
$E_u^*(k_{\text{tot}})$	PCE function of class u (function of total effective density)	[]

A.2 Solution procedure for two vehicle classes

Step 1. Solve for $k_{\text{tot}} \in [0, k_{\text{tot,max}}]$:

$$\left[\frac{p_2}{p_1} V_1(k_{\text{tot}}) + V_2(k_{\text{tot}}) \right] k_{\text{tot}} - \left[\frac{p_2}{p_1} E_2^*(k_{\text{tot}}) V_1(k_{\text{tot}}) + V_2(k_{\text{tot}}) \right] k_{\text{mix}} = 0. \quad (\text{A.1})$$

Step 2. Use the fundamental diagram to find the speeds:

$$v_1 = V_1(k_{\text{tot}}) \text{ and } v_2 = V_2(k_{\text{tot}}). \quad (\text{A.2})$$

Step 3. Find k_1 and k_2 by using:

$$k_1 = \frac{v_2}{\frac{p_2}{p_1} v_1 + v_2} k_{\text{mix}} \text{ and } k_2 = k_{\text{mix}} - k_1. \quad (\text{A.3})$$

Step 4. Find s_1 and s_2 based on the reciprocal relations:

$$s_1 = 1/k_1 \text{ and } s_2 = 1/k_2. \quad (\text{A.4})$$

A.3 Solution procedure for U vehicle classes

Step 1. Solve for $k_{\text{tot}} \in [0, k_{\text{tot,max}}]$:

$$\sum_{u=1}^U \left\{ p_u [k_{\text{tot}} - E_u(k_{\text{tot}}) k_{\text{mix}}] \prod_{\substack{i=1 \\ i \neq u}}^U V_i(k_{\text{tot}}) \right\} = 0. \quad (\text{A.5})$$

Step 2. Use the fundamental diagram to find the speeds:

$$v_u = V_u(k_{\text{tot}}), \forall u. \quad (\text{A.6})$$

Step 3. Find k_1 and all other k_u by using:

$$k_1 = \frac{p_1 \prod_{i=2}^U v_i}{\sum_{u=1}^U [p_u \prod_{i=1, i \neq u}^U v_i]} k_{\text{mix}} \text{ and } k_u = \frac{p_u v_1}{p_1 v_u} k_1. \quad (\text{A.7})$$

Step 4. Find s_u based on the reciprocal relations:

$$s_u = 1/k_u, \forall u. \quad (\text{A.8})$$

Summary

Road traffic is important to everybody in the world. People travel and commute every day. For those who travel by cars (or other types of road vehicles), traffic congestion is a daily experience. One essential goal of traffic researchers is to reduce traffic congestion and to improve the whole traffic system operation and the environment. To achieve this goal, we have to first understand prevailing traffic situations, then perform pro-active traffic control and management. The estimation of traffic states in the past, in the present and in the future plays an important role in traffic management and control systems. This thesis focuses on the development of traffic state estimation approaches, which provide such traffic state information.

In road networks, traffic states refer to typical quantities, such as travel times, traffic speeds, traffic flow and density. These quantities reflect the traffic conditions. Based on these data, we are able to find out when a traffic jam starts, or where a traffic accident occurs.

However, it is not feasible to get the full picture of traffic states from the current monitoring systems. Due to cost and technical constraints, we can only obtain spatially and temporally discretised traffic data. These traffic data are collected mainly from point-based sensors, such as inductive loops, radars, and cameras. Alternatively, traffic information might be observed by probe vehicles with a selected penetration rate. In all cases, the detected data usually contain errors and noise, which might hinder further analyses. Based on these constraints, this thesis aims to develop a traffic state estimation procedure to solve the foregoing problems and to provide accurate and complete traffic state information. In this procedure, both traffic flow models and the available observation data are used to estimate the most probable traffic states within a data-assimilation framework.

Our approach is formulated using a moving observer perspective, resulting in a Lagrangian formulation of traffic state estimation. In the Lagrangian coordinate system, coordinates move with the vehicles. The Lagrangian formulated first-order traffic flow model is applied to describe the evolution of traffic state variables. The proposed Lagrangian formulation of traffic state estimation offers both theoretical and computational advantages over the conventional Eulerian formation. Moreover, this approach can capture the dynamics of multiple vehicle classes by implementing a multi-class traffic flow model. In this thesis, data pre-processing methods are also developed to

improve the quality of the observation inputs. Both Eulerian and Lagrangian sensing data are incorporated into the state estimation. The online technique, known as the Extended Kalman Filter (EKF), is applied for data assimilation: this combines traffic model prediction with observation input correction. Importantly, the Lagrangian concept is not restricted to the EKF method with the first-order traffic flow model, but can also be applied to other data-assimilation techniques in combination with more involved macroscopic traffic flow models. A series of experimental studies based on both synthetic and real-world data have been performed to test the proposed methodology. These studies have validated both the mixed-class and the multi-class traffic state estimation methods. The results have demonstrated that the Lagrangian traffic state estimation outperforms the Eulerian approaches in the EKF-based framework, and the multi-class approach further improves the performance of state estimation compared with the mixed-class case. In summary, Lagrangian multi-class state estimation can provide accurate class-specific traffic information for class-specific control applications and traffic management.

Samenvatting

Wegverkeer is belangrijk voor iedereen ter wereld. Mensen reizen elke dag. Dege-
nen die reizen per auto (of een ander type motorvoertuig) ervaren dagelijks congestie.
Essentiële doelen van verkeersonderzoekers zijn het reduceren van verkeerscongestie,
het verbeteren van het functioneren van het gehele verkeerssysteem en het reduceren
van de effecten van dit systeem op het milieu. Om dit doel te bereiken moeten we
eerst gangbare verkeerssituaties begrijpen om vervolgens proactief verkeersregelingen
en -management toe te passen. Het schatten van verkeerstoestanden in het verleden,
het heden, en de toekomst speelt een belangrijke rol in de verkeersmanagement- en
regelingssystemen. Dit proefschrift focust op de ontwikkeling van aanpakken die deze
verkeerstoestanden schatten.

Voor wegennetwerken geldt dat verkeerstoestanden verwijzen naar typische groothe-
den als reistijd, verkeerssnelheid, verkeersstroom en dichtheid. Deze grootheden geven
verkeerscondities weer. Gebaseerd op deze data kunnen we te weten komen waar een
file start, of waar een verkeersincident gebeurt.

Het is echter niet mogelijk om het complete beeld van de verkeerstoestanden te verkrij-
gen van de huidige monitoringssystemen. Vanwege de kosten en technische beperkin-
gen kunnen we enkel in ruimte en tijd gediscrèteerde verkeersdata verkrijgen. Deze
data worden hoofdzakelijk verzameld met behulp van op een vast punt geïnstalleerde
sensoren, zoals inductieloops, radar, en camera's. Verkeersinformatie kan echter ook
geobserveerd worden met behulp van meetvoertuigen die gemengd worden met het
overige verkeer gebaseerd op een geselecteerd ratio. In alle gevallen bevat de gede-
tecteerde data meestal afwijkingen en ruis, welke verdere analyse kunnen belemmeren.
Gebaseerd op deze beperkingen streeft dit proefschrift naar de ontwikkeling van een
procedure die de verkeerstoestand schat om de genoemde problemen op te lossen en
te voorzien in accurate en complete verkeerstoestandsinformatie. In deze procedure
worden zowel verkeersmodellen als beschikbare observatiedata gebruikt om de meest
waarschijnlijke verkeerstoestanden te schatten binnen een data-assimilatie kader.

Onze aanpak is geformuleerd vanuit het perspectief van een bewegende waarnemer,
resultierend in een Lagrangiaanse formulering van de schatting van de verkeersto-
estand. In het Lagrangiaanse coördinatensysteem bewegen de coördinaten met de voer-
tuigen mee. Het Lagrangiaans geformuleerde eerste orde verkeersstroommodel is
toegepast om de ontwikkeling van variabelen die de verkeerstoestand weergeven te

beschrijven. De voorgestelde Lagrangiaanse formulering van de schatting van de verkeerstoestand biedt zowel theoretische als rekenkundige voordelen ten opzichte van de conventionele Eulerse formulering. Deze aanpak kan bovendien de dynamiek van meerdere voertuigklassen meenemen door het implementeren van een verkeersstroommodel met meerdere klassen. In dit proefschrift worden ook diverse data voorwerkingsmethoden ontwikkeld om de kwaliteit van de observatie-input te verbeteren. In de toestandsschatting worden zowel Eulerse als Lagrangiaanse data meegenomen. Voor de data assimilatie wordt gebruik gemaakt van een online techniek bekend als het Extended Kalman Filter (EKF). Deze techniek combineert de voorspelling van systeemtoestanden gebaseerd op verkeersstroommodellen met correctie van de geobserveerde input. Belangrijk is dat het Lagrangiaanse concept niet beperkt is tot de EKF-methode met het eerste orde verkeersstroommodel, maar ook toegepast kan worden op andere data assimilatie technieken in combinatie met meer geavanceerde macroscopische verkeersstroommodellen. Een serie van experimentele studies gebaseerd op zowel synthetische als realistische data is uitgevoerd om de voorgestelde methodologie te testen. In deze studies zijn de procedures voor het schatten van verkeerstoestanden gevalideerd voor zowel een representatie van het verkeer met een gemengde klasse als met meerdere klassen. Deze resultaten laten zien dat de Lagrangiaanse schatting van de verkeerstoestand de Eulerse aanpak overtreft gegeven het op het EKF gebaseerde kader, en dat de aanpak met meerdere klassen de prestatie van de toestandsschatting verder verbetert vergeleken met de aanpak met een gemengde klas. Lagrangiaanse toestandsschatting met meerdere klassen kan voorzien in accurate klasse-specifieke verkeersinformatie voor klasse-specifieke regelingstoepassingen en verkeersmanagement.

(Dutch translation provided by Olga Huibregtse)

(Summary in Chinese)

概述

道路交通与每个人都息息相关。对于每天出行和工作的人们，交通堵塞已经成为他们的日常体验。交通研究者致力于缓解交通堵塞，提高整个交通系统的性能和改善居住环境。为了达到这个目的，我们必须首先了解交通状态，进而实施前瞻性的交通控制和管理。因此，过去、当下和将来的交通状态估计对于整个交通管理都至关重要。本文将着眼于研究交通状态估计的新方法，以提供相应的交通信息。

通常在道路网络中，交通状态是指与交通相关、反映交通特征的量：诸如行程时间、交通速度、交通流量和密度等等。这些量反映交通状况：比如何时发生交通堵塞，何处发生了交通事故等等。

然而，人们通常无法直接从交通监测系统中获得完整的交通状态信息。由于成本和技术的限制，人们只能得到时空上离散的交通数据。在目前的交通监测系统中，主要由基于定点的侦测器收集交通数据，例如：感应线圈、雷达、摄像机，或者由一定比例的浮动车提供交通信息。这些测量数据中往往包含误差和噪音，不利于进一步的分析。基于这些限制，本文致力于研究交通状态估计的方法以解决上述问题，并获得准确完整的交通状态信息。此方法在数据同化的研究框架中，通过交通流模型和实际观测到的数据来估计最可能的交通状态。

具体而言，我们的方法从移动观察者的视角，提出一个基于拉格朗日坐标系的交通状态估计模型。在拉格朗日坐标体系下，坐标轴随着车辆一起运动。基于拉格朗日坐标系的一阶交通流模型用于描述交通状态变量之间的关系。新提出的拉格朗日模型比传统的欧拉模型在交通状态估计中有着理论和计算上面的优势。同时，新的方法采用多车流交通流模型描述不同车流种类的动态变化。本文也提出了不同的数据预处理方法，用于提高交通观测输入的品质。欧拉和拉格朗日形式的交通数据也被应用到状态估计中。数据同化的方法是拓展卡尔曼滤波法。此方法具有实效性的优点，它将交通模型预测和观察纠正结合。值得一提的是，拉格拉日的概念不局限于基于一阶交通流模型的卡尔曼滤波法，它可拓展到其他的结合更高阶交通流模型的数据同化科技中。一系列仿真和实际数据的实验对新的研究方法进行了测试。这些实验校验了新提出的单车流和多车流交通状态估计模型。结果证明，在拓展卡尔曼滤波法的框架下，基于拉格朗日坐标系的交通状态估计法优于基于欧拉坐标系的方法，多车流的方法相对

于单车流的方法能进一步提高状态估计的性能。总而言之, 基于拉格朗日的多车流交通状态方法能够为多车流的控制应用和交通管理提供有利的数据支持。

TRAIL Thesis Series

The following list contains the most recent dissertations in the TRAIL Thesis Series. For a complete overview of more than 100 titles please see the TRAIL website: www.rsTRAIL.nl.

The TRAIL Thesis Series is a series of the Netherlands TRAIL Research School on transport, infrastructure and logistics.

Yuan, Y., *Lagrangian Multi-class Traffic State Estimation*, T2013/5, March 2013, TRAIL Thesis Series, the Netherlands

Schreiter, T., *Vehicle-class Specific Control of Freeway Traffic*, T2013/4, March 2013, TRAIL Thesis Series, the Netherlands

Zaerpour, N., *Efficient Management of Compact Storage Systems*, T2013/3, February 2013, TRAIL Thesis Series, the Netherlands

Huibregtse, O.L., *Robust Model-Based Optimization of Evacuation Guidance*, T2013/2, February 2013, TRAIL Thesis Series, the Netherlands

Fortuijn, L.G.H., *Turborotonde en turboplein: ontwerp, capaciteit en veiligheid*, T2013/1, January 2013, TRAIL Thesis Series, the Netherlands

Gharehgozli, A.H., *Developing New Methods for Efficient Container Stacking Operations*, T2012/7, November 2012, TRAIL Thesis Series, the Netherlands

Duin, R. van, *Logistics Concept Development in Multi-Actor Environments: Aligning stakeholders for successful development of public/private logistics systems by increased awareness of multi-actor objectives and perceptions*, T2012/6, October 2012, TRAIL Thesis Series, the Netherlands

Dicke-Ogenia, M., *Psychological Aspects of Travel Information Presentation: A psychological and ergonomic view on travellers' response to travel information*, T2012/5, October 2012, TRAIL Thesis Series, the Netherlands

Wismans, L.J.J., *Towards Sustainable Dynamic Traffic Management*, T2012/4, September 2012, TRAIL Thesis Series, the Netherlands

Hoogendoorn, R.G., *Empirical Research and Modeling of Longitudinal Driving Behavior under Adverse Conditions*, T2012/3, July 2012, TRAIL Thesis Series, the Netherlands

Carmona Benitez, R., *The Design of a Large Scale Airline Network*, T2012/2, June 2012, TRAIL Thesis Series, the Netherlands

Schaap, T.W., *Driving Behaviour in Unexpected Situations: a study into the effects of drivers compensation behaviour to safety-critical situations and the effects of mental workload, event urgency and task prioritization*, T2012/1, February 2012, TRAIL Thesis Series, the Netherlands

Muizelaar, T.J., *Non-recurrent Traffic Situations and Traffic Information: determining preferences and effects on route choice*, T2011/16, December 2011, TRAIL Thesis Series, the Netherlands

Cantarelli, C.C., *Cost Overruns in Large-Scale Transportation Infrastructure Projects: a theoretical and empirical exploration for the Netherlands and Worldwide*, T2011/15, November 2011, TRAIL Thesis Series, the Netherlands

Vlies, A.V. van der, *Rail Transport Risks and Urban Planning: solving deadlock situations between urban planning and rail transport of hazardous materials in the Netherlands*, T2011/14, October 2011, TRAIL Thesis Series, the Netherlands

Pas, J.W.G.M. van der, *Clearing the Road for ISA Implementation? Applying adaptive policymaking for the implementation of intelligent speed adaptation*, T2011/13, October 2011, TRAIL Thesis Series, the Netherlands

Zegeye, S.K., *Model-Based Traffic Control for Sustainable Mobility*, T2011/12, October 2011, TRAIL Thesis Series, the Netherlands

Mahr, T., *Vehicle Routing under Uncertainty*, T2011/11, September 2011, TRAIL Thesis Series, the Netherlands

Pel, A.J., *Transportation Modelling for Large-scale Evacuations*, T2011/10, July 2011, TRAIL Thesis Series, the Netherlands

Zheng, F., *Modelling Urban Travel Times*, T2011/9, July 2011, TRAIL Thesis Series, the Netherlands

Vlassenroot, S.H.M., *The Acceptability of In-vehicle Intelligent Speed Assistance (ISA) Systems: from trial support to public support*, T2011/8, June 2011, TRAIL Thesis Series, the Netherlands

Kroesen, M., *Human Response to Aircraft Noise*, T2011/7, June 2011, TRAIL Thesis Series, the Netherlands

Nielsen, L.K., *Rolling Stock Rescheduling in Passenger Railways: applications in short-term planning and in disruption management*, T2011/6, May 2011, TRAIL Thesis Series, the Netherlands

Qing, O., *New Approach to Fusion of Heterogeneous Traffic Data*, T2011/5, May 2011, TRAIL Thesis Series, the Netherlands

About the author

Curriculum Vitae

Yufei Yuan was born in Guilin, China, on 15th January, 1984. He started bachelor study in Engineering Mechanics in Shanghai Jiao Tong University since 2002, where he received the Excellent Academic Scholarship for three consecutive academic years. After graduating in 2006, he continued master study with scholarship at the department of Transport and Planning, Delft University of Technology, in the Netherlands. In July 2008, he finished his master project entitled: Coordination of ramp metering control in motorway networks, which was sponsored by the Dutch ministry of Transport.

In October 2008, he started PhD research with Professor Serge Hoogendoorn. During his PhD, he was involved in a European FP7 project: Control for coordination of distributed systems, and several contract research projects for the Dutch Ministry of Transport, including “File-proof Amsterdam”, “CHECK”, “Dynamax In-Car” and “Praktijkproef Amsterdam” projects. His research interest lies on traffic state estimation and prediction, data processing, traffic flow theory and simulation.

Publications

Journal Articles

1. Yuan, Y., J.W.C. van Lint, R.E. Wilson, F.L.M. van Wageningen-Kessels, and S.P. Hoogendoorn (2012) Real-time Lagrangian traffic state estimator for freeways. *IEEE Transactions on Intelligent Transportation Systems*, 13(1):59-70.
2. Van Wageningen-Kessels, F.L.M., Y. Yuan, S.P. Hoogendoorn, J.W.C. van Lint, and C. Vuik (2012) Discontinuities in the Lagrangian formulation of the kinematic wave models. *Transportation Research Part C: Emerging Technologies*. In press, available online.
3. Yuan, Y., J.W.C. van Lint, F.L.M. van Wageningen-Kessels, and S.P. Hoogendoorn (2012) Network-wide Lagrangian traffic state estimation using loop and floating car data. *Journal of Intelligent Transportation Systems: Technology, Planning, and Operations*. In press, available online.
4. Yuan, Y., R.E. Wilson, J.W.C. van Lint, and S.P. Hoogendoorn (2012) Estimation of multi-class and multi-lane counts from aggregate loop detector data. *Transportation Research Record*. In press.

Peer-reviewed Conference Contributions

5. Yuan, Y., W. Daamen, S.P. Hoogendoorn, and J.L.M. Vrancken (2009) Coordination of Agent-based Control in a Freeway Network, In: *Proceedings of CCDC (Chinese Control and Decision Making Conference)*, Guilin, China.
6. Yuan, Y., W. Daamen, S.P. Hoogendoorn, and J.L.M. Vrancken (2009) Coordination Concepts for Ramp Metering Control in a Freeway Network, In: *Proceedings of IFAC-CTS*, Redondo Beach, USA.
7. Yuan, Y., J.L.M. Vrancken, W. Daamen, S.P. Hoogendoorn (2009) Coordinated Ramp Metering: A case study with the HERO algorithm, In: *Proceedings of ITS World Congress 2009*, Stockholm, Sweden.
8. Wang, Y., Y. Yuan, and J.L.M. Vrancken (2009) Traffic Parameter Estimation on Motorway Networks by Combination of Filtering Techniques, In: *Proceedings of SMC2009*.
9. Yuan, Y., J.W.C. van Lint, T. Schreiter, S.P. Hoogendoorn, and J.L.M. Vrancken (2010) Automatic speed-bias correction with flow-density relationships, In: *Proceedings of the 2010 IEEE International Conference on Networking, Sensing and Control* (pp. 1-7). Chicago: IEEE.

10. Yuan, Y., F.L.M. van Wageningen-Kessels, J.W.C. van Lint, S.P. Hoogendoorn, and J.L.M Vrancken (2010) A new Lagrangian traffic state estimator for freeway networks, In: *Proceedings of 11th TRAIL Congress (DVD)* (pp. 1-5). Rotterdam: Research School on Transport, Infrastructure and Logistics.
11. Schreiter, T., J.W.C. van Lint, Y. Yuan, and S.P. Hoogendoorn (2010) Propagation Wave Speed Estimation of Freeway Traffic with Image Processing Tools, In: *Proceedings of TRB 89th annual meeting* (pp. 1-17). Washington D.C.: Transportation Research Board.
12. Schreiter, T., Y. Yuan, J.W.C. van Lint, JWC, and S.P. Hoogendoorn (2010) The “Adaptive Smoothing Method” with spatially varying kernels: ASM-svK. *Extended abstracts in TRISTAN VII* (pp. 685-688). Trondheim: Norwegian University of Science and Technology.
13. Van Wageningen-Kessels, F.L.M., Y. Yuan, S.P. Hoogendoorn, J.W.C. van Lint, and C. Vuik (2010) Discontinuities in the Lagrangian formulation of the kinematic wave models. In: *Proceedings of Traffic Flow Theory and Characteristics Summer Conference: does traffic data support traffic models?* Annecy, France.
14. Yuan, Y., J.W.C. van Lint, S.P. Hoogendoorn, J.L.M. Vrancken, and T. Schreiter (2011) Freeway traffic state estimation using extended Kalman filter for first-order traffic model in Lagrangian coordinates, In: *Proceedings of the 2011 IEEE International Conference on Networking, Sensing and Control* (pp. 121-126). Delft: IEEE.
15. Yuan, Y., F.L.M. van Wageningen-Kessels, J.W.C. van Lint, and S.P. Hoogendoorn (2011) Two modelling and discretization choices for Lagrangian multi-class first-order traffic flow model and their related (dis-)advantages, In: *Proceedings of the 9th International Conference on Traffic and Granular Flow '11* (pp. 125-133). Moscow: Russian Academy of Sciences. Proceedings to be published.
16. Yuan, Y., J.W.C. van Lint, F.L.M. van Wageningen-Kessels, and S.P. Hoogendoorn (2011) Lagrangian traffic state estimation for freeway networks, In: *Proceedings of the 2nd International Conference on Models and Technologies for ITS* (pp. 1-4). Leuven: Universiteit Leuven. Extended abstract, available online.
17. Van Wageningen-Kessels, F.L.M., Y. Yuan, S.P. Hoogendoorn, J.W.C. van Lint, and C. Vuik (2011) Sinks and sources in Lagrangian coordinates: derivation, interpretation and simulation results. In: *Proceedings of TRB 90th annual meeting* (pp. 1-16). Washington D.C.: Transportation Research Board.
18. Yuan, Y., R.E. Wilson, J.W.C. van Lint, and S.P. Hoogendoorn (2012) Estimation of multi-class and multi-lane counts from aggregate loop detector data. In: *Proceedings of TRB 91st annual meeting* (pp. 1-14). Washington D.C.: Transportation Research Board.

Technical Reports

19. Hoogendoorn, S.P., W. Daamen, and Y. Yuan (2009) Ex-ante evaluatie gecoördineerde toeritdosering Amsterdam : Tuning, evaluatie en aanpassing van het HERO algoritme. Delft: Technische Universiteit Delft.
20. Van Lint, J.W.C., T. Schreiter and Y. Yuan (2009) Technische en functionele haalbaarheid CHECK-algoritme voor de productie van statistische verkeersgegevens en indicatoren. Delft: Technische Universiteit Delft.
21. Van Wageningen-Kessels, F.L.M. and Y. Yuan (2011) Technical documentation Fastlane: Description of model equations and simulation. Technical report, TU Delft, 2008, updated in 2011.
22. Netten, B.D., A. Hegyi, W.J. Schakel, I.R. Wilmink, J.H. Hogema, M. Wang, J. Baan, I. Passchier, C. van Leeuwen, B. van Arem, T. Schreiter and Y. Yuan (2012). Dynamax In Car: preparation for a field test. Den Haag: TNO.
23. Yuan, Y. (2012) Technical documentation: Fileschatter. Technical report, TU Delft.