

Fast Model Predictive Control Approaches for Road Traffic Control

Han, Yu

DOI

[10.4233/uuid:4e441a1c-b5cb-4c2a-8dc4-edd64736013f](https://doi.org/10.4233/uuid:4e441a1c-b5cb-4c2a-8dc4-edd64736013f)

Publication date

2017

Document Version

Final published version

Citation (APA)

Han, Y. (2017). *Fast Model Predictive Control Approaches for Road Traffic Control*. [Dissertation (TU Delft), Delft University of Technology]. TRAIL Research School. <https://doi.org/10.4233/uuid:4e441a1c-b5cb-4c2a-8dc4-edd64736013f>

Important note

To cite this publication, please use the final published version (if applicable). Please check the document version above.

Copyright

Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

Takedown policy

Please contact us and provide details if you believe this document breaches copyrights. We will remove access to the work immediately and investigate your claim.

Fast Model Predictive Control Approaches for Road Traffic Control

Yu Han

This thesis is a result from a project funded by
China Scholarship Council



Cover illustration: Yu Han and Bo Huang

Fast Model Predictive Control Approaches for Road Traffic Control

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
donderdag 7 dec 2017 om 15:00 uur
door

Yu HAN

Master of Science in System Engineering
geboren te Shandong, China

Dit proefschrift is goedgekeurd door de:

Promotor: Prof. dr. ir. S.P. Hoogendoorn

Copromotor: Dr. ir. A. Hegyi

Samenstelling van de promotiecommissie:

Rector Magnificus,	voorzitter
Prof. dr. ir. S.P. Hoogendoorn,	promotor
Dr. ir. A. Hegyi,	copromotor

onafhankelijke leden:

Prof. dr. ir. B. van Arend,	Technische Universiteit Delft
Prof. dr. ir. H. Hellendoorn,	Technische Universiteit Delft
Prof. L. Leclercq,	Université de Lyon
Dr. S. Ahn,	University of Wisconsin-Madison
Dr. M. Ramezani,	University of Sydney

TRAIL Thesis Series no. T2017/13, the Netherlands Research School TRAIL

TRAIL

P.O. Box 5017

2600 GA Delft

The Netherlands

Phone: +31 (0) 15 278 6046

Fax: +31 (0) 15 278 4333

E-mail: info@rsTRAIL.nl

ISBN: 978-90-5584-230-8

Copyright © 2017 by Yu Han

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

Preface

The four-year studying experience in TU Delft is precious and unforgettable for me. During these four years, I have had a completely new view on traffic operations and management, which is a highly practical relevant topic. This experience paves the way for me to contribute to traffic flow management in reality, which is always my aspiration. I feel so lucky to meet so many nice people in this period. It is my great pleasure to take this opportunity to thank all the people who helped me and accompanied me during my PhD study.

First of all, I would like to thank the China Scholarship Council. This dissertation would not exist without its financial support.

I am extremely indebted to my excellent supervision team, Serge Hoogendoorn, Andreas Hegyi, and Yufei Yuan, for giving me sufficient guidance and support. Serge, my promoter, thanks a lot for your continuous enthusiasm not only on my research, but also on my personal development. Your research ideas always inspire me and your support always gives me confidence. I am very grateful to Andreas, my co-promoter, for coaching me towards an independent researcher. Your research attitude has strongly influenced me. It is incredible that every time you made hundreds of comments on my paper and nearly every comment makes sense. I have been growing up from addressing your valuable comments. I thank my daily supervisor, Yufei, for tirelessly reading my paper, correcting my English writing mistakes, and helping me to deal with technical difficulties. I could not have wished a better supervision team!

My special thanks goes to Mohsen Ramezani for hosting me during my visit to the University of Sydney. I like a lot the paper that we wrote together, and I hope it will get good impact on the research community. I am grateful to Markos Papageorgiou for his contribution to the paper in Chapter 3 of this thesis. I thank Claudio Roncoli for his contribution to the papers in Chapter 3 and Chapter 6. I also thank Jeroen van der Gun, Goof van de Weg, and Femke van Wageningen-Kessels for commenting on the paper in Chapter 2.

I would like to thank my colleagues in the department of Transport and Planning for making a great academic environment. I have benefited from the DrinX meeting, the Spark meeting, the TrafCon meeting, and so on. I thank Mehdi, Goof, and Jeroen for many useful academic discussions. Thank Jeroen for translating my English Summary

into Dutch. My special thanks goes to Priscilla and Conchita for their kind help with the logistics.

I would like to thank Sue Ahn, Ludovic Leclercq, Mohsen Ramezani, Bart van Arem, and Hans Hellendoorn for serving in my PhD committee.

Beside of those who directly influenced my research, I am also grateful to my colleagues and friends for making my life more enjoyable. Many thanks to my office-mates, Freddy, Xiao, Hugo, Kai, and many others who stayed shortly, for casual talks and helps. Thank Lin and Xavier for taking coffee breaks with me when we were bored from work. Thank Pengling, Yao and Yihong for the days of drinking together. Thank Ding and Vincent for going dinner together with me every now and then. Thank other Chinese colleagues, Meng, Yusen, Yaqing, Fei, Mo, Xiaoxia, for inviting me for dinners and also thank all other Chinese colleagues who joined the dinners and made the atmosphere wonderful.

I am grateful to Delft Chinese basketball team. I had so many happy moments when we trained together and played games together. I specially thank some of the teammates, Bairong, Hao, and Yiran, for drinking with me.

Finally, I would like to thank my family, for their unconditional support. Special thanks to my father, who encouraged me to do this PhD research when I was hesitating. This thesis is dedicated to you.

Yu Han
Beijing, November 2017

Contents

Preface	i
List of Figures	xiv
List of Tables	xv
1 Introduction	1
1.1 Background	2
1.2 Introduction to dynamic traffic control	2
1.2.1 Ramp metering	3
1.2.2 Variable speed limits	5
1.2.3 Route guidance and intersection signal control	7
1.3 Model predictive control for road traffic	8
1.3.1 Traffic flow models	10
1.4 Research objectives	11
1.5 Main Contributions	12
1.6 Thesis outline	13
2 A new extended discrete first-order model to reproduce jam waves	15
2.1 Introduction	16
2.2 Description and analysis of jam waves	18
2.2.1 Practical Description	18
2.2.2 Theoretical analysis	19
2.3 Model development	20
2.3.1 Empirical analysis	20

2.3.2	A new discrete first-order model	22
2.4	Model validation	26
2.4.1	Simulation test on a synthetic freeway stretch	27
2.4.2	Simulation tests for a real life freeway stretch	28
2.5	Conclusions	30
3	Resolving freeway jam waves by variable speed limits	33
3.1	Introduction	34
3.2	Model description	36
3.2.1	The cell-transmission model	36
3.2.2	Extended versions of the CTM	37
3.2.3	The new extended CTM	39
3.3	The optimal control formulation	44
3.3.1	The linear property of the model	44
3.3.2	Objective function	46
3.3.3	The minimum VSL constraint	46
3.4	Benchmark problem	47
3.4.1	Reproducing of a jam wave	48
3.4.2	The set up of controllers	50
3.4.3	Control performance	52
3.5	Conclusion and future research	58
4	Validation of an extended discrete first-order model with VSLs	59
4.1	Introduction	60
4.2	Model description	62
4.2.1	The new extended discrete first-order model	62
4.3	Reproducing the mechanisms of VSLs	66
4.3.1	Reproducing the SPECIALIST algorithm	66
4.3.2	Reproducing the mainstream metering approach	68
4.4	Model calibration and validation	70
4.4.1	Calibration of the proposed model	71
4.4.2	Calibration of the METANET model	73
4.4.3	Model validation	77
4.5	Conclusions	87

5	MFD for hierarchical ramp metering in freeways	89
5.1	Introduction	90
5.2	Modeling of the freeway MFD	92
5.2.1	Relationship between network production and outflow	94
5.2.2	Modeling the density heterogeneity	96
5.2.3	The effect of capacity drop on freeway MFD	98
5.2.4	Functional form of the freeway MFD model	100
5.3	Control design	102
5.3.1	The upper level controller: The MPC approach	103
5.3.2	The lower level controller: local ramp metering	105
5.4	Numerical experiments	106
5.4.1	Case study 1-The extended CTM as the process model	106
5.4.2	Case study 2-METANET as the process model	111
5.5	Conclusion	112
6	An extended LQMPC approach for multi-destination urban networks	117
6.1	Introduction	118
6.2	Route-specific network model	120
6.2.1	Definition of model variables	120
6.2.2	Updating of traffic states	122
6.3	The classical LQMPC approach	126
6.4	An extended LQMPC approach	128
6.4.1	Forward simulation	129
6.4.2	Minimum flow constraint on each OD pair	131
6.4.3	Reproducing spillback	135
6.4.4	Destination constraint	135
6.4.5	Evaluation step of the extended LQMPC	137
6.5	Case study	138
6.6	Conclusions	143

7 Conclusions	145
7.1 Main findings and conclusions	146
7.2 Implications for practice	148
7.3 Recommendations for future research	149
Bibliography	151
Summary	163
Samenvatting	165
TRAIL Thesis Series	167
About the author	169

List of Figures

1.1	Two cases to explain ramp metering, (a) without and (b) with ramp metering. Grey areas indicate congestion zones. q^{dch} and q^{cap} represent the queue discharge rate and the free-flow capacity respectively, and $q^{\text{dch}} < q^{\text{cap}}$. f^{nc} and f^{rm} represent the off-ramp flows with and without ramp metering, and $f^{\text{nc}} < f^{\text{rm}}$	4
1.2	Illustration of traffic evolution under the SPECIALIST. The left figure is the time-space graph and the right figure is the fundamental diagram.	6
1.3	The mainstream metering approach used in (Carlson et al., 2010a). The on-ramp is perceived as a potential bottleneck. q_{in} and q_{out} are the inflow and outflow of the VSL control area. f_{VSL} is the VSL-induced capacity.	7
1.4	The MPC scheme for traffic control. The figure is taken from Hegyi (2004).	9
1.5	The outline of the thesis.	14
2.1	Visualization of a jam wave from empirical data and theoretical analysis. Empirical data were obtained from loop detectors of Dutch freeway A13. There are four lanes on this freeway stretch.	20
2.2	Results of the empirical study. (a) The targeting motorway stretch. (b) The speed contour plot of the analyzing period. (c) The density-flow plot (fundamental diagram) of the analyzing period. (d-e) Data analysis of the proposed site. The plots represent the relation between density $\rho_{(i+1)}$ and flux $q_{(i \rightarrow i+1)}$, $8 \leq i \leq 16$	23
2.3	Description of the proposed model. (a) shows the demand and the supply function of the proposed model. (b) represents the different fundamental diagram of two consecutive cells. (c) shows the acceleration process of two neighbouring jam cells.	26
2.4	The hypothetical freeway stretch, demand profile, and simulation results.	29
2.5	The real life freeway stretch, demand profiles, and simulation results. .	31

3.1	A freeway stretch divided into discrete cells.	37
3.2	The triangular-shaped FD of the CTM (a), and the demand and the supply parts (b).	38
3.3	The demand (a) and supply (b) parts of the FD of the model from (Roncoli et al., 2015a).	39
3.4	An example to explain the discharging state. Cell $i + 1$ is in the discharging state.	41
3.5	The left column, figures (a,c,e,g) show the relation between $\rho_i(k)$ and $Q_{i+1}(k)$, which represents the capacity drop assumption. The right column shows the demand and supply of cell $i + 1$, (b) for condition $\rho_i(k) \leq \rho_i^{cr}$, (d) for condition $\rho_i(k) = \rho_i^J$, (f) for condition $\rho_i^{cr} < \rho_i(k) < \rho_i^J$, (h) for condition $\rho_i^{cr} < \rho_i(k) < \rho_i^J$ and cell $i + 1$ is under VSL control with the value $V_{i+1}(k)$	43
3.6	A graphical representation of the freeway stretch for the benchmark problem.	48
3.7	The traffic demand at the origin (a), and the density downstream of the freeway stretch (b).	50
3.8	The speed (km/h) contour plot (a), and the flow (veh/h) contour plot (b) of the simulation without VSL control.	51
3.9	The calibrated fundamental diagram of controller 3	51
3.10	The density (veh/km/lane) contour plots (without VSL control) from (a) the METANET model, (b) Roncoli's model, and (c) the new extended CTM.	53
3.11	The outflow of the freeway stretch from the simulation of the METANET model (a), Roncoli's model (b), and the new extended CTM (c).	54
3.12	The speed contour plots (with VSL control) of controllers 1, 2, 3 (a,c,e), and the flow contour plots of controllers 1, 2, 3 (b,d,f).	56
3.13	The VSL control signals (km/h) generated by controller 1 (a), controller 2 (b), and controller 3 (c).	57
4.1	(a) The triangular-shaped FD of the CTM. (b) The the sending flow and receiving flow functions.	63
4.2	The left figure depicts the capacity drop assumption of the proposed model. As the density of cell $i - 1$ increases, the capacity of the cell i decreases. The right figure depicts the sending flow function of cell i	64
4.3	An example to explain the discharging state. Cell $i + 1$ is in the discharging state.	64

4.4	The depiction of the receiving flow function of the proposed model.	65
4.5	The quadrangular fundamental diagram of the proposed model. ρ_{cr}^V is the VSL-induced critical density and f_{VSL} is the VSL-induced capacity.	66
4.6	Illustration of traffic evolution under the SPECIALIST. The left figure is the time-space graph and the right figure is the fundamental diagram.	67
4.7	A synthetic case to show that the model reproduces the SPECIALIST algorithm. (a) and (b) are the speed contour plot and the flow contour plot without VSLs. (c) and (d) are the speed contour and flow contour with VSLs. Some of the traffic states in Fig. 4.6 are highlighted in (c) and (d). (e) and (f) are the throughput of the stretch with and without VSLs from the theory and the simulation.	69
4.8	The mainstream metering approach used in (Carlson et al., 2010a). The on-ramp is perceived as a potential bottleneck. q_{in} and q_{out} are the inflow and outflow of the VSL control area. f_{VSL} is the VSL-induced capacity.	70
4.9	A synthetic case to show that the model reproduces the mainstream metering approach. (a) and (b) are the speed contour plot and the flow contour plot without VSLs. (c) and (d) are the speed contour and flow contour with VSLs. In (c), the VSLs are implemented at cells 10 and 11. The speed at upstream of the VSL-control area also reduces, because the VSL-control area has a high density that propagates upstream.	71
4.10	A graphical representation of the freeway stretch for the SPECIALIST field test. Traffic goes from the left to the right. Number in the figure represents detector number. The data we obtained are detected flows and the time mean speeds of the detectors from 4 to 25. This part of the stretch is homogeneous except for a gas station located in the middle of this part, which has a negligible flow compared with the mainstream flow). The distance between neighboring detectors are between 325 m and 650m.	72
4.11	The comparison of the real-data for the calibration and the results reproduced from the calibrated models in data set 1. Figures of the first row are plotted from real data, and second row third row are plotted from the simulation results of the proposed model and the extended METANET model.	76
4.12	The comparison of the real-data for the calibration and the results reproduced from the calibrated models in data set 2. (a) is the VSL control scheme of the data. (b), (c) and (d) are the speed contour plots from real data, the proposed model and the METANET model.	77

4.13	The comparison of the real-data for the validation and the results reproduced from the models in case 1. Figures of the left column are plotted from real data, and the middle column and the right column are plotted from the simulation results of the proposed model and the extended METANET model.	78
4.14	Comparison between the real flow data and the validated model results of each detector in case 1. The blue solid line represents the results from real data. The red dashed line and black dash-dotted line represent the results from the proposed model and the extended METANET model respectively.	81
4.15	Comparison between the real speed data and the validated model results of each detector in case 1. The blue solid line represents the results from real data. The red dash line and black dash-dot line represent the results from the proposed model and the extended METANET model.	82
4.16	The comparison of the real-data for the validation and the results reproduced from the models in case 2. Figures of the first row are plotted from real data, and the second row and the third row are plotted from the simulation results of the proposed model and the extended METANET model.	83
4.17	Comparison between the real flow data and the validated model results of each detector in case 2. The blue solid line represents the results from real data. The red dash line and black dash-dot line represent the results from the proposed model and the extended METANET model respectively.	85
4.18	Comparison between the real speed data and the validated model results of each detector in case 2. The blue solid line represents the results from real data. The red dash line and black dash-dot line represent the results from the proposed model and the extended METANET model respectively.	86
5.1	The sketch of Dutch freeway A13-L. The freeway stretch is about 16km in length and includes 6 on-ramps and 6 off-ramps.	92
5.2	Aggregated loop detector data of April 25, 2013: (a) weighted average flow versus weighted average density; (b) total outflow versus weighted average density. The shapes in (a) and (b) are unimodal, with less scatter in the left part and high scatter in the right part. Note that in (b), some data points from 6:00 to 9:00 have lower values of the total outflow compared to the traffic condition from 21:00 to 24:00 due to longer average trip lengths, see Fig. 5.3.	94

5.3	The whisker plot of P/G aggregated for all of the weekdays during March and April 2013. On each box, the central red mark indicates the median, and the bottom and top blue edges of the box indicate the 25th and 75th percentiles, respectively. The whiskers extend to the most extreme data points. In the rectangle area, the values of P/G are relatively time-invariant, which indicates that the average trip lengths are nearly time-invariant. The average trip lengths from 6:00 to 7:30 are longer, which results in lower values of total outflows, see Fig. 5.2 (b).	95
5.4	Aggregated loop detector data of April 16, 2013: (a) weighted average flow versus weighted average density; (b) density heterogeneity versus weighted average density. Red arrows in the figures represent the direction of time in the hysteresis loop.	96
5.5	(a) The estimation results of the density heterogeneity model. The blue line represents (5) and the green lines represent (8). The red points are the estimated density heterogeneity at each time instance. (b) The comparison between real data and model estimation. Data are obtained from April 16, 2013. Note that the x axis in the two figures are different, where in (a) it is $\bar{\rho}$ and in (b) it is $\bar{\rho}^{cr}$	99
5.6	The layout of the synthetic freeway stretch that is used to model the effect of capacity drop on MFD. Cell number are marked at the upstream of each on-ramp. Circled numbers represent the order of the on-ramps.	99
5.7	(a) and (b): The density contour plots of scenario 1 and 2 in the synthetic case. The density of the jam head in (a) is around 35 (veh/km/lane) and in (b) is around 50 (veh/km/lane). (c): The $\bar{\rho}$ - σ curves of the two scenarios. (d): The $\bar{\rho}$ - \bar{q} curves of the two scenarios. In (c) and (d), when $\bar{\rho}$ of the two scenario are around 73 (veh/km), the values of σ in the two scenarios are similar but the value of \bar{q} in scenario 1 is substantially higher than \bar{q} in scenario 2. (e) and (f) show the accuracy of the MFD model (10) with and without η . The solid lines represent the MFD of the simulation. The dotted lines and dashed lines respectively represent the MFDs with and without η in (10).	101
5.8	The estimated MFDs for (a) April 2 and (b) March 27. MFD-1 is the MFD with η and MFD-2 is the one without η . The proposed freeway MFD model can capture the complex hysteresis patterns.	103
5.9	Time scales of the process model (top) and the prediction model (bottom).	104
5.10	The layout of the synthetic freeway stretch. Cell numbers are marked at the upstream of each on-ramp and off-ramp. The order of on-ramps are shown as circled numbers.	107
5.11	The demand profile of the simulation.	107

- 5.12 (a): The estimated MFDs in scenario 2 and 3. Note that the average density is calculated over three lanes. (b): The accuracy of the prediction of σ . Blue lines represents the evolution of σ in the process. Black lines represent the value of σ estimated by (8) which has a dynamic value of $\bar{\rho}^{cr}$. Red lines represent the values of σ predicted by the prediction model, which assumes a constant value of $\bar{\rho}^{cr}$ in the prediction horizon. 108
- 5.13 The density (veh/km/lane) contour plots of the scenario without control and all five control scenarios in case study 1. On-ramps 3 and 5 are the main bottlenecks. Evidently, the proposed MFD-based ramp metering control strategy (i.e. scenario 2) efficiently reduces traffic density at on-ramp locations. Scenario 1 represents the theoretical upper bounds of MPC performance. 110
- 5.14 The queue lengths (on left) and ramp flows (on right) of on-ramps 1-5 of the five control scenarios in case study 1. 114
- 5.15 The density contour plots of the scenario without control and all three control scenarios in case study 2. 115
- 5.16 The queue lengths (on left) and ramp flows (on right) of on-ramps 1-5 of the three control scenarios in case study 2. 116
- 6.1 The depiction of cells and links. Red arrows represent links, and black dashed lines represent the boundaries between cells. Origin cells have no incoming link and destination cells have no outgoing link. Normal cells are connected by one incoming link and one outgoing link. . . . 120
- 6.2 The depiction of merging and diverging cells. Red arrows represent links. The incoming links of merging cells are called merging links and the outgoing links of diverging cells are called diverging links. . . 121
- 6.3 An example intersection to explain the model variables. The figure is taken from Zhu & Ukkusuri (2015). f_1, f_2, \dots, f_8 are the flows of intersection links. Red dots in the figure represent conflicting points of flows. The intersection is indexed as intersection 1. For simplicity, we assume that there are four phases in a cycle, and $W_1^1 = \{1, 3\}$, $W_1^2 = \{2, 4\}$, $W_1^3 = \{5, 7\}$, $W_1^4 = \{6, 8\}$. Note that there can be more phases in a cycle as long as the movements in each cycle do not conflict. u_1^1, u_1^2, u_1^3 and u_1^4 are the green time fraction of each phase so the traffic light cycle constraint is: $u_1^1 + u_1^2 + u_1^3 + u_1^4 \leq 1$. Green time fraction constraints are: $f_1 \leq u_1^1 \cdot Q_1, f_2 \leq u_1^2 \cdot Q_2, f_3 \leq u_1^1 \cdot Q_3, f_4 \leq u_1^2 \cdot Q_4, f_5 \leq u_1^3 \cdot Q_5, f_6 \leq u_1^4 \cdot Q_6, f_7 \leq u_1^3 \cdot Q_7, f_8 \leq u_1^4 \cdot Q_8$ 122

- 6.4 Example networks to explain the usage of the LQMPC. Network 1 is a single destination network and network 2 is a multiple destination network. Red arrows in network 1 indicate the bifurcations, where traffic can make their route choices. Grey area in network 2 represents congestion. If the receiving flow of the gray area is 0, and the turn fraction of link 2 is higher than 0, then according to (18), both f_1 and f_2 are 0. 129
- 6.5 The flowchart of the control approach. There are two blocks in this flowchart, where the upper block is the traffic control part and the lower part is the control performance evaluation part. 130
- 6.6 An example to show the c-links (red arrows) of an OD pair that has a single route. Dashed rectangles represent cells and arrows represent links. In Network 1, link 2 is the c-link of OD pair O1-D1. In Network 2, links 1 and 2 are the c-links of O1-D1. 131
- 6.7 An example to show the c-links of an OD pair that has multiple routes. Bold lines represent the roadway with two directions. In Network 3, links 1, 2, 3 are the c-link of OD pair O1-D1. In Network 4, links 1, 2, 3 and 4, 5, 6 are the c-links of the OD pair. 133
- 6.8 The cumulative curves at the origin and the destination of an OD pair. The red line is the cumulative curve at the origin and the dashed blue line is the cumulative curve at the destination. $N_O^{o,d}(t)$ and $N_D^{o,d}(t)$ are the cumulative number of vehicles at the origin and the destination. t is the current time that the controller runs, and $TF_{o,d}$ is the free flow travel time of the OD pair. The difference between $N_O^{o,d}(t + K_p - TF_{o,d})$ and $N_D^{o,d}(t)$ is the maximum number of vehicles that can travel to destination d from origin o from t to $t + K_p$. For destination d , the maximum number of arriving vehicles is $\sum_o^o (N_O^{o,d}(t + K_p - TF_{o,d}) - N_D^{o,d}(t))$. . . 136
- 6.9 A conceptual explanation about the boundaries of the solution spaces of the original LQMPC (the red line), the extended LQMPC (the blue line), and the non-linear optimization (the black line). 138
- 6.10 Sketch of the synthetic network. The bold line indicates freeway and other lines indicate urban arterials. 'BN1' and 'BN2' represent bottlenecks. The circles represent signalized intersections. 139
- 6.11 The depiction of the movements. Red circled numbers in the right intersection are the index of roadway sections at the upstream of the intersection. Blue numbers represent the index of some of the links. . . 139

- 6.12 Simulation results of the case study. Line colors represent the results from different scenarios. Black: scenario 1, system optimal. Blue: scenario 2, the proposed control approach. Red: scenario 3, the non-optimizing approach. Pink: scenario 4, the original LQMPC. (a), (b), and (c) are the outflows of scenarios 4 and 1, scenarios 1 and 2, and scenarios 2 and 3. The solid lines, dashed lines, and dotted lines represent the outflows of destinations B, D, and F respectively. (d) is the turn fraction of link 1 at the left intersection in Fig. 6.11. (e) and (f) are the queue lengths of section 1, 2 and 3 in Fig. 6.11. The solid lines, dashed lines, and dotted lines represent the queue lengths of road sections 1, 2, and 3 respectively. 141

List of Tables

3.1	The calibration results of controllers 2 and 3	52
3.2	The total travel delay and computation time of three controllers	55
4.1	Calibrated parameters of the proposed model.	73
4.2	Calibrated parameters of the METANET model	75
4.3	Quantitative results of the models calibration and validation. In the table, the first case shows the calibration results of the proposed model with the quadrangular FD and the third case shows the calibration results of the proposed model with the triangular FD.	80
5.1	The estimation and testing results of different MFD models. The field data from April are used to estimate the parameters in (10), i.e. the training set, while the data from March are used to test the performance. In estimation of the MFD without σ and η , b_1 and b_2 are equal to 0. In estimation of the MFD with σ and without η , b_2 is equal to 0.	102
5.2	The performance of each control scenario in case study 1.	111
5.3	The performance of each control scenario in case study 2.	112
6.1	The demand pattern of the synthetic case.	138

Chapter 1

Introduction

The first chapter of the thesis introduces the research scope and highlights the main contributions. We first present the background of this thesis and highlight the importance of road traffic control. Then, traffic control measures and traffic control approaches that are used in this thesis are briefly introduced with a short literature review in each part. The research objectives and contributions are discussed thereafter. Finally, the outline of the thesis is presented.

1.1 Background

Due to the growing transportation demand and the urbanization trend, traffic congestion has become a global issue that has a significant impact on our society's productivity. Traffic congestion generates various of problems such as reducing capacity of traffic infrastructures, inducing travel delays and potentially unsafe conditions. The construction of new infrastructure to alleviate congestion and accommodate higher traffic demand is not always viable due to economical and environmental concerns. Instead, road traffic management, which can make a better use of existing traffic infrastructure, has been extensively investigated in the literature and applied to the field to combat traffic congestion. This dissertation proposes and develops new method for efficient traffic network control.

In the literature, various studies have reported that dynamic traffic control is a good solution to reduce congestion. The various approaches are demonstrated through simulation analysis (e.g. Papageorgiou (1990); Smith & Ghali (1990); Lo et al. (2001)) and field test experiments (e.g. Papageorgiou et al. (1997); Hegyi & Hoogendoorn (2010); Hoogendoorn et al. (2013)). Nowadays, most of the dynamic traffic control systems in the field are responsive control systems, in which traffic control centers receive data from sensors and apply control policies that respond to the prevailing traffic conditions. However, the use of advanced control techniques like predictive control systems can considerably improve the reduction in total time spent (Kotsialos et al., 2002b; Hegyi, 2004; Bellemans et al., 2006). Although traffic predictive control systems have been extensively investigated in the literature, few have been implemented into the field. Generally, traffic predictive control systems face two main challenges towards field application: (i) High computation complexity, and (ii) low prediction accuracy, i.e., a high mismatch between the predicted traffic dynamics and real traffic dynamics. The computation complexity problem has been addressed in many ways, e.g., to use decentralized or distributed control schemes (e.g. Tettamanti & Varga (2010); Camponogara & De Oliveira (2009)), to simplify the non-linear predictive control systems to linear or quadratic control systems that can be efficiently optimized (e.g. Muralidharan & Horowitz (2015); Roncoli et al. (2015b)), to develop more efficient computational algorithms (e.g. Kotsialos et al. (1999); Zegeye et al. (2012); van de Weg et al. (2015)). This thesis mainly focuses on the second avenue and develops accurate traffic flow models that can be applied to efficient predictive control systems.

1.2 Introduction to dynamic traffic control

Dynamic traffic control is the most commonly used strategy to improve the operation efficiency of road infrastructure. In general, dynamic traffic control utilizes the measurements of traffic conditions over time to estimate the time-varying traffic state and computes dynamic traffic control signals (traffic control measures) to influence traffic

behavior and generates the effects that lead to higher network throughput, lower travel delays, and lower fuel consumptions, etc. The total time spent (TTS) is a widely used quantity to describe the operation efficiency of road infrastructure by summing the time that vehicles spend in a traffic network.

There are some typical traffic problems in traffic networks that decrease the traffic operation efficiency and increase the TTS, for example, capacity drop, spillback, underutilization of parallel routes, etc. These traffic problems are often suppressed by different control measures such as ramp metering, variable speed limits, route guidance, and intersection control. These control measures have been not only extensively investigated in the literature, but also widely used in the field since they are available in many road networks in the world. This thesis focuses on developing advanced control strategies using those measures, that have good potentials to be applied to the field. In this section we give an introduction of the traffic problems those control measures intend to solve and the control methods to apply those measures.

1.2.1 Ramp metering

Ramp metering is a freeway traffic control measure that has not only been extensively investigated in the literature but also widely applied in real life. Ramp metering regulates inflows from on-ramps by traffic signals. In some countries like the USA or the Netherlands, the ramp metering allows one car per green per lane. In other countries, there are implementations that allow two or more cars per green (Frejo, 2015). It has been demonstrated that ramp metering improves the merging behavior of traffic flow which may have a significant positive influence on traffic safety due to less lane changes and reduced driver stress (Papageorgiou & Kotsialos, 2000). Moreover, ramp metering may substantially reduces the total time spent of drivers if it is effectively implemented. Generally, freeway traffic systems benefit from ramp metering control by preventing the occurrence of two traffic phenomena:

1. Capacity drop, which is known as the phenomenon that the outflow of a traffic congestion (queue discharge rate) is significantly lower (between 10 to 30 %) than the free-flow capacity of the same location. If ramp metering is effectively implemented, the flow downstream of the on-ramp can be as high as the free-flow capacity. Fig. 1.1 shows the comparison of the flows at the downstream of the on-ramp with and without ramp metering. q^{dch} and q^{cap} represent the queue discharge rate and the free-flow capacity respectively, and $q^{\text{dch}} < q^{\text{cap}}$.
2. Spillback, which is the phenomenon that the congestion of a downstream link affects the possible outflow volume of the upstream link. As shown in Fig. 1.1 (a), the congestion originates from the on-ramp and propagates to the off-ramp, which reduces the flow that leaves the freeway from the off-ramp. If the ramp metering is effectively implemented, the spillback can be completely prevented

or avoided for as long as possible. Fig. 1.1 shows the comparison of the off-ramp flows with and without ramp metering. f^{nc} and f^{rm} represent the off-ramp flows with and without ramp metering, and $f^{nc} < f^{rm}$.

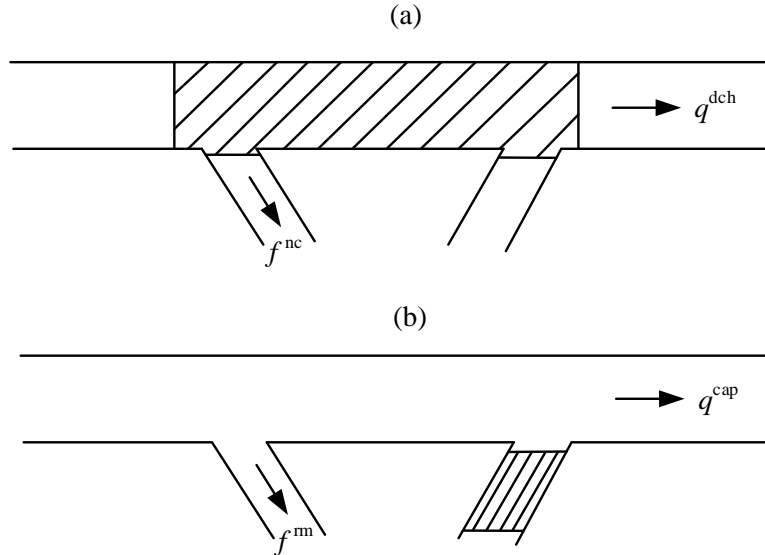


Figure 1.1: Two cases to explain ramp metering, (a) without and (b) with ramp metering. Grey areas indicate congestion zones. q^{dch} and q^{cap} represent the queue discharge rate and the free-flow capacity respectively, and $q^{dch} < q^{cap}$. f^{nc} and f^{rm} represent the off-ramp flows with and without ramp metering, and $f^{nc} < f^{rm}$.

Apart from avoiding the capacity drop and spillback, traffic systems can get benefit from the ramp metering indirectly by changing travelers route choice behavior so as to achieve a desired traffic flow distribution in the overall network, e.g., avoidance of the rat-running phenomenon, increased or decreased utilization of underutilized or overloaded, respectively, parallel arterials etc (Papageorgiou & Kotsialos, 2000).

Previous works have proposed several ramp metering strategies which can be categorized into fixed-time strategies (Wang & May, 1973), reactive control strategies (Masher et al., 1975), and optimal control strategies (Papamichail et al., 2010). Papageorgiou & Kotsialos (2000) presents a thorough literature review of the ramp metering-related studies in the last century. In recent years, advanced coordinated ramp metering methods using model predictive control (MPC) approaches that are based on macroscopic traffic flow models are becoming more and more popular. Studies of Kotsialos & Papageorgiou (2004); Hegyi et al. (2005a); Papamichail et al. (2010) presented different optimal approaches of ramp metering that are based on the METANET model. Gomes & Horowitz (2006) presented an optimal freeway ramp metering approach that is based on the asymmetric cell transmission model, which can be efficiently solved by linear programming. Han et al. (2015b) presented a linear quadratic MPC approach for ramp metering based on an extended cell transmission model (CTM) which takes the capacity drop into account. This thesis focuses on the use of MPC strategies of

ramp metering to reduce the total time spent of drivers. In Chapter 5, a hierarchical ramp metering approach (with an MPC approach at the upper level) is proposed, and it is compared with other MPC controllers of ramp metering in previous works.

1.2.2 Variable speed limits

Over the past decades, variable speed limits (VSLs) have emerged as a popular control measure for traffic flows on freeways. Nowadays, most of the implemented VSL systems are used to increase safety by warning on decreasing the speed limits upstream of congested areas. Nevertheless, VSLs systems that aim to improve the traffic operation efficiency have been extensively investigated in the literature. Earlier studies have assumed that the freeway capacity can be raised if the speeds across vehicles in different lanes is harmonized by VSLs (Zackor, 1972). Following this line, Cremer (1979) proposed a quantitative model for the VSL-induced fundamental diagram, in which the free flow capacity was assumed to increase under certain VSL rates. It has been found that VSLs can increase utilization of the shoulder lane on a long homogeneous freeway stretch (Duret et al., 2012). For shorter freeway stretches, although VSLs can homogenize the speed of individual vehicles, no evidence has shown that VSLs can increase the free flow capacity. In fact, later investigations could not identify any capacity increase that could be attributed to VSLs (Smulders, 1990; Soriguera et al., 2017).

In the literature, two mechanisms of VSLs are commonly used to improve the traffic operation efficiency. Hegyi et al. (2008) proposed the SPECIALIST algorithm that is based on the temporal transition states of VSLs to resolve freeway jam waves. Carlson et al. (2010a) proposed the mainstream metering approach that is based on the assumption of the VSL-induced capacity reduction to prevent the activation of on-ramp bottlenecks. In the following, we briefly introduce the two mechanisms of VSLs.

The SPECIALIST algorithm is designed based on the shock wave theory to resolve freeway jam waves. As has been presented in (Hegyi et al., 2008), traffic jams that have an upstream moving head and tail are known as jam waves (also known as wide moving jams in some studies, e.g., (Kerner & Rehborn, 1996)). The queue discharge rate of a jam wave is typically around 30 percent lower than the free flow capacity (Schönhof & Helbing, 2007). Fig. 5.4 is used to explain the theory of SPECIALIST. The time-space graph in the left figure shows the traffic states on a road stretch and their propagation over time. The density-flow diagram in the right figure shows the corresponding density and flow values for these states. According to the shock wave theory, the boundary (front) between two states in the left figure has the slope as the slope of the line that connects the two states in the right figure. The SPECIALIST theory uses this basic relationship to resolve jam waves.

Area 2 represents a short jam wave that propagates upstream and which is surrounded by traffic in free-flow (area 1 and 6). As soon as the jam wave is detected, the speed

limits upstream of the jam wave are switched on, where traffic state changes to state 3. The inflow of area 2, which is the flow of state 3, is lower than the outflow of area 2, which is the flow of state 1, so the width of the jam wave is narrowing as time advances. To resolve the jam wave, the required length of the speed-limited stretch depends on the density and flow associated with state 2 and the physical length of the detected jam. When the jam wave is resolved there remains an area with the speed limits active (state 4) with a moderate density (higher than in free-flow, lower than in the jam wave). It is assumed that traffic leaving area 4 has a higher flow and a higher speed than state 4, which is represented by state 5. A basis assumption in the SPECIALIST theory is that the traffic from area 4 can flow out more efficiently than a queue discharging from full congestion as in the shock wave (flow of state 2), which results in the increasing of the total throughput. In a later research, the analysis of the data from the SPECIALIST field test confirms this assumption (Hegyi & Hoogendoorn, 2010).

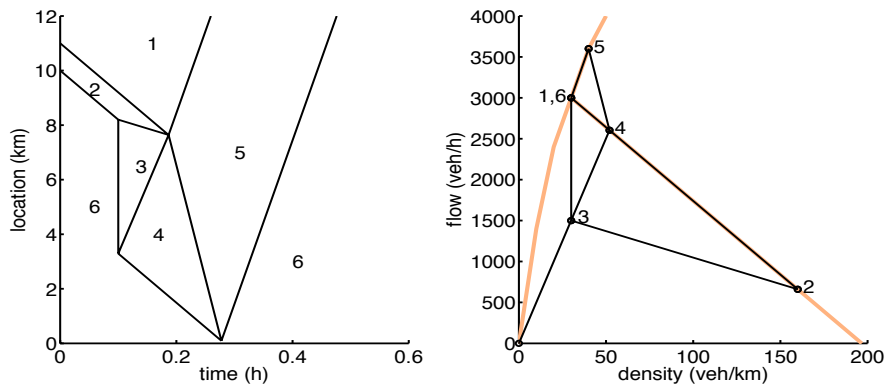


Figure 1.2: Illustration of traffic evolution under the SPECIALIST. The left figure is the time-space graph and the right figure is the fundamental diagram.

Empirical evidences in (Papageorgiou et al., 2008) show that sufficiently low VSLs lead to a lower flow (capacity) than in the unrestricted case. This forms the basis of the mainstream metering approach proposed by (Carlson et al., 2010a). Fig. 1.3 explains the theory of the mainstream metering approach. If the arriving demand is higher than the VSL-induced capacity, the VSL application area becomes an active mainstream bottleneck that limits the area's outflow to values that corresponds to the VSL-induced capacity. This approach may be applied at the upstream of a potential bottleneck (e.g. an on-ramp merging area) to avoid its activation and the related throughput reduction as a result of the capacity drop. More specifically, the mainstream flows are regulated by VSLs such that capacity flow can be established at the downstream bottleneck. Then the final mainstream throughput is maximized, leading to a decrease of the total time spent.

There are various theories and algorithms to determine the appropriate values of the VSLs, for example, analytical approaches ((Hegyi et al., 2008; Mahajan et al., 2015; Chen et al., 2014a)) and model predictive control approaches ((Hegyi et al., 2005b; Carlson et al., 2010a; Roncoli et al., 2015b)). Analytical approaches are usually efficient in computation and easy to be implemented. However, they are not easy to be

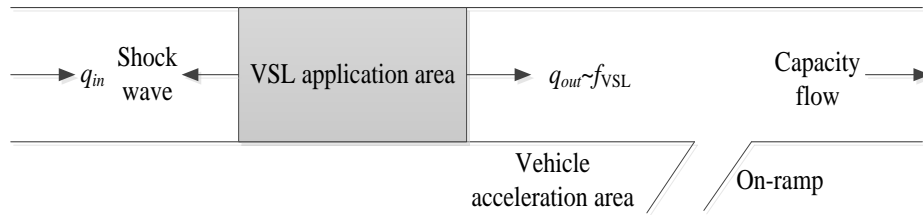


Figure 1.3: The mainstream metering approach used in (Carlson et al., 2010a). The on-ramp is perceived as a potential bottleneck. q_{in} and q_{out} are the inflow and outflow of the VSL control area. f_{VSL} is the VSL-induced capacity.

adapted to new situations (e.g., changes in infrastructure). MPC strategies of VSLs that are based on the METANET model have been extensively investigated in the literature. A noticeable problem for METANET-based MPC of VSLs is that the solution of the optimization may depend on the selection of the initial guess trajectory, in the sense that if the initial guess is not appropriately chosen, the solution of the optimization might get stuck at a local minimum. Studies of Muralidharan & Horowitz (2015); Roncoli et al. (2015b) have developed linear MPC strategies of VSLs. The control strategies have been demonstrated to be efficient in terms of computation speed and reducing total time spent by the designed case studies. However, the type of traffic jam considered in their case studies is only the standing queue.

This thesis focuses on the development of MPC strategies of VSLs to resolve free-way jam waves. Chapter 3 presents an MPC strategy that is based on a linear traffic flow model, which can reproduce both VSL mechanisms that have mentioned above. The accuracy of the traffic flow model in reproducing traffic dynamics under the VSL control is validated in Chapter 4.

1.2.3 Route guidance and intersection signal control

The field of intersection traffic control has been studied and developed in a variety of ways during the past decades. The initial purpose of intersection control is to ensure traffic safety by giving rights to conflicting direction of flows at different times. From the perspective of improving the traffic operation efficiency, different intersection control strategies are developed to reduce the total time spent of drivers. In the literature, intersection traffic control strategies are categorized into local control strategies (e.g. Webster & Cobbe (1963), Van Katwijk (2008), Varaiya (2013)) and coordinated control strategies (e.g. Sims & Dobinson (1980), Robertson & Bretherton (1991), Diakaki et al. (2002)). For many of the network-wide coordinated signal control strategies, fixed turning fractions at intersections are assumed in order to reduce the complexity. As a matter of fact, intersection control problem and dynamic traffic assignment problem are interrelated (Peeta & Ziliaskopoulos, 2001). On one hand, intersection delays significantly influence travelers route choice behavior and on the other hand, travelers route choice behavior influence the arriving flows at intersections. Thus, integrating

route choice control and intersection control may further improve the traffic operation efficiency.

Drivers route choice behavior can be controlled by dynamic route guidance. Practically, dynamic route guidance control actions are often performed via the roadside variable message signs (VMS). Drivers make their route choice decisions based on the information that are shown by the VMS. Recently, with the development of autonomous vehicles, some studies have explored traffic routing strategies within a connected vehicle environment (Zhu & Ukkusuri, 2015).

The integration of route choice control and intersection control increases the complexity of the model to describe the traffic system, and thus results in complicated optimization problems that is difficult to be solved efficiently. Therefore, many studies regarding the optimization of traffic signals and route choices used heuristic algorithms to solve the problem, which do not always obtain reasonable optimization results (Ceylan & Bell, 2004; Teklu et al., 2007). The linear quadratic model predictive control approach presented by Le et al. (2013) is a successful application to efficiently optimize turn fractions and traffic signals for single destination networks. Due to the fact that the LQMPC has a discrete linear prediction model, traffic state in each discrete segment of the traffic network is aggregated and route choice behavior of traffic in each OD pair is not considered. The controller pushes out traffic flow as much as possible, regardless of traffic desired origin and destination relations. Thus, the LQMPC approach may be ineffective when applied to multi-destination traffic networks, because the desired demand might end up at wrong destinations.

This thesis focuses on developing a real-time tractable optimization strategy of traffic signals and route choices that can be used to multi-destination networks. In Chapter 6, route choices and intersection signals are optimally computed using an extended linear quadratic model predictive control approach.

1.3 Model predictive control for road traffic

Model predictive control, also known as receding horizon, rolling horizon, or moving horizon control, has been extensively investigated in the traffic control area (Hegyi et al., 2005a; Carlson et al., 2010a). MPC for traffic systems utilizes a traffic model to predict traffic state evolution based on the current state of the system, and determines the optimal control actions that result in the optimum value of an objective function. This feature enables the controller to take advantage of potentially larger future gains at a current (smaller) cost, so as to avoid myopic control actions. After optimization, the control value of the first sample of the optimal control action is applied to the process. The remaining part of the control signal is recalculated in the finite rolling horizon scheme.

The MPC scheme for traffic control is shown in Fig. 1.4. The traffic process is usually represented by a traffic flow model which should be accurately enough to represent the

reality. In each control step, the MPC controller predicts the traffic dynamics based on the traffic states (e.g., speed, flow and density in the various links and segments of the network) that are measured from the process model. The MPC controller can take the disturbances, i.e., the uncontrollable inputs of the traffic system, into account to compute the optimal control signals. The disturbances include the changes of traffic demand, the variation of boundary conditions (e.g., the shock waves or congestion entering at the destinations of the network), and the occurrence of accidents that influence the traffic anywhere in the network. The optimal control signals consist of the control values of the traffic control measures, such as ramp metering rates, speed limits, route guidance signals, or intersection signals. The optimal control signals are implemented to the process model to solve or prevent the traffic problems.

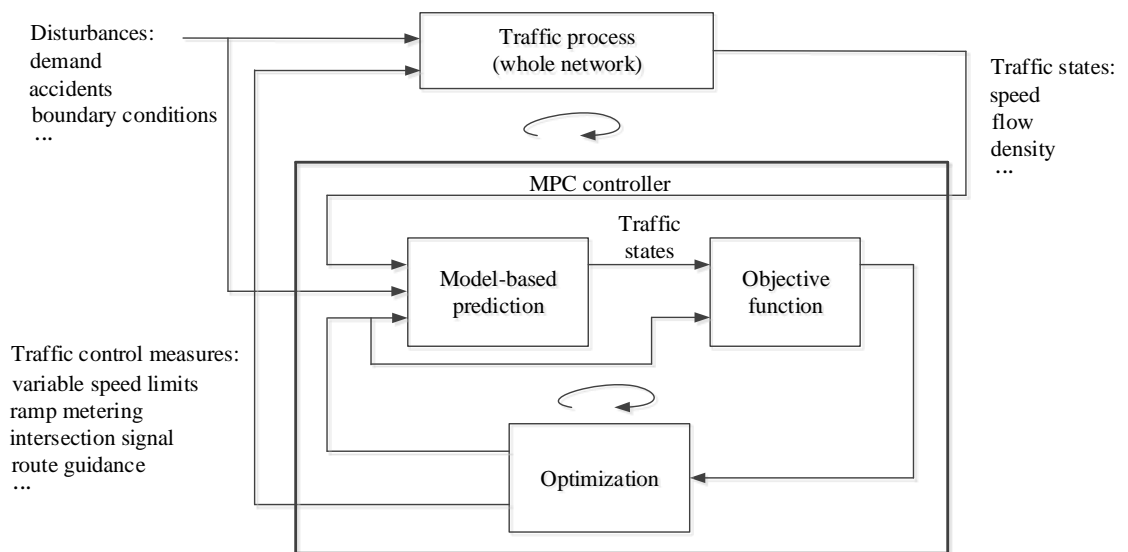


Figure 1.4: The MPC scheme for traffic control. The figure is taken from [Hegyi \(2004\)](#).

The advantages and disadvantages of MPC for traffic control have been discussed in the literature ([Hegyi, 2004](#); [Burger et al., 2013](#)). It has been commonly recognized that the feedback structure of the MPC significantly reduces the adverse effects of unpredictable disturbances. Besides, the MPC approach is easy tuning and the controller can be easily adapted to new situations (e.g., changes in infrastructure) by adapting the prediction model. Different control measures can be easily integrated in an MPC control system. For example, studies of [Hegyi et al. \(2005a\)](#); [Carlson et al. \(2010b\)](#); [Roncoli et al. \(2015b\)](#) have applied MPC controllers for integrated control measures.

It has also been commonly recognized that MPC with prediction models and objective functions that lead to non-linear and non-convex optimization problems are computationally complex and may get stuck in local optimum. Linear MPC approaches have been presented to overcome the shortcomings of non-linear and non-convex MPCs ([Muralidharan & Horowitz, 2015](#); [Roncoli et al., 2015b](#)). However, these prediction models do not accurately reproduce some traffic phenomena such as the propagation of jam waves, which may significantly deteriorate the control performance. In this thesis, the influence of prediction model accuracy to the control performance (in terms of

reducing the total time spent), and the influence of prediction model complexity to the computation time of the MPC are explored in Chapters 3, 5 and 6.

Many of the previous studies have applied same models as the prediction model and the process model that represents the reality, which eliminates the influence of the prediction model inaccuracy to the control performance. However in real life applications, there is a significant mismatch between the dynamics of the prediction model and the field process. Hence, to demonstrate the effectiveness of the proposed MPC controllers, a process model that is different from the prediction model needs to be applied. In this thesis, the proposed MPC controllers are tested by traffic flow models that have enough difference from the prediction model.

1.3.1 Traffic flow models

Macroscopic traffic flow models play an important role in developing efficient traffic control strategies for traffic networks. In the literature, two types of macroscopic traffic flow models, namely the cell transmission model (CTM) and the METANET model, and many of their variants, have been extensively used to develop model-based traffic control strategies.

The cell-transmission model is developed by [Daganzo \(1994\)](#), and it is consistent with the hydrodynamic theory of LWR model ([Lighthill & Whitham, 1955](#); [Richards, 1956](#)), so we also call it as discrete LWR model or discrete first-order model in this thesis. The cell-transmission model and many of its variants are solved by a minimum demand (sending capacity) and supply (receiving capacity) method, based on the Eulerian (space-time) formulation ([Lebacque, 1996](#)), for calculating the fluxes at discrete cell boundary. By knowing the current state of the cells and parameters of the fundamental diagram, the fluxes between cells can be calculated, and the cell state can be predicted over time. One advantage of these models is that they can easily be deployed to solve a linear optimization problem for traffic control. The maximization of the network throughput can be represented by a linear function of fluxes between cells. The fluxes that are determined by the minimum of the demand and the supply functions, can be represented by linear inequality constraints. Therefore, the optimal control scheme can be determined by solving a linear programming problem. For example, [Ziliaskopoulos \(2000\)](#) applied the cell-transmission model for network optimization based on a linear programming problem. [Gomes & Horowitz \(2006\)](#) developed a linear model predictive control strategy based on an asymmetric cell transmission model. For the purpose of developing efficient control strategies, it is important that the models of the cell-transmission type can reproduce the relevant traffic phenomena, such as the propagation of shockwaves and spillbacks. However, there are still many traffic phenomena that such kind of models could not reproduce, such as jam waves, traffic oscillations, and traffic hysteresis.

The METANET model is a second-order traffic flow model, which reproduces traffic flow patterns by describing the spatial temporal evolution of speed ([Kotsialos et al.,](#)

2002a). The METANET model and its variants have been applied for optimal control of VSLs, ramp metering, route guidance etc (Hegyi et al., 2005a; Carlson et al., 2010b; Le et al., 2013). The METANET model reproduces some traffic phenomena, e.g. the propagation of jam waves, more accurately than the cell-transmission model. Nevertheless, the METANET-based optimization problems are non-linear and non-convex, which has a higher computation load than the cell-transmission model-based optimizations.

Recently, the macroscopic fundamental diagram (MFD), which links the accumulation (weighted sum of links densities and lengths) and the production (weighted sum of links flows and lengths) of an urban region, provides an efficient tool for expressing aggregated dynamics of urban traffic networks (Daganzo, 2007; Geroliminis & Daganzo, 2008). The MFD has been utilized to solve optimal perimeter control problems in (Geroliminis et al., 2013; Haddad et al., 2013; Ramezani et al., 2015). Although a unimodal and low-scatter MFD was observed in a homogeneous region of a city (i.e. small spatial density heterogeneity), existence of a well-defined MFD in general cases is still an open question, which might undermine the benefits of MFD in modeling traffic dynamics.

In short, a well-developed traffic flow model should have a good trade-off in reproducing congestion levels, reproducing traffic phenomena, and developing efficient control strategies. This thesis improves the accuracy of some prediction models in the MPC, while keeping the model complexity limited. Chapter 2 focuses on the jam wave propagation phenomenon and proposes an extended CTM which keeps a linear feature and reproduces the propagation of jam waves accurately. The model is compared with the same type of models in previous works. Chapter 5 proposes a freeway MFD model which can accurately express the aggregated traffic dynamics of freeway networks. Both models are applied to efficient model predictive control strategies.

Although many model predictive control approaches in the literature have been demonstrated to be effective in reducing the TTS through simulations, few have validated the prediction model in reproducing the behavior under traffic control. The predicting traffic dynamics under traffic control needs to be validated because the accuracy of the prediction model is essential to the control performance. This thesis takes a step towards this direction (Han et al., 2017a). In Chapter 4, the prediction model that is embedded in the model predictive controller proposed in Chapter 3 is validated using field data.

1.4 Research objectives

The objective of this thesis is to develop real-time model predictive control strategies for control measures, such as ramp metering, variable speed limits, route guidance, and intersection signals to improve the operation efficiency of traffic networks. The

main objective is broken down into two sub-objectives, and are related to modeling and control aspects. Specifically:

Regarding the modeling aspect, the objective is to develop accurate traffic flow models that can be applied to efficient model predictive control strategies on freeways. To this end, this thesis mainly focus on developing and validating discrete first-order models and aggregated models, i.e., the macroscopic fundamental diagram, which have linear formulation and simple solution method.

In the literature, recently proposed model predictive control strategies that are based on discrete first-order models has significantly improved the computation speed of the controllers. However, it is found that existing discrete first-order models cannot accurately reproduce capacity drop and the propagation of jam waves. To this end, one of the objectives of this thesis is to develop a new extended discrete first-order traffic flow model for a better reproduction of capacity drop and the propagation of jam waves.

Apart from discrete first-order traffic flow models, the macroscopic fundamental diagram is another tool that has been widely applied to efficient model predictive control strategies. While the MFD has been extensively explored in urban traffic, the modeling of freeway MFD has not been sufficiently discussed. This thesis intends to develop a freeway MFD model that accurately reproduces the aggregated freeway traffic dynamics.

Regarding the control aspect, the objective is to develop real-time model predictive control strategies to address a certain traffic problems using different control measures. For example, using variable speed limits to suppress freeway jam waves, applying ramp metering to resolve on-ramp bottlenecks, and employing route guidance and intersection signals to manage flow distributions over traffic networks. The traffic problems addressed in this thesis have been considered by other model predictive control strategies of same control measures in previous works. This thesis intends to increase the computation speeds of the controllers, while keep a comparable or obtain a better control performance in terms of reducing the total time spent of traffic.

1.5 Main Contributions

The main contributions of the thesis is highlighted in this section. In line with the research objectives presented in the previous section, the contribution of this thesis are categorized into the development of traffic flow models and real-time model predictive control strategies.

Regarding modeling, we propose an extended discrete first-order model to reproduce the propagation of jam waves on freeways (see Chapter 2). The model makes a few assumptions in determining the boundary flows between segments, and the assumptions are supported by empirical findings. To show the improvement of the model, the

proposed model is compared with other discrete first order models that incorporate the capacity drop in reproducing the propagation of jam waves.

We validate the prediction model embedded in an MPC controller (proposed in Chapter 3) of VSLs using field traffic data (see Chapter 4). The model is calibrated and quantitatively compared with the extended METANET model, which has been extensively used for MPC approaches of VSLs.

We propose a freeway MFD model to predict the aggregated traffic state of freeway networks (See Chapter 5). The effect of density heterogeneity and capacity drop on characteristics of freeway MFD are investigated based on field traffic data. In addition, for the first time, we present a model to predict the evolution of the density heterogeneity in a freeway network. The accuracy of the MFD model and the heterogeneity model are validated using field data.

Regarding control, this thesis proposes several model predictive control strategies for traffic control measures such as variable speed limits, ramp metering, route guidance, and intersection signals, that are real time tractable. In the literature, there is no specific research aimed to resolve freeway jam waves through a linear MPC formulation. This thesis fills this gap by developing a linear-quadratic MPC (See Chapter 3) based on an extended discrete first-order traffic flow model proposed in Chapter 2. The proposed control approach is compared in a benchmark case study with a second-order model-based MPC approach (Hegyi et al., 2005a) and a first-order model-based MPC approach (Roncoli et al., 2015b), in terms of computation time and total travel delay reduction.

In addition, we propose a hierarchical control approach of ramp metering based on freeway MFD dynamics (see Chapter 5). For the first time, we present a ramp metering control strategy based on the macroscopic fundamental diagram. The MFD-based controller requires less computation efforts, which benefits a real-time application. The presented controller is tested with different models (e.g. CTM and METANET) as the process model and compared with other optimal control approaches.

Furthermore, we propose an extended linear quadratic model predictive control (LQMPC) approach of integrated routing and urban intersection control for multi-destination networks (see Chapter 6). Compared to the state of the art, the previous LQMPC approach runs fast enough for a real time application, however, the approach cannot be applied to multi-destination networks because it cannot preserve correct OD relations. In this thesis, we use a heuristic method to preserve correct OD relations in the extended LQMPC.

1.6 Thesis outline

This thesis consists of 7 chapters that are briefly described in the follows. Chapter 1 is the introduction part of this thesis. The following five chapters correspond to

the specific research objectives that are proposed in Section 1.4. Chapter 7 is the conclusion part. Note that each chapter is a completely stand-alone research paper including an abstract, introduction, methodology, results, and conclusions with its own notations. A depiction of the thesis outline is shown in Fig. 1.5.

Chapter 1 introduces the research scope and highlights the main contributions.

Chapter 2, 3, and 4 focus on the modeling and control of freeway jam waves. Specifically, Chapter 2 proposes an extended discrete first-order model, which intends to accurately reproduce the capacity drop and the propagation of jam waves. Chapter 3 develops a linear MPC strategy that is based on the proposed model to resolve freeway jam waves. Chapter 4 validates the prediction model that is used in the MPC strategy in Chapter 3.

Chapter 5 addresses on-ramps bottlenecks on freeway networks. A hierarchical ramp metering control strategy that is based on a proposed MFD model is presented in this Chapter. The presented control approach is tested by using different traffic flows models, including the model proposed in Chapter 2, as the process models.

Chapter 6 focuses on the management of flow distribution over multiple destination traffic networks. An extended linear quadratic model predictive control strategies is presented for the integration of routing and intersection control in multi-destination networks.

Chapter 7 concludes this thesis by summarizing the main contributions and discussing the potential field applications, and presenting the future research directions.

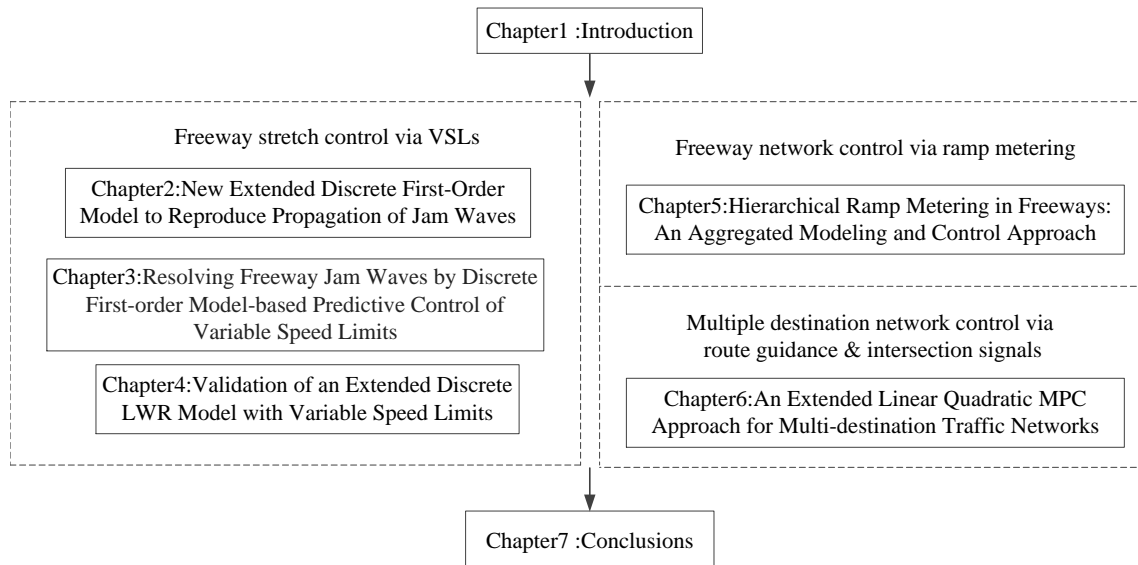


Figure 1.5: The outline of the thesis.

Chapter 2

A new extended discrete first-order model to reproduce the propagation of jam waves

In this chapter we propose an extension of the discrete LWR model of the cell transmission model type, to reproduce the capacity drop and the propagation of jam waves. Recently studies have tried to incorporate the capacity drop into discrete first-order traffic flow models for traffic optimization purpose. It is found that the inflow to a discharging cell predicted by these models might have been overestimated, which will influence the propagation of a jam wave. Empirical analysis has been carried out to confirm this assumption. Furthermore, it is found that the extent of the flow reduction depends on the state difference between the targeting cell and its upstream cell. Based on these findings, a new mathematical model formulation is given. Simulations with both a hypothetical freeway stretch and a real life freeway stretch are performed to test the behavior of the proposed model. The previously mentioned models are also simulated for comparison. Simulation results indicate that the proposed model has a better performance in reproducing jam waves. In addition, the proposed model can be used in a linear model predictive control framework, and formulated as a linear optimization problem, which may be beneficial for a real-life, real-time application.

This chapter is an edited version of the article:

Han, Y., Yuan, Y., Hegyi, A., Hoogendoorn, S. P. (2016). New Extended Discrete First-Order Model to Reproduce Propagation of Jam Waves. *Transportation Research Record: Journal of the Transportation Research Board*, (2560), 108-118.

2.1 Introduction

Traffic flow models play an important role in evaluating the impact of traffic congestion and developing efficient traffic control strategies for traffic networks. In the last decades, several freeway traffic flow models have been presented for both simulation and control purposes. Dynamic freeway traffic flow modeling started from [Lighthill & Whitham \(1955\)](#); [Richards \(1956\)](#), which developed the first dynamic traffic flow model (LWR model), based on a hyperbolic partial differential equation. The LWR model mainly contains two equations, one is the vehicle conservation equation, and the other one is known as the fundamental diagram. Based on this representation, [Daganzo \(1994\)](#) proposed a discretized version of the LWR model, known as the cell transmission model (CTM). Afterwards, many variants of the discrete first-order model have been presented in ([Gomes & Horowitz, 2006](#); [Muralidharan et al., 2009](#); [Roncoli et al., 2015a](#); [Han et al., 2015b](#)).

The cell transmission model and many of its variants are solved by a minimum demand (sending capacity) and supply (receiving capacity) method, based on the Eulerian (space-time) formulation, for calculating the fluxes at discrete cell boundary ([Lebacque, 1996](#)). By knowing the current state of the cells and parameters of the fundamental diagram, the fluxes between cells can be calculated, and the cell state can be predicted over time. One advantage of these models is that they can easily be deployed to solve a linear optimization problem for traffic control. The maximization of the network throughput can be represented by a linear function of fluxes between cells. The fluxes that are determined by the minimum of the demand and the supply functions, can be represented by linear inequality constraints. Therefore, the optimal control scheme can be determined by solving a linear programming problem. For example, [Ziliaskopoulos \(2000\)](#) applied the cell transmission model for network optimization based on a linear programming problem. [Gomes & Horowitz \(2006\)](#) developed a linear model predictive control strategy based on an asymmetric cell transmission model.

For the purpose of developing efficient control strategies, it is important that discrete first-order models can reproduce the relevant traffic phenomena, such as the propagation of shockwave and spillback. However, there are still many traffic phenomena that a discrete first-order model could not reproduce, such as jam waves, traffic oscillations, and traffic hysteresis.

Different higher-order models have been developed to overcome the shortcomings of first-order models. For a review of different models readers can refer to ([van Wageningen-Kessels et al., 2015](#)). In higher-order models, such as the METANET model proposed in [Kotsialos et al. \(2002a\)](#), the propagation of a jam wave can be reproduced by describing the spatial temporal evolution of speed. An important drawback of high-order model is that drivers may be affected by the traffic condition behind them ([Daganzo, 1995](#)). This criticism has been addressed by the Aw-Rascle-Zhang model, which has the anisotropic property ([Aw & Rascle, 2000](#)). But still, due to the non-linearity of high-order traffic flow models, the optimization problems based on

this kind of models are often difficult to be solved efficiently and accurately.

In short, a well-developed traffic flow model should have a good trade-off in reproducing congestion levels, understanding traffic phenomena, and developing efficient control strategies. Previous freeway traffic flow models in literatures have some shortcomings, either in describing traffic phenomena or in developing efficient control strategies. At the moment, a freeway traffic flow model that can reproduce relevant traffic characteristics for control and that can be efficiently solved is still lacking.

In general, two types of traffic jams on freeways can be identified. As has been presented in [Hegyi et al. \(2008\)](#), traffic jams with the head fixed at the bottleneck are known as standing queues, and jams that have an upstream moving head and tail are known as jam waves (also known as wide moving jam in some studies, e.g. ([Kerner & Rehborn, 1996](#))). A standing queue can be formed at a busy on-ramp, a lane drop bottleneck, or any kind of infrastructural bottlenecks. It can be reproduced by many discrete first-order models, as will be shown in section 4 of this chapter. A jam wave generates various negative effects. For example, it reduces freeway capacity, triggers standing queue, and results in travel delays and unsafe traffic conditions. Therefore, it is important for a traffic flow model to reproduce jam waves. As we know, the formation of a jam wave is stochastic and not easy to predict ([Schönhof & Helbing, 2007](#)). By [Laval & Leclercq \(2010\)](#) a possible explanation is provided for the spontaneous appearance of stop-and-go waves as a result of very small disturbances. However their explanation is in terms of a microscopic car-following model, and it is not trivial to translate it into a macroscopic (and linear) description. For these reasons, we will focus on the reproduction of jam wave propagation after it has occurred, which follows known patterns.

The outflow of a jam wave is lower than the free flow capacity, as a result of the capacity drop phenomenon. Recently, some studies have tried to incorporate the capacity drop into first-order traffic flow models. [Leclercq et al. \(2011, 2016\)](#) proposed analytical models to predict the capacity drop at an active merge bottleneck. [Schreiter et al. \(2010\)](#) assumed two constant capacity values for the cell transmission model, one for free flow condition and the other for congested condition. [Srivastava & Geroliminis \(2013\)](#) also proposed an extended LWR model with two capacity values, and they provided a memory-based methodology to choose the appropriate value of capacity for the numerical solution. The research shed light on different ways of modeling the capacity drop. However due to the non-linearity, these models have not been utilized for traffic control.

Some discrete first-order models which incorporated with the capacity drop, have been developed into traffic optimizations. [Muralidharan et al. \(2009\)](#) incorporated the capacity drop into a link-node cell transmission model, which uses a discontinuous fundamental diagram. The capacity drop was assumed to be triggered on only at the bottleneck cell, with a constant value. [Roncoli et al. \(2015a\)](#) proposed an optimization-oriented first-order model, which also represented the capacity drop by a reduced demand function. They assumed that the extent of the demand reduction of a cell has a

linear relation with the vehicle density of the same cell. In our previous work, a similar assumption has been made (Han et al., 2015b). We also assumed that the extent of the capacity drop of a cell has a linear relation, but with the density of the upstream cell. In this way the capacity drop at lane-drop bottleneck can also be reproduced. These models have been proposed for traffic optimizations, however, none of them have been validated in reproducing the propagation of jam waves.

The main contribution of this chapter is that we compare the performance of different discrete first-order models that can be used in an optimization framework (i.e. can be formulated as a linear or quadratic problem) in reproducing jam waves. After analyzing the shortcomings of these models, a new discrete first-order model is developed and validated. Besides the ability in reproducing the propagation of jam waves, the proposed model keeps the linear property of previous ones, which makes it applicable into a linear optimization problem for traffic control.

The organizing of the chapter is as follows. In Section 2.2 we practically describe the traffic flow features of the jam wave and theoretically analyze the reason why existing discrete first-order traffic flow models could not reproduce the propagation of a jam wave accurately. It is found that the supply function of the classical discrete first-order models might not be applicable to the moving jam condition. In Section 2.3, an empirical study is performed to check if the real data matches the theoretical findings. After checking the data, in Section 2.4, two reasonable assumptions are made. Based on these two assumptions, the detailed mathematical formulation of the first-order traffic flow model is given. In Section 2.5, the model is validated by both a synthetic case and a real life case. In the last section, we present the main conclusion and propose the future research direction.

2.2 Description and analysis of jam waves

2.2.1 Practical Description

A jam wave is a common traffic flow phenomenon on freeways. It usually originates from a traffic breakdown, which is caused for instance by a vehicle suddenly braking or lane changing in a high-demand traffic flow situation. Therefore, since so many (partially unobservable) factors influence the formation process, the formation of a jam wave is considered stochastic and difficult to be reproduced. From different empirical studies, some common patterns of its propagation can be distilled. For example, the propagation speed of jam waves is relatively stable, typically between 15-20 km/h (Kerner & Rehborn, 1996). The flow downstream of a jam wave is significantly lower (around 30 percent lower) than the free flow capacity, which is caused by the capacity drop phenomenon (Schönhof & Helbing, 2007). It can propagate for a long time period and distance, and resolves only with the traffic demand decreasing (Kerner & Rehborn, 1996). In order to illustrate these features, we consider the time-space contour plots of

a moving jam on the A13 freeway in the Netherlands that took place on 04-Sep-2012. As shown in Fig. 2.1 (a), the speed reduction started at detector 4 and propagated to detector 16. It propagated for about 5 km and lasted for about 15 minutes, meaning that the propagation speed is around 20km/h. The reason why the jam is resolved might be attributed to the decreased demand upstream of the jam, as shown in the dashed rectangle area in Fig. 2.1 (b). It can be found that the capacity drop phenomenon is associated with the jam wave. As shown by the solid triangle area in Fig. 2.1 (b), the queue discharge rate (red area, around 6000 veh/h) of the congestion is much less than the free flow capacity (yellow areas, around 8000 veh/h).

2.2.2 Theoretical analysis

The example of the last section shows that the capacity drop associated with a jam wave. In reality the inflow to a jam wave is always slightly lower than the capacity flow, since the critical state is not stable. The outflow of a jam will be higher than the inflow if the capacity drop is not taken into consideration in a traffic flow model, and the propagation of the jam will be resolved. Therefore, incorporating the capacity drop is the requirement for a traffic flow model to reproduce the propagation of jam wave. For example, the density-flow diagram of a freeway stretch is represented by Fig. 2.1 (d). State 2 represents the jam state, and 1 represents the traffic state out of the jam. Thus according to the shock wave theory (Lighthill & Whitham, 1955; Richards, 1956), if the inflow to the jam equals the outflow of the jam, the propagation of the jam wave is shown in the time-space graph of Fig. 2.1 (c).

Recently, some approaches have incorporated the capacity drop into discrete first-order traffic flow models (Gomes & Horowitz, 2006; Muralidharan et al., 2009; Roncoli et al., 2015a; Han et al., 2015b). These models are the variants of the cell transmission model, and are solved by the minimum demand and supply method. In addition, they have been used as the underlying models for linear optimization problems, which is consistent with the objective of this chapter. For traffic control applications, the freeway networks are always divided into cells with a typical length of several hundred meters, which on one hand guarantees a fast computation and on the other hand, is convenient for traffic state estimations. However, when a cells length becomes large, its traffic state might be inhomogeneous, which might result in inaccurate flow predictions since the fundamental diagram represents the density-flow relations for homogeneous traffic states. For instance, when a cell is in a mixed state with a less congested front part and a more congested back part (we define it as the discharging state), its average density is lower than the density of the back part. Under this circumstance, the inflow to a discharging cell might be overestimated by the previous models, which will influence the propagation of a jam wave.

For example, the space and time are divided discretely, shown as i , $i + 1$, and t_1 , t_2 in Fig. 2.1(c). At time step t_2 , cell $i + 1$ is in a discharging state. The boundary flux $q_{(i \rightarrow i+1)}$ is determined based on the average density of cell $i + 1$, because the receiving

capacity is dominant in determining the flux. However, since the back part of cell $i + 1$ is more congested than the front part, calculating the flux based on an average density might overestimate the flux, and this will influence the jam wave propagation since the congestion accumulation of cell i is slowing down. We further investigate this assumption through empirical analysis and simulations in the next sections.

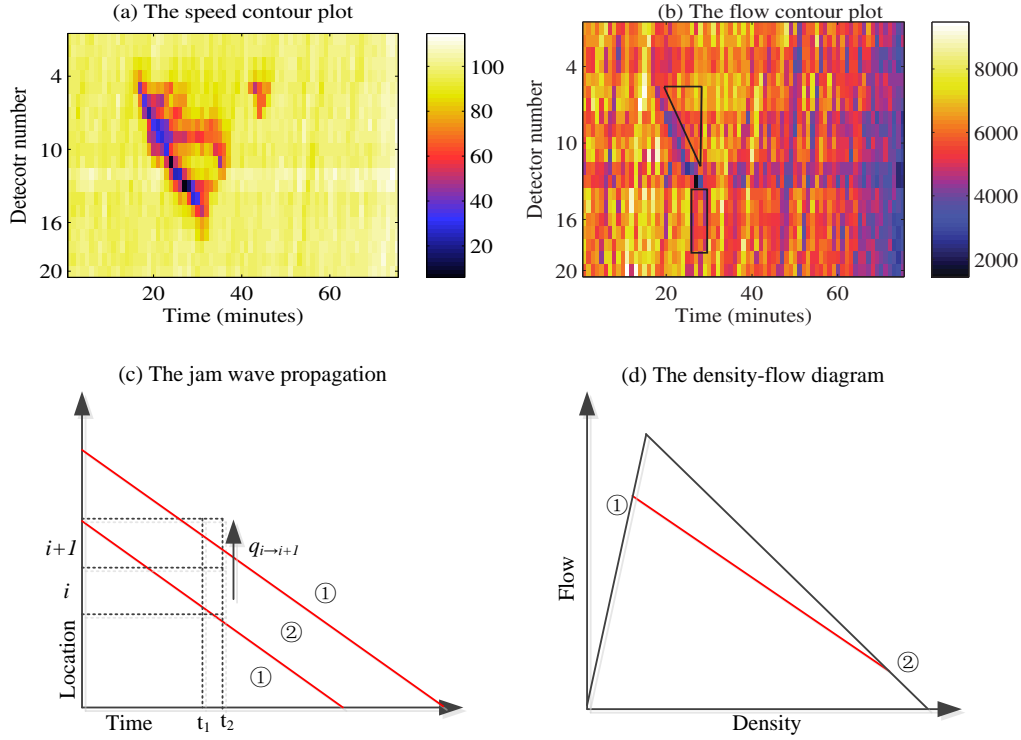


Figure 2.1: Visualization of a jam wave from empirical data and theoretical analysis. Empirical data were obtained from loop detectors of Dutch freeway A13. There are four lanes on this freeway stretch.

2.3 Model development

2.3.1 Empirical analysis

To understand the key mechanism that a discrete first-order model requires to capture to be able to correctly reproduce jam waves, empirical analysis has been performed. The site we choose is part of British motorway M42 (Wilson, 2008). We have chosen this site since there are enough detectors on a homogeneous freeway stretch, which can provide sufficient data for analysis. As shown in Fig. 2.2 (a), there are nine detectors (8-16) placed every one hundred meters on the homogeneous freeway stretch with same geometric structure. The detectors recorded the speed and detected time of every single vehicle, so we can calculate the harmonic mean speed to approximate the space

mean speed, which is more accurate than the time mean speed often provided by other databases.

Data were obtained on day 03-Oct-2008, between 12:00 to 18:00 for the analysis. In this period, both the free flow state, a state close to capacity, and congestion state can be found. Additionally, several moving jams have formed in this time period, as shown in Fig. 2.2 (b). To see if the supply function of the cell transmission model is applicable for a discharging cell, we should compare the supply of a receiving cell between when it is in discharging state with when it is not. The following shows the process of selecting appropriate data for analysis.

Assume that the freeway stretch is divided into cells by the detectors, and that we let the detectors represent the downstream boundary of cells. For cells 8-16, the sending cell is represented by i , and the receiving cell is represented by $i + 1$. If cell $i + 1$ is in discharging state, the following conditions should be satisfied.

$$\rho_{i+1} \geq \rho_{cr} \quad (2.1a)$$

$$\rho_i < \rho_{i+1} + \delta_1 < \rho_{i+2} + \delta_2, \delta_1 \geq \delta_2 > 0 \quad (2.1b)$$

$$v_i > v_{i+1} + \xi_1 > v_{i+2} + \xi_2, 0 < xi_1 \leq xi_2 \quad (2.1c)$$

For the cell transmission model, the supply function of a receiving cell is dominant in determining the flux when the density of the receiving cell is beyond the critical value. Thus, the first requirement is that the density of cell $i + 1$, ρ_{i+1} , is higher than the critical density, ρ_{cr} , formulated as condition (2.1a). To determine the value of the critical density, the flow-density plot (fundamental diagram) of the freeway stretch is plotted, and shown in Fig. 2.2 (c). From the figure we can see that the critical density is higher than 100 veh/km. To make sure all the selected states are in congested condition, we only collected the data that $\rho_{i+1} \geq 120$. Note that for British motorway M42, speed limit control is activated when jam wave occurred, with the value 40 or 50 mph typically. This causes a kind of double structure (two free flow branches) of the fundamental diagram shown in Fig. 2.2 (c). The speed limit will not have any influence on our analysis, because we only collect data for congestion condition.

Since a jam wave propagates from the downstream to the upstream cell, if cell $i + 1$ is in discharging state, then cell $i + 2$ should be in the discharging state or free flowing state, and cell i should be in the congestion accumulating state. Thus, the density of cell $i + 1$ should be higher than that of cell i , and lower than the density of cell $i + 2$, shown as (2.1b). The speed of cell $i + 1$ should be higher than that of cell $i + 2$, and lower than the density of cell i , shown as (2.1c). δ and ξ are adjustable parameters which are relevant to the width of the moving jam. δ_1 and δ_2 are set to 10, and ξ_1 and ξ_2 are set to 3 in this case, based on an experiential test.

If all conditions in equation (2.1) are satisfied, the points $(\rho_{i+1}, q_{(i \rightarrow i+1)})$ are plotted on the left picture of Fig. 2.2 (d), which are highlighted as red circles. $q_{(i \rightarrow i+1)}$ is the boundary flux from cell i to cell $i + 1$. If cell $i + 1$ is not in discharging state, the

following requirements should be satisfied.

$$\rho_{i+1} \geq \rho_{cr} \quad (2.2a)$$

$$\rho_i < \rho_{i+1} < \rho_{i+2} \quad (2.2b)$$

$$v_i > v_{i+1} > v_{i+2} \quad (2.2c)$$

If the above conditions are satisfied, the points $(\rho_{i+1}, q_{(i \rightarrow i+1)})$ are plotted on the left picture of Fig. 2.2 (d), which are highlighted as red stars. The red star scatters are linearly fitted, as shown in the figure, and can be perceived as the congestion branch of a fundamental diagram. The result clearly shows that for a same density value, the flux $q_{(i \rightarrow i+1)}$ which under conditions in (2.1), are lower than those under conditions in (2.2). We will refer to this phenomenon as the supply drop.

The influence of the extent of supply drop is also investigated. We recall the capacity drop assumption made in our previous work (Han et al., 2015b). The extent of the capacity drop for a free flowing cell depends on the density difference between this cell and the upstream cell. This inspires us to make a similar assumption that the supply drop level also depends on the density difference between two consecutive cells.

To investigate this assumption, the red circles in the left picture of Fig. 2.2 (d) are divided into two groups. If $\rho_i - \rho_{i+1} \geq 30$, points $(\rho_{i+1}, q_{(i \rightarrow i+1)})$ are plotted as blue circles. If $10 < \rho_i - \rho_{i+1} < 30$, points $(\rho_{i+1}, q_{(i \rightarrow i+1)})$ are still represented by red circles, which are shown in the right picture of Fig. 2.2 (d). It is obvious that the value of $q_{(i \rightarrow i+1)}$ are lower when $10 < \rho_i - \rho_{i+1} < 30$ than when $\rho_i - \rho_{i+1} \geq 30$. Therefore, from the analysis of this case, it can be observed that the extent of the supply drop has a positive relation with the density difference between the sending cell and the receiving cell.

2.3.2 A new discrete first-order model

The model development process is divided into 4 steps. At step 1, the new supply function of a discharging cell is introduced with a homogeneous freeway stretch. Based on the new supply function, the equations in determining the flux between cells are proposed at step 2. Step 3 extends the proposed model from homogeneous freeway stretches to generic situations. Step 4 elaborates the node models for merging and diverging, so that the proposed model can be applied to a network.

Step 1. Two reasonable assumptions are made in the last sub-section. a). The existence of the supply drop. Under the same density condition, the supply of a discharging cell is lower than a non-discharging cell. b). The extent of the supply drop has a positive relation with the density difference between the sending cell and the receiving cell. Based on these two assumptions, a new supply function is proposed,

$$S_{i+1}^{(1)} = \beta_{i+1}^{(1)}(\rho_{i+1}^J - \rho_{i+1}), \rho_i \leq \rho_{i+1} \quad (2.3a)$$

$$S_{i+1}^{(2)} = a + b(\rho_i - \rho_{i+1}), \rho_i > \rho_{i+1} \quad (2.3b)$$

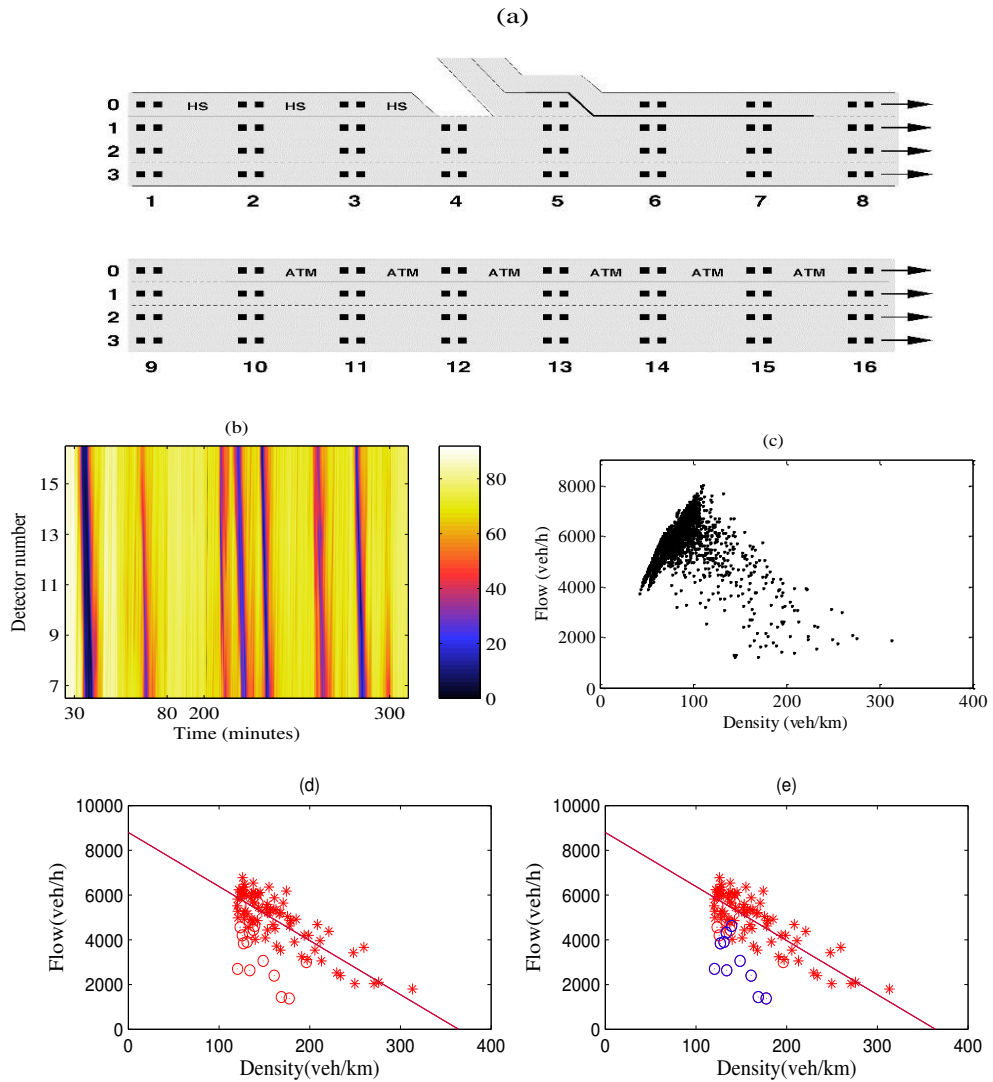


Figure 2.2: Results of the empirical study. (a) The targeting motorway stretch. (b) The speed contour plot of the analyzing period. (c) The density-flow plot (fundamental diagram) of the analyzing period. (d-e) Data analysis of the proposed site. The plots represent the relation between density $\rho_{(i+1)}$ and flux $q_{(i \rightarrow i+1)}$, $8 \leq i \leq 16$.

where, $S_{i+1}^{(1)}$ is the supply of cell $i + 1$ when it is in non-discharging state, which is represented by $\rho_i - \rho_{i+1}$. ρ_{i+1}^J is the jam density of cell $i + 1$. $S_{i+1}^{(2)}$ is the supply cell $i + 1$ when it is in discharging state, represented by $\rho_i \geq \rho_{i+1}$. $\beta_{i+1}^{(1)}$ is the congestion wave speed of the cell transmission model, as indicated from Fig. 2.3 (a). ρ_{i+1}^J is the jam density of cell $i + 1$. The supply of a discharging cell is assumed to have a linear relationship with the density difference between the sending cell and the receiving cell, so coefficient a and b are constant. If $\rho_i = \rho_{(i+1)}$, based on (2.3a), $S_{i+1}^{(2)} = a$, and based on (2.3b), $S_{i+1}^{(1)} = \beta_{i+1}^{(1)}(\rho_{i+1}^J - \rho_{i+1})$. When $\rho_i = \rho_{i+1}$, $S_{i+1}^{(1)} = S_{i+1}^{(2)}$. Thus, $a = \beta_{i+1}^{(1)}(\rho_{i+1}^J - \rho_{i+1})$. For homogeneous cells, $\rho_i^J = \rho_{i+1}^J$. Thus, $a = \beta_{i+1}^{(1)}(\rho_i^J - \rho_i)$.

If $\rho_i > \rho_{i+1}$, and $\rho_{i+1} = \rho_{i+1}^{cr}$, the capacity drop phenomenon happens because the receiving cell is in the free flowing state. According to our previous assumptions, the queue discharge rate of cell $i + 1$, c'_{i+1} is,

$$c'_{i+1} = c_{i+1} \left(1 - \alpha_{i+1} \frac{\rho_i - \rho_i^{cr}}{\rho_i^J - \rho_i^{cr}} \right) \quad (2.4)$$

where, α_{i+1} is the maximum extent of the capacity drop of cell $i + 1$. ρ_i^{cr} is the critical density of cell i .

In (4), if $\rho_i = \rho_i^J$, then $c'_{i+1} = c_{i+1}(1 - \alpha_{i+1})$. In (2.3a), if $\rho_{i+1} = \rho_{i+1}^{cr}$, $c'_{i+1} = (c_{i+1} \cdot (1 - \rho_{i+1}))/v_{i+1}$, (v_{i+1} is the free flow speed of cell $i + 1$), and $\rho_i = \rho_i^J$, then $S_{i+1}^{(2)} = c_{i+1}(1 - \alpha_{i+1})$. Thus, $b = (c_{i+1}(1 - \alpha_{i+1}))/(\rho_{i+1}^J - \rho_{i+1}^{cr})$. b is denoted as $\beta_{i+1}^{(2)}$. It represents the congestion wave speed of a cell when it is in discharging state, as indicated in Fig. 2.3 (a). Therefore, equation (2.3b) can be reformulated as,

$$S_{i+1}^{(2)} = \beta_{i+1}^{(1)}(\rho_i^J - \rho_i) + \frac{c_{i+1}(1 - \alpha_{i+1})}{\rho_{i+1}^J - \rho_{i+1}^{cr}} \cdot (\rho_i - \rho_{i+1}) \quad (2.5)$$

Step 2. Step 1 gives two supply functions of a receiving cell for over-critical condition. For free flow condition, the supply of a receiving cell equals to the queue discharge rate. To summarize, the supply of the receiving cell S_{i+1} , is determined by the following equation,

$$S_{i+1} = \min(S_{i+1}^{(1)}, S_{i+1}^{(2)}, c'_{i+1}) \quad (2.6)$$

The demand of cell i is determined by the minimum between the actual vehicle demand of this cell and the queue discharge rate of this cell.

$$D_i = \min(\rho_i \cdot v_i, c'_i) \quad (2.7)$$

where, v_i is the free flow speed of cell i .

The same as the cell transmission model, $q_{(i \rightarrow i+1)}$ is determined by the minimum value between D_i and $S_{(i+1)}$,

$$q_{(i \rightarrow i+1)} = \min(D_i, S_{i+1}) \quad (2.8)$$

To illustrate the approach, the demand and supply function can be represented by Fig. 2.3 (a). The demand function is represented by the blue line, and the supply function is represented by the red line. The physical meaning of the proposed model can be explained in this way: suppose that two consecutive cells, cell i and cell $i + 1$ are in jam state at time t_0 . The acceleration process of cell $i + 1$ can be represented by the blue line of Fig. 2.3 (c), while the state change of cell i can be represented by the red line.

Step 3. The proposed model is introduced with a homogeneous freeway stretch. It can be extended to generic situations. If two consecutive cells have different geometric structure and different fundamental diagrams, for example, with the on-ramp or lane drop in between, then the supply of the receiving cell depends on the mapped state difference. For example, as shown in the Fig. 2.3 (b), the black lines denote the fundamental diagram of sending cell i , and the solid red lines denote the fundamental diagram of receiving cell $i + 1$. To get the state difference, first we need to map the state of cell i to cell $i + 1$. The mapped state is represented by ρ'_i as shown in Fig. 2.3 (b). It can be calculated by the following equation,

$$\rho'_i = \frac{c'_{i+1} + \beta_{i+1}^{(2)} \rho_{i+1}^{\text{cr}'} - \beta_{i+1}^{(1)} \rho_{i+1}^{\text{J}}}{\beta_{i+1}^{(2)} - \beta_{i+1}^{(1)}} \quad (2.9)$$

Replace ρ_i with ρ'_i in (5), $S_{i+1}^{(2)}$ can be calculated correspondingly.

Step 4. For the proposed discrete models, the state of each cell is updated every time step, by the following equation,

$$\rho_i(k+1) = \rho_i(k) + \frac{T}{L_i} [q_{(i-1 \rightarrow i)}(k) - s_i(k) + r_i(k)] \quad (2.10)$$

where, $\rho_i(k)$ is the density of cell i at time step k . T is the time duration of one time step, L_i is the length of cell i . $s_i(k)$ is outflow of cell i towards the off-ramp. $r_i(k)$ is the inflow to cell i from the on-ramp. $r_i(k)$ is determined by the following equation,

$$r_i(k) = \min \left(\frac{\omega_i(k)}{T}, u_i, \max \left(\frac{u_i}{c_{i-1} + u_i} \cdot S_i(k), S_i(k) - \frac{\rho_i(k) l_i(k)}{T} \right) \right) \quad (2.11)$$

where, $\omega_i(k)$ is the vehicle number of on-ramp i . u_i is the saturation flow of on-ramp i . On-ramp flux $r_i(k)$ shares the supply $S_{i+1}(k)$ with the mainstream flux $q_{(i \rightarrow i+1)}(k)$ proportionally to the capacity or saturation flow, if there are enough demand from both the mainstream and the on-ramp. If the demand from the mainstream is not enough to share the supply, the remaining supply will be distributed to the mainstream flow, and vice versa. $\omega_i(k)$ evolves as the following equation:

$$\omega_i(k+1) = \omega_i(k) + T(d_i(k) - r_i(k)) \quad (2.12)$$

where, $d_i(k)$ is the demand of on-ramp i at time step k .

Since the model is solved by the minimum value between the demand function and the supply function, and both the demand function and the supply function are piecewise linear, the proposed model can be formulated as a linear optimization problem for traffic control. A referential example can be found in (Han et al., 2015b).

From the data analysis we know that even though the supply of a cell and the state difference between the targeting cell and the downstream cell have positive relation, their relation is much complex than a linear relation. We assumed a linear relation just for the convenience of future optimizations. However, we believe that the linear relation is sufficient for the proposed model to reproduce the propagation of a jam wave satisfactorily. Specific model validations are performed in the next section.

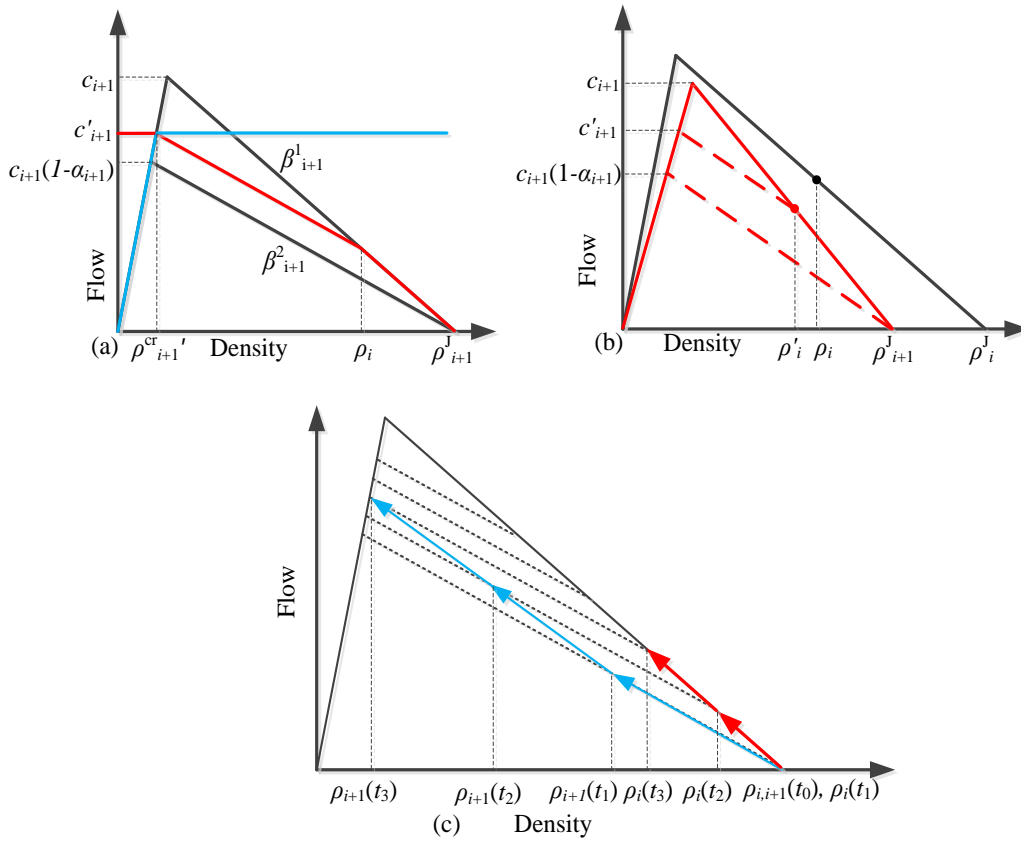


Figure 2.3: Description of the proposed model. (a) shows the demand and the supply function of the proposed model. (b) represents the different fundamental diagram of two consecutive cells. (c) shows the acceleration process of two neighbouring jam cells.

2.4 Model validation

To confirm that the proposed model is able to reproduce the propagation of a jam wave satisfactorily, simulations of both a synthetic freeway stretch and a real life freeway

stretch are performed. Some of the previous mentioned discrete first-order traffic flow models are also simulated for comparison. The first one is the model from (Roncoli et al., 2015a) (short as model R). The second one is the model from (Muralidharan & Horowitz, 2012) (short as model M). These models are all the variants of the cell transmission model, and they have been used as the underlying models for a linear optimization problems, which is consistent with the objective of this work (Muralidharan & Horowitz, 2015; Roncoli et al., 2015b).

2.4.1 Simulation test on a synthetic freeway stretch

A hypothetical homogeneous freeway stretch is proposed to test the performance of the proposed models. The two-lanes freeway stretch, is divided into 17 cells of, equal length of 0.6 km, as shown in Fig. 2.4 (a). The free flow speed is set to 108 km/h, and the capacity is set to 2000 veh/h/lane, so the critical density can be calculated as the division of the capacity and the free flow speed. The wave speed is set to 18 km/h, so the jam density can be known correspondingly. The maximum extent of capacity drop is set to 35% for model R and the proposed model, based on empirical findings (Yuan et al., 2015). The constant capacity drop of model M is set to 15% (Muralidharan & Horowitz, 2015).

The demand profile of the simulation is shown in Fig. 2.4 (b). Demand decreases from 3800veh/h to 3000 veh/h at time step 280, to show the influence of demand to the propagation of jam waves. At time step 180, the density of cell 16 is manually set to the jam density, 250veh/km. Then the congestion propagation of the previous mentioned models are simulated. The simulation results are shown in Fig. 2.4 (c-j).

As shown in the density contour plots of Fig. 2.4, the cell transmission model and model R show similar results. The jams are resolved, before the traffic demand decreasing. For model M, the jam evolves into a standing queue, because a constant amount of the capacity drop is assumed. The proposed model also shows dispersion, but it is much less than the other models. Before the traffic demand decreasing, traffic jam still exist and is wider than the start. Thus, the proposed model works better in reproducing jam wave in this hypothetical freeway stretch.

The flow contour plots indicate how the capacity drop is reproduced by these models. For the cell transmission model, the capacity drop is not taken into account, so the outflow from congestion is always the free flow capacity. For model M, the capacity drop only takes place at cell 16. When cell 16 has an equal inflow and outflow, its density stays constant. That is why the congestion head does not move upstream. For model R, even though the capacity drop is assumed to happen for every cell, we can see from Fig. 2.4 (h) that the capacity drop actually only happens at cell 16. The proposed model, by contrast, reproduces the capacity drop better. From Fig. 2.4 (j) it can be seen that the outflow of every cell in jam is less than the free flow capacity. We have also investigated the models behavior as the temporal step tends to zero. The

simulation results are similar to the above case, that the dispersion of previous models are conspicuous while the proposed model is much less.

2.4.2 Simulation tests for a real life freeway stretch

The traffic flow model proposed in the last section is now applied to a particular freeway stretch for model validation. The chosen site is a stretch of Dutch freeway A13, connecting the city Rotterdam and the city Delft. We choose this site because there are not many on-ramps and off-ramps attached to this stretch. In the Netherlands, flows of the on-ramps and the off-ramps are hardly detected. Therefore, the demand of the on-ramp and the splitting ratio of the off-ramp are roughly estimated by the flow difference between the upstream detector and downstream detector of the on-ramp or the off-ramp. If there are too many ramps involved, the estimation errors might be magnified and leads to inaccurate predictions.

The considered stretch has a length of 7.8 km. A graphical representation of the freeway stretch is shown in Fig. 2.5 (a). There are 19 detectors included in the targeting stretch, which are represented by black dots. Each segment in the figure represents a cell, with a length of 150 m. On-ramps (represented by R1, R2), off-ramps (represented by S1, S2, S3), and lane drop (represented by a red line), divide the freeway stretch into 7 sections. The calibrated capacity (veh/h) of each section are also shown in Fig. 2.5 (a). Other parameters are assumed to be constant for all seven sections for convenience. The free flow speed v is set to 108 km/h, the wave speed β_1 is set to 19.5 km/h, the maximum extent of the capacity drop is set to 35%.

On the day 08-Oct-2012, a jam wave occurred on A13 during the morning peak. The jam was formed at a bottleneck downstream of the targeting stretch and propagated through it. The speed contour plot of the targeting stretch, from 7:30 to 9:30, is shown in Fig. 2.5 (e). In the simulation, the formation of the jam wave is reproduced by changing the boundary condition. The outflow of the downstream boundary of the freeway stretch is changed to 2000 veh/h manually at time period 8:22-8:25.

Demand profiles of the mainstream and on-ramps are shown in Fig. 2.5 (b-c). Since there are not many vehicles using on-ramp R2, the demand of on-ramp 2 is assumed to be a constant value. The total demand of the considered time period is calculated by the difference between the flow aggregation of the upstream detector and the downstream detector. Fig. 2.5 (d) is the splitting ratio of three off-ramps. The same as on-ramp 2, off-ramp S2 is also not busy. Thus, the off-ramp splitting ratio of S2 is assumed to be constant as well.

Speed contour plots of the simulations are shown in Fig. 2.5 (f-i). The cell transmission model and model R show similar result, that the jam wave does not propagate for a sufficient long time and distance, and it is resolved spontaneously when the jam propagates upstream. For model M, the head of the congestion forms to a standing queue again. Only the proposed model reproduces a satisfactory result that the moving

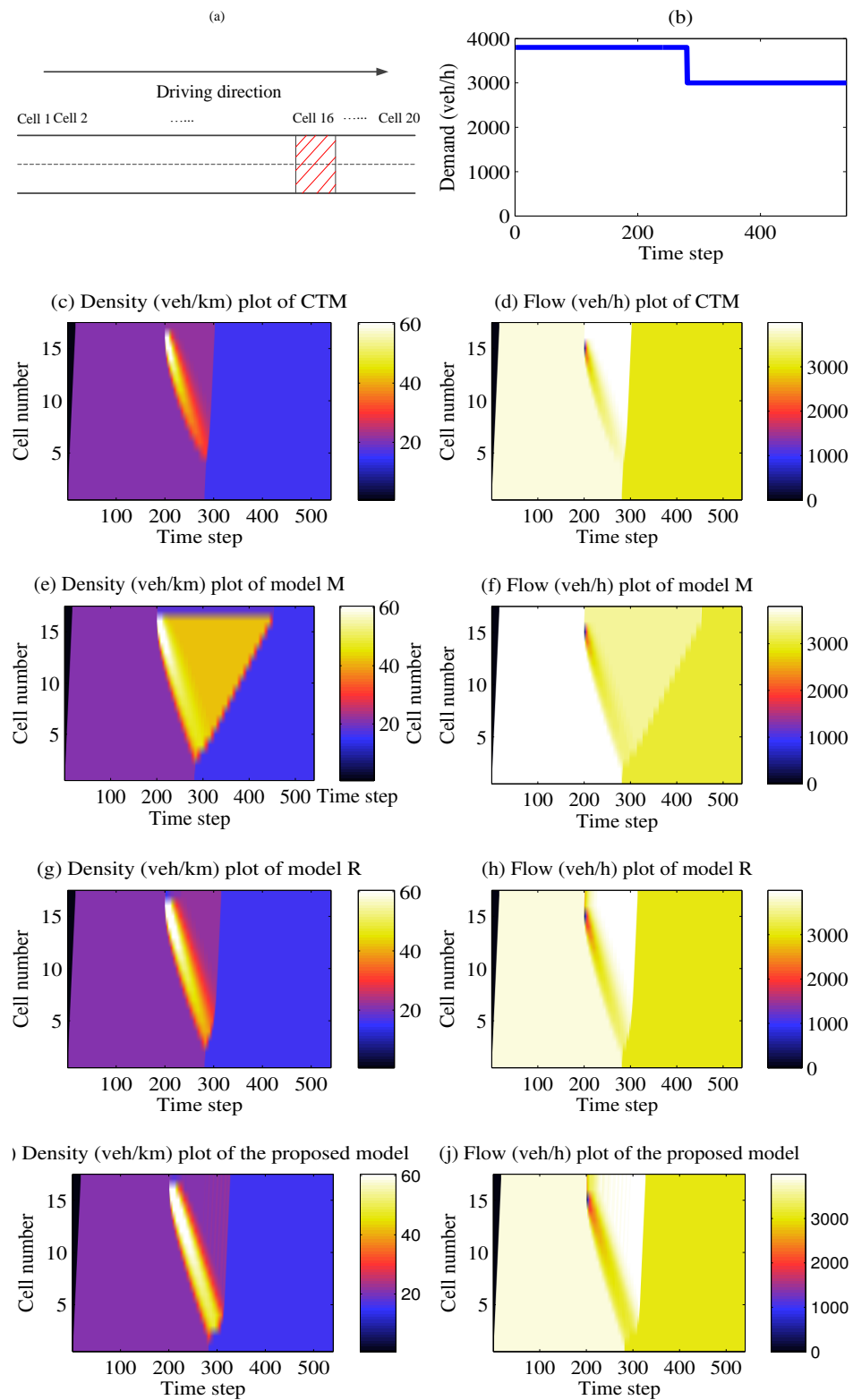


Figure 2.4: The hypothetical freeway stretch, demand profile, and simulation results.

jam propagates through the freeway stretch, which is in line with reality. Note that the proposed model is sensitive to the demand and the value of α . Therefore, the small dispersion shown in the downstream end of Fig. 2.5 (f) might be attribute to the small detecting error of the demand.

The simulation results are also tested quantitatively. The Root-mean-square error (RMSE) is chosen as the quantitative criteria for evaluating the simulation results. In this case only the time period from 8:22 to 8:53, when the jam wave was taken place at the stretch is selected for calculation. The RMSE is calculated based on the following equation. The RMSE of the proposed model, model R, CTM and model M are 26.98, 34.77, 34.86, and 41.8. Thus, quantitative results also show that the proposed model have a better performance in reproducing jam waves.

$$\text{RMSE} = \sqrt{\frac{\sum_{t=52}^{t=83} (\tilde{v}_t - v_t)^2}{32}} \quad (2.13)$$

2.5 Conclusions

In this chapter we propose a linear discrete first-order traffic flow model that aims to reproduce the propagation of jam waves. First, theoretical and empirical analyses are performed to confirm that the supply function of the cell transmission model is not applicable for the jam wave condition. Two assumptions are made based on theoretical and empirical findings. The first one is a new concept, the supply drop, which means that the supply of a cell should be decreased when the cell is in discharging state. The second one is about the extent of the supply drop. It is assumed that the extent depends on the state difference between two neighboring cells. Based on these two assumptions, a linear supply function of a discharging-stated cell is proposed. After that, the mathematical formulation of the proposed model is elaborated.

Model validation is performed by data from both simulation of a synthetic freeway stretch and a real life freeway stretch. From the synthetic case, it is shown that the proposed model is able to better reproduce the propagation of jam waves and the capacity drop. For the real life case, qualitatively, it is found the jam wave reproduced by the proposed model propagates longer in time and space than that reproduced by the other models, which fits reality better. Quantitatively, the root-mean-square error of the proposed model is much lower than that of the other models. It could be concluded that the proposed model has a better performance in reproducing the propagation of jam waves.

Besides the advancement in reproducing traffic phenomena, the proposed model keeps the advantage of the cell transmission-based models in respect of computation effi-

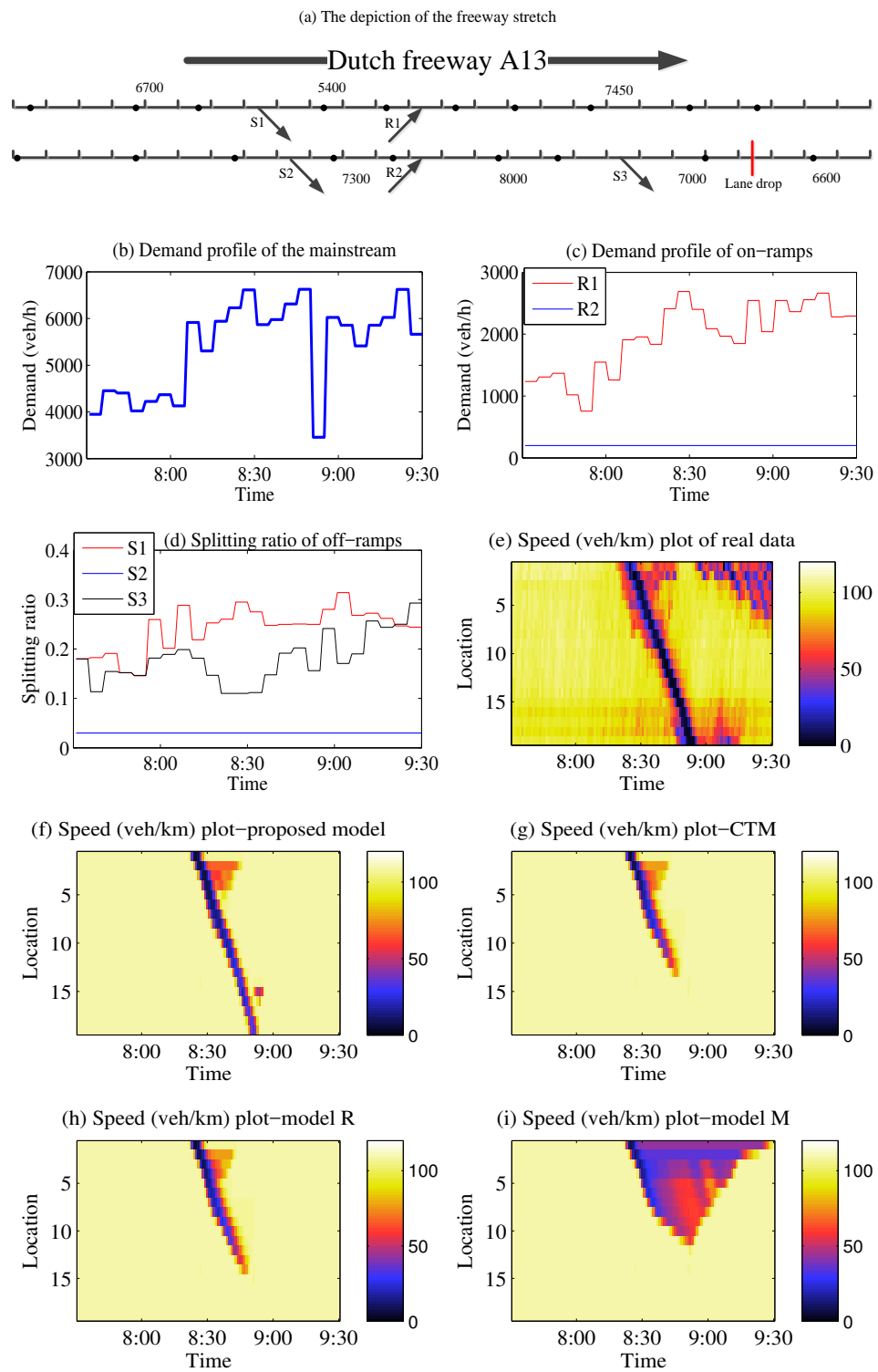


Figure 2.5: The real life freeway stretch, demand profiles, and simulation results.

ciency. The linear property of the proposed model makes it suitable for a linear optimization problem for traffic control purpose. This will be the focus of our future research.

Chapter 3

Resolving freeway jam waves by discrete first-order model-based predictive control of variable speed limits

In this chapter we develop a fast model predictive control (MPC) approach for variable speed limit coordination to resolve freeway jam waves. Existing MPC approaches that are based on the second-order traffic flow models suffer from high computation load due to the non-linear and non-convex optimization formulation. In recent years, simplified MPC approaches which are based on discrete first-order traffic flow models have attracted more and more attention because they are beneficial for real-time applications. In literature, the type of traffic jam resolved by these approaches is limited to the standing queue in which the jam head is fixed at the bottleneck. Another type of traffic jam known as the jam wave, has been neglected by the discrete first-order model-based MPC approaches. To fill this gap, we develop a fast MPC approach based on a more accurate discrete first-order model. The model keeps the linear property of the classical discrete first-order model, meanwhile takes traffic flow features of jam waves propagation into consideration. A classical non-linear MPC and a recently proposed linear MPC are compared with the proposed MPC in terms of computation speed and jam wave resolution by a benchmark problem. Simulation results show that the proposed MPC resolves the jam wave with a real-time feasible computation speed.

This chapter is an edited version of the article:

Han, Y., Hegyi, A., Yuan, Y., Hoogendoorn, S., Papageorgiou, M., Roncoli, C. (2017). Resolving freeway jam waves by discrete first-order model-based predictive control of variable speed limits. *Transportation Research Part C: Emerging Technologies*, 77, 405-420.

3.1 Introduction

Due to increasing demand, traffic jams happen frequently at current freeway networks. Traffic jams generate many negative effects such as reducing freeway capacity, inducing travel delays and potentially unsafe traffic conditions. Infrastructure expansion is an effective way to combat traffic jams. However, adding additional lanes and new freeways are not always viable due to economical or environmental concerns. Instead, traffic control measures are playing more and more important roles in improving traffic operation efficiency.

To implement effective traffic control measures to resolve traffic jams, the features of traffic jams have to be investigated. In general, two types of traffic jams on freeways can be identified. As has been presented in (Hegyi et al., 2008), traffic jams with the head fixed at the bottleneck are known as standing queues, and jams that have an upstream moving head and tail are known as jam waves (also known as wide moving jams in some studies, e.g., (Kerner & Rehborn, 1996)). A standing queue can be formed at a busy on-ramp, a lane drop bottleneck, or any kind of infrastructural bottleneck. According to empirical studies in literature (Cassidy & Bertini, 1999; Chung et al., 2007; Srivastava & Geroliminis, 2013), the outflow of a standing queue is typically 10-15 percent lower than the free flow capacity, which is known as the capacity drop. Jam waves usually originate from breakdowns, and in most cases from standing queues. Possible triggers for the formation of a jam wave could be a vehicle suddenly braking or a lane-changing vehicle in a high-demand traffic flow situation. Therefore, since so many (partially unobservable) factors influence the formation process, the formation of a jam wave is considered stochastic and difficult to be reproduced. However, from different empirical studies, some common patterns of its propagation can be distilled. For example, the propagation speed of jam waves is relatively constant, typically between 15-20 km/h (Kerner & Rehborn, 1996). The queue discharge rate of a jam wave is typically around 30 percent lower than the free flow capacity (Schönhof & Helbing, 2007). It can propagate for a long time period and distance, and resolves only with the traffic demand decreasing (Kerner & Rehborn, 1996). Jam waves are considered more difficult to reproduce than standing queues, because the bottleneck of jam waves moves dynamically rather than fixed at specific locations.

The common feature of standing queues and jam waves is that both of them are associated with a capacity drop. Thus, the aim of many freeway traffic control methods are often to prevent the activation of bottlenecks, so as to avoid the occurrence of the capacity drop and maximize the outflow of the bottleneck. Variable speed limit (VSL) and ramp metering are commonly used control measures on freeway networks to regulate the flow and delay the onset of congestion. Ramp metering regulates the inflow from the on-ramp to the mainstream of the freeway to delay or prevent the activation of the bottleneck. The performance of ramp metering control is limited by available storage space at on-ramps (Carlson et al., 2011). Normally VSLs do not suffer from the storage capacity restriction since the mainstream of freeways are wider and longer

than on-ramps. VSLs may be used to reduce the mainstream flow, so as to resolve traffic jams or delay or prevent the activation of bottlenecks. There are two mechanisms of VSLs in reducing mainstream flows that are explained as follows. First, when lower VSLs are applied on a freeway stretch at under-critical densities, there is a transition to a new traffic state, which serves the same flow at lower speed and high density. During this transition, the density increase leads to a temporary reduction of the mainstream flow (Carlson et al., 2010a). Second, sub-critical VSLs lead to a fundamental diagram that has lower capacity with lower VSLs.

There are various theories and algorithms to determine the appropriate values of the VSLs. SPECIALIST is an analytical approach for VSL control to resolve jam waves based on the shock wave theory (Hegyi et al., 2008). While the approach was successfully tested in practice (Hegyi & Hoogendoorn, 2010), its disadvantage is that due to its feed-forward structure, disturbances during activation of the VSLs are not handled. Later on, the algorithm has been extended to a feedback structure, which is known as COSCAL v2 (Mahajan et al., 2015), to overcome the deficiency of the SPECIALIST algorithm. Analytical approaches are usually efficient in computation and easy to be implemented. However, they are not easy to be adapted to new situations (e.g., changes in infrastructure). Traffic model predictive control (MPC) is another control approach which has attracted the attention of many researchers. It predicts the evolution of traffic dynamics and calculates the optimal control scheme for the time period in which the relevant traffic dynamics occurs. This feature enables the controller to take advantage of potentially larger future gains at a current (smaller) cost, so as to avoid a myopic control actions. The coordination of VSL by MPC to suppress jam waves has been investigated by (Hegyi et al., 2005b), where the design is based on a non-linear second-order model METANET (Kotsialos et al., 2002a). Another advantage of MPC is that various of control measures can be easily integrated into one control system. Consequently, VSL control is often integrated with other control measures such as ramp metering by MPC (Hegyi et al., 2005a; Carlson et al., 2010a; Roncoli et al., 2015b).

Optimal control formulations based on the discrete macroscopic second-order model, METANET, have become a popular choice since the model has been calibrated and validated with a reasonable accuracy. However, the non-linear and non-convex optimization formulation of METANET-based MPC might result in high computation load, especially if the optimization is solved by the standard SQP algorithm (Hegyi et al., 2005b). Researchers have managed to solve this problem by developing more efficient computational algorithms, such as the feasible-direction algorithm (Kotsialos et al., 1999), the parameterization algorithm (Zegeye et al., 2012; van de Weg et al., 2015). A noticeable problem for METANET-based MPC of VSL is that the solution of the optimization may depend on the selection of the initial guess trajectory, in the sense that if the initial guess is not appropriately chosen, the solution of the optimization might get stuck at a local minimum. In recent years, linear or quadratic optimizations for traffic control have attracted more and more attention. The origin of this type of

approach dates back to (Gazis & Potts, 1963), who suggested a store-and-forward modeling approach to traffic control problems. Following this line, Papageorgiou (1995) proposed a linear optimization approach for integrated traffic control based on a store-and-forward model. Later on, Lo (1999); Ziliaskopoulos (2000) embedded the cell transmission model into a linear optimization problem for dynamic traffic assignment purpose. Since the capacity drop was not considered, the only gain of the controller was coming from increasing the off-ramp flows. Recent research has incorporated the capacity drop into linear MPC for VSL and ramp metering control (Muralidharan & Horowitz, 2015; Roncoli et al., 2015b). The controllers have been demonstrated to be efficient in terms of computation speed and reducing total travel delay by the designed case studies. However, the type of traffic jam considered in their case studies is only the standing queue.

As far as we are concerned, there is no specific research in literature which aimed to resolve freeway jam waves through a linear MPC formulation. To fill this gap, in this chapter we develop a linear-quadratic MPC based on an extended discrete first-order traffic flow model. The extended model keeps the linear property of the classical discrete first-order model, meanwhile takes traffic flow features of jam waves propagation into consideration. The performance of the proposed controller is compared with a classical non-linear MPC and one recently proposed linear MPC in terms of resolving jam waves and the computation speed for a benchmark problem.

The rest of this chapter is set up as follows. Section 3.2 describes the mechanism of the extended model and the models for comparison. Section 3.3 elaborates on the formulation of the MPC approach. In Section 3.4 the presented MPC approach is compared with other MPC approaches for a benchmark problem, in terms of the capability of resolving jam waves, the resulting total travel time, and the required computation time. Section 3.5 presents the conclusions and some topics for future research.

3.2 Model description

In this section, the mechanism of several discrete first-order models in literature are briefly described. The classical cell transmission model is described in Section 3.2.1. Then some extended cell transmission models which have incorporated the capacity drop are summarized in Section 3.2.2. The new extended CTM and the VSL assumption are described in Section 3.2.3.

3.2.1 The cell-transmission model

The origin of discrete first-order models is the well-known cell transmission model (CTM) developed by (Daganzo, 1994), which is consistent with the hydrodynamic theory of LWR model (Lighthill & Whitham, 1955; Richards, 1956). The roadway

stretch is divided into cells, which are increasingly numbered from upstream to downstream, as shown in Fig. 3.1. In the following, cell i is referred to as an upstream cell, and cell $i + 1$ is interpreted as a downstream cell. Traffic states of the roadway stretch are represented by the densities of cells at discrete time steps indexed by k . The density of cell i at time step $k + 1$, which is represented as $\rho_i(k + 1)$ [veh/km], is updated as:

$$\rho_i(k + 1) = \rho_i(k) + \frac{T}{L_i}[f_{i-1}(k) - f_i(k)], \quad (3.1)$$

where, T is the duration of the discrete time steps. L_i is the length of cell i . f_{i-1} and f_i [veh/h] are the downstream boundary flows of cell $i - 1$ and cell i .

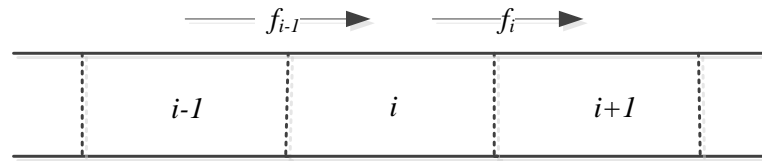


Figure 3.1: A freeway stretch divided into discrete cells.

Lebacque (1996) showed that CTM is the first-order discrete Godunov approximation of the LWR model. According to the Godunov scheme, the flow $f_i(k)$ between cell i and cell $i + 1$ is determined by the minimum of the sending flow (represented by $S_i(k)$) and the receiving flow (represented by $R_{i+1}(k)$), also known as the demand of sending cell and the supply of receiving cell. $f_i(k)$ is represented as,

$$f_i(k) = \min(S_i(k), R_{i+1}(k)), \quad (3.2)$$

The demand of sending cell, $S_i(k)$, is represented as:

$$S_i(k) = \min(\rho_i(k)v, c), \quad (3.3)$$

which states that the outflow of cell i is restricted by the vehicles present in the sending cell and the capacity of the sending cell. The supply of receiving cell, $R_{i+1}(k)$, is represented as:

$$R_{i+1}(k) = \min(c, \beta(\rho^J - \rho_{i+1}(k))) \quad (3.4)$$

which states that the inflow to cell $i + 1$ is restricted by capacity of the receiving cell and the amount of space in the receiving cell. The CTM assumes a triangular-shaped fundamental diagram (FD) for the above equations, which is characterized by the free flow speed v , the capacity c , the congestion wave speed β , and the jam density ρ^J . The FD and the demand and supply functions are depicted in Fig. 3.2. ρ^{cr} denotes the critical density, which separates the FD into free-flow branch and congestion branch.

3.2.2 Extended versions of the CTM

It is commonly known that the classical CTM does not take the capacity drop phenomenon into consideration. Thus, many researchers have presented extensions of

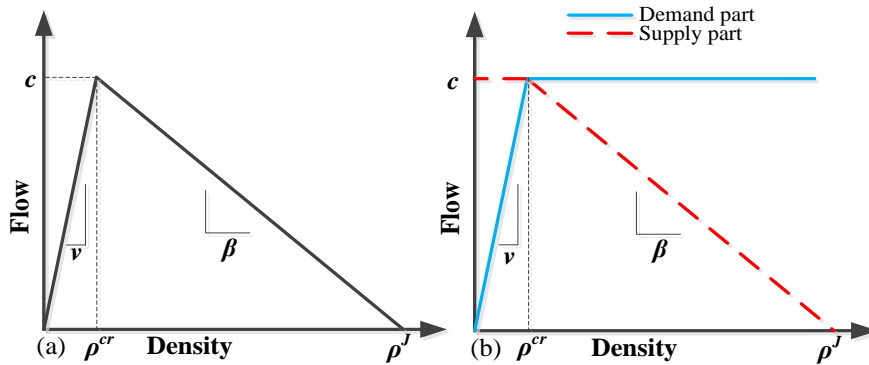


Figure 3.2: The triangular-shaped FD of the CTM (a), and the demand and the supply parts (b).

the CTM to incorporate the capacity drop. Most of those approaches try to incorporate the capacity drop by changing the demand function, which is first proposed by (Lebacque, 2003). Schreiter et al. (2010) assumed two constant capacity values for the cell transmission model, one for free-flow condition and the other for congested condition. Srivastava & Geroliminis (2013) extended the LWR model by defining a FD with two capacity values. A memory-based approach was proposed to choose the appropriate capacity value, so as to distinguish congested and uncongested states. However due to non-linearities, these models have not been used for model-based traffic control. Muralidharan (2012) incorporated the capacity drop into a link-node cell transmission model, which used a discontinuous fundamental diagram. The capacity drop was assumed to be triggered only at the bottleneck cell, and with a constant value. The model has been developed into a mixed-integer linear programming for network traffic optimization. However, since bottlenecks have to be predefined at fixed locations, this approach is not suitable for the jam wave type of congestion under which bottlenecks move dynamically. Roncoli et al. (2015a) proposed an optimization-oriented first-order model, which also represented the capacity drop by a reduced demand function. They assumed that the extent of the demand reduction of a cell has a linear relation with the vehicle density of the same cell. When the density of a cell is beyond the critical value, the demand decreases linearly with increasing density. Srivastava et al. (2014) used a similar assumption for urban traffic to capture the headway behavior. The empirical study by (Yuan et al., 2015) has shown that the magnitude of capacity drop has a negative relation with the speed in congestion. Thus, the state-dependent capacity drop assumption by (Roncoli et al., 2015a) is considered to be more reasonable than the assumptions with a constant extent of capacity drop. Han et al. (2015b) used a similar capacity drop assumption for lane drop bottlenecks. In that study, the demand reduction which represented the capacity drop, was assumed to take place at the downstream cell of the cell with over-critical density, rather than in the cell with over-critical itself.

The model from (Roncoli et al., 2015a) has considered both lateral flows and longitudinal flows. Stronger capacity drop maybe reproduced when there are lateral flows due

to lane changes or on-ramp traffic. In this chapter we do not take lane changing into consideration since the focus is on jam waves for which the lateral flows do not play an important role. In the following we only introduce their assumptions on longitudinal flows. In their model, the capacity drop was characterized by variable c^{jam} , which represented the flow discharge rate of a cell in jam state. As shown in Fig. 3.3 (a), the demand reduces gradually as the density increases at the over-critical part. The supply remains the same as the original cell transmission model, which is shown in Fig. 3.3 (b).

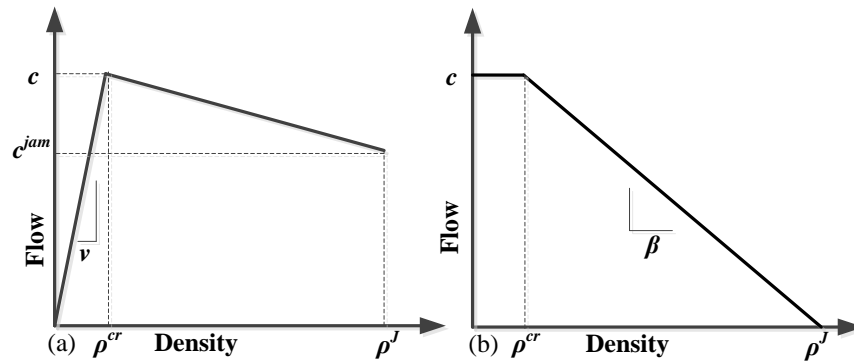


Figure 3.3: The demand (a) and supply (b) parts of the FD of the model from (Roncoli et al., 2015a).

It is found that the type of traffic jam considered by most of the above mentioned models is only the standing queue. In (Han et al., 2016a), it has been tested that most of the above mentioned models can not reproduce the propagation of jam waves accurately. Thus, we have proposed a new extension of the discrete first-order model to reproduce the propagation of jam waves more accurately, which will be briefly introduced in the next section.

3.2.3 The new extended CTM

As has been discussed in (Han et al., 2016a), the cell transmission model and many of its variants have difficulties in reproducing the propagation of jam waves accurately. One reason is that capacity drop has not been adequately incorporated. The capacity drop assumption that we propose is a linear reduction of the outflow of cell $i + 1$ as a function of the density of cell i in case the densities are over-critical. For simplicity, in this chapter we assume that all of the cells have same properties. However, note that it is possible to make the model parameters cell-dependent. For a general presentation of the model, readers are referred to (Han et al., 2016a). Without loss of generality, the index of cell number remains in the following equations. According on the capacity drop assumption, the maximum outflow of cell $i + 1$ can be represented as:

$$Q_{i+1}(k) = \min \left(c_{i+1}, c_{i+1} \left(1 - \alpha \frac{\rho_i(k) - \rho_i^{cr}}{\rho_i^J - \rho_i^{cr}} \right) \right), \quad (3.5)$$

where, $Q_{i+1}(k)$ is the maximum outflow of cell $i + 1$ at time step k . α denotes the maximum extent of the capacity drop. The capacity drop model is depicted by the sub-figures (a,c,e,g) of Fig. 3.5. In this way, the capacity drop can not only be reproduced at on-ramp bottlenecks, but also other infrastructural bottlenecks such as lane drops (Han et al., 2015b). When cell i is in an over-critical state and cell $i + 1$ is in an under-critical state, $Q_{i+1}(k)$ represents the flow discharge rate of cell $i + 1$, which is lower than the free-flow capacity c_{i+1} . Accordingly, the critical density of cell $i + 1$ under such a condition, represented by $\rho_{i+1}^{\text{cr-dch}}(k)$, is lower than ρ_i^{cr} because the free-flow speed does not change. $\rho_{i+1}^{\text{cr-dch}}(k)$ is calculated as:

$$\rho_{i+1}^{\text{cr-dch}}(k) = \frac{Q_{i+1}(k)}{v_{i+1}}. \quad (3.6)$$

Thus, the demand of the proposed model is represented as:

$$S_{i+1}(k) = \begin{cases} \rho_{i+1}(k)v_{i+1} & \text{if } \rho_{i+1}(k) < \rho_{i+1}^{\text{cr-dch}}(k) \\ Q_{i+1}(k) & \text{if } \rho_{i+1}(k) \geq \rho_{i+1}^{\text{cr-dch}}(k). \end{cases} \quad (3.7)$$

The other aspect that plays a role in the proper reproduction of jam wave propagation is the supply function when a cell is in a discharging state. According to the cell transmission model, the supply function (3.4) is dominating in determining the flow if both the sending cell and the receiving cell have over-critical densities. In other words, the boundary flow between two over-critical cells only depends on the average density of the receiving cell. However in practice, traffic states at the over-critical part can be inhomogeneous if a cell has a long length (typically several hundred meters when applied to traffic control). For example, if the jam head is in the middle of cell $i + 1$, then cell $i + 1$ is in such a state that the back part is congested while the front part is not, which is shown in Fig. 3.4. We define such a state as the discharging state. Note that the discharging state may be avoided by making the cell size small. However, the proposed model is applied to a linear optimization problem for traffic control purpose. If a smaller cell length is used, the linear optimization framework results in higher number of cells, control variables, and constraints. This increases the computation time and may render the control problem too complex to be solved in real-time. From this point of view, our way of modeling gives good trade-off between modeling accuracy and controller performance (in terms of computation speed).

Suppose traffic states at the downstream of cell $i + 1$ are free-flow states. The discharging state of cell $i + 1$ is captured by the condition $\rho_i(k) > \rho_{i+1}(k) > \rho_{i+1}^{\text{cr-dch}}(k)$ in the proposed model. The inequality is explained as follows. According to (5), if $\rho_i(k) \leq \rho_i^{\text{cr}}$, then $Q_{i+1}(k) = c_{i+1}$, and if $\rho_i(k) > \rho_i^{\text{cr}}$, then $Q_{i+1}(k) < c_{i+1}$. According to (6), if $Q_{i+1}(k) = c_{i+1}$, then $\rho_{i+1}^{\text{cr-dch}}(k) = \rho_i^{\text{cr}}$; if $Q_{i+1}(k) < c_{i+1}$, then $\rho_{i+1}^{\text{cr-dch}}(k) < \rho_i^{\text{cr}}$. Now, given $\rho_i(k) > \rho_{i+1}^{\text{cr-dch}}(k)$, let's consider two cases: first when $\rho_{i+1}^{\text{cr-dch}}(k) < \rho_i^{\text{cr}}$, and second when $\rho_{i+1}^{\text{cr-dch}}(k) = \rho_i^{\text{cr}}$. If $\rho_{i+1}^{\text{cr-dch}}(k) < \rho_i^{\text{cr}}$, then $Q_{i+1}(k) < c_{i+1}$ and $\rho_i(k) > \rho_i^{\text{cr}}$. If $\rho_{i+1}^{\text{cr-dch}}(k) = \rho_i^{\text{cr}}$, then $\rho_i(k) < \rho_i^{\text{cr}}$. Thus if $\rho_i(k) > \rho_{i+1}^{\text{cr-dch}}(k)$ is given, then $\rho_i(k) > \rho_i^{\text{cr}}$ and thus cell i is in an over-critical state. If the density of cell $i + 1$ is lower than the

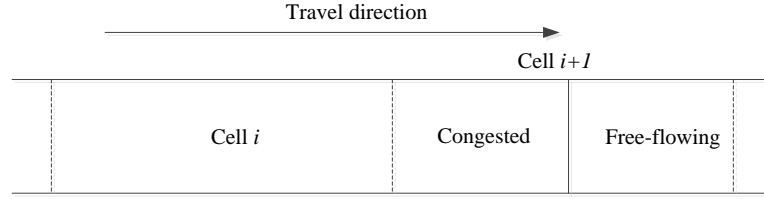


Figure 3.4: An example to explain the discharging state. Cell $i + 1$ is in the discharging state.

density of cell i , but higher than the critical discharging density, $\rho_{i+1}^{\text{cr-dch}}(k)$, then cell $i + 1$ should be in a mixed state with a congested upstream part and free-flow downstream part. According to the definition in Fig. 3.4, cell $i + 1$ is in the discharging state. Under the discharging state, it is assumed that the supply of cell $i + 1$ is overestimated by (3.4) because the average density $\rho_{i+1}(k)$ is lower than the density upstream part of cell $i + 1$. Empirical analysis in Han et al. (2016a) has confirmed that this assumption is reasonable. Thus, we assume two supply functions $R_{i+1}^{(1)}(k)$ and $R_{i+1}^{(2)}(k)$, which apply for the discharging state and the non-discharging state respectively when cell $i + 1$ is over-critical. The non-discharging supply function is represented as,

$$R_{i+1}^{(1)}(k) = \beta_{i+1}^{(1)} (\rho_{i+1}^J - \rho_{i+1}(k)), \quad \text{if } \rho_i(k) \leq \rho_{i+1}(k) < \rho_{i+1}^J, \quad (3.8)$$

where, $\beta_{i+1}^{(1)}$ represents the slope of the line which connects the critical state and the jam state. It expresses the congested states of traffic in a jam or getting in a jam. When cell $i + 1$ is in the discharging state, the supply from (8) would overestimate the real value. It was also shown in (Han et al., 2016a) that the extent of the supply overestimation has a positive relation with the density difference between the sending cell and the receiving cell. To keep the linear property of the model, we assume a linear relation between the density difference $\rho_i(k) - \rho_{i+1}(k)$ and the supply reduction $R_{i+1}^{(1)}(k) - R_{i+1}^{(2)}(k)$. The discharging supply function is represented as,

$$R_{i+1}^{(2)}(k) = R_{i+1}^{(1)}(k) - (\beta_{i+1}^{(1)} - \beta_{i+1}^{(2)}) \cdot (\rho_i(k) - \rho_{i+1}(k)), \quad \text{if } \rho_{i+1}^{\text{cr-dch}}(k) \leq \rho_{i+1}(k) < \rho_i(k). \quad (3.9)$$

where, $\beta_{i+1}^{(2)}$ represents the slope of the line which connects the state of a discharging cells and the state of its upstream cell. It expresses the congested states of traffic leaving a jam wave, which in free-flow will lead to the queue discharge rate that is lower than the free-flow capacity. When traffic of cell $i + 1$ is getting in a jam, the traffic state lies on the line with the slope $\beta_{i+1}^{(1)}$. When traffic in cell $i + 1$ is leaving a jam, the traffic state lies on the line with the slope $\beta_{i+1}^{(2)}$. Thus, the line of $\beta_{i+1}^{(1)}$ represents the deceleration of traffic whereas $\beta_{i+1}^{(2)}$ represents the acceleration of traffic. In real traffic, deceleration is usually faster than acceleration, thus $\beta_{i+1}^{(1)}$ is higher than

$\beta_{i+1}^{(2)}$. The overall supply function of cell $i + 1$, $R_{i+1}(k)$ is formulated as:

$$R_{i+1}(k) = \begin{cases} Q_{i+1}(k) & \text{if } \rho_{i+1}(k) < \rho_{i+1}^{\text{cr-dch}}(k) \\ R_{i+1}^{(2)}(k) & \text{if } \rho_{i+1}^{\text{cr-dch}}(k) \leq \rho_{i+1}(k) < \rho_i(k) \\ R_{i+1}^{(1)}(k) & \text{if } \rho_i(k) \leq \rho_{i+1}(k) \leq \rho_{i+1}^{\text{J}}. \end{cases} \quad (3.10)$$

The above equations can be further explained by considering the following three conditions:

1. $\rho_i(k) \leq \rho_i^{\text{cr}}$. Under this condition, neither capacity drop nor supply reduction happens for cell $i + 1$. Thus $Q_{i+1}(k) = c_{i+1}(k)$, and the demand and supply of cell $i + 1$ are the same as which from the CTM, which are shown as sub-figure (b) of Fig. 3.5.
2. $\rho_i(k) = \rho_i^{\text{J}}$. Cell i is in the jam state. Thus, the demand of cell $i + 1$ is represented by the blue line of sub-figure (d) in Fig. 3.5, which has a maximum flow $Q_{i+1}(k) = c_{i+1}(k)(1 - \alpha)$ and a critical density $\rho_{i+1}^{\text{cr-dch}}(k) = c_{i+1}(k)(1 - \alpha)/v_{i+1}$. Under this condition, ρ_i can not be smaller than ρ_{i+1} , so the supply of cell $i + 1$ at the over-critical part is represented by $R_{i+1}^{(2)}$, which equals to $\beta_{i+1}^{(2)}(\rho_i^{\text{J}} - \rho_{i+1}(k))$ according to the third equation of (3.10). The supply of cell $i + 1$ at the under-critical part equals to F_{i+1} , and the overall supply of cell $i + 1$ is represented by the red line of sub-figure (d) in Fig. 3.5.
3. $\rho_i^{\text{cr}} < \rho_i(k) < \rho_i^{\text{J}}$. The demand of cell $i + 1$ is calculated by (3.6) and (3.7), and represented by the blue line in Fig. 3.5 (f). Under this condition, the supply at the over-critical part is represented by $R_{i+1}^{(1)}(k)$ when $\rho_i(k) \leq \rho_{i+1}(k)$, and is represented by $R_{i+1}^{(2)}(k)$ when $\rho_i(k) > \rho_{i+1}(k)$. The line of $R_{i+1}^{(1)}(k)$ and $R_{i+1}^{(2)}(k)$ intersect at $\rho_i(k) = \rho_{i+1}(k)$. The supply of cell $i + 1$ at the under-critical part equals to $Q_{i+1}(k)$, and the overall supply of cell $i + 1$ is represented by the red line in Fig. 3.5 (f).

Variable speed limits (VSLs) are used as the control measure in this chapter. Thus, the effect of VSLs have to be taken into account by the model. The influence of VSLs on the shape of FD was investigated by (Papageorgiou et al., 2008). Some findings of the reported investigations can be summarized as: (i) Speed limits decrease the slope of the flow-occupancy diagram at under-critical occupancies. (ii) The VSL-affected flow-occupancy curve crosses the no-VSL curve, shifting the critical occupancy to higher values in the flow-occupancy diagram. (iii) There is no conclusive result whether VSLs will increase the flow capacity or not, since a slight increase is only visible at some locations. Since the effect of VSLs on aggregated traffic flow features is complex, simplified VSL model assumptions were often made. Hegyi et al. (2005a) assumed the VSLs impact was to replace the left part of the flow-occupancy curve by a straight line with slope corresponding to the displayed VSL. Carlson et al. (2010a) incorporated a

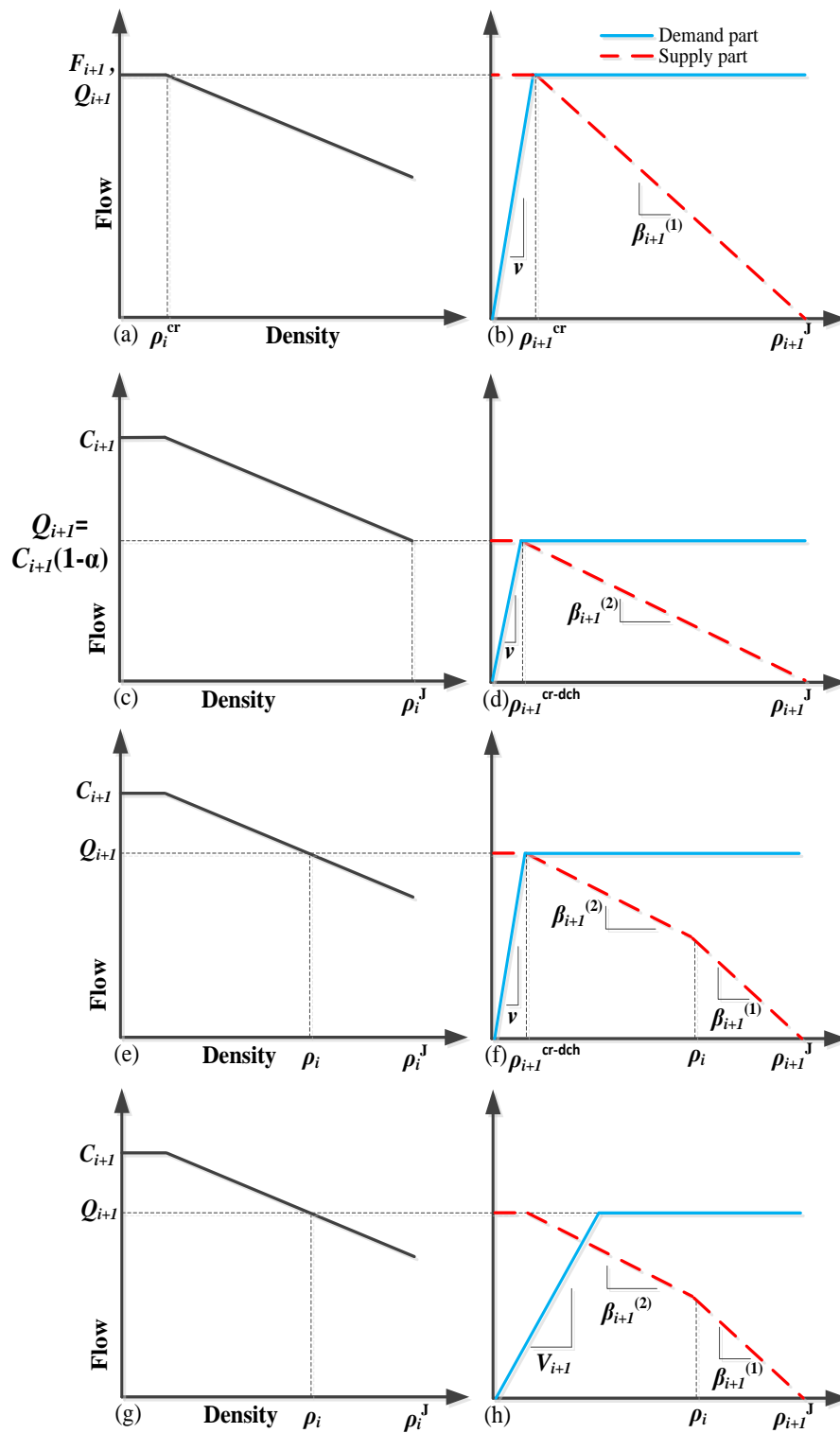


Figure 3.5: The left column, figures (a,c,e,g) show the relation between $\rho_i(k)$ and $Q_{i+1}(k)$, which represents the capacity drop assumption. The right column shows the demand and supply of cell $i + 1$, (b) for condition $\rho_i(k) \leq \rho_i^{cr}$, (d) for condition $\rho_i(k) = \rho_i^J$, (f) for condition $\rho_i^{cr} < \rho_i(k) < \rho_i^J$, (h) for condition $\rho_i^{cr} < \rho_i(k) < \rho_i^J$ and cell $i + 1$ is under VSL control with the value $V_{i+1}(k)$.

VSL model into the METANET model based on the empirical findings of (Papageorgiou et al., 2008).

For VSL models incorporated in discrete first-order models, the approach is often that the demand function is modified while the supply function is left unchanged (Muralidharan, 2012; Hadiuzzaman & Qiu, 2013). The VSL value in this chapter is, however, derived from the optimization outcome. Specifically, the optimization problem as formulated in Section 3.3 will result in the optimal flows over the cell boundaries. When these optimal flows are lower than the flows that would follow from the model, then reduced VSLs are needed to enforce these reduced flows. The value of the optimal VSLs can be calculated from the optimal flows and the corresponding densities.

In other applications of this model, when the VSLs are already known, the traffic evolution can be calculated based on a VSL-based demand function. As shown in Fig. 3.5 (g-h), if a VSL signal is applied to cell $i + 1$ with the value $V_{i+1}(k)$, the demand of the cell is the minimum of $V_{i+1}(k)\rho_{i+1}(k)$ and $Q_{i+1}(k)$, which is represented by the blue line in Fig. 3.5 (h). The new extended CTM incorporated with VSLs can be used for model calibration and validation.

3.3 The optimal control formulation

This section shows the MPC formulation based on the model proposed in the previous section. MPC-based controllers predict traffic flow evolutions by traffic flow models, and determine the optimal control signals based on current system states. The length of prediction horizon is represented by N_p , and the length of control horizon is represented as N_c . The duration of a control time step is represented as T_c . After optimization the first sample of the optimal control signals is applied to the process. The remaining part of the control signal is recalculated in the rolling horizon scheme. For a detailed description of MPC approach, readers are referred to (Camacho & Bordons, 2012). MPC-based controllers are formulated into linear or non-linear optimizations, depending on the model properties. The MPC approaches that are formulated into linear optimization problems can be solved very efficiently (Roncoli et al., 2015b). The proposed model has a linear property, which will be described in Section 3.3.1. The MPC controller in this chapter is formulated as a linear-quadratic optimization problem. The objective function of the proposed MPC is presented in Section 3.3.2.

3.3.1 The linear property of the model

In literature, extended cell transmission models have been reformulated as linear optimization problems for traffic control (Gomes & Horowitz, 2006; Muralidharan & Horowitz, 2015; Roncoli et al., 2015b). The most important feature of this model type

is the linear property. Traffic states can be predicted by the following linear equation,

$$\rho_i(k+n) = \rho_i(k) + \sum_{j=k}^{k+n-1} \frac{T}{L_i} [f_{i-1}(j) - f_i(j)], \quad (3.11)$$

where, flows between cells, $f_i(j)$, $j = k : k+n-1$, are decision variables of the model. The calculation of $f_i(j)$ are delegated to the controller, and only upper bounds are specified to $f_i(j)$ as follows:

$$f_i(j) \leq \rho_i(j)v_i(j) \quad (3.12)$$

$$f_i(j) \leq c_i \left(1 - \alpha \frac{\rho_{i-1}(j) - \rho_{i-1}^{cr}}{\rho_{i-1}^J - \rho_{i-1}^{cr}} \right) \quad (3.13)$$

$$f_i(j) \leq c_{i+1} \left(1 - \alpha \frac{\rho_i(j) - \rho_i^{cr}}{\rho_i^J - \rho_i^{cr}} \right) \quad (3.14)$$

$$f_i(j) \leq \beta_{i+1}^{(1)} (\rho_{i+1}^J - \rho_{i+1}(j)) \quad (3.15)$$

$$f_i(j) \leq \beta_{i+1}^{(1)} (\rho_i^J - \rho_i(j)) + \beta_{i+1}^{(2)} (\rho_i(j) - \rho_{i+1}(j)). \quad (3.16)$$

The modeling approach for flows is based on the demand and supply functions in Equation (3.7) and (3.10). More specifically, equations (3.12) and (3.13) correspond to two terms of the demand function (3.7) and equations (3.14), (3.15), (3.16) correspond to three terms of the supply function (3.10). Thus, the lines of the demand and the supply function in Fig. 3.5 (b), (d), (f) represent the upper bounds of $f_i(j)$, corresponding to different values of $\rho_i(j)$. In case no control actions are applied, the actual flow equals the minimum between the demand and supply flows and the solution of $f_i(j)$ lies in the lines of the upper bounds. If the solution of $f_i(j)$ is lower than the minimum of the demand and supply flows, then cell i at time step j is under VSL control. The displayed VSL value, $V_i(j)$, is calculated by:

$$V_i(j) = \frac{f_i(j)}{\rho_i(j)}. \quad (3.17)$$

The above equations show the linear property of the model, which enables the formulation of a linear optimization problem. Thus the state variables are calculated in one step for the whole prediction horizon. In the formulation of non-linear optimizations, the state variables have to be calculated sequentially and iteratively, which might result in a higher computation load.

3.3.2 Objective function

The objective function of a controller needs to reflect the level of congestion in the freeway. Total travel time (TTT) captures the aggregate performance of traffic conditions on all users in the freeway. Minimizing TTT is equivalent to minimize the total number of vehicles in the entire network. The number of vehicles in the network is represented by a vector $X(k)$, where $X(k) = [L_1\rho_1(k), L_2\rho_2(k), \dots, L_I\rho_I(k)]'$. The evolution of $X(k)$ is represented by a linear matrix formulation:

$$X(k+n) = X(k) + \sum_{j=k}^{k+n-1} B * U(j) + \sum_{j=k}^{k+n-1} d(j), \quad (3.18)$$

where, $U(j) = [f_1(j), f_2(j), \dots, f_I(j)]'$ is a decision vector. B is a matrix that contains the freeway structure information. The non-zero elements in matrix B are the duration of the discrete time step. $d(j)$ is the traffic demand vector. To approximate the minimization of the TTT, the objective of the proposed controller is to minimize a quadratic function of the number of vehicles in the network, which is shown as:

$$\min \sum_{j=k}^{k+N_p-1} [X(j)'PX(j) + GU(j)], \quad (3.19)$$

where, P is a matrix with all 1 elements, so the term $X(j)'PX(j)$ aims to minimize the quadratic function of the number of vehicles in the network in the prediction horizon. G is a vector of the cell lengths multiplied by a small negative value. The term $GU(j)$ aims to maximize the flows, so as to address the so-called holding back problem. Holding back may occur for a congested state, where the VSLs cannot have an influence. The second term in (3.19) (which was initially suggested in (Papageorgiou, 1995)) mitigates this undesirable occurrence by rewarding the solutions where the flows take higher values. Note that this second term in (3.19) has a physical meaning, as it is proportional to the total travel distance (TTD). Thus, (3.19) may be perceived as a weighted combination of TTT minimization and TTD maximization. A more detailed explanation of this objective function is in (Le et al., 2013). In that paper, the linear quadratic objective function is applied to an urban signal control problem. The objective function (3.19) under the constraints in Section 3.3.1 can be solved by an appropriate solver (e.g., CPLEX). The outputs of the controller are optimal flows between cells.

3.3.3 The minimum VSL constraint

Typical real-world speed limit systems have some pre-defined minimum speed that can be displayed. The optimization variables of the proposed controller are flows, thus the minimum VSL constraints cannot be applied directly, instead minimum flow constraints are applied to achieve the same purpose. However, minimum flow constraints might cause infeasibility due to conflicting constraints. For example, the flows

prediction made at time step k is denoted by $f_i(j|k)$ (the same meaning as $f_i(j)$ in (3.12-3.16)). If the minimum flow constraints restrict $f_i(j|k) \geq A_i(j|k)$, but the demand and supply constraints restrict $f_i(j|k) \leq B_i(j|k)$, then the optimization problem will not have any solution if $A_i(j|k) > B_i(j|k)$.

To overcome this problem, here we introduce a method to apply the minimum flow constraints based on a forward prediction. The basic idea is to find possibly conflicting flows by predicting future traffic states, then exclude them from the minimum flow constraints. Therefore, before running the optimization at each control step, the traffic states are predicted by the prediction model of the controller. The flows from the forward prediction are denoted as $\hat{f}_i(j|k)$. The following equation is used as the minimum flow constraints for the optimization:

$$f_i(j|k) \geq A_i(j|k), \text{ if } f_i(j|k) \geq A_i(j|k) \quad (3.20a)$$

$$f_i(j|k) \geq 0, \text{ if } f_i(j|k) < A_i(j|k). \quad (3.20b)$$

The values of $A_i(j|k)$ are set based on the predicted density of every cell at every time step. For any cell at which VSLs can be applied, $A_i(j|k)$ are set to the value that $A_i(j|k) = \text{VSL}_{min} \times \rho_i(j|k)$. It can be guaranteed that the optimization problem with the above constraint has a solution, because $\hat{f}_i(j|k)$ is a feasible solution of the optimization problem with the above constraints. Note, however, that in this way, it cannot be ensured that every $f_i(j|k)$ under VSL control is restricted by the minimum VSL constraints. For flows in the prediction that are lower than $A_i(j|k)$, the minimum flow constraint cannot be guaranteed. This can be perceived as a shortcoming of linear optimizations. But still, the proposed method provides a way to realize the minimum flow constraints to some extent.

3.4 Benchmark problem

A benchmark problem is set up to test the performance of the proposed controller. A classical non-linear MPC and one recently proposed linear MPC, are compared with the proposed MPC in terms of resolving jam waves and the computation speed. The benchmark network we choose is a simple homogeneous freeway stretch, which is sufficient to test the relevant contributions we present in this chapter. The 6km freeway stretch contains three lanes, and it is divided into 20 cells, which is depicted in Fig. 3.6. Note that in practice, deceleration areas are needed for VSLs for safety reasons. Deceleration areas for VSLs are not considered in this chapter because the transition period is much shorter than the normal control period. For example, the SPECIALIST theory also did not consider deceleration areas (Hegyi et al., 2008). However in practice, the SPECIALIST field test imposed a very short deceleration period for safety. The field test results show that the deceleration did not have much effect (Hegyi & Hoogenboom, 2010). To reproduce a jam wave, the second-order model METANET is used as the process model to represent the real world. The process of reproducing the jam

wave is presented in Section 3.4.1. The prediction models of the proposed controller and the linear MPC for comparison are calibrated with the simulation results of the METANET. The calibration of the models' parameters and the setup of the controllers are introduced in Section 3.4.2. The performance of the controllers are shown in Section 3.4.3.

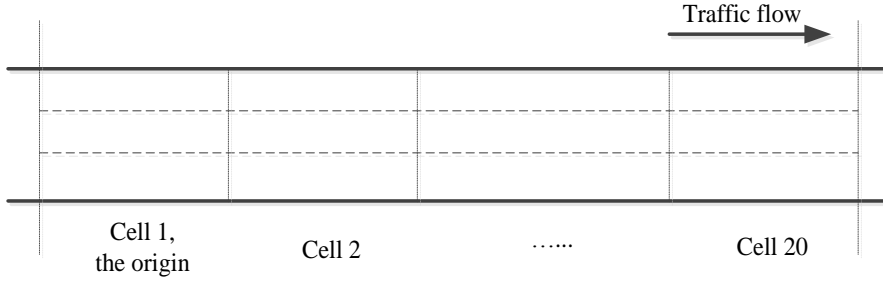


Figure 3.6: A graphical representation of the freeway stretch for the benchmark problem.

3.4.1 Reproducing of a jam wave

The second-order macroscopic traffic flow model METANET, which can reproduce capacity drop and the propagation of jam waves with a reasonable accuracy, is used as the process model to represent the real world. The METANET model is briefly introduced in the following.

In the METANET model, the freeway network is divided into segments which have a uniform geometry structure. Each segment i is characterized by the traffic density $\rho_i(k)$ [veh/lane/km], the mean speed $v_i(k)$ [km/h], and the outflow $q_i(k)$ [veh/h], where k is the index of the discrete time step. The following equations describe the evolution of the freeway stretch over time. The outflow of each segment equals to the density times the mean speed and the number of lanes of that segment (represented by λ_i):

$$q_i(k) = \rho_i(k)v_i(k)\lambda_i, \quad (3.21)$$

The density of a segment follows the vehicle conservation law, which is represented as:

$$\rho_i(k+1) = \rho_i(k) + \frac{T}{L_i\lambda_i} (q_{i-1}(k) - q_i(k)), \quad (3.22)$$

The mean speed of segment i at time step $k+1$ depends on the mean speed at time step k , the speed of the inflow of vehicles, and the density downstream, which is represented as:

$$v_i(k+1) = v_i(k) + \frac{T}{\tau} (V(\rho_i(k)) - v_i(k)) + \frac{T}{L_i} v_i(k) (v_{i-1}(k) - v_i(k)) - \frac{\vartheta T}{\tau L_i} \frac{\rho_i(k+1) - \rho_i(k)}{\rho_i(k) + \kappa}, \quad (3.23)$$

where, τ , ϑ , κ are model parameters. $V(\rho_i(k))$ denotes the desired speed that drivers try to achieve, which is:

$$V(\rho_i(k)) = v_{\text{free},i} \cdot \exp\left[-\frac{1}{a_m} \left(\frac{\rho_i(k)}{\rho_{\text{cr},i}}\right)^{a_m}\right], \quad (3.24)$$

where, a_m is a model parameter. $v_{\text{free},i}$ represents the free flow speed of segment i , and $\rho_{\text{cr},i}$ denotes the critical density at which the flow is maximal. The effect of VSL follows the assumption of (Hegyi et al., 2005a), that the desired speed is the minimum of (3.24) and the speed limit displayed on variable message sign:

$$V(\rho_i(k)) = \min\left(VSL_i(k), v_{\text{free},i} \cdot \exp\left[-\frac{1}{a_m} \left(\frac{\rho_i(k)}{\rho_{\text{cr},i}}\right)^{a_m}\right]\right). \quad (3.25)$$

The first segment of the freeway stretch is perceived as the origin, and modeled by simple queue model. The number of vehicles in the first segment is calculated as:

$$\omega_1(k+1) = \omega_1(k) + T(d(k) - q_1(k)), \quad (3.26)$$

where, $d(k)$ is the demand of the origin. $q_1(k)$ is determined by:

$$q_1(k) = \min\left[d(k) + \frac{\omega_1(k)}{T}, \lambda_1 V(\rho_{\text{cr},1}) \rho_{\text{cr},1}\right], \quad (3.27)$$

where, the first term is the available traffic in time period k . The second term is the capacity flow. For the upstream boundary condition, the speed of the origin segment is assumed to be the same as the second segment, so $v_1(k) = v_2(k)$. For the downstream boundary condition, the density downstream of the freeway stretch is defined according to the simulation scenario.

The parameters of METANET model are set up as follows. In (3.22), $L_i=0.3$ km, $\lambda_i=3$. $T=5$ s. In (3.23), τ is set to 18 s, κ is set to 40 veh/km/lane, and ϑ is set to 30 km^2/h . These parameters are referred to (Kotsialos et al., 1999). In (3.24), $\rho_{\text{cr}}=27.6$ veh/km/lane, $a_m=2.5$, $v_{\text{free}}=108$ km/h. These parameters are selected based on the experience of a reasonable shape of the fundamental diagram. The input of the system are the traffic demand at the origin and the density downstream of the freeway stretch. Traffic demand profile is shown in Fig. 3.7 (a). The downstream density equals to the critical value constantly, except for a pulse from time step 380 to 400 which is used to trigger the jam wave. The pulse is chosen large enough to cause an upstream-propagating jam wave. The downstream density profile is shown in Fig. 3.7 (b).

The entire simulation process contains 1440 time steps, which correspond to 2 hours. Simulation results by the METANET model are shown in Fig. 3.8. The jam wave propagates from the downstream end to the upstream end in about 200 time steps. The flow contour plot shows a conspicuous capacity drop, that the outflow of the jam (the yellow area) is much lower than the flow after the jam wave is resolved (the white area).

As presented, three model predictive controllers are compared in this benchmark problem. To tune the parameters of N_p and N_c , we follow the rule that N_p should be larger

than the typical travel time from the controlled cells to the exit of the network, and N_c should have a good trade-off between the computational effort and the performance. The prediction horizon and control horizon are set to 10 minutes, and the duration of control step T_c is set to 10 seconds. For a fair comparison the parameters are set to same values for three controllers. The controllers are assumed to be activated at time step 420 when the jam wave has formed and deactivated if the density of every cell in the freeway stretch is lower than the critical value so there will be no risk to form a new jam. In the following we briefly describe the mechanisms of three controllers.

1. Controller 1, the METANET-based MPC. The optimal control formulation we refer to (Hegyi et al., 2005b), thus both the process model and the prediction model in controller 1 are the METANET model. The optimization problem is solved by MATLAB implementation of the SQP algorithm (fmincon). The lowest VSL value is set to 35 km/h.
2. Controller 2, a model predictive controller, which uses (3.19) as the objective function and the model from (Roncoli et al., 2015a) (referred to as model R in the following) as the prediction model. Model R has taken lane-changing into consideration. Since lateral flows are not considered in this chapter, we only use their longitudinal flow model which is introduced in Section 2.2. $A_i(j)$ are set to values based on the predicted density of every cell at every time step.
3. Controller 3, a model predictive controller, which uses (3.19) as the objective function and the new extended CTM as the prediction model. $A_i(j)$ are set to values based on the predicted density of every cell at every time step. The linear quadratic formulation of controllers 2 and 3 are solved by CPLEX from MATLAB toolbox.

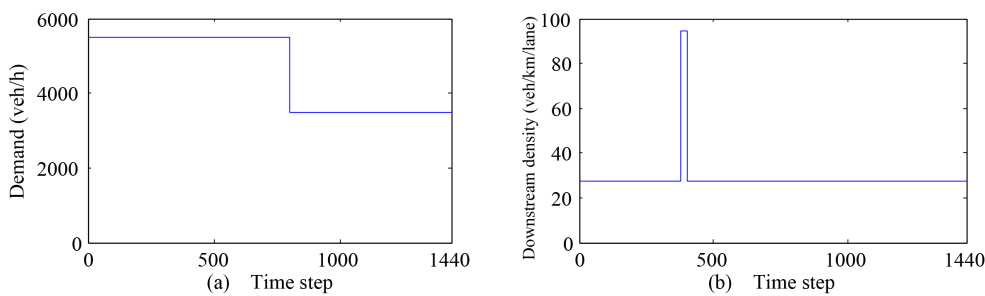


Figure 3.7: The traffic demand at the origin (a), and the density downstream of the freeway stretch (b).

3.4.2 The set up of controllers

The model parameters of controllers 2 and 3 are calibrated with the simulation results from the METANET model. The calibration is performed based on an optimization-based method to best match the model parameters with the simulation results of the

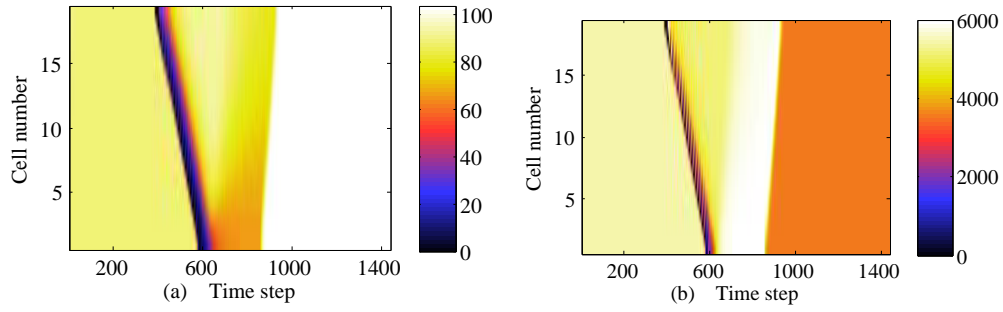


Figure 3.8: The speed (km/h) contour plot (a), and the flow (veh/h) contour plot (b) of the simulation without VSL control.

METANET model. The root mean square error (RMSE) of density is chosen as the performance indicator of the calibration, which is shown as the following equation:

$$\min \frac{1}{(I-1) \times (N-k+1)} \sqrt{\sum_{i=2}^I \sum_{j=k}^N (\rho_i(j) - \hat{\rho}_i(j))^2}, \quad (3.28)$$

where, $\hat{\rho}_i(j)$ is the density of cell i at time step j from the METANET simulation, and $\rho_i(j)$ is the simulated density from the prediction models of controllers 2 and 3. Data are extracted from time step 401 to 700, thus $N = 700$, $k = 401$. The optimization problem is solved by MATLAB implementation of the SQP algorithm (fmincon). Multi-starting points have been tried to avoid the optimization results getting stuck in a local optimum. The calibration results of controllers 2 and 3 are shown in Table 3.1. The parameter which characterizes the extent of capacity drop in the model of model R is c^{jam} , while in the new extended CTM it is α . Since they represent the same physical meaning, for consistency both of them are represented by α in the table. If α is known, $\beta^{(2)}$ can be calculated accordingly. The values of parameters of controller 3 are also shown in Fig. 3.9

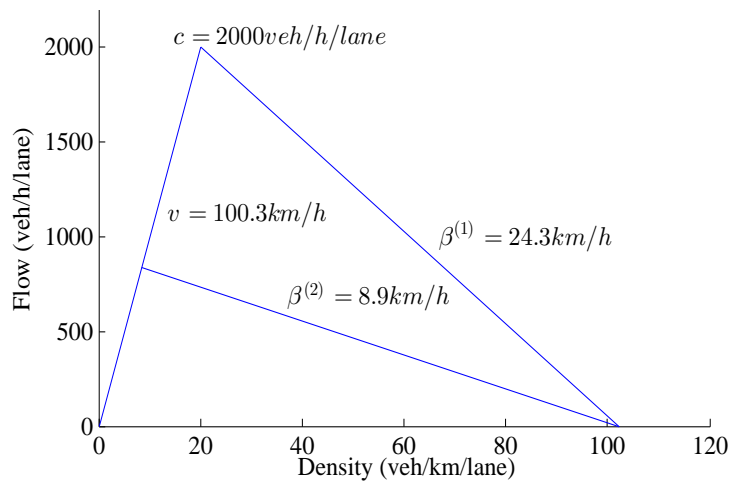


Figure 3.9: The calibrated fundamental diagram of controller 3

Table 3.1: The calibration results of controllers 2 and 3

	v (km/h)	α	c (veh/h)	$\beta^{(1)}$ (km/h)	RMSE
Controller 2	100.32	0.58	6000	24.33	10.30
Controller 3	100.75	0.79	6000	23.9	8.41

The RMSE shows that the new extended CTM has a more accurate simulation result than model R. Fig. 3.10 shows the density contour plots which are simulated by the predictions models. Both the new extended CTM and model R show dispersions of the jam wave, but qualitatively speaking, the dispersion of the new extended CTM is less.

Fig. 3.11 shows the outflow of the freeway stretch simulated by those three models. The METANET model reproduces a substantial capacity drop, where the outflow of the network stays around 5200 veh/h till time step 650. The outflow reproduced by model R gradually increases, and recovers to the free-flow capacity value at time step around 550. The outflow reproduced by the new extended CTM stays in the range of 4700 to 5500 veh/h, from time step 450 to 650. The oscillation of flow happen due to the numerical issue, which can be briefly explained as follows: Suppose cell i is in a discharging state, and cell $i + 1$ is in free-flow state. Under this circumstance, the outflow of the network is dominated by the outflow of cell $i + 1$, f_{i+1} , because cell $i + 1$ is the most upstream cell of all downstream free-flow cells. Since cell i is in the discharging state, the vehicle density ρ_i is decreasing and according to (4.1), f_{i+1} is also decreasing. That is the reason for the increasing of the outflow. When ρ_i decreases to an under-critical value, cell $i - 1$ becomes the discharging cell and the outflow of the network is dominated by f_i . That's the reason for the sudden decreasing of the outflow. This process repeats as the discharging cell moves upstream, causing the oscillation of flow. The numerical issue can be perceived as a drawback of the model, nevertheless, the new extended CTM still reproduces capacity drop during jam waves propagation, which is the main target traffic control aims to avoid.

3.4.3 Control performance

The simulation results of three controllers are shown in Fig. 3.12. It can be seen that the jam wave is resolved by controllers 1 and 3. VSL control signals are applied to the upstream of the jam area to limit the inflow to the jam, creating such a condition that the outflow of the jam is higher than the inflow, so that the jam wave will be resolved. Fig. 3.13 shows the VSL signals generated by the controllers. The white color indicates areas where no VSL signal was applied. At the beginning of traffic control, the VSL control area generated by controller 1 is longer than that by controller 3, so the jam wave is resolved faster by controller 1. Controller 2 narrows the width of the jam, but it is unable to resolve it. In order to exclude the possibility that the length of the simulated stretch was causing the inability of controller 2 to resolve the jam, we have tried a scenario which the freeway stretch was extended to 40 cells. The result was that the jam wave was still not resolved by the VSL control even when the length

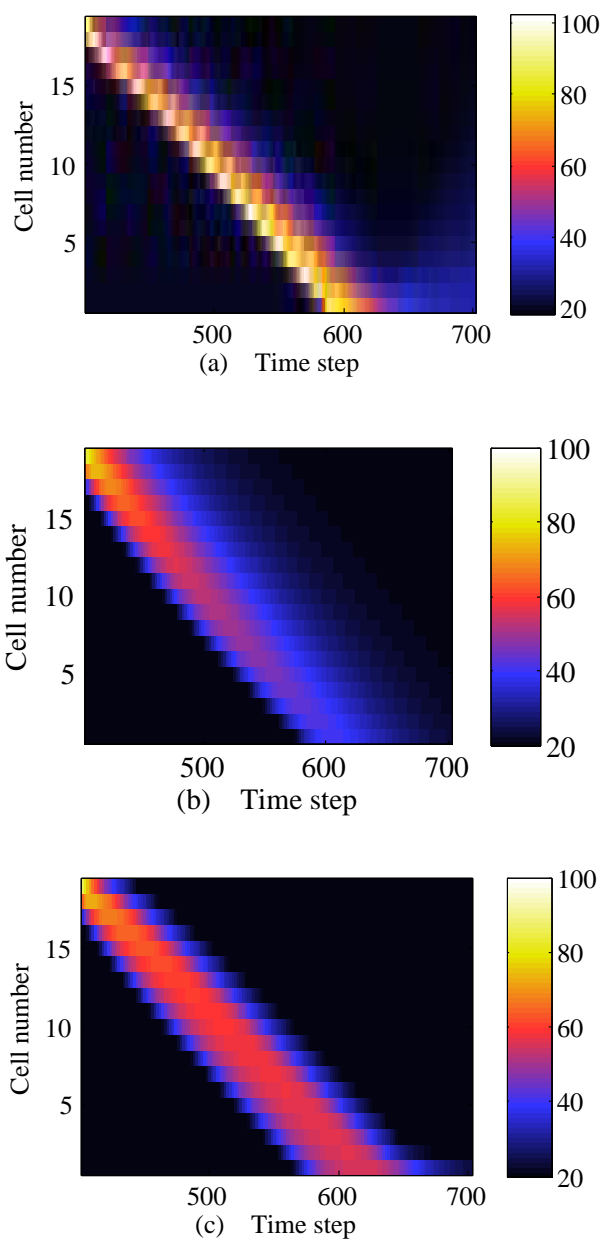


Figure 3.10: The density (veh/km/lane) contour plots (without VSL control) from (a) the METANET model, (b) Roncoli's model, and (c) the new extended CTM.

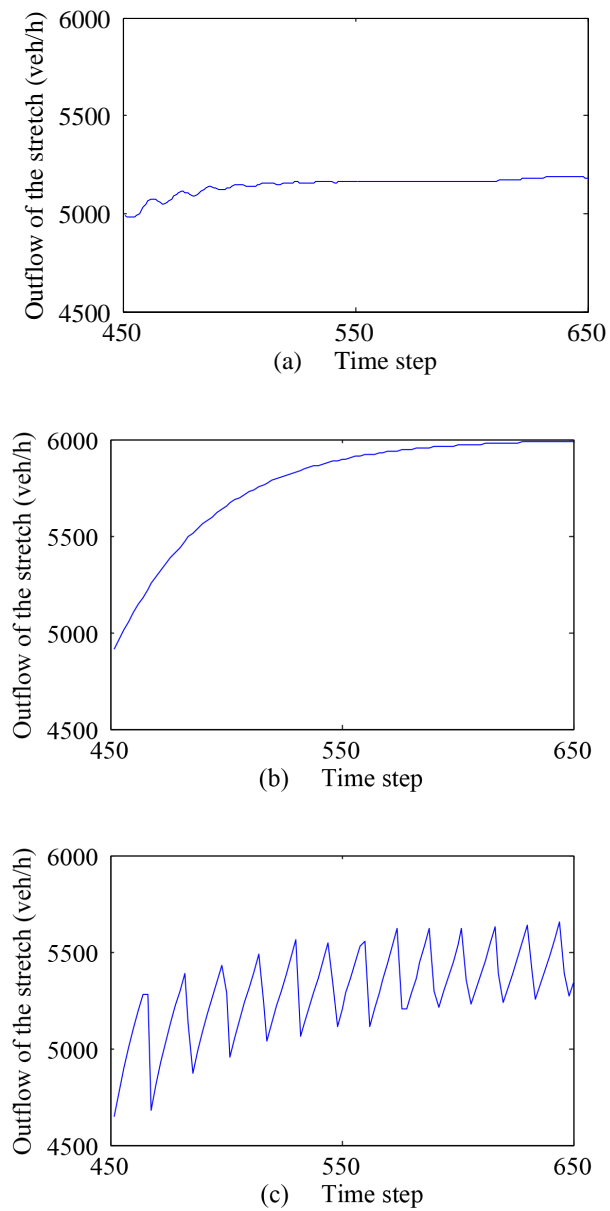


Figure 3.11: The outflow of the freeway stretch from the simulation of the METANET model (a), Roncoli's model (b), and the new extended CTM (c).

of the freeway stretch was doubled. The reason is that the VSL signal generated by controller 2 is too short in length and time to resolve the jam wave. This could be attributed to the fact that the prediction model of controller 2 has underestimated the extent of the capacity drop, thus required VSLs (in length and time) to resolve the jam is underestimated. For controllers 2 and 3, most of the control signals are higher than the minimum VSL value, which are shown in Fig. 3.13 (b-c). A few control signals which are lower than the minimum VSL value is because of that the flow bound cannot be set to cells which have lower values of flows, as explained in Section 3.3. In this case, the proportion of the VSL signals which are lower than the desired minimum VSL value (35 km/h) is less than 2%, and the minimum value of contradicted signals is higher than 30 km/h.

More results are summarized in Table 3.2. It can be seen that all of the three controllers reduce the total travel delay compared to the no-control scenario. Total travel delay is calculated from the difference between the total travel time in the simulations and the free flow total travel time. Controller 1 achieves less total travel delay, but at a cost of a higher computation time which is not real-time feasible. Controller 2 has a less computation time, but it does not reduce the total travel delay as much as controllers 1 and 3. In addition, since the jam wave is not resolved, the traffic upstream of the network still suffers bad influences. Controller 3 has a more balanced performance, because it reduces total travel delay substantially compared to controller 2 while it can maintain a low computation time. The computation time of controllers 2 and 3 are both possible for a real time application. The computer with an E5-1620 processor and 16 GB RAM is used for the optimizations. Note that one of the reason which causes the longer computation time of controller 1 is that the non-linear optimization problem is solved by a standard SQP algorithm. As mentioned earlier, there are some algorithms which could increase the computation speed of the non-linear optimization. More comparisons with different kinds of algorithms remain a subject for future research.

Table 3.2: The total travel delay and computation time of three controllers

Scenarios	Total travel delay (h)	Computation time
No control	242	-
The non-linear MPC (Controller 1)	112.7	6-8 min
The linear MPC (Controller 2)	136.1	6 s
The proposed MPC (Controller 3)	116.2	9 s

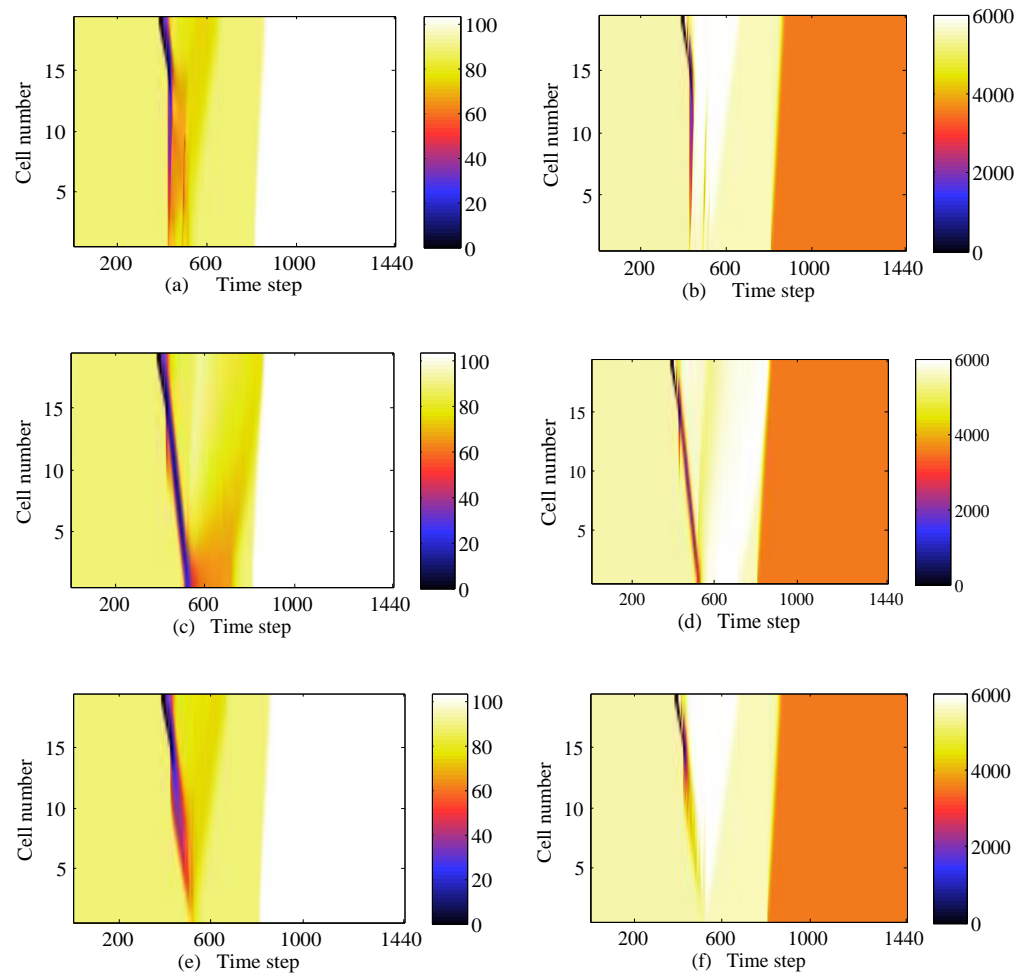


Figure 3.12: The speed contour plots (with VSL control) of controllers 1, 2, 3 (a,c,e), and the flow contour plots of controllers 1, 2, 3 (b,d,f).

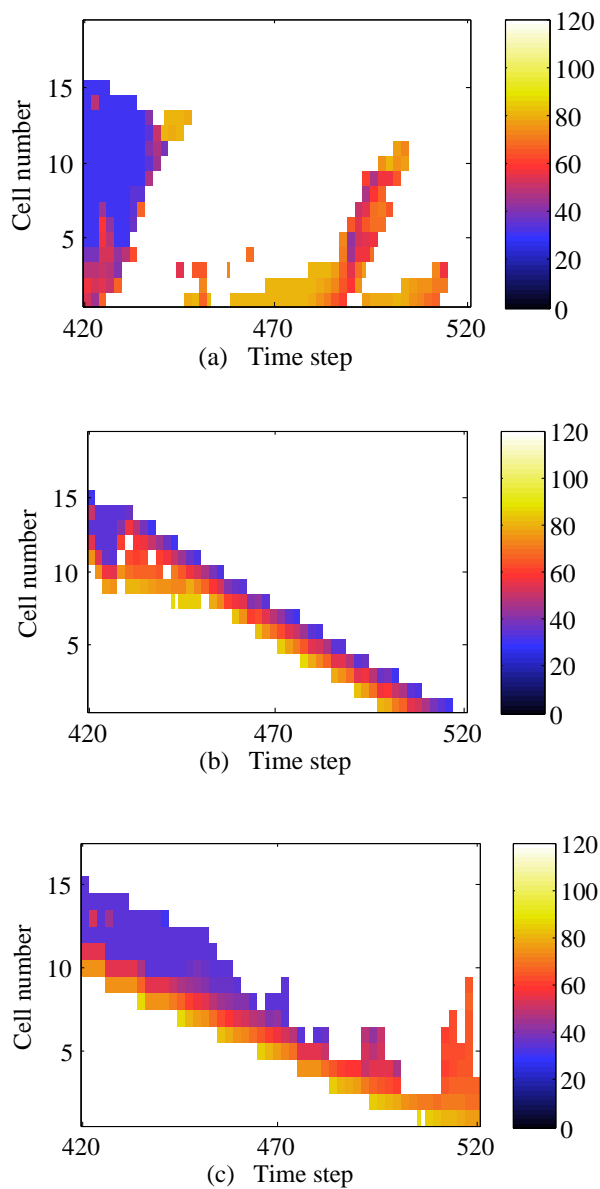


Figure 3.13: The VSL control signals (km/h) generated by controller 1 (a), controller 2 (b), and controller 3 (c).

3.5 Conclusion and future research

In this chapter we apply a linear MPC for optimal VSL control to resolve jam waves on freeways. The proposed MPC is developed based on a new extended discrete first-order model which takes the capacity drop into account and is able to reproduce jam wave propagation accurately. The proposed MPC controller is applied to a benchmark freeway stretch to test the performance in terms of computation speed and jam wave resolution. The second-order model METANET that can reproduce the capacity drop phenomenon, is used as the process model to represent the real world. Two other controllers, one non-linear MPC based on the METANET model and the other linear MPC based on a first-order model, were also tested in the benchmark problem for comparison.

In the benchmark problem, the proposed controller shows a better performance in resolving the jam wave and reducing the total travel delay than the compared linear MPC. Meanwhile, the proposed controller keeps a real time feasible computation time, which has a significant improvement comparing to the METANET-based MPC. This feature enables the proposed controller to be implemented in practice.

This chapter also proposes a way to implement the minimum flow constraints to linear MPC formulations. A forward prediction is performed before the optimization process, so as to ensure that the minimum flow constraints do not conflict with the demand and supply constraints.

Future research will consider: implementing safety constraints into the controller. For safety, normally the VSL variation between the neighboring cells should be lower than a certain value. At this moment it is not trivial to implement such constraints into the controller. The proposed controller also needs to be tested for integrated control measures (such as ramp metering, route guidance, etc.) in a large-scale network.

Chapter 4

Validation of an extended discrete first-order model with variable speed limits

This chapter validates the prediction model embedded in a model predictive controller (MPC) of variable speed limits (VSLs). The MPC controller was designed based on an extended discrete first-order model with a triangular fundamental diagram. In our previous work, the extended discrete first-order model was designed to reproduce the capacity drop and the propagation of jam waves, and it was validated with reasonable accuracy without the presence of VSLs. As VSLs influence traffic dynamics, the dynamics including VSLs needs to be validated, before it can be applied as a prediction model in MPC. For conceptual illustrations, we use two synthetic examples to show how the model reproduces the key mechanisms of VSLs that are applied by existing VSL control approaches. Furthermore, the model is calibrated by use of real traffic data from Dutch freeway A12, where the field test of a speed limit control algorithm (SPECIALIST) was conducted. In the calibration, the original model is extended by using a quadrangular fundamental diagram which keeps the linear feature of the model and represents traffic states at the under-critical branch more accurately. The resulting model is validated using various traffic data sets. The accuracy of the model is compared with a second-order traffic flow model. The performance of two models is comparable: both models reproduce accurate results matching with real data. Flow errors of the calibration and validation are around 10%. The extended discrete first-order model-based MPC controller has been demonstrated to resolve freeway jam waves efficiently by synthetic cases. It has a higher computation speed comparing to the second-order model-based MPC.

This chapter is an edited version of the article:

Han, Y., Y. Yuan, A. Hegyi, and S. P. Hoogendoorn. (2017). Validation of an extended discrete first-order model with variable speed limits. *Transportation Research Part C: Emerging Technologies*, 83,117.

4.1 Introduction

Over the past decades, variable speed limits (VSLs) have emerged as a popular control measure for traffic flows on freeways. The main objective of VSLs is to improve traffic operation efficiency. For this purpose, earlier studies have assumed that the freeway capacity can be raised if the speeds across vehicles in different lanes is harmonized by VSLs (Zackor, 1972). Following this line, Cremer (1979) proposed a quantitative model for the VSL-induced fundamental diagram, in which the free flow capacity was assumed to increase under certain VSL rates. It has been found that VSLs can increase utilization of the shoulder lane on a long homogeneous freeway stretch (Duret et al., 2012). For shorter freeway stretches, although VSLs can homogenize the speed of individual vehicles, no evidence has shown that VSLs can increase the free flow capacity. In fact, later investigations could not identify any capacity increase that could be attributed to VSLs (Smulders, 1990; Soriguera et al., 2017).

Recently, some studies have focused on resolving jam waves or preventing the onset of congestion by using analytical approaches. A well-known example of this type of approach is the SPECIALIST algorithm (Hegyi et al., 2008), which was developed based on the Kinematic Wave theory (Lighthill & Whitham, 1955; Richards, 1956). It aims to resolve freeway jam waves and achieve a higher outflow by limiting the arriving flow to the jam via VSLs. While the approach has been successfully tested in practice (Hegyi & Hoogendoorn, 2010), its disadvantage is that its feed-forward structure does not correctly handle demand changing. Many of the failures are due to significant demand increase during the VSL control (Han et al., 2015a). The algorithm was later extended to a feedback structure, which is known as COSCAL v2, to better handle disturbances (Mahajan et al., 2015). Chen et al. (2014a) recently proposed several analytical VSL schemes based on the kinematic wave theory, to increase the discharge rate of both recurrent and non-recurrent bottlenecks. In those approaches, a triangular fundamental diagram with stable states on the congestion branch was assumed. However in reality, traffic states on the congestion branch are unstable, which might undermine the implementation potential of those approaches.

Another avenue of the VSL control is the model-based optimization method. Traffic model predictive control (MPC) predicts the evolution of traffic dynamics and calculates the optimal control scheme for the time period in which the relevant traffic dynamics occur. This feature enables the controller to take advantage of potentially larger future gains at a current (smaller) cost, so as to avoid myopic control actions. Hegyi et al. (2005b) and Carlson et al. (2010a) presented MPC approaches that are based on a non-linear second-order traffic flow model METANET (Kotsialos et al., 2002a). In Hegyi et al. (2005b), VSLs were applied to suppress freeway moving jams. The application of VSLs upstream of a moving jam temporarily decreases the mainstream arriving flow to resolve the moving jam. The assumed VSL effect was to replace the left part of the flow-density curve by a straight line with the slope corresponding to the displayed VSL values. In Carlson et al. (2010a), the main objective was to control

the free-flow traffic upstream of a bottleneck to prevent the bottleneck activation. It was assumed that sufficiently low VSLs lead to lower capacity in the fundamental diagram than in no-VSL cases. Thus, the application of VSLs upstream of a bottleneck permanently reduces the mainstream arriving flow, which helps to avoid the bottleneck activation. The incorporated VSL model was based on the empirical findings from (Papageorgiou et al., 2008).

In recent years, linear or quadratic MPC approaches of VSLs that are based on extended discrete first-order models have become a popular choice because they can achieve a low computation time, which is beneficial for real-time applications (Muralidharan & Horowitz, 2015; Roncoli et al., 2015b; Hadiuzzaman & Qiu, 2013). Due to the fact that the prediction models of those controllers are not able to reproduce the propagation of jam waves accurately, the type of traffic jam that those controllers addressed is limited to standing queue. Recently, Han et al. (2017b) presented a linear quadratic MPC which is based on a new extended discrete first-order model to resolve freeway jam waves. The new extended discrete first-order model is solved by the minimum sending flow and receiving flow method, in which the boundary flow between two cells are determined by the minimum between the sending flow (demand) and the receiving flow (supply) (Lebacque, 1996). For the VSL model incorporated in the new extended first-order model, it was assumed that lower values of VSLs reduce the sending flow of a cell and the receiving flow of a cell keeps unchanged. With this assumption VSLs can be used to regulate flows from the mainstream of freeways. In that paper, the presented controller was compared with a METANET-based MPC in terms of computation speed and resolving freeway jam waves. Simulation results show that the new extended discrete first-order model-based MPC is able to resolve freeway jam waves efficiently, and it has less computation cost than the METANET-based MPC.

Although many MPC approaches of VSLs have been demonstrated to be effective in resolving traffic jams or reducing total travel delay through simulations, few have validated the prediction model in reproducing the behavior of VSLs. The MPC approach presented in (Han et al., 2017b) was based on a new extended discrete first-order model with a triangular fundamental diagram, which was validated with reasonable accuracy without the presence of VSLs. As VSLs influence traffic dynamics, the prediction model of the MPC forecasts the evolution of traffic states under VSLs based on the effects of VSLs, and the control performance relies on the accuracy of the forecasting. Thus, the predicting traffic dynamics under VSLs need to be validated since the accuracy of the prediction model is essential to the control performance. This chapter moves towards this direction.

For conceptual illustrations, we use two synthetic examples to show how the model reproduces the key mechanisms of VSLs that are applied by existing VSL control approaches in the literature. In addition, model calibration and validation are performed by using real traffic data from Dutch freeway A12, where the field test of a speed limit control algorithm called SPECIALIST was conducted. In the calibration and validation part, the prediction model is extended by using a quadrangular fundamental diagram

which keeps the linear feature of the model and represents traffic states at the under-critical branch more accurately. An extended METANET model with VSLs, which is commonly used for MPC of VSLs, is also calibrated and validated with the same data for comparison (Hegyi et al., 2005a). The results for both models are evaluated qualitatively and quantitatively.

The remaining of the chapter is organized as follows. Section 4.2 presents the extended first-order model and the incorporated VSL model. In Section 4.3, two synthetic cases are presented to show how the model behaves in reproducing the key mechanisms of VSLs that are used by existing VSL approaches in the literature. In Section 4.4, model calibration and validation are performed to demonstrate that the prediction model can accurately forecast the effects of VSLs to jam waves. Section 4.5 presents the conclusions.

4.2 Model description

4.2.1 The new extended discrete first-order model

Our model is an extended version of the cell transmission model (CTM) developed by Daganzo (1994), which is consistent with the hydrodynamic theory of the LWR model (Lighthill & Whitham, 1955; Richards, 1956). In the original CTM, the roadway stretch is divided into cells, which are increasingly numbered from upstream to downstream. For each cell i at time step k the exit flow $f_i(k)$ is expressed as the minimum between the upstream sending flow $S_i(k)$ and downstream receiving flow $R_{i+1}(k)$ by applying the Godunov scheme (Godunov, 1959). For convenience the time index k is omitted in the following. The CTM assumes a triangular-shaped fundamental diagram (FD), which is characterized by the free-flow speed v [km/h], the capacity c [veh/h], the congestion wave speed β [km/h], and the jam density ρ^J [veh/km]. The duration of a time step should be smaller than the free-flow travel time of a cell for numerical stability (this condition is known as the Courant-Friedrichs-Lewy (CFL) condition). The fundamental diagram and the sending flow and receiving flow functions of the CTM are depicted in Fig. 4.1.

The CTM is not able to reproduce some traffic phenomena such as the capacity drop, traffic oscillations, and traffic hysteresis. The capacity drop, which is known as the phenomenon that the outflow of a traffic jam is lower than the free-flow capacity, is crucial for the development of traffic operation strategies. The main physical explanations for such a phenomenon are lower speeds for merging vehicles combined with bounded accelerations, e.g. (Cassidy & Rudjanakanoknad, 2005; Laval & Daganzo, 2006; Leclercq et al., 2016), and the impacts of driver behaviors, e.g. (Cassidy & Ahn, 2005; Chen et al., 2014b).

An extended CTM was proposed to reproduce the capacity drop and the propagation of jam waves. In that model, the capacity drop is represented by a modified sending

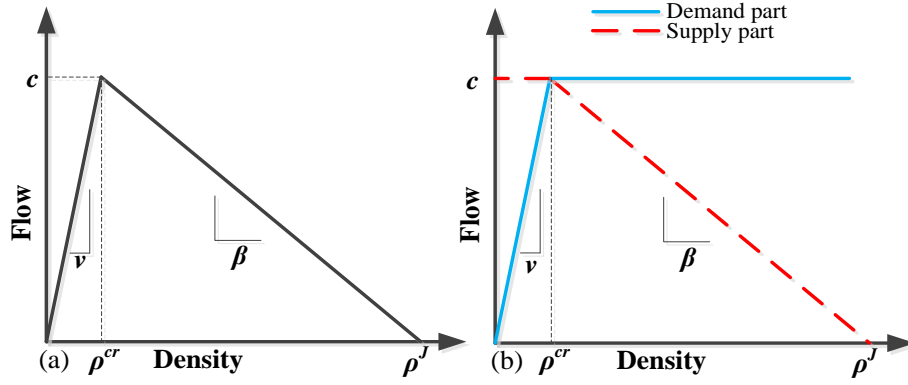


Figure 4.1: (a) The triangular-shaped FD of the CTM. (b) The the sending flow and receiving flow functions.

flow function. The maximum flow of cell i , Q_i , is assumed to reduce linearly as the density of the upstream cell $i - 1$ increases at the over-critical part of the FD, as shown in Fig. 4.2. Such assumption is made based on the empirical study of Yuan et al. (2015), which shows that the magnitude of capacity drop has a negative relation with the speed in congestion. For simplicity, in this chapter we assume that all cells have same properties. However, note that it is possible to make the model parameters cell-dependent. For a full presentation of the model, readers are referred to (Han et al., 2016b). Without loss of generality, the index of cell number remains in the following equations. According to the capacity drop assumption, the maximum outflow of cell $i + 1$ can be represented as:

$$Q_i = \min \left(c_i, c_i \left(1 - \alpha \frac{\rho_{i-1} - \rho_{i-1}^{cr}}{\rho_{i-1}^J - \rho_{i-1}^{cr}} \right) \right), \quad (4.1)$$

where, α denotes the maximum extent of the capacity drop. The sending flow of cell i , S_i , is represented by the following equations,

$$S_i = \min(\rho_i v_i, Q_i) \quad (4.2)$$

The function of the receiving flow is also modified, to calculate the flows more accurately under the so-called discharging state. The discharging state happens when the jam head is between two cell boundaries. For example, if the jam head is in the middle of cell $i + 1$, then cell $i + 1$ is in such a state that the back part is congested while the front part is not, which is shown in Fig. 4.3. For a discharging cell $i + 1$, if its receiving flow is determined by the average density, it will be overestimated since only the upstream part of cell $i + 1$ is congested. Thus, we assume that the receiving flow of a discharging cell has a reduction, R_{i+1}^{diff} , compared to the original CTM. This assumption is demonstrated by empirical analysis in Han et al. (2016b). R_{i+1}^{diff} is represented as:

$$R_{i+1}^{\text{diff}} = \left(\beta_{i+1}^{(1)} - \beta_{i+1}^{(2)} \right) \cdot (\rho_i - \rho_{i+1}) \quad (4.3)$$

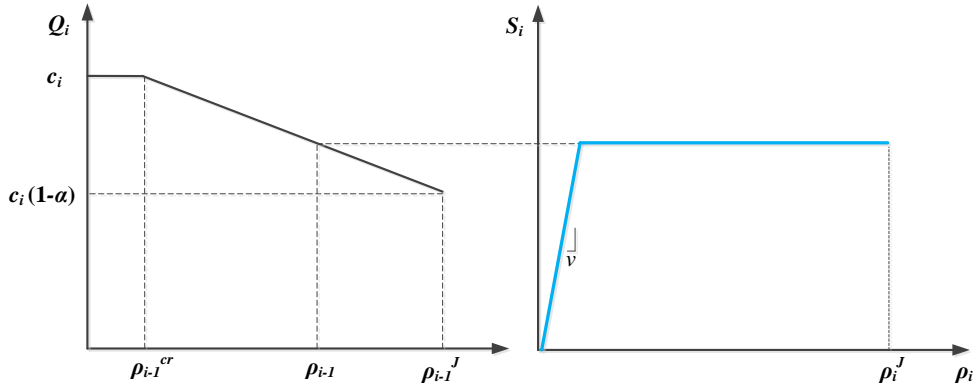


Figure 4.2: The left figure depicts the capacity drop assumption of the proposed model. As the density of cell $i - 1$ increases, the capacity of the cell i decreases. The right figure depicts the sending flow function of cell i .

where, $\beta_{i+1}^{(1)}$ and $\beta_{i+1}^{(2)}$ are depicted in Fig. 4.4. $\beta_{i+1}^{(1)}$ represents the slope of the line which connects the critical state and the jam state. It expresses the congested states of traffic in a jam or getting in a jam. $\beta_{i+1}^{(2)}$ represents the slope of the line which connects the state of a discharging cells and the state of its upstream cell. It expresses the congested states of traffic leaving a jam wave, which in free-flow will lead to the queue discharge rate that is lower than the free-flow capacity. When traffic of cell $i + 1$ is getting in a jam, the traffic state lies in the line with the slope $\beta_{i+1}^{(1)}$. When traffic in cell $i + 1$ is leaving a jam, the traffic state lies in the line with the slope $\beta_{i+1}^{(2)}$. Thus, the line of $\beta_{i+1}^{(1)}$ represents the deceleration of traffic whereas $\beta_{i+1}^{(2)}$ represents the acceleration of traffic. In real traffic process, traffic's deceleration is faster than its acceleration, thus $\beta_{i+1}^{(1)}$ is higher than $\beta_{i+1}^{(2)}$. The overall receiving flow function of cell $i + 1$, R_{i+1} is formulated as:

$$R_{i+1} = \min \left(Q_{i+1}, \beta^1(\rho_{i+1}^J - \rho_{i+1}) - R_{i+1}^{\text{diff}}, \beta^1(\rho_{i+1}^J - \rho_{i+1}) \right). \quad (4.4)$$

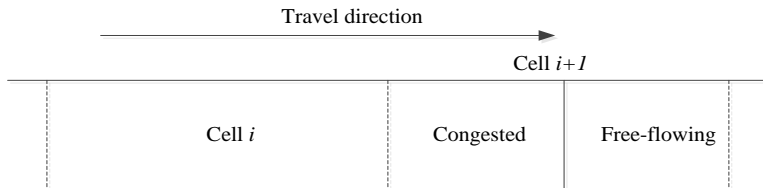


Figure 4.3: An example to explain the discharging state. Cell $i + 1$ is in the discharging state.

The proposed model is applied for optimal VSL control by solving an optimization problem which is based on a linear quadratic objective function and several linear constraints (Han et al., 2017b). The objective function was a weighted combination of total time spent minimization and total travel distance maximization. Linear constraints which correspond to the sending flow and receiving flow function, represent

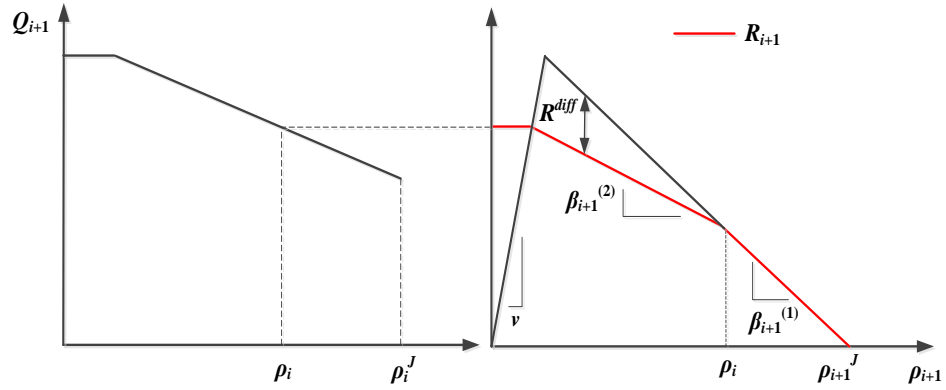


Figure 4.4: The depiction of the receiving flow function of the proposed model.

the upper bounds of the solutions. In case no control actions are implemented, the solution of flows are given by the prescribed upper bounds. When the solution of the optimization problem returns an optimal flow value which is lower than the prescribed upper bounds, then reduced VSLs are needed to enforce these reduced flows. The value of the optimal VSLs can be calculated as the ratio of the optimal flows and the corresponding densities.

In other applications of this model (e.g., model validation), when the values of VSLs are already known, the traffic evolution can be calculated based on a VSL-based sending flow function which is represented as,

$$S_i = \min(\rho_i v_i, Q_i, \rho_i V_i), \quad (4.5)$$

where, V_i is the optimal VSL value of cell i . In practice, drivers are expected to drive at the optimal speed but they usually do not fully comply with the displayed speed limit and their target speed is usually higher than what is displayed. The relation between the optimal speed (or the actual driving speed) and the displayed speed is expressed as,

$$V_i^{\text{disp}} = \frac{V_i}{1 + \theta_1}, \quad (4.6)$$

where, $1 + \theta_1$ is the non-compliance factor. Note that the left-hand side of the FD could be modeled via any concave piecewise-linear function other than one single line shown in Fig. 4.2. If more lines are used to model the under-critical traffic situations, a better approximation of the critical speed can be obtained. For the controller proposed in (Han et al., 2017b), a triangular FD was applied to make a fair comparison between the MPC that is based on the proposed model and another discrete LWR model-based MPC (Roncoli et al., 2015a). In the calibration and validation part of this chapter (Section 4), we extend the original model by using a quadrangular FD which represents traffic states at the under-critical branch more accurately. Results will demonstrate that the accuracy of the model improves by using the quadrangular FD. The MPC controller that is based on the quadrangular FD has more linear constraints than the one based on the triangular FD. Nevertheless, since both FDs have linear features,

the computation time of the controllers will not result in significant differences. The quadrangular fundamental diagram is shown in Fig. 4.5. The sending flow function of the proposed quadrangular fundamental diagram is represented by the following equation:

$$S_i = \min(\rho_i v, Q_i, \rho_i V_i(k), c - \gamma(\rho^{cr} - \rho_i)), \quad (4.7)$$

where, γ is the slope of the red line, as shown in Fig. 4.5. In the figure, ρ_{cr}^V is the VSL-induced critical density and f_{VSL} is the VSL-induced capacity.

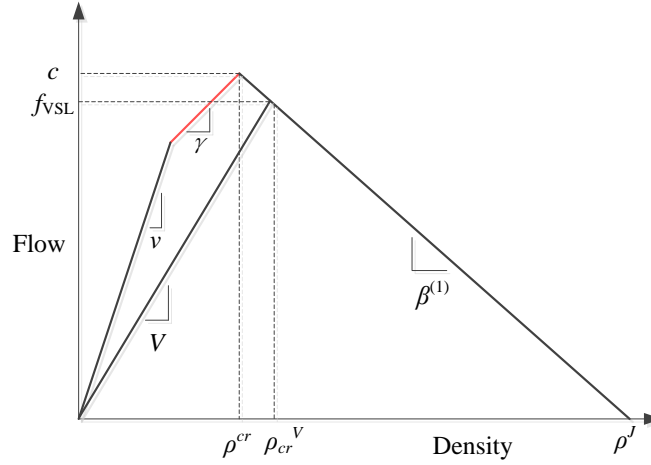


Figure 4.5: The quadrangular fundamental diagram of the proposed model. ρ_{cr}^V is the VSL-induced critical density and f_{VSL} is the VSL-induced capacity.

4.3 Reproducing the mechanisms of VSLs

In the literature, the mechanisms of VSLs are applied to different VSL control approaches. For example, Hegyi et al. (2008) proposed the SPECIALIST algorithm which was based on the temporal transition states of VSLs to resolve freeway jam waves. Carlson et al. (2010a) proposed the mainstream metering approach which was based on the assumption of the VSL-induced capacity reduction to prevent the activation of on-ramp bottlenecks. In this section we use two synthetic cases to show that the model is able to reproduce the mechanisms that are relevant for these well-known VSL approaches.

4.3.1 Reproducing the SPECIALIST algorithm

The SPECIALIST algorithm is designed based on the shock wave theory to resolve freeway jam waves. Fig. 4.6 is used to explain the theory of SPECIALIST. The time-space graph in the left figure shows the traffic states on a road stretch and their propagation over time. The density-flow diagram in the right figure shows the corresponding density and flow values for these states. According to the shock wave theory, the

boundary (front) between two states in the left figure has the slope as the slope of the line that connects the two states in the right figure, which satisfies the Rankine-Hugoniot conditions (LeVeque, 1992). The SPECIALIST theory uses this basic relationship to resolve jam waves.

Area 2 represents a short jam wave that propagates upstream and which is surrounded by traffic in free-flow (area 1 and 6). As soon as the jam wave is detected, the speed limits upstream of the jam wave are switched on, where traffic state changes to state 3. The inflow of area 2, which is the flow of state 3, is lower than the outflow of area 2, which is the flow of state 1, so the width of the jam wave is narrowing as time advances. To resolve the jam wave, the required length of the speed-limited stretch depends on the density and flow associated with state 2 and the physical length of the detected jam. When the jam wave is resolved there remains an area with the speed limits active (state 4) with a moderate density (higher than in free-flow, lower than in the jam wave). It is assumed that traffic leaving area 4 has a higher flow and a higher speed than state 4, which is represented by state 5. A basis assumption in the SPECIALIST theory is that the traffic from area 4 can flow out more efficiently than a queue discharging from full congestion as in the shock wave (flow of state 2), which results in the increasing of the total throughput. In a later research, the analysis of the data from the SPECIALIST field test confirms this assumption (Hegyi & Hoogendoorn, 2010). For a full presentation of the concept, readers are referred to (Hegyi et al., 2008).

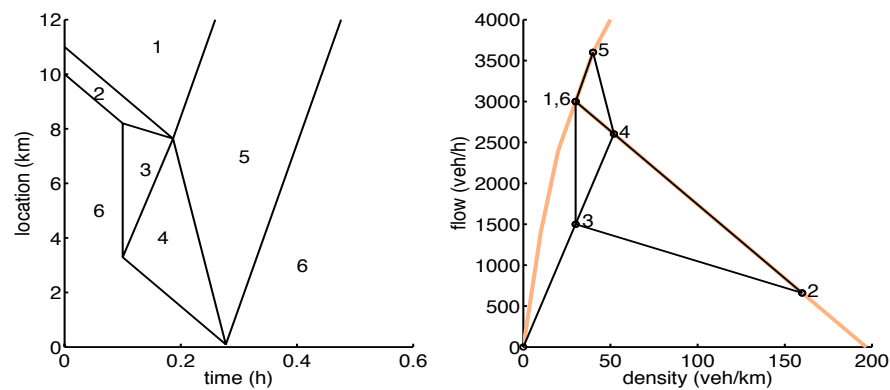


Figure 4.6: Illustration of traffic evolution under the SPECIALIST. The left figure is the time-space graph and the right figure is the fundamental diagram.

The proposed model is applied to reproduce the process in which a jam wave is resolved by the SPECIALIST algorithm. The set up of the synthetic case is explained as follows. A homogeneous freeway stretch with 3 lanes, which is divided into 20 cells, is assumed in this example. The cell length is set to 300m and the duration of a discrete time step is set to 10 seconds. The free flow speed is set to 108 km/h, which satisfies the CFL condition. Other model parameters are taken from (Han et al., 2017b). θ_1 is set to 0 so we assume that all drivers comply with the speed limits. The jam wave is created by changing the boundary condition at the downstream of the stretch. As shown in Fig. 4.7 (a) and (b), a jam wave is created at the downstream of the stretch

and propagated through the stretch if no control action is taken. Fig. 4.7 (f) shows the throughput of the stretch without VSLs, in which the red line represents the throughput from the SPECIALIST theory and the blue line represents the throughput from the simulation. It can be seen that without VSLs, the throughput of the stretch is lower than the capacity after the jam wave being activated. The outflow reproduced by the proposed model shows oscillation behavior, which is attributed to a numerical issue, which can be explained as follows. Suppose cell i is in a discharging state, and cell $i + 1$ is in free-flow state. Under this circumstance, the outflow of the network is dominated by the outflow f_{i+1} of cell $i + 1$, because cell $i + 1$ is the most upstream cell of all downstream free-flow cells. Since cell i is in the discharging state, the vehicle density ρ_i is decreasing and according to (1), f_{i+1} is also decreasing. That is the reason for the increasing outflow. When ρ_i decreases to an under-critical value, cell $i - 1$ becomes the discharging cell and the outflow of the network is dominated by f_i . That's the reason for the sudden decrease of the outflow. This process repeats as the discharging cell moves upstream, causing the oscillation of the flow. Note that the fluctuation of flow looks big because the time step is small. The fluctuations are around 200-300 veh/h, which is less than 1 veh/time step. This would have limited effect to traffic flow.

At time step 200, VSLs are switched on to resolve the jam wave. The value of the VSL is set to 60 km/h and the VSL control area is calculated based on the SPECIALIST algorithm. According to the SPECIALIST theory, the maximum outflow of area 5 should be tuned to a value that is a bit lower than the free flow capacity and the speed of state 5 should be according to the fundamental diagram and the chosen outflow. In our discrete LWR model, the outflow of the VSL area is calculated by equation (1), in which ρ_{i-1} equals to the VSL-induced critical density. Fig. 4.7 (c) shows the speed contour plot of the simulation under VSLs and the VSL-controlled area 3 and 4 are highlighted. Fig. 4.7 (d) shows the simulated flow contour plot under VSLs and traffic states 1, 4, 5, and 6 are highlighted in the figure. It can be seen that area 5 has a higher flow than area 1 and the corresponding area in the no-control case, thus the desired results that the SPECIALIST resolves the jam wave and gets a higher throughput are reproduced by the proposed model. Fig. 4.7 (e) shows the throughput of the stretch with VSLs, which also indicates that the simulation captures a higher throughput which is close to the capacity achieved by the VSLs.

4.3.2 Reproducing the mainstream metering approach

Empirical evidences in (Papageorgiou et al., 2008) show that sufficiently low VSLs lead to a lower flow (capacity) than in the unrestricted case. This forms the basis of the mainstream metering approach proposed by (Carlson et al., 2010a). Fig. 4.8 explains the theory of the mainstream metering approach. If the arriving demand is higher than the VSL-induced capacity, the VSL application area becomes an active mainstream bottleneck that limits the area's outflow to values that corresponds to the VSL-induced capacity. This approach may be applied at the upstream of a potential bottleneck (e.g.

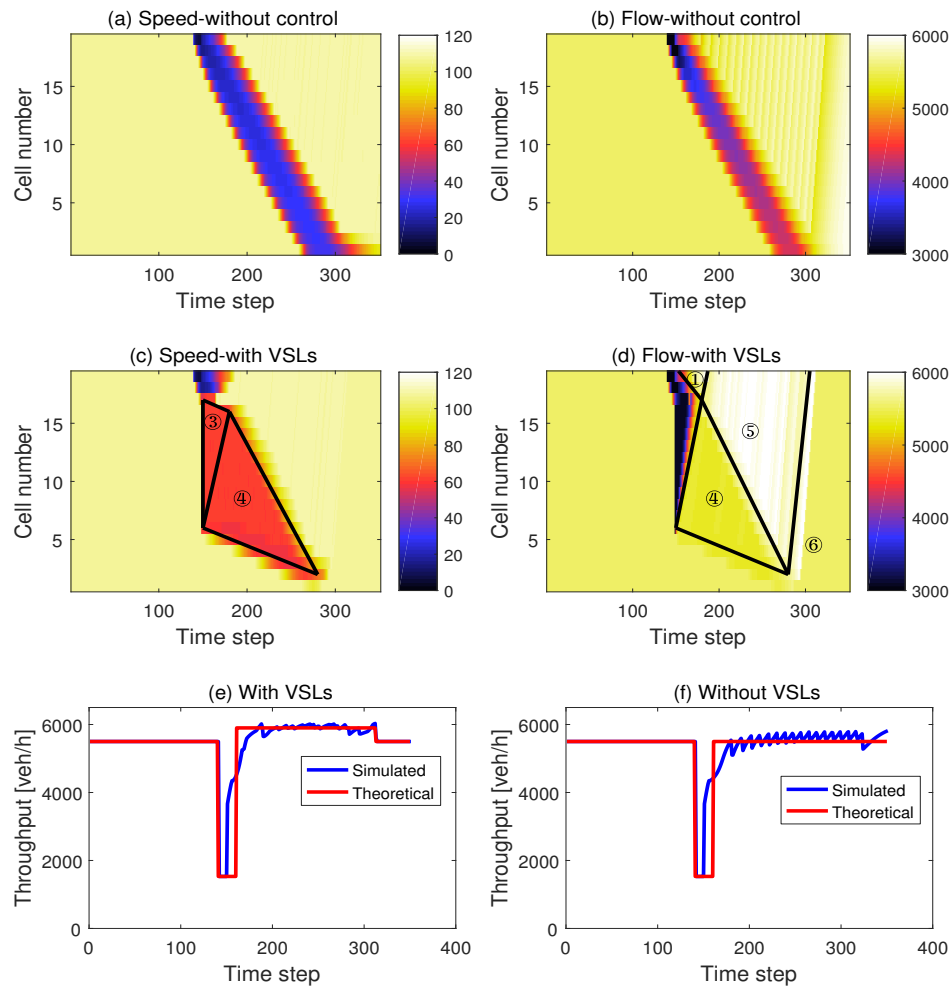


Figure 4.7: A synthetic case to show that the model reproduces the SPECIALIST algorithm. (a) and (b) are the speed contour plot and the flow contour plot without VSLs. (c) and (d) are the speed contour and flow contour with VSLs. Some of the traffic states in Fig. 4.6 are highlighted in (c) and (d). (e) and (f) are the throughput of the stretch with and without VSLs from the theory and the simulation.

an on-ramp merging area) to avoid its activation and the related throughput reduction as a result of the capacity drop. More specifically, the mainstream flows are regulated by VSLs such that capacity flow can be established at the downstream bottleneck. Then the final mainstream throughput is maximized, leading to a decrease of the total time spent.

We use another synthetic case to show how the model behaves in reproducing the mainstream metering effect. The synthetic freeway stretch has the same structure as the freeway stretch in Fig. 4.8. The freeway stretch is divided into 20 cells and an on-ramp bottleneck is assumed at the upstream of cell 15. The model parameters are the same as the previous example. The mainstream demand is set to 5400 [veh/h] and the on-ramp demand is set to 700 [veh/h] for time steps 50-250. Thus, in this time period the on-ramp becomes an active bottleneck since the demand is higher than the capacity

which is 6000 [veh/h]. Traffic dynamics of the non-control scenario is simulated, and the speed and flow contour plots are shown in Fig. 4.9. At time step 50, the on-ramp bottleneck is activated and since then the throughput of the freeway stretch reduces to values lower than the free-flow capacity as a result of the capacity drop.

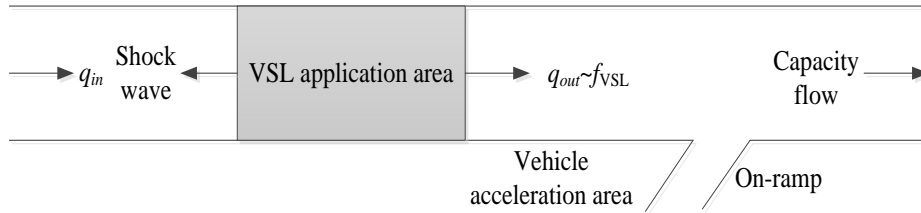


Figure 4.8: The mainstream metering approach used in (Carlson et al., 2010a). The on-ramp is perceived as a potential bottleneck. q_{in} and q_{out} are the inflow and outflow of the VSL control area. f_{VSL} is the VSL-induced capacity.

To avoid the activation of the bottleneck, VSLs are implemented at the upstream of the bottleneck. The speed limit value is set to 56 km/h such that f_{VSL} equals to the difference between the free flow capacity and the on-ramp flow. The speed and flow contour plots of the control scenario are shown in Fig. 4.9 (c) and (d). From Fig. 4.9 (c) it can be seen that an active bottleneck is created at the upstream of the VSL-implemented area (cell 10 and 11), which helps to avoid the activation of the on-ramp bottleneck. Fig. 4.9 (d) shows that the throughput of the freeway stretch in the VSL control period equals to the free-flow capacity, which is higher than the throughput of the corresponding time period in the no-control case. Thus, the desired mainstream metering effect is reproduced by the proposed model.

4.4 Model calibration and validation

To show the performance of the model in reproducing real traffic, the traffic flow model introduced in the previous section is applied to a freeway stretch to calibrate its parameters and validate its accuracy. For comparison purposes, an extended METANET model incorporated with VSLs is also calibrated and validated (Hegyi et al., 2005a). In the calibration part, we use two data sets which include data with and without VSLs to represent comprehensive traffic dynamics. And next the calibrated models are applied to two representative cases for model validations, in which one is a successful case and the other is a failed application from the SPECIALIST field test.

The site for model calibration and validation is Dutch freeway A12, where the SPECIALIST field test has been carried out (Hegyi & Hoogendoorn, 2010). The layout of the tested site is sketched in Fig. 4.10. The tested freeway stretch contains 33 loop detectors, and the data we obtained are detected flows and the time mean speeds of the detectors from 4 to 25. This part of the stretch is homogeneous except for a gas station located in the middle of this part, which has a negligible flow compared with

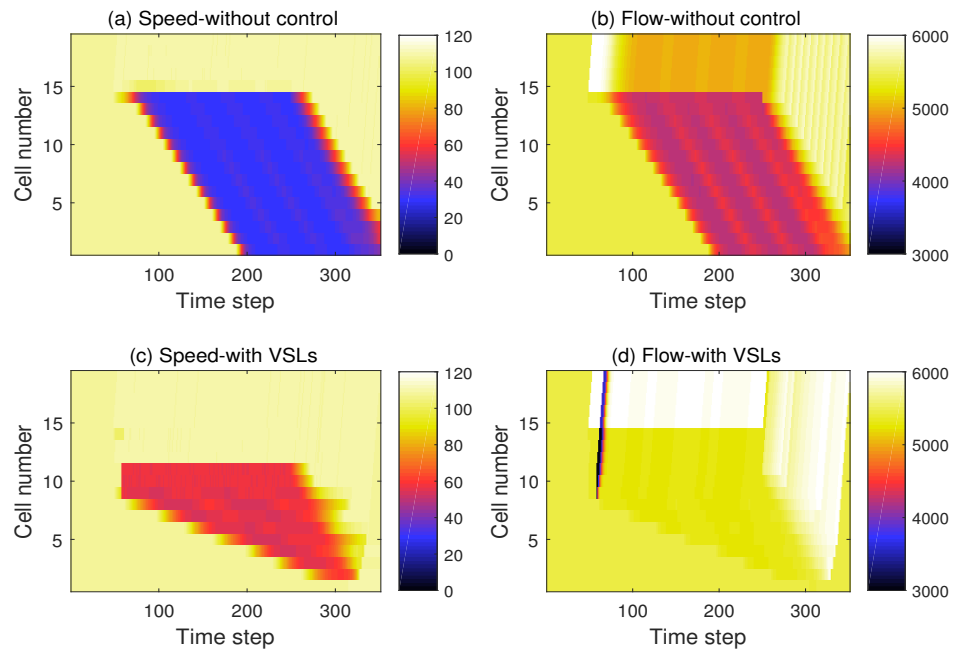


Figure 4.9: A synthetic case to show that the model reproduces the mainstream metering approach. (a) and (b) are the speed contour plot and the flow contour plot without VSLs. (c) and (d) are the speed contour and flow contour with VSLs. In (c), the VSLs are implemented at cells 10 and 11. The speed at upstream of the VSL-control area also reduces, because the VSL-control area has a high density that propagates upstream.

the mainstream flow). The distance between neighboring detectors are between 325 m and 650m. It is assumed that the stretch is divided into cells and the middle of two neighboring detectors is perceived as the boundary of a cell. The density of a cell is estimated through dividing the flow by the time mean speed.

4.4.1 Calibration of the proposed model

To calibrate the model parameters which account for the propagation of traffic jams (e.g., $\beta^{(1)}$ and $\beta^{(2)}$), data that involve traffic jams have to be taken into consideration. For the same reason, data that involve VSLs also need to be included to calibrate the drivers' compliance parameter θ_1 .

Most of the traffic jams at the target site originate from further downstream of the freeway stretch. If a traffic jam is considered to be resolvable by the SPECIALIST algorithm, the VSL control will be activated upstream of the jams. The first data set we selected is from the afternoon peak of January 18, 2010, when jaw wave was triggered at the downstream end of the stretch and propagated through the section from detectors 25 to 14. Data from detectors 24 to 15 are used for the calibration. Note that data from detectors 14 and 25 are excluded because in our data the detection near ramps are often inaccurate. Data from the upstream section (from detector 4 to 10) are not included in the calibration because the influence from the on-ramp and the off-ramp of the gas

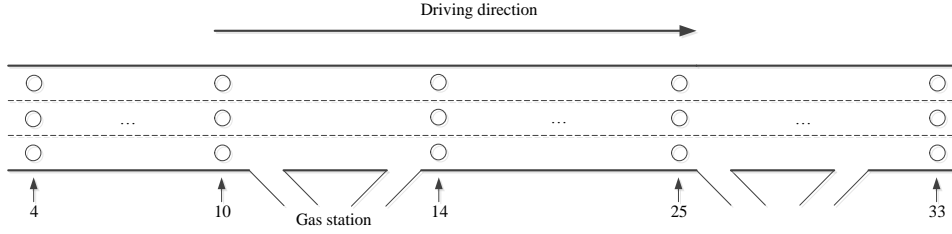


Figure 4.10: A graphical representation of the freeway stretch for the SPECIALIST field test. Traffic goes from the left to the right. Number in the figure represents detector number. The data we obtained are detected flows and the time mean speeds of the detectors from 4 to 25. This part of the stretch is homogeneous except for a gas station located in the middle of this part, which has a negligible flow compared with the mainstream flow). The distance between neighboring detectors are between 325 m and 650m.

station are tried to be avoided, even though it is considered to be minor. The flow, speed and density contour plots of the data are shown in the first row of Fig. 4.11. The second data set we chose is from October 12, 2009, which include data with VSLs. As shown in Fig. 4.12 (b), a jam wave was detected at 16:00 and the VSLs were switched on. The VSL scheme is shown in Fig. 4.12 (a). The same as the data set 1, data from detectors 24 to 15 are used for the calibration.

The density and flow of detector 15 are fed in the model as demand. Initial densities, $\rho_i(1), i = 15, \dots, 24$, are also fed in the model as the initial condition. The vehicle density of the most downstream cell (cell 24) is fed in the model as the downstream boundary condition. When the traffic state at the downstream end of the stretch was in breakdown, the exit flow of cell 23 is fed in the model to trigger the jam wave.

The calibration of model parameters aims to enable the proposed model to reproduce real traffic conditions with the highest accuracy. The parameter calibration problem is formulated as a non-linear optimization problem which aims to minimize the discrepancy between the model predictions and the real data. The following formula is used to represent the performance indicator J ,

$$\begin{aligned}
 J = \min \varphi^f \sqrt{\frac{\sum_{k=1}^K \sum_{i=I_1}^{I_2} [f_i(k) - \hat{f}_i(k)]^2}{K(I_2 - I_1 + 1)}} + \varphi^v \sqrt{\frac{\sum_{k=1}^K \sum_{i=I_1}^{I_2} [v_i(k) - \hat{v}_i(k)]^2}{K(I_2 - I_1 + 1)}} \\
 + \varphi^\rho \sqrt{\frac{\sum_{k=1}^K \sum_{i=I_1}^{I_2} [\rho_i(k) - \hat{\rho}_i(k)]^2}{K(I_2 - I_1 + 1)}}, \quad (4.8)
 \end{aligned}$$

where, $f_i(k)$, $v_i(k)$ and $\rho_i(k)$ are the flow, speed and density calculated by the model at cell i (In this case $I_1=15$ and $I_2=24$). $\hat{f}_i(k)$ and $\hat{v}_i(k)$ are the flows and speeds measured from the detectors. $\hat{\rho}_i(k)$ is the density estimated as the ratio of $\hat{f}_i(k)$ and $\hat{v}_i(k)$ based on the fundamental relation. φ^f , φ^v and φ^ρ are set to values \bar{f}^{-1} , \bar{v}^{-1} , and $\bar{\rho}^{-1}$

respectively to normalize the results, where \bar{f} , \bar{v} , and $\bar{\rho}$ are the average measured flow, density, and speed. The non-linear and non-convex optimization problem is solved by MATLAB implementation of the SQP algorithm (fmincon). To ensure that a global optimum is achieved, 100 starting points which are based on empirical initial guess and random noises are used. Parameters that lead to the minimum value of the cost function are selected and shown in Tab. 4.1. Based on the calibrated parameters and according to the fundamental diagram in Fig. 4.5, the values of parameters $\beta^{(2)}$ and ρ^J are deduced, where $\beta^{(2)} = -11.1$ [km/h] and $\rho^J = 79.3$ [veh/km/lane]. Note that the calibrated jam density is lower than normal values (between 110 to 150 veh/km/lane) because there is a significant amount of trucks in the freeway. A graphical representation of the results obtained in the model calibration from data set 1 can be seen in Fig. 4.11, where the comparison of flow, speed and density contour plots between the calibrated model and real measurement are shown. The speed contour plots reproduced by the proposed model in data set 2 is shown in Fig. 4.12 (c). Quantitatively, the root mean square error (RMSE) of flow, speed and density are shown in Tab. 4.3. The proposed model with a triangular fundamental diagram is also calibrated, and the quantitative results are shown in Tab. 4.3. It can be seen that the proposed model with a quadrangular fundamental diagram improves the accuracy in reproducing speeds and densities. As a referential scenario, we also calibrated the standard cell transmission model (Daganzo, 1994). It can be seen that the standard CTM is less accurate than the proposed model with both triangular and quadrangular fundamental diagrams.

Table 4.1: Calibrated parameters of the proposed model.

Parameters	v [km/h]	c [veh/h]	$\beta^{(1)}$ [km/h]	ρ_{cr} [veh/km/lane]	α	γ	θ_1
Value	110.1	2186	-38.2	21.8	0.66	90.4	0.43

4.4.2 Calibration of the METANET model

In the literature, the second-order traffic flow model METANET has been extended to incorporate VSLs for traffic control purposes (Hegyi et al., 2005a; Carlson et al., 2010a). While the METANET model has been calibrated and validated with reasonable accuracy, its performance with the VSL control has not been explored. Thus, in this section, we calibrate the METANET model with the same data as presented in the previous section. For a better understanding, the extended METANET model is shortly described as follows.

In the METANET model, the freeway stretch is divided into segments and each segment i is characterized by the traffic density $\rho_i(k)$ [veh/lane/km], the mean speed $v_i(k)$ [km/h], and the outflow $q_i(k)$ [veh/h], where k is the index of the discrete time step. The following equations describe the evolution of the freeway stretch over time. The outflow of each segment equals to the density times the mean speed and the number of

lanes of that segment (represented by λ_i):

$$q_i(k) = \rho_i(k)v_i(k)\lambda_i, \quad (4.9)$$

The density of a segment follows the vehicle conservation law, which is represented as:

$$\rho_i(k+1) = \rho_i(k) + \frac{T}{L_i\lambda_i}(q_{i-1}(k) - q_i(k)), \quad (4.10)$$

The mean speed of segment i at time step $k+1$ depends on the mean speed at time step k , the speed of the inflow of vehicles, and the density downstream, which is represented as:

$$v_i(k+1) = v_i(k) + \frac{T}{\tau}(V(\rho_i(k)) - v_i(k)) + \frac{T}{L_i}v_i(k)(v_{i-1}(k) - v_i(k)) - \frac{\vartheta T}{\tau L_i} \frac{\rho_{i+1}(k) - \rho_i(k)}{\rho_i(k) + \kappa}, \quad (4.11)$$

where, τ , ϑ , κ are model parameters. $V(\rho_i(k))$ denotes the desired speed that drivers try to achieve, which is:

$$V(\rho_i(k)) = v_i^{\text{free}} \cdot \exp \left[-\frac{1}{a_m} \left(\frac{\rho_i(k)}{\rho_{cr,i}} \right)^{a_m} \right], \quad (4.12)$$

where, a_m is a model parameter. v_i^{free} represents the free flow speed of segment i , and $\rho_{cr,i}$ denotes the critical density at which the flow is maximal. The effects of VSLs follows the assumption of (Hegyi et al., 2005a), in which the desired speed is the minimum of two quantities: the desired speed based on the experienced density (4.12), and the desired speed caused by the speed limit displayed on the variable message sign

$$V(\rho_i(k)) = \min \left((1 + \theta_2)\text{VSL}_i(k), v_i^{\text{free}} \cdot \exp \left[-\frac{1}{a_m} \left(\frac{\rho_i(k)}{\rho_{cr,i}} \right)^{a_m} \right] \right). \quad (4.13)$$

where, $\text{VSL}_i(k)$ is the speed limit imposed on cell i at time k , and $1+\theta_2$ is the non-compliance factor that expresses that drivers usually do not fully comply with the displayed speed limit and their target speed are usually higher than what is displayed.

The METANET model is calibrated with the same data and the same procedure as presented in the previous section. The value of the calibrated model parameters are shown in Tab. 4.2. A graphical representation of the results obtained in the model calibration from data set 1 can be seen in Fig. 4.11, where the comparison of flow, speed and density contour plots between the calibrated model and real measurement are shown. The speed contour plots reproduced by the proposed model in data set 2 is shown in Fig. 4.12 (d). For the calibration results of data set 1, it can be seen that both the proposed model and the METANET model reproduced the propagation of jam waves. Quantitatively, the root mean square error (RMSE) of flow, speed and density are shown in Tab. 4.3. The accuracy of two models is comparable. The METANET model has a better performance in reproducing the flow and the speed, and the proposed model is more accurate in reproducing the density.

Table 4.2: Calibrated parameters of the METANET model

Parameters	v_f [km/h]	a_m	ρ_{cr} [veh/km/lane]	τ	ϑ	κ	θ_2
Value	127.1	2.4	24.2	13.9	29.8	0.09	0.42

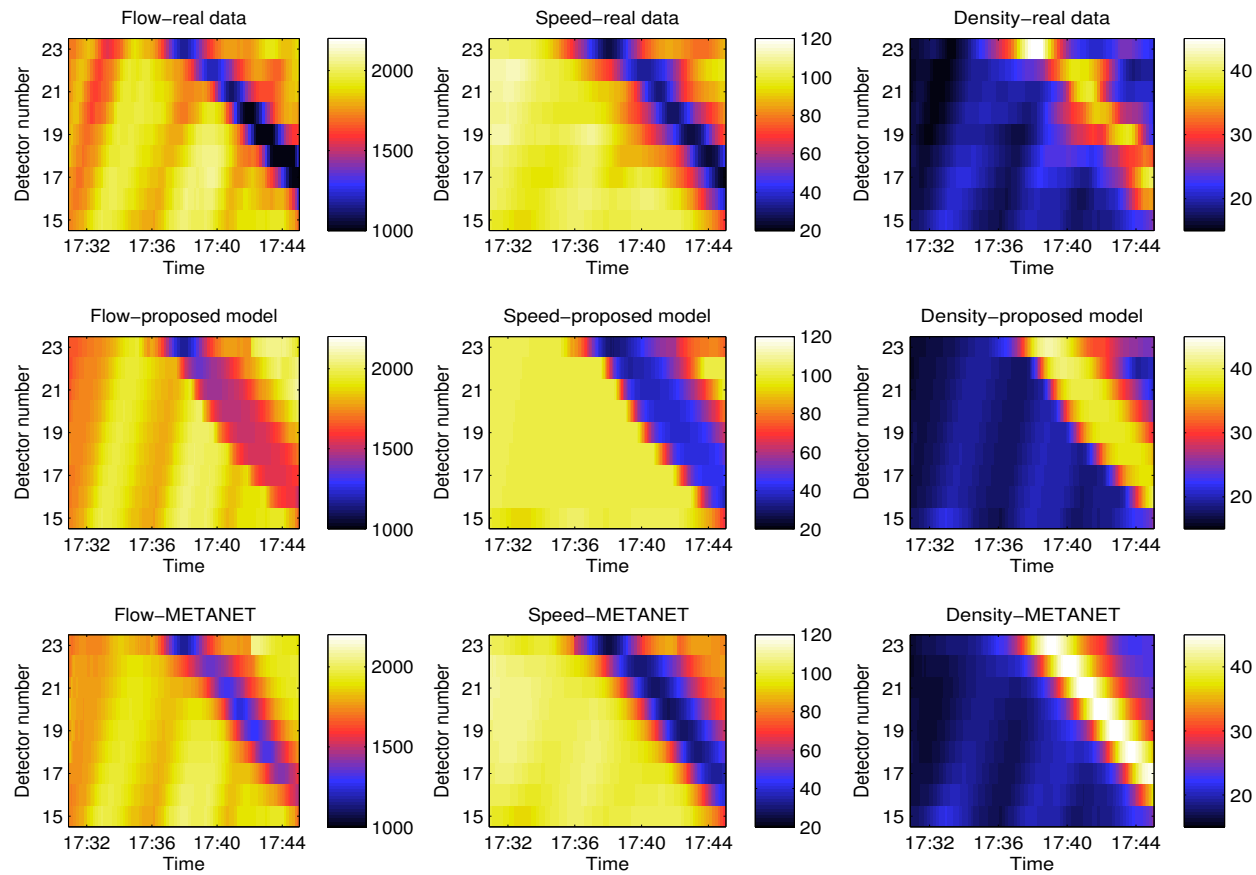


Figure 4.11: The comparison of the real-data for the calibration and the results reproduced from the calibrated models in data set 1. Figures of the first row are plotted from real data, and second row third row are plotted from the simulation results of the proposed model and the extended METANET model.

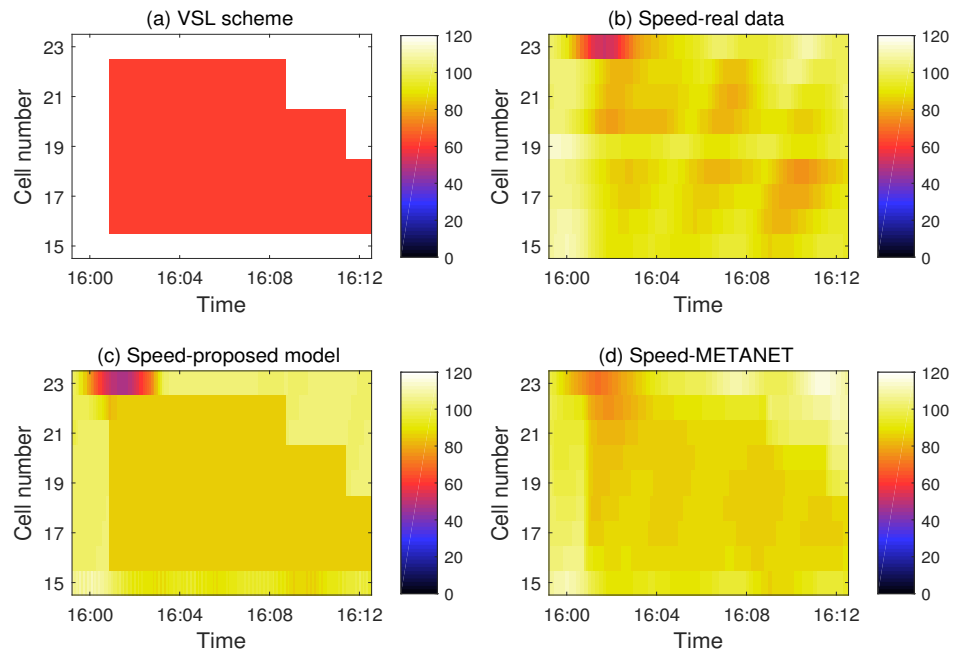


Figure 4.12: The comparison of the real-data for the calibration and the results reproduced from the calibrated models in data set 2. (a) is the VSL control scheme of the data. (b), (c) and (d) are the speed contour plots from real data, the proposed model and the METANET model.

4.4.3 Model validation

The purpose of the model validation here is to demonstrate that the proposed model is able to accurately reproduce traffic dynamics, including jam wave resolution under the VSL control. To this end, data that are used for the validation should involve both jam waves and the VSL control. To explore if the model would overestimate or underestimate the effects of VSLs, we chose two representative cases, where one is a successful case and the other is a failed application from the SPECIALIST field test for the validation.

The parameters resulting from the calibration process are applied to the same freeway stretch but for different days. Data of the first case, the successful case, are obtained from February 26, 2010. A moving jam was detected at 12:10 and the VSL control was switched on. The flow and speed contour plots of the freeway stretch are shown in the left column of Fig. 4.13. The models are fed with boundary data, which are the same process as presented in the model calibration.

The simulated speed and flow contour plots of the proposed model and the extended METANET model are shown as the middle column and the right column in Fig. 4.13 respectively. From Fig. 4.13, it can be seen that both the proposed model and the extended METANET model reproduce the results that the jam wave is resolved by the VSL control, which is consistent with the results from real data. *This demonstrates*

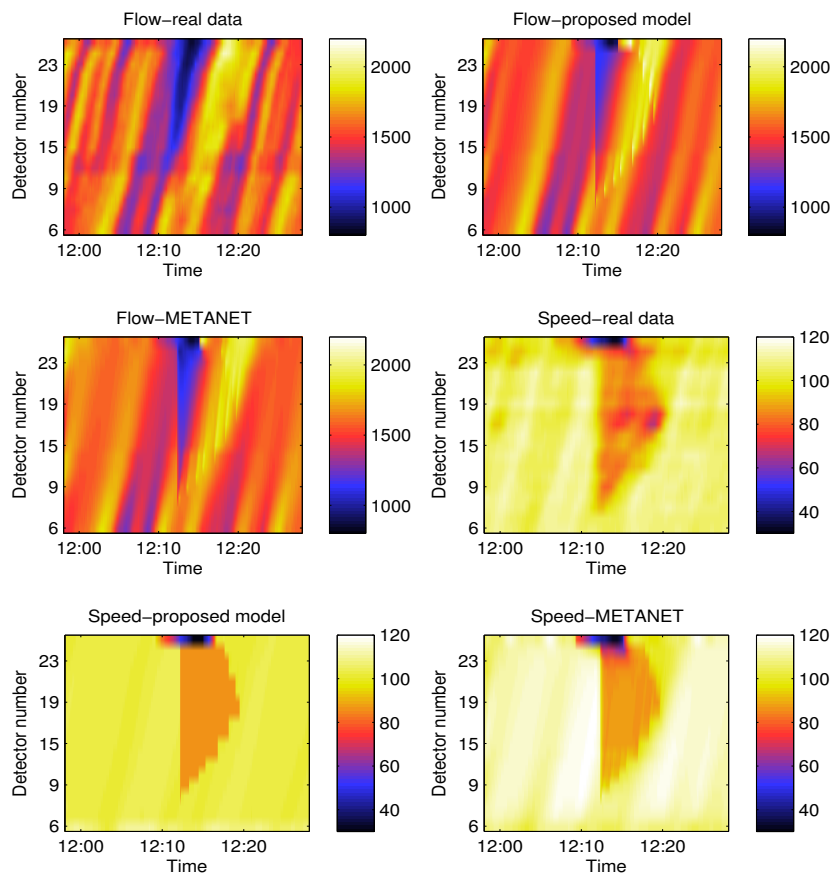


Figure 4.13: The comparison of the real-data for the validation and the results reproduced from the models in case 1. Figures of the left column are plotted from real data, and the middle column and the right column are plotted from the simulation results of the proposed model and the extended METANET model.

that the models do not underestimate the VSL effects. The comparison between the real flow and speed data and the simulated flow and speed data can be seen in Fig. 4.14 and Fig. 4.15. The expected traffic state transition under VSLs can be observed in Fig. 4.14 and Fig. 4.15. At 12:12 the VSLs were activated and the speed and flow considerably reduced in the area with active speed limits (from detector 9 to 24), which corresponds to state 3 in Fig. 4.6. At 12:18 VSLs were switched off and high flows (around 2000 veh/h) getting out of the VSL-controlled area were obtained, which corresponds to state 5 in Fig. 4.6. Both the proposed model and the extended METANET model reproduce the traffic states transition accurately. Quantitative results which include the RMSE of flow, speed and density are shown in Tab. 4.3. The two models have similar performance, and the flow, speed, and density errors of the two models are around 10%. These results indicate that both the proposed model and the extended METANET model can reproduce the effects of VSLs to the evolution of jam waves with a similar accuracy in this case.

Table 4.3: Quantitative results of the models calibration and validation. In the table, the first case shows the calibration results of the proposed model with the quadrangular FD and the third case shows the calibration results of the proposed model with the triangular FD.

Cases	J (10^{-2})	Flow RMSE (veh/h/lane)	Flow error (percent)	Speed RMSE (veh/h/lane)	Speed error (percent)	Density RMSE (veh/h/lane)	Density error (percent)
1. Calibration proposed model	30.3	143.2	11.9	7.4	11.5	2.7	12.0
2. Calibration METANET	30.1	135.4	10.9	6.0	8.4	3.1	13.0
3. Calibration triangular FD	30.9	139.7	11.4	7.6	11.5	2.8	12.2
4. Calibration standard CTM	31.6	152.4	13.1	7.8	13.4	2.8	12.8
5. Validation 1 proposed model	26.2	118.5	8.9	6.3	7.1	1.9	11.8
6. Validation 1 METANET	26.0	109.5	8.0	8.6	9.1	1.6	9.4
7. Validation 2 proposed model	44.7	163.4	10.0	12.1	18.9	4.7	18.5
8. Validation 2 METANET	51.7	202.4	12.2	11.8	18.6	5.8	21.7
9. Validation 2 standard CTM	60.5	219.7	14.5	18.3	40.6	5.9	20.1

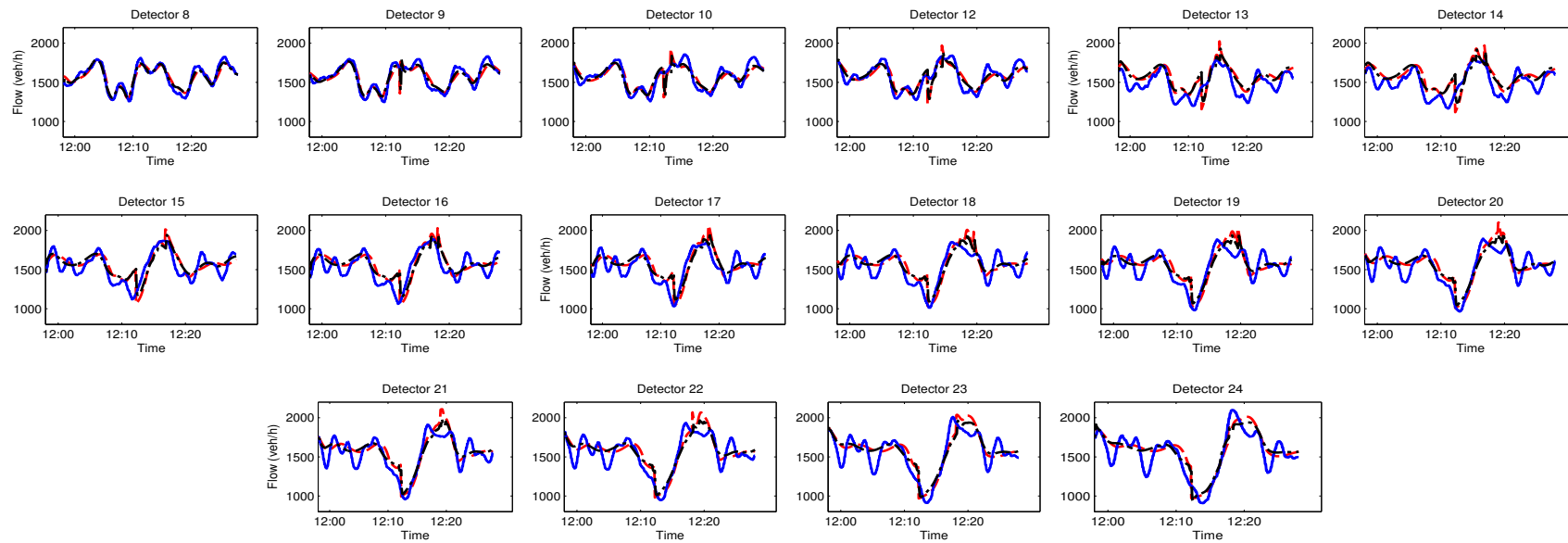


Figure 4.14: Comparison between the real flow data and the validated model results of each detector in case 1. The blue solid line represents the results from real data. The red dashed line and black dash-dotted line represent the results from the proposed model and the extended METANET model respectively.

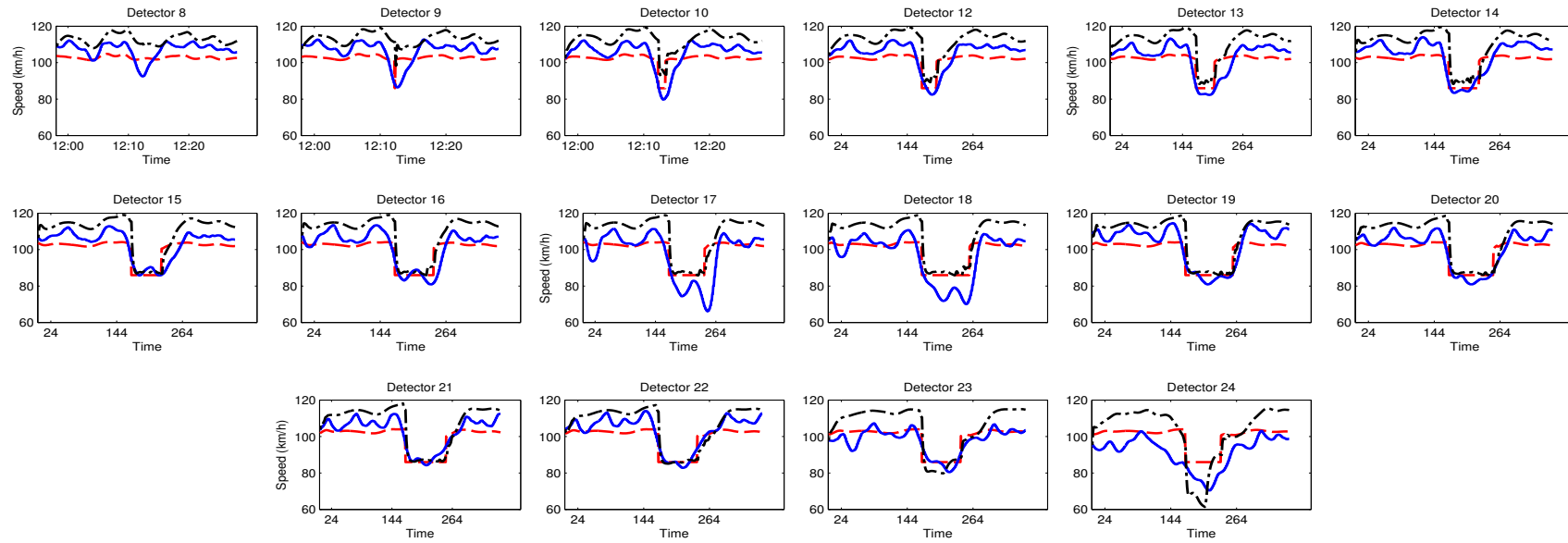


Figure 4.15: Comparison between the real speed data and the validated model results of each detector in case 1. The blue solid line represents the results from real data. The red dash line and black dash-dot line represent the results from the proposed model and the extended METANET model.

The data from the failed application case were obtained from October 30, 2010. A moving jam was detected at 7:20 and the VSL control was switched on. The speed and flow contour plot of the freeway stretch is shown in the left column of Fig. 4.16. Traffic jam originated from detector 20, which may be attributed to a traffic accident. VSLs were activated at the upstream of the traffic jam, but the control was unable to resolve the jam. The reason of the failure might be attributed to the increasing traffic demand during the VSL control period.

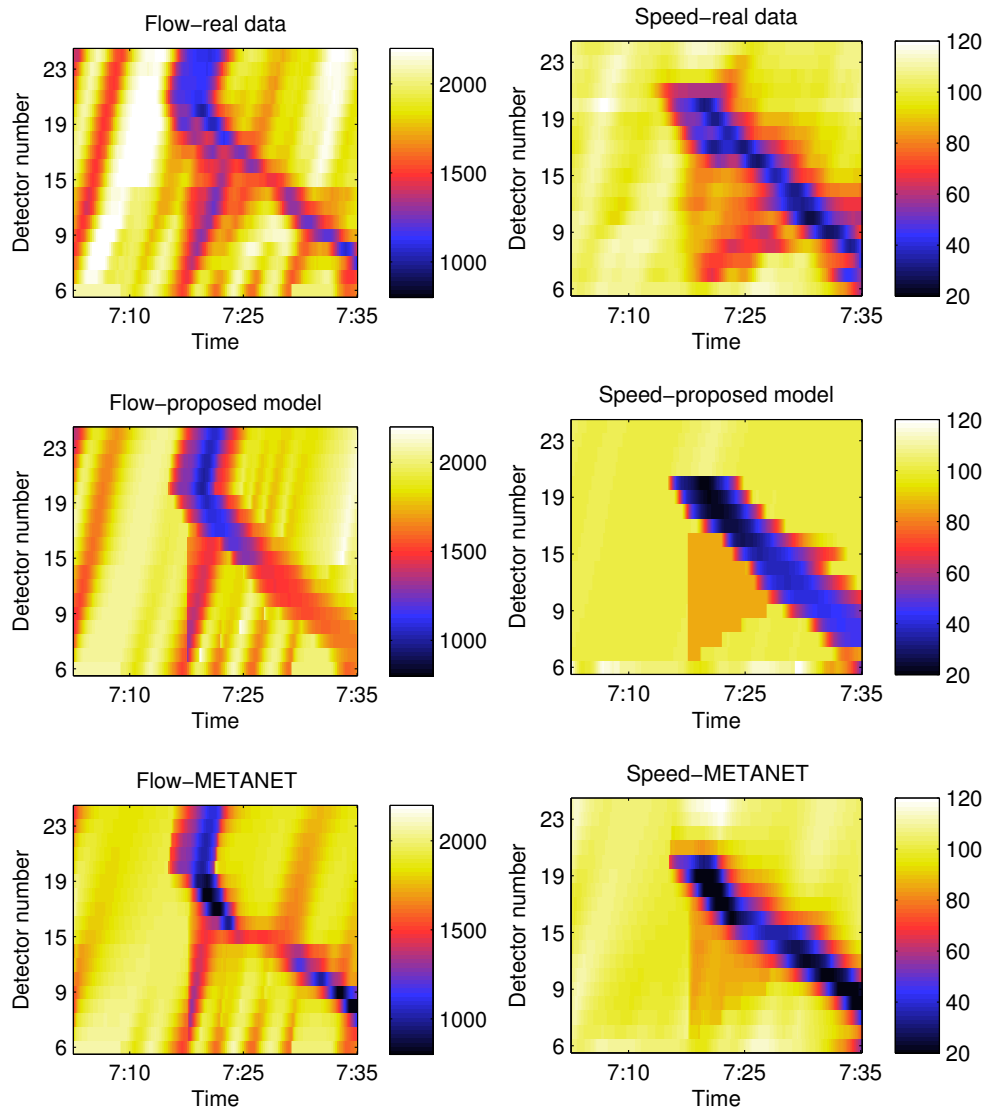


Figure 4.16: The comparison of the real-data for the validation and the results reproduced from the models in case 2. Figures of the first row are plotted from real data, and the second row and the third row are plotted from the simulation results of the proposed model and the extended METANET model.

Boundary data are fed in the models. In addition, the exit flow of cell 20 during the period of traffic breakdown is also fed in the models to trigger the onset of the traffic jam. The simulated speed and flow contour plots are shown in the middle column and the right column of Fig. 4.16. In the simulation results, it can be seen that both the

proposed model and the extended METANET model reproduce the jam wave that is not resolved by the VSL control, which is consistent with the results from real data. *This demonstrates that the VSL models do not overestimate the VSL effects.* The comparison between real flow and speed data and simulated flow and speed data can be seen in Fig. 4.17 and Fig. 4.18. At 7:18 the VSLs were activated and the speed and flow considerably reduced in the area with active speed limits. The VSLs failed to resolve the jam wave, thus low flow and speed, which indicate the jam wave, still existed at detectors (detector 8-20) upstream of the VSL-controlled area after the VSL control. Both the proposed model and the extended METANET model reproduce these facts. The RMSE of flow and speed are shown in Tab. 4.3. Two models have comparable results that the flow error is around 10% and the speed and density errors are around 20 %.

As presented, the standard CTM cannot accurately reproduce the propagation of jam waves. We ran the simulation of the standard CTM in this case, and the quantitative results are shown in Tab. 4.3. It can be seen that the speed error is as high as 40%, which is not as accurate as the proposed model and the METANET model.

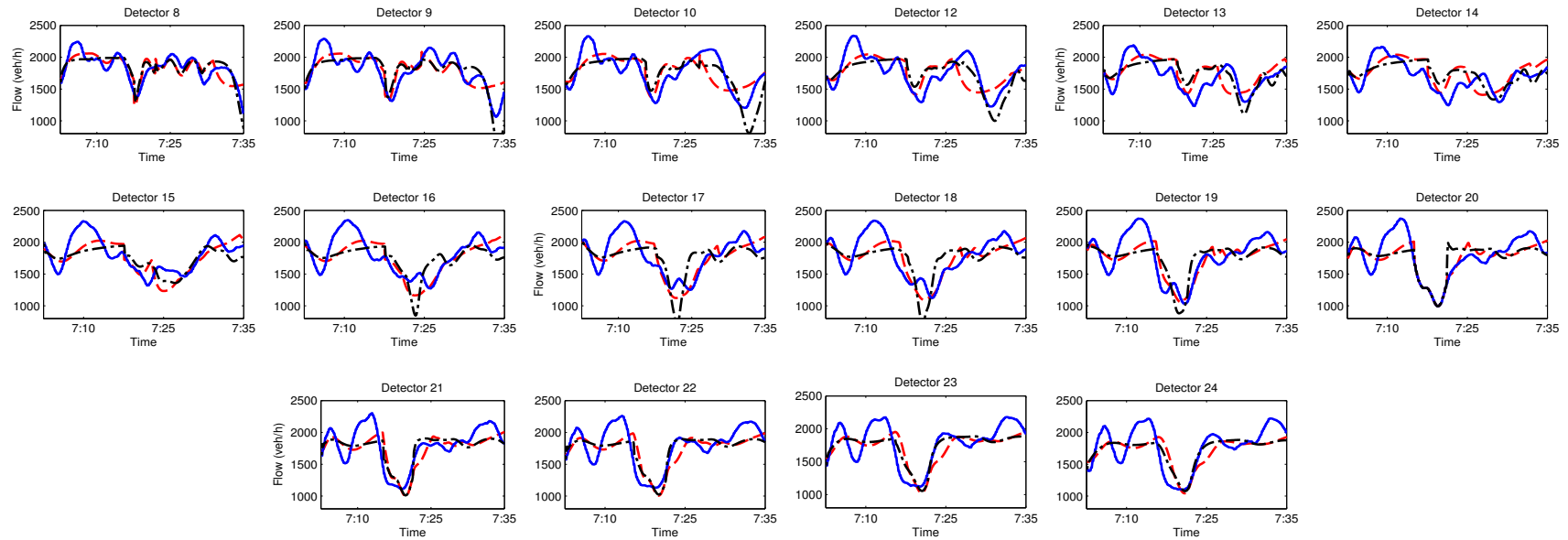


Figure 4.17: Comparison between the real flow data and the validated model results of each detector in case 2. The blue solid line represents the results from real data. The red dash line and black dash-dot line represent the results from the proposed model and the extended METANET model respectively.

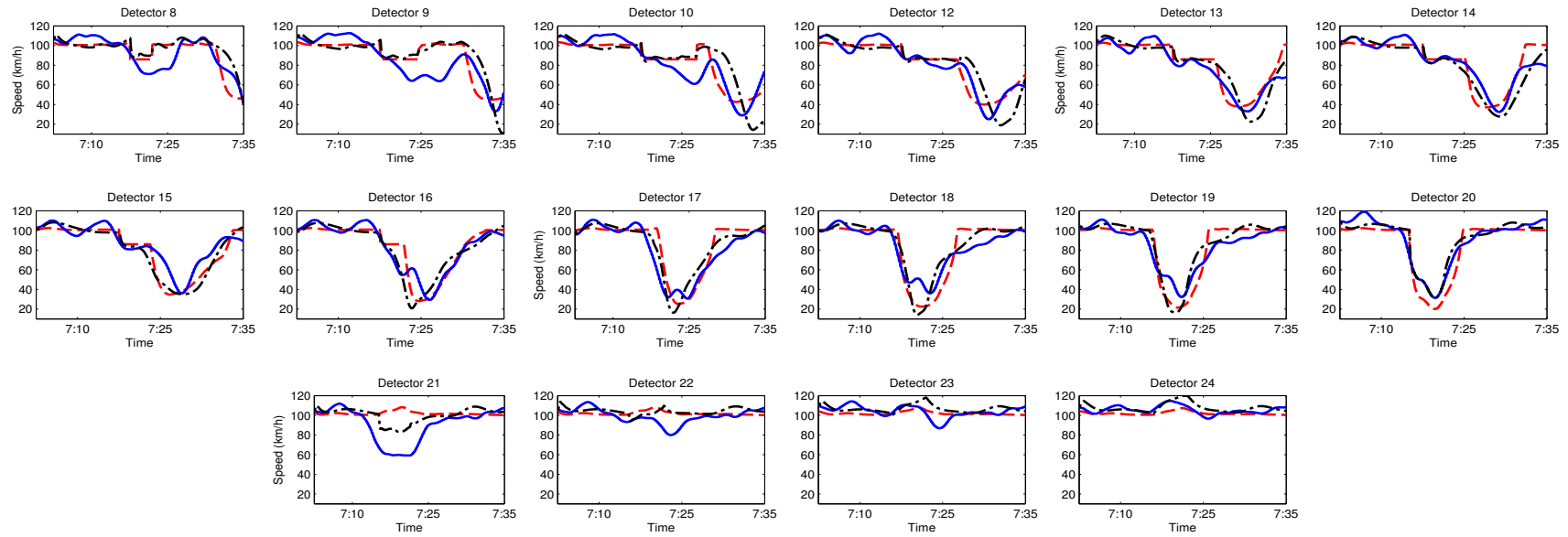


Figure 4.18: Comparison between the real speed data and the validated model results of each detector in case 2. The blue solid line represents the results from real data. The red dash line and black dash-dot line represent the results from the proposed model and the extended METANET model respectively.

4.5 Conclusions

In this chapter, we have validated the prediction model of a MPC of VSLs that is based on a new extended discrete LWR model. We have tested two synthetic cases to show that the model is able to reproduce the key mechanisms that are applied by existing VSL control algorithms. Simulation results show that the model can reproduce the traffic states transition of VSLs used by the SPECIALIST algorithm and the gating effect applied by the mainstream metering control approach. Furthermore, we have calibrated and validated the model with real data to demonstrate its accuracy in terms of predicting the evolution of jam waves under the VSL control. As the target for the model validation, we have chosen a successful case and a fail application from the SPECIALIST field test to show if the model underestimates or overestimate the effects of VSLs. The proposed model has been compared with an extended METANET model which also incorporates with VSLs. The performance of two models are comparable: both models reproduce traffic states transition and jam waves evolution accurately. Quantitatively, both models reproduce accurate results which have a good match with real data. The quantitative results are in line with other model validation works in the literature. Besides, similar quantitative errors between the prediction model and the process model are also obtained in the previous chapter which used the proposed model for a MPC controller to resolve freeway jam waves, which gives confidence of the model validation results. The MPC controller has been demonstrated to be computationally more efficient. Model validation results have demonstrated that the proposed MPC can make accurate prediction on the evolution of jam waves under the VSL control, which is a step towards the real life application.

Chapter 5

Hierarchical ramp metering in freeways: An aggregated modeling and control approach

Recent studies on aggregated traffic modeling of urban networks have shown the existence of a well-defined and low-scatter Macroscopic Fundamental Diagram (MFD) that links network aggregated flow and density. However, the MFD of freeway networks typically exhibits high scatter and hysteresis loops that challenge the monitoring capability and control performance of MFD-based congestion management schemes in freeways. This chapter investigates the effect of density heterogeneity and capacity drop on characteristics of freeway MFD based on field traffic data. In addition, we introduce a model to capture the evolution of density heterogeneity which is essential to reproduce the dynamics of freeway MFD accurately.

The proposed model is incorporated in a hierarchical control structure for coordinated ramp metering on a freeway stretch with multiple congestion bottlenecks and on- and off-ramps. At the upper level, a model predictive control (MPC) approach is developed to optimize total inflow from on-ramps to the freeway stretch, where the prediction model is formulated based on the proposed MFD model. The lower level controller distributes the optimal total inflows to each on-ramp of the freeway based on local traffic state feedback. The proposed ramp metering framework shows desirable performance to reduce the total time spent and eliminate congestion. The control approach is compared with other coordinated ramp metering controllers based on the MPC framework with different prediction models (e.g. CTM and METANET). Simulation outcomes highlight that the MFD-based hierarchical controller (i) is better able to overcome the modeling mismatch between the prediction model and the plant (process model) in the MPC framework and (ii) requires less computation effort than other nonlinear controllers.

This chapter is an edited version of the article:

Han, Y., Ramezani, M., Hegyi, A., Yuan, Y., Hoogendoorn, S. Hierarchical ramp metering in freeways: An aggregated modeling and control approach. *Transportation Research Part B: Methodological*, submitted for review.

5.1 Introduction

The macroscopic fundamental diagram (MFD), which links the accumulation (weighted sum of links densities and lengths) and the production (weighted sum of links flows and lengths) of an urban region, provides an efficient tool for expressing aggregated dynamics of urban traffic networks (Daganzo, 2007; Geroliminis & Daganzo, 2008). Although a unimodal and low-scatter MFD was observed in a homogeneous region of a city (i.e. small spatial density heterogeneity), existence of a well-defined MFD in general cases is still an open question, which might undermine the benefits of MFD in modeling traffic dynamics. For example, studies of Mazlounian et al. (2010); Geroliminis & Sun (2011b); Gayah & Daganzo (2011) have shown that urban networks with heterogeneous distribution of density exhibit network flows smaller than those that approximately meet homogeneous conditions. Thus, recent studies focus more on modeling and controlling dynamics of heterogeneity in urban networks (Ramezani et al., 2015; Yildirimoglu et al., 2015).

The MFD has been utilized to introduce elegant city-scale traffic control strategies, e.g. Perimeter control, to improve mobility and decrease delays in large-scale urban networks. In the literature, different control approaches have been used to solve the perimeter control problems. For example, classical feedback control approaches have been implemented by Keyvan-Ekbatani et al. (2012, 2015); Kouvelas et al. (2017); Ampountolas et al. (2017), and the model predictive control (MPC) approach has been used to solve the optimal control problems in (Geroliminis et al., 2013; Haddad et al., 2013; Ramezani et al., 2015). Moreover, methods of MFD estimation are studied in (Leclercq et al., 2014; Laval & Castrillón, 2015). Recently, studies of Haddad & Shraiber (2014); Haddad (2015) designed robust perimeter controllers to systematically take into account uncertainties in MFD-based dynamics. MFD-based controllers require less detailed information on traffic states compared to microscopic/mesosopic model-based traffic controllers. The MFD substantially reduces the complexity of traffic models because the packets of vehicular traffic are considered as a single continuum entity. Furthermore, MFD-based controllers need less computation effort that is beneficial for real-time applications.

Although MFD-based controllers have been investigated in urban networks, they have not been scrutinized for freeway traffic control. It has been found that a well-defined and low-scatter MFD does not always exist for freeways (Daganzo, 2011). Cassidy et al. (2011) showed that a well-defined and low-scatter freeway MFD can be observed only when traffic is in stable regime (either congested or non-congested) in all lanes and on all links. Studies of Ji et al. (2010); Saberi & Mahmassani (2012); Geroliminis & Sun (2011a) have shown that freeway MFD exhibits hysteresis pattern during the onset and offset of congestion, which represents an adverse effect on the freeway overall performance. In these studies, the heterogeneity of density distribution was perceived as a main factor that causes the hysteresis pattern. Knoop & Hoogendoorn (2013) defined the spatial spread of density as a function of the weighted variance of

the densities in all sections, and fitted a third-order polynomial functional form to field traffic data obtained from a Dutch freeway. It was found that by including the density heterogeneity, the prediction accuracy of the MFD improves significantly. In Geroliminis & Sun (2011a), it was found that non-equilibrium states in individual detectors' measurements in freeways, which include transition flows and the capacity drop, also influence the production of the MFD. The traffic patterns in freeway networks have distinct characteristics, e.g., route choice and congestion propagation, that challenge the traffic congestion modeling and management. Thus, understanding and modeling the dynamics of density heterogeneity is crucial for developing MFD-based traffic control schemes. This chapter presents a systematic analysis to quantify the effect of determining traffic state factors on MFD and offers a tractable coordinated freeway congestion control scheme.

Ramp metering control is a common operation measure in freeways to regulate flows from on-ramps. There is an extensive literature on ramp metering control. Early works attempted to alleviate local traffic problems using local ramp metering (Papageorgiou et al., 1991). It has been identified that local ramp metering approaches are not always efficient to ameliorate global traffic conditions of the overall traffic network (Papageorgiou & Kotsialos, 2000). Since then, various advanced coordinated ramp metering methods using MPC approaches that are based on macroscopic traffic flow models have been presented. MPC framework contains a prediction model that predicts the evolution of traffic dynamics and estimates the optimal control scheme for the time period in which the relevant traffic dynamics occur. This feature enables the controller to take advantage of potentially larger future gains at a current (smaller) cost, so as to avoid myopic control actions. Studies of Kotsialos & Papageorgiou (2004); Hegyi et al. (2005a); Papamichail et al. (2010) presented different optimal approaches of ramp metering that were based on the METANET model. Gomes & Horowitz (2006) presented an optimal freeway ramp metering approach that was based on the asymmetric cell transmission model, which can be efficiently solved by linear programming. Han et al. (2015b) presented a linear quadratic MPC approach for ramp metering based on an extended cell transmission model (CTM) which took the capacity drop into account. For a detailed description of METANET and CTM, readers are referred to Kotsialos & Papageorgiou (2004); Daganzo (1994). The aforementioned studies used same models as the prediction model and the process model which represents the reality. However in real life applications, there is a significant mismatch between the dynamics of the prediction model and the field process. Hence, developing a traffic model that demonstrates robustness to modeling mismatch in the MPC framework will increase the confidence that the proposed approach will also accurately work in practice, where there is always a model mismatch. This is a challenging task and is still not sufficiently discussed in the literature.

The objectives of this chapter are twofold, and are related to modeling and control aspects. Regarding the modeling aspect, we introduce a model to capture the evolution of density heterogeneity which is essential to reproduce the dynamics of freeway MFD

accurately. By analyzing field data, we find that the density heterogeneity is (i) a linear function of the average density when no congestion occurs in the freeways and (ii) is a non-linear function when congestion occurs. The presented model captures this phenomenon and estimates the density heterogeneity accurately. Furthermore, we present a quantitative measure to represent the effect of capacity drop on the MFD. The MFD is represented by an exponential functional form, which is fitted to and tested with field data. Regarding the control aspect, we present a hierarchical control approach for freeway ramp metering. At the upper level, a MPC approach is developed to optimize total inflows from on-ramps to the freeway stretch, where the prediction model is formulated based on the proposed MFD model. At the lower level, the controller distributes the optimal total inflow to each on-ramp of the freeway based on local traffic state feedback. The presented controller is tested with different models (e.g. CTM and METANET) as the process model and compared with other optimal control approaches. The results demonstrate that the MFD-based controller achieves desirable performance to reduce the total time spent. It is of particular significance that the simulation results show that the proposed controller performs better with an inherent model mismatch.

The remaining of the chapter is set up as follows. The modeling of the freeway MFD is presented in Section 5.2. In Section 5.2.1, we explore the relation between network production and network outflow through field data. Sections 5.2.2 and 5.2.3 present the density heterogeneity model and the capacity drop model respectively. The accuracy of the MFD is tested with field data in Section 5.2.4. Section 5.3 describes the control design, where the upper level MPC controller and the lower level controller are presented. In Section 5.4, numerical experiments to test the performance of the proposed controller are presented. A summary and the discussion of future work conclude the chapter in Section 5.5.

5.2 Modeling of the freeway MFD

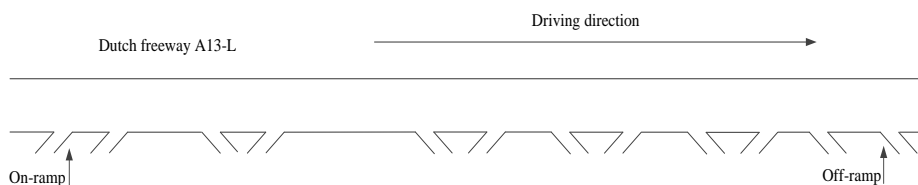


Figure 5.1: The sketch of Dutch freeway A13-L. The freeway stretch is about 16km in length and includes 6 on-ramps and 6 off-ramps.

The MFD considered in this chapter represents the aggregated traffic behavior of a freeway stretch, which has one mainstream origin and destination, and several on-ramps and off-ramps with corresponding origins and destinations demands. The field data of the Dutch freeway A13-L are used for empirical analysis. As shown in Fig. 5.1,

the freeway stretch is about 16 km in length including 6 on-ramps and 6 off-ramps. 38 detector stations are placed at the freeway mainstream, while the distances between the detector range from 300 m to 600 m. The data consist of detected flows and time mean speeds, respectively denoted as q_s (veh/h) and v_s (km/h), of lane s at each detector station during a minute, from March and April, 2013.

The freeway stretch is divided into segments, where we consider that the middle of two neighboring detectors is the boundary of a segment. Speeds over lanes in each segment i are harmonically averaged to approximate the space mean speed v_i ,

$$v_i = \left(\frac{\sum_s q_s \frac{1}{v_s}}{\sum_s q_s} \right)^{-1}. \quad (5.1)$$

The detected flows over lanes in each segment are also aggregated, and the aggregated flow of segment i is denoted as q_i . q_i and v_i are aggregated for every 5 minutes. Then, the average density of each segment, i.e. ρ_i , is estimated as the ratio of q_i and v_i based on the fundamental relation. The MFD links the weighted averaged flow, \bar{q} , and the weighted average density, $\bar{\rho}$, which are estimated as,

$$\bar{q} = \frac{\sum_{i=1}^I q_i l_i}{\sum_{i=1}^I l_i}, \quad (5.2)$$

$$\bar{\rho} = \frac{\sum_{i=1}^I \rho_i l_i}{\sum_{i=1}^I l_i}, \quad (5.3)$$

where l_i is the length of segment i and I is the number of segments in the freeway stretch. Note that the denominators of (2) and (3) are the total length of the freeway stretch, and the numerators of (2) and (3) are respectively the so-called production (represented by P) and accumulation (represented by A) of the freeway stretch. The MFD links P and A , or interchangeably \bar{q} and $\bar{\rho}$.

Fig. 5.2 (a) shows the freeway MFD which is obtained from loop detector data on April 25, 2013. It can be seen that the freeway MFD is unimodal, with less scatter in the left part and high scatter in the right part. The left part of the MFD represents the aggregated traffic states of the freeway when no congestion occurs, which often happens during off-peak hours. The right part of the MFD represents the aggregated traffic states of the freeway when congestion occurs, which often happens during peak hours.

For traffic control purposes on freeways, we are interested in the total outflow (network exit function) of the freeway stretch (represented by G , see Fig. 5.2 (b)) which is essential to predict traffic dynamics through mass (the number of vehicles) conservation equations. The total outflow of the freeway stretch includes the flows that leave

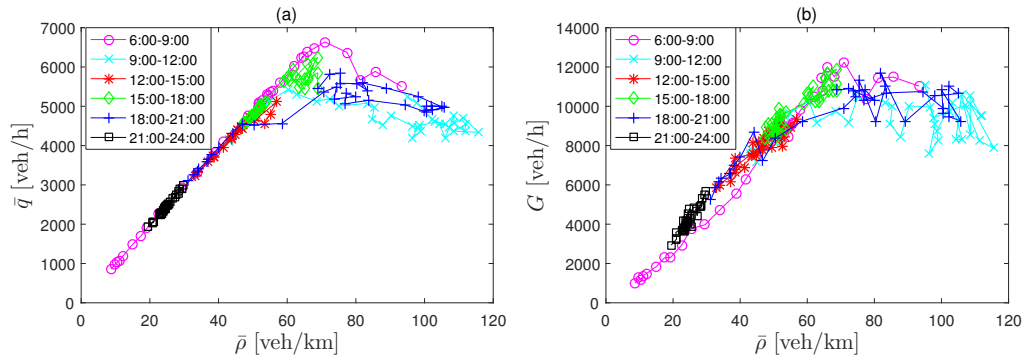


Figure 5.2: Aggregated loop detector data of April 25, 2013: (a) weighted average flow versus weighted average density; (b) total outflow versus weighted average density. The shapes in (a) and (b) are unimodal, with less scatter in the left part and high scatter in the right part. Note that in (b), some data points from 6:00 to 9:00 have lower values of the total outflow compared to the traffic condition from 21:00 to 24:00 due to longer average trip lengths, see Fig. 5.3.

the freeway stretch at the downstream end and the off-ramps. The relation between the production and the total outflow of the freeway stretch is investigated in the next section.

5.2.1 Relationship between network production and outflow

The analysis of field data of Yokohama in Geroliminis & Daganzo (2008) shows that the ratio of the production and the trip completion rate (network outflow) is relatively constant over time. The ratio can be perceived as the vehicular average trip length (represented by \bar{x}) in the network. This relationship can be readily adapted to the freeway system. For example, suppose a pair of on-ramp and off-ramp are close to each other and are located in the middle of a freeway stretch. We consider two situations that occur in the freeway stretch: Situation 1, no traffic is using either the off-ramp or the on-ramp, and all the vehicles are going out of the freeway stretch at the downstream end; Situation 2, all of the vehicles at the upstream of the off-ramp are going out the freeway stretch through the off-ramp and the same amount of vehicles are entering the freeway stretch from the on-ramp and going out at the downstream end. For these two situations, the following relationships hold: $A_1 \approx A_2$, $P_1 \approx P_2$, $\bar{x}_1 \approx 2\bar{x}_2$, and $G_1 = \frac{1}{2}G_2$. For the same accumulation, the productions of the two situations are the same, but the total outflow of situation 2 is two times higher than that of situation 1 while the average trip length in situation 2 is half of that in situation 1. Thus, the total outflow is not only a function of the accumulation, but also relevant to the average trip length.

Data from the Dutch freeway A13-L are explored to investigate the relation between network production and outflow. Since there is no detector placed at off-ramps, off-ramp flows are estimated as the flow difference of detectors directly upstream and

downstream of the off-ramps. The ratio between P and G is shown as a box plot in Fig. 5.3, where each box represents data of a time slot of 5 minutes that are obtained from all weekdays in March and April, 2013. On each box, the central mark indicates the median, and the bottom and top edges of the box indicate the 25th and 75th percentiles, respectively. The whiskers in each time slot extend to the most extreme data point.

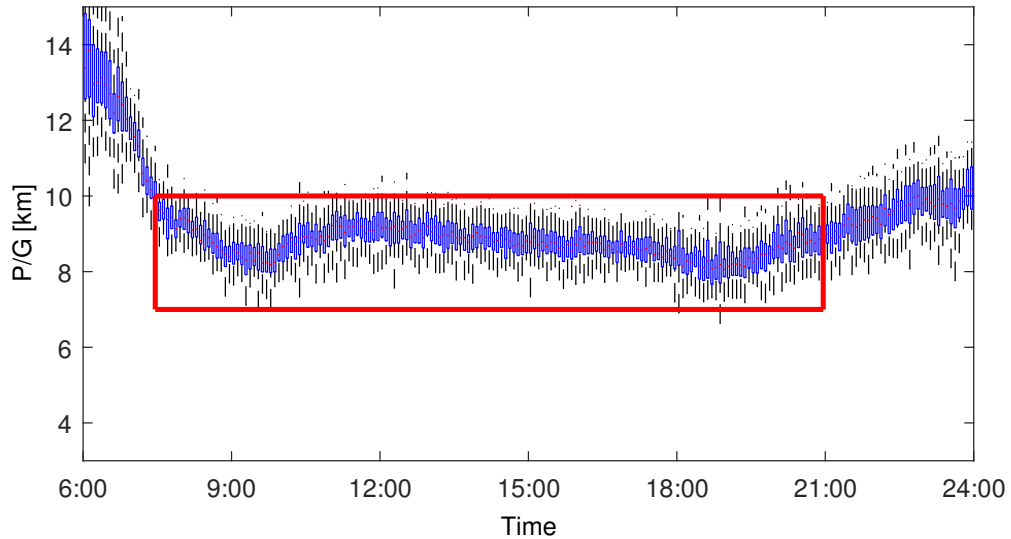


Figure 5.3: The whisker plot of P/G aggregated for all of the weekdays during March and April 2013. On each box, the central red mark indicates the median, and the bottom and top blue edges of the box indicate the 25th and 75th percentiles, respectively. The whiskers extend to the most extreme data points. In the rectangle area, the values of P/G are relatively time-invariant, which indicates that the average trip lengths are nearly time-invariant. The average trip lengths from 6:00 to 7:30 are longer, which results in lower values of total outflows, see Fig. 5.2 (b).

Fig. 5.3 demonstrates that the average trip length is relatively time-invariant from 7:30 to 21:00, which is in the range 8.5 ± 1 km. Traffic volumes at late evening and early morning are low and less travelers exit the freeway stretch through off-ramps, which result in longer average trip lengths. At peak hours (around 9:00 and 18:00) the average trip lengths are slightly shorter because some travelers might exit the freeway stretch through off-ramps once congestion occurs downstream. If the average trip length is constant over time, then it is physically equivalent to use G and P as the performance indicator in the MFD. The MFD which links the total outflow and the weighted average density is shown in Fig. 5.2 (b). It can be seen that the scatter of two MFDs have similar features.

Note that if the average trip length is time-variant, to accurately predict the total outflow of the freeway stretch, the dynamics of the average trip length also needs to be predicted. The average trip length can be predicted based on dynamic OD matrix, and this is beyond the scope of this chapter. A general assumption for the rest of this chapter is that the average trip length is time-invariant.

5.2.2 Modeling the density heterogeneity

A major factor of MFD scatter, which has been identified in the literature (e.g. [Mazlounian et al. \(2010\)](#)), is the spatial density heterogeneity. In this chapter, we use the standard deviation of density per lane to represent the density heterogeneity, σ . This is formulated as,

$$\sigma = \sqrt{\frac{1}{I} \sum_{i=1}^I \left(\frac{\rho_i}{\lambda_i} - \frac{\sum_{i=1}^I \frac{\rho_i}{\lambda_i}}{I} \right)^2} \left(\frac{\text{veh}}{\text{km} \cdot \text{lane}} \right), \quad (5.4)$$

where λ_i is the number of lanes in segment i . Empirical studies of freeway networks show that for same values of accumulations, higher density heterogeneity results in lower network production ([Geroliminis & Sun, 2011a](#); [Knoop & Hoogendoorn, 2013](#)). To demonstrate the significance of the density heterogeneity in MFD modeling, [fig. 5.4](#) shows the aggregated loop detector data of April 16, 2013, in which (a) represents the weighted average flow as a function of the average density and (b) represents the heterogeneity as a function of the average density. [Fig. 5.4 \(a\)](#) and (b) are perceived as the combination of two parts, (i) data points on a straight line and (ii) a more complex pattern of hysteresis loops. For the same value of the average density, the points on the straight line have lower values of density heterogeneity and higher values of average flow than other points. The points can be divided into two groups, where one is from the morning peak (within the time period 9:00-12:00) and the other is from the evening peak (within the time period from 18:00-21:00). In this example, generally, data from the morning peak have lower values of the density heterogeneity and higher values of average flow than data from the evening peak (for same values of the average density). The points from the evening peak exhibit hysteresis pattern, which in [Fig. 5.4 \(a\)](#) is in a clockwise shape and in [Fig. 5.4 \(b\)](#) is in a counter-clockwise shape. All these observations indicate that higher density heterogeneity results in lower network production.

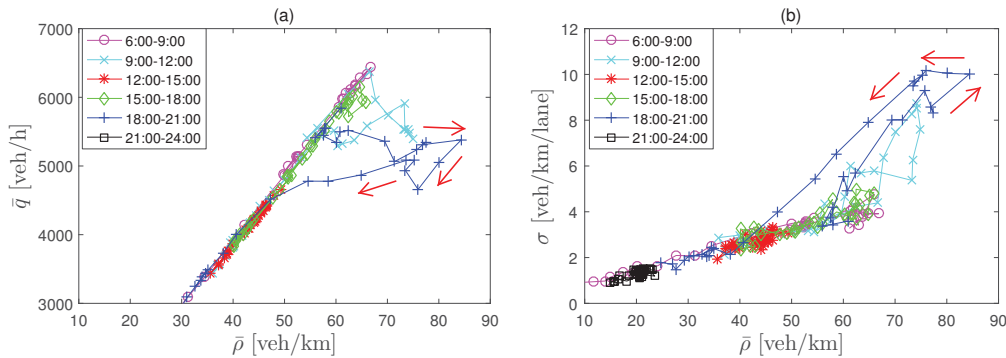


Figure 5.4: Aggregated loop detector data of April 16, 2013: (a) weighted average flow versus weighted average density; (b) density heterogeneity versus weighted average density. Red arrows in the figures represent the direction of time in the hysteresis loop.

Ramezani et al. (2015) developed a traffic congestion control strategy to reduce the heterogeneity in urban networks. They assumed that a well-defined relationship between the average occupancy and the variance of occupancy exists in a homogeneous sub-region of the urban network. However in a heterogeneous freeway network, the density heterogeneity and the average density are strongly correlated with a complex pattern, see Fig. 5.4 (b). Hence, characteristics of the dynamics of density heterogeneity, which may have a significant influence to the control performance, are needed to be further investigated and properly modeled.

The relation between σ and $\bar{\rho}$ can be perceived as the combination of two parts represented by a linear function and a non-linear function. If no congestion occurs in the freeway, the data points fall into the linear function which is represented as,

$$\sigma(t) = s_1 \cdot \bar{\rho}(t), \text{ if } \forall i \in [1, 2, \dots, I], \rho_i(t) \leq \rho_i^{\text{cr}} \quad (5.5)$$

where i is the index of segments of freeways and s_1 is the positive model parameter. ρ_i^{cr} is the critical density of segment i . It is assumed that if the density of every segment in the freeway is lower than the critical density, then no congestion occurs in the freeway. Thus if no congestion occurs in the freeway, the density heterogeneity increases linearly as the average density increases. Once congestion occurs, the density heterogeneity evolves in a non-linear way which can be represented by high-order polynomial functions.

The value of the average density when congestion occurs in the freeway is not always the same because it depends on different traffic situations, for example, the activation of different bottlenecks, the blockade of different off-ramps, and the occurrence of moving jams under different traffic loadings. Thus, the threshold of the model switching to the non-linear function is not constant. If no congestion is detected, the model is represented by the linear function (5.5) and once congestion is detected, the model switches to the non-linear function (5.8). Note that depending on different traffic situations, the scatter points of the density heterogeneity are widely dispersed in σ - $\bar{\rho}$ plane. Thus, we propose a non-linear function that includes a time varying parameter, $\bar{\rho}^{\text{cr}}(t)$, to represent the maximum average density that can be represented by the linear function under the traffic situation at time t . $\bar{\rho}^{\text{cr}}(t)$ is formulated as,

$$\bar{\rho}^{\text{cr}}(t) = \sum_{i=1}^I \frac{\tilde{\rho}_i(t) l_i}{L}, \quad (5.6)$$

where L is the total length of the freeway stretch, and

$$\tilde{\rho}_i(t) = \begin{cases} \rho_i(t), & \text{if } \rho_i(t) \leq \rho_i^{\text{cr}} \\ \rho_i^{\text{cr}}, & \text{if } \rho_i(t) > \rho_i^{\text{cr}} \end{cases} \quad (5.7)$$

The non-linear function is formulated as,

$$\sigma(t) = s_2 \cdot (\bar{\rho}(t) - \bar{\rho}^{\text{cr}}(t))^3 + s_3 \cdot (\bar{\rho}(t) - \bar{\rho}^{\text{cr}}(t))^2 + s_4 \cdot (\bar{\rho}(t) - \bar{\rho}^{\text{cr}}(t)) + s_5, \quad (5.8)$$

if $\exists i \in [1, 2, \dots, I], \rho_i(t) > \rho_i^{\text{cr}}$

where s_2 , s_3 , s_4 , and s_5 are model parameters.

To investigate the proposed heterogeneity model, we test the data from April 16, 2013. The time varying parameter $\bar{\rho}^{\text{cr}}(t)$ is estimated at every time instance based on (5.6). The parameters s_1 to s_5 estimation is formulated as a non-linear optimization problem that aims to minimize the root mean square error (RMSE) between the model estimations and real training data. The estimation results are shown in Fig. 5.5 (a), where the blue line represents the linear function (5.5) and the green lines represent the non-linear function (5.8) for different values of $\bar{\rho}^{\text{cr}}$. Each of the green lines represents the non-linear heterogeneity function at each time instance in the afternoon peak, which corresponds to the time instants of the data points from 18:00 to 21:00 in the counter-clockwise hysteresis loop in Fig. 5.4 (b). The comparison between real data and model estimation is shown in Fig. 5.5 (b). It can be seen that the model well reproduces the scatters and hysteresis of the density heterogeneity. If the model is embedded into the MPC framework, $\bar{\rho}^{\text{cr}}$ is assumed to be constant during the control horizon, as its dynamics is difficult to predict because the state of individual segment, ρ_i , is not forecast by the MFD. The accuracy of the model with this assumption is tested in Section 5.4.

5.2.3 The effect of capacity drop on freeway MFD

In addition to the density heterogeneity, it has been found that the freeway MFD is also affected by the non-equilibrium states in individual detectors' measurements, which include transition flows and capacity drop (Geroliminis & Sun, 2011a). In this Section, we explore the effect of the capacity drop on the weighted average flow of a freeway stretch through a synthetic case. The simulation model of the synthetic case is an extended cell transmission model (CTM) that is capable of reproducing the capacity drop and propagation of jam waves (Han et al., 2016b). The extended CTM reproduces different extent of the capacity drop by assuming that the outflow of a jam is lower if the density of the jam head is higher. A simulation test is conducted on a synthetic 3-lane freeway stretch which is 16 km in length and divided into 40 cells. Four on-ramps are placed on the freeway stretch, which are potential bottlenecks. The layout of the synthetic freeway stretch is shown in Fig. 5.6.

To demonstrate the influence of capacity drop on freeway MFD, we consider two scenarios where $\bar{\rho}$ and σ are equal or similar as much as possible, but with different extents of the capacity drop. According to the mechanism of the model that the extent of capacity drop depends on the densities of the jam heads, we create two jams with different jam heads densities, which are shown in Fig. 5.7 (a) and (b) respectively. In scenario 1, the congestion is triggered at on-ramp 2 and the density at the head of the congestion is about 35 [veh/km/lane]. In scenario 2, the congestion is triggered at on-ramp 3, and the density at the head of the congestion is about 50 [veh/km/lane]. The relationship between \bar{q} and $\bar{\rho}$, and σ and $\bar{\rho}$ of the two scenarios are plotted in Fig. 5.7 (c) and (d) respectively. It can be seen that when the average density is around 60 [veh/km], two scenarios have equal density heterogeneity. However, the average flow

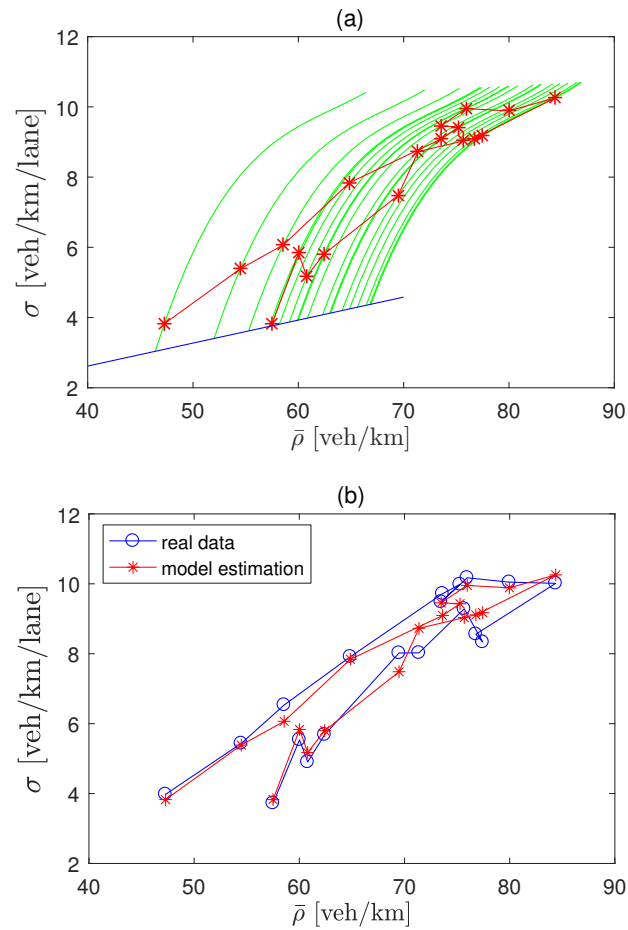


Figure 5.5: (a) The estimation results of the density heterogeneity model. The blue line represents (5) and the green lines represent (8). The red points are the estimated density heterogeneity at each time instance. (b) The comparison between real data and model estimation. Data are obtained from April 16, 2013. Note that the x axis in the two figures are different, where in (a) it is $\bar{\rho}$ and in (b) it is $\bar{\rho}^{ct}$.

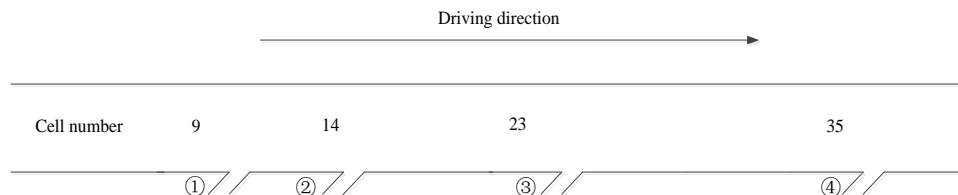


Figure 5.6: The layout of the synthetic freeway stretch that is used to model the effect of capacity drop on MFD. Cell number are marked at the upstream of each on-ramp. Circled numbers represent the order of the on-ramps.

of scenario 2 is significantly lower (5.5%) than the one of scenario 1, as a result of higher capacity drop.

It is not trivial to quantify the influence of capacity drop on the average flow of MFD. Empirical studies suggested that the extent of the capacity drop is larger if the speed in the jam is lower (Yuan et al., 2015). Thus, the speed in the jam is considered as a factor to quantify the effect of the capacity drop. Besides, if the number of traffic jams is large, the influence of the capacity drop is also supposed to be large. Accordingly, we introduce a variable η to quantify the effect of the capacity drop, which is represented as,

$$\eta = \sqrt{\sum_{j=1}^J (v^{\text{cr}} - v_j^{\text{jam}})^2}, \quad (5.9)$$

where j is the index of traffic jams, and J is the number of traffic jams occurring in a freeway stretch. v_j^{jam} is the speed in jam j , and v^{cr} is the critical speed associated with the critical density and the free-flow capacity. We use the square root of the summation because as the number of jam increases, the density heterogeneity is also expected to increase, which will dilute the influence of η . If the speeds in an area are below the critical value and both upstream and downstream of the area have free flow speed, then the area is perceived as a jam area. v_j^{jam} can be calculated as the average speed in the jam area. Note that first-order traffic flow models (e.g., CTM) consider the density as the state variable. If these models are used, equation (5.9) can be reformulated by using ρ_i^{jam} and ρ^{cr} because there is a one-to-one relationship between those variables.

5.2.4 Functional form of the freeway MFD model

In (Ramezani et al., 2015), an exponential functional form is found to be a good estimation of MFD of heterogeneous urban networks. In this chapter we use the same functional form, which is formulated as,

$$\bar{q} = (D_3 \cdot \bar{\rho}^3 + D_2 \cdot \bar{\rho}^2 + D_1 \cdot \bar{\rho}) \cdot \left(a \cdot e^{b_1 \sigma + b_2 \eta} + (1 - a) \right) \quad (5.10)$$

where, a , b_1 , b_2 , D_1 , D_2 and D_3 are estimated parameters. The function is a product of a third-order polynomial function which represents a low scatter MFD (i.e. the upper MFD envelope that is associated with the most homogeneous traffic condition) and an exponential function which represents the heterogeneity and capacity drop effects. Note that the choice of the third-order polynomial function is a balance between complexity and accuracy. A second-order function is not able to represent the low scatter MFD because the critical density will be half of the jam density, which is not realistic. Higher order functions increase the complexity of the MFD model.

To examine the accuracy of the proposed MFD model, the parameters of function (5.10) are estimated based on real data which are obtained from Dutch freeway A13-L. The field data cover March and April 2013. Data from April are used to estimate

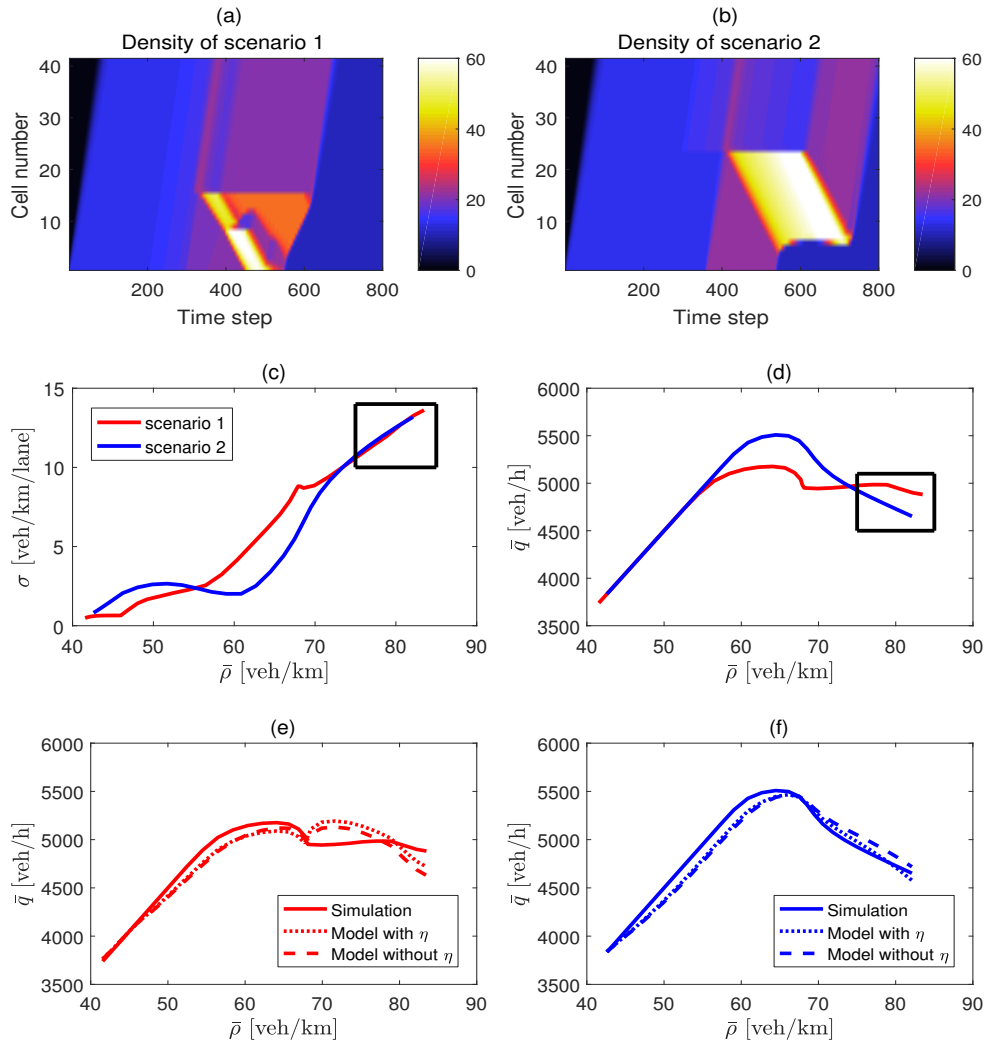


Figure 5.7: (a) and (b): The density contour plots of scenario 1 and 2 in the synthetic case. The density of the jam head in (a) is around 35 (veh/km/lane) and in (b) is around 50 (veh/km/lane). (c): The $\bar{\rho}$ - σ curves of the two scenarios. (d): The $\bar{\rho}$ - \bar{q} curves of the two scenarios. In (c) and (d), when $\bar{\rho}$ of the two scenario are around 73 (veh/km), the values of σ in the two scenarios are similar but the value of \bar{q} in scenario 1 is substantially higher than \bar{q} in scenario 2. (e) and (f) show the accuracy of the MFD model (10) with and without η . The solid lines represent the MFD of the simulation. The dotted lines and dashed lines respectively represent the MFDs with and without η in (10).

the parameters in (5.10), while the data from March are used to test the performance. $\bar{\rho}$, σ and η are calculated respectively based on (5.3), (5.4), and (5.9). The MFDs with and without η are estimated and tested respectively. The MFD curves of one of the days from the estimation data set (i.e. in April) and from the testing data set (i.e. in March) are shown in Fig. 5.8. The root mean square error (RMSE) of two MFDs are shown in Tab. 5.1. It can be seen that for both MFDs, the estimation error and the testing error are below 5 %. Thus, both MFDs represent the average flow accurately, and the MFDs with η further improves the accuracy. The estimation and testing results of the MFD without σ and η are also shown in Tab. 5.1. It can be seen that the errors are significantly higher when the MFD model does not address the effect of heterogeneity and capacity drop. Note that unlike the synthetic case in which we consider different densities of jam heads, the day-to-day pattern of traffic congestion is similar in this field data. This is the reason why the accuracy of the MFD with the introduction of η does not improve significantly. To highlight the significance of η , we estimate the parameters of the MFDs based on simulation data in the previous section. The estimation results are shown in Fig. 5.7 (e) and (f). It can be seen that both MFDs represent the critical point accurately, while the MFD with η reproduces \bar{q} more accurately at the part where both $\bar{\rho}$ and σ have similar values (the rectangle area).

Table 5.1: The estimation and testing results of different MFD models. The field data from April are used to estimate the parameters in (10), i.e. the training set, while the data from March are used to test the performance. In estimation of the MFD without σ and η , b_1 and b_2 are equal to 0. In estimation of the MFD with σ and without η , b_2 is equal to 0.

	Estimation RMSE (veh/h)	Estimation error percentage (%)	Testing RMSE (veh/h)	Testing error percentage (%)
MFD with both σ and η	160.7	4.28	176	4.65
MFD with σ and without η	167.5	4.55	184.8	4.87
MFD without σ and η	273.4	6.31	333.8	8.58

5.3 Control design

In this section, a hierarchical ramp metering approach is designed to improve the free-way traffic operation efficiency. At the upper level of the controller, a model predictive control (MPC) approach is applied to determine the optimal total inflow from the on-ramps to the freeway stretch. The upper level MPC approach is based on the proposed MFD model. The lower level controller locally distributes the optimal total inflow that is determined by the upper level controller to each on-ramp of the freeway.

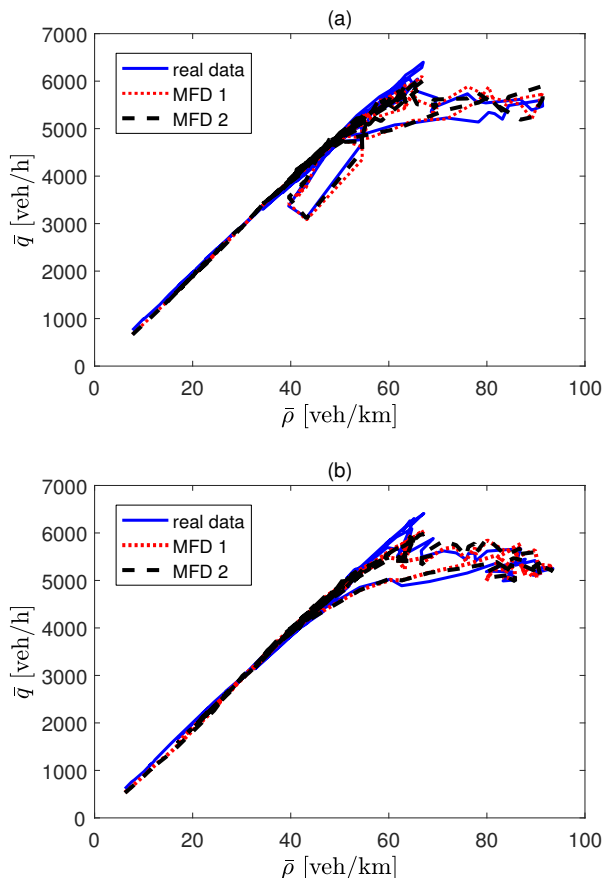


Figure 5.8: The estimated MFDs for (a) April 2 and (b) March 27. MFD-1 is the MFD with η and MFD-2 is the one without η . The proposed freeway MFD model can capture the complex hysteresis patterns.

5.3.1 The upper level controller: The MPC approach

The MPC approach has largely been used in freeway traffic control problems (e.g. Hegyi et al. (2005a); Haddad et al. (2013)). The MPC approach utilizes a traffic model to predict traffic states evolutions based on the current states of the system, and determine the optimal control actions that result in the optimum value of an objection function. This feature enables the controller to take advantage of potentially larger future gains at a current (smaller) cost, so as to avoid a myopic control actions. After optimization, the control value of the first sample of the optimal control action is applied to the process. The remaining part of the control signal is recalculated in the finite rolling horizon scheme. For a detailed description of the MPC approach, readers are referred to (Camacho & Bordons, 2012).

In this chapter, the proposed MFD model is used as the prediction model of the MPC approach. The model predicts the dynamics of the freeway stretch accumulation A and the total number of vehicles R at all on-ramps. The dynamics of A is formulated as,

$$A(k+1|t) = A(k|t) + (U(k|t) - G(k|t) + f_1(k|t)) \cdot T_c, \quad (5.11)$$

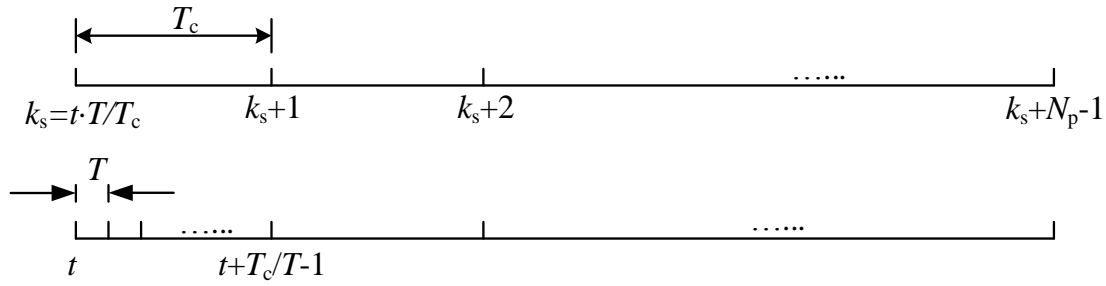


Figure 5.9: Time scales of the process model (top) and the prediction model (bottom).

where f_1 is the flow from the freeway mainstream origin, t is the discrete time index for the process model and k is the discrete time index for the prediction model, see Fig. 5.9. For simplicity, we assume that the duration of the discrete time step of the prediction model is the same as the duration of the control time step, i.e. T_c . We assume that $\frac{T_c}{T}$ is an integer, and the controller is activated at time step t of the process model, when the time step of prediction model, $k_s = \frac{t \cdot T}{T_c}$, is an integer. The time scales depiction of the process model and the prediction model are shown in Fig. 5.9. U is the control variable, which represents the total flows that enter the freeway through all the on-ramps. G is the total outflow of the freeway estimated by the freeway MFD. Based on the empirical analysis in Section 5.2.1, we assume that the average trip length is time-invariant and \bar{q} and G are readily transformable. If the total outflow of the freeway stretch is not measurable, then G is formulated as,

$$G(k|t) = \frac{\bar{q}(k|t) \cdot L}{\bar{x}}, \quad (5.12)$$

and we estimate the parameters in (5.10) based on the measured data of the weighted average densities and the weighted average flows. If the total outflow is measurable, then G is formulated as,

$$G(k|t) = \left(D_3' \left(\frac{A(k|t)}{L} \right)^3 + D_2' \left(\frac{A(k|t)}{L} \right)^2 + D_1' \left(\frac{A(k|t)}{L} \right) \right) \times \left(a' \cdot e^{(b_1' \sigma(k|t) + b_2' \eta(t))} + (1 - a') \right), \quad (5.13)$$

and we estimate the parameters a' , b_1' , b_2' , D_1' , D_2' , and D_3' based on the measured data of the accumulations and the total outflows. $\eta(t)$ is calculated according to (5.9) and is assumed to be constant over the prediction horizon, whereas the dynamic of σ is formulated as,

$$\sigma(k|t) = \begin{cases} s_1 \cdot \frac{A(k|t)}{L}, & \text{if } \frac{A(k|t)}{L} \leq \bar{\rho}^{\text{cr}}(t); \\ s_2 \left(\frac{A(k|t)}{L} - \bar{\rho}^{\text{cr}}(t) \right)^3 + s_3 \left(\frac{A(k|t)}{L} - \bar{\rho}^{\text{cr}}(t) \right)^2 + s_4 \left(\frac{A(k|t)}{L} - \bar{\rho}^{\text{cr}}(t) \right) + s_5, & \\ \text{if } \frac{A(k|t)}{L} > \bar{\rho}^{\text{cr}}(t), & \end{cases} \quad (5.14)$$

where s_1 , s_2 , s_3 , s_4 , and s_5 are model parameters. Based on the measurements from the field, $\bar{\rho}^{\text{cr}}(t)$ is calculated according to (5.6) and is assumed to be constant over the prediction horizon.

The number of vehicles on the on-ramps, i.e. R , is modeled as,

$$R(k+1|t) = R(k|t) + \left(\sum_{o=2}^O d_o(k|t) - U(k|t) \right), \quad (5.15)$$

where o is the index of origins of the freeway ($o=1$ is the freeway mainstream origin, and $o = 2, 3, \dots, O$ are on-ramps), and $d_o(k)$ is the demand of origin o at time step k . We define the MPC objective function to minimize the total time spent (TTS) of vehicles on the freeway and on-ramps during the prediction horizon. The overall optimization problem is,

$$\begin{aligned} \min_{U(k|t)} \quad & \sum_{k=\frac{t}{T_c}}^{\frac{t}{T_c} + N_p - 1} (A(k|t) + R(k|t)) \cdot T_c, \\ \text{subject to} \quad & \\ & (10)-(13), \end{aligned} \quad (5.16)$$

$$A(k|t), R(k|t), \sigma(k|t), P(k|t), U(k|t) \geq 0, \forall k = k_s, k_s + 1, \dots, k_s + N_p - 1.$$

5.3.2 The lower level controller: local ramp metering

The lower level control approach is devised to distribute the optimal total inflow calculated by the MPC upper level controller to each on-ramp of the freeway stretch. Based on the optimal total inflow, U , obtained from the MPC, the total amount of flows that are needed to be restricted by the on-ramps, $Q(t)$, is

$$Q(t) = \sum_{o=2}^O f_o(t) - U \left(\frac{t \cdot T}{T_c} | t \right) \quad (5.17)$$

where $f_o(t)$ is the on-ramp o flow without ramp metering control and is estimated by a simple node model as,

$$f_o(t) = \min \left(q_o, \frac{n_o(t)}{T_c}, S_o(t) \cdot \frac{q_o}{q_o + c_o} \right) \quad (5.18)$$

where q_o is the capacity of on-ramp o and c_o is the capacity of the mainstream at the downstream of on-ramp o . $n_o(t)$ is the queue length of on-ramp o . $S_o(t)$ is the receiving capacity at the downstream of on-ramp o , which is calculated according to the CTM based on the measurement of density at the downstream of on-ramp o .

To distribute $Q(t)$ to each on-ramp of the freeway stretch, we divide the freeway stretch into sections which include a number of segments. An on-ramp is located in the middle of each section such that flows around the section can be regulated by ramp metering control. $Q(t)$ is distributed to each on-ramp based on the weight of each section. The weight of section s , i.e. w_s , is

$$w_s = \begin{cases} 0, & \text{if } \forall i \in I_s, \rho_i \leq \rho_i^{\text{cr}} \\ \frac{\sum_{i \in I_s} \rho_i l_i}{\sum_{i \in I_s} l_i} \cdot (\max_{i \in I_s}(\rho_i) - \min_{i \in I_s}(\rho_i)), & \text{if } \exists i \in I_s, \rho_i > \rho_i^{\text{cr}} \end{cases} \quad (5.19)$$

where I_s is the set of segments in section s . $\max_{i \in I_s}(\rho_i)$ is the maximum measured density in section s , and $\min_{i \in I_s}(\rho_i)$ is the minimum measured density in section s . The term $\sum_{i \in I_s} \rho_i l_i / \sum_{i \in I_s} l_i$ represents the weighted average density of section s and the term $\max_{i \in I_s}(\rho_i) - \min_{i \in I_s}(\rho_i)$ represents the density variance. The weight w_s can be interpreted as the combination of the average density and density variance of section s . Then, the ramp metering rate of on-ramp o , which is located in section s , is

$$r_o(t) = f_o(t) - Q(t) \cdot \frac{w_s}{\sum w_s} \quad (5.20)$$

5.4 Numerical experiments

The presented hierarchical control approach is applied to several numerical experiments to test the performance of the proposed MFD based ramp metering. For comparison purposes, MPC controllers that are based on different prediction models are also applied to the numerical experiments. To scrutinize the robustness of the controllers, we use two different models as the process model, namely the extended CTM and the METANET model. The performances in terms of total time spent (TTS) and the computation time of the controllers are evaluated.

5.4.1 Case study 1-The extended CTM as the process model

In this case study, the extended CTM presented by Han et al. (2016b) is used as the process model. The 3-lane synthetic freeway stretch is 16 km in length and it is divided into 40 cells. It includes 5 on-ramps and 5 off-ramps, as shown in Fig. 5.10. The parameters of the process model are set as the following. The free-flow speed v is 100 [km/h]. The congestion wave speed β is set to 20 [km/h]. The capacity c is 2000 [veh/h/lane]. The maximum extent of the capacity drop α is set to 30%. The capacities of all the on-ramps are set to 2000 [veh/h]. The time duration of a simulation step T is set to 10 [s]. Off-ramp splitting ratio is 0.2 for each off-ramp at every time step. The demand profiles of the mainstream origin and each on-ramp are shown in Fig. 5.11. The simulation period is 6 hours (6:00-12:00) that contains 2 hours of peak time (7:00-9:00). The density contour plot of the simulation is shown in Fig. 5.13 (a). It can be seen that two traffic jams are triggered at on-ramps 3 and 5 at the peak time, due to increasing demand at the on-ramps.

Five control scenarios are investigated in this case study. The objective of each control scenario is the same, which is minimizing the TTS in the prediction horizon. The prediction horizon N_p is set to one hour. The control sampling time T_c is set to 2 minutes and the first control sample of the optimization is applied to the process. The set up of five scenarios are:

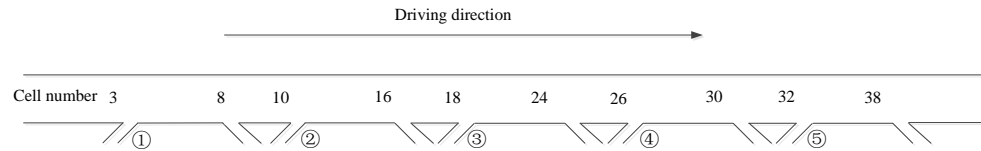


Figure 5.10: The layout of the synthetic freeway stretch. Cell numbers are marked at the upstream of each on-ramp and off-ramp. The order of on-ramps are shown as circled numbers.

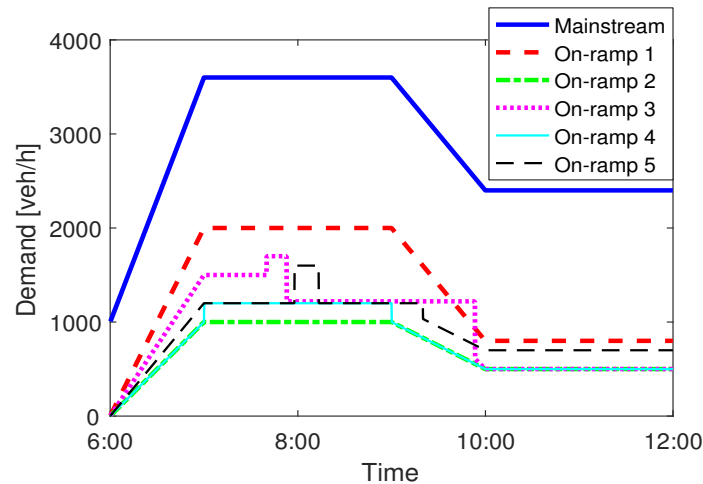


Figure 5.11: The demand profile of the simulation.

1. Scenario 1. The system optimal scenario, where in this scenario, the extended CTM is used as the prediction model, so the process and the prediction models are the same. This scenario is used as the upper achievable bound of control performance.
2. Scenario 2. The proposed control approach is applied in this scenario. That is, the proposed MFD model is used as the prediction model. The total outflow is measurable from the simulation, thus (5.13) is used to represent the dynamics of G . The parameters in (5.13) and (5.14) are estimated with the data from the simulation of the process model. The MFD of the process model (i.e. the extended CTM model) and the estimated MFD are shown in Fig. 5.12 (a).
3. Scenario 3. The same control approach as scenario 2 is applied, except that the prediction model is the MFD without the heterogeneity and capacity drop reduction, i.e. the exponential term in (5.13) is equal to one. This scenario intends to explore the significance of the density heterogeneity modeling in the performance of the controller. The estimated MFD (that is incapable of reproducing hysteresis loop) is shown as the black dotted line in Fig. 5.12 (a).
4. Scenario 4. The original CTM is used as the prediction model in this scenario. The parameters of the original CTM are set to the same as the extended CTM (the process model), except that the parameter which represents the extent of the

capacity drop is set to 0.

5. Scenario 5. The second-order model METANET is used as the prediction model in this scenario. The METANET model is estimated with the data from the simulation of the process model. The parameter estimation problem is formulated as a non-linear optimization problem which aims to minimize the discrepancy between the simulation results of the process model and the prediction model. The non-linear and non-convex optimization problem is solved by MATLAB implementation of the SQP algorithm. To ensure that a global optimum is achieved, 100 starting points which are based on empirical initial guess and random noises are used.

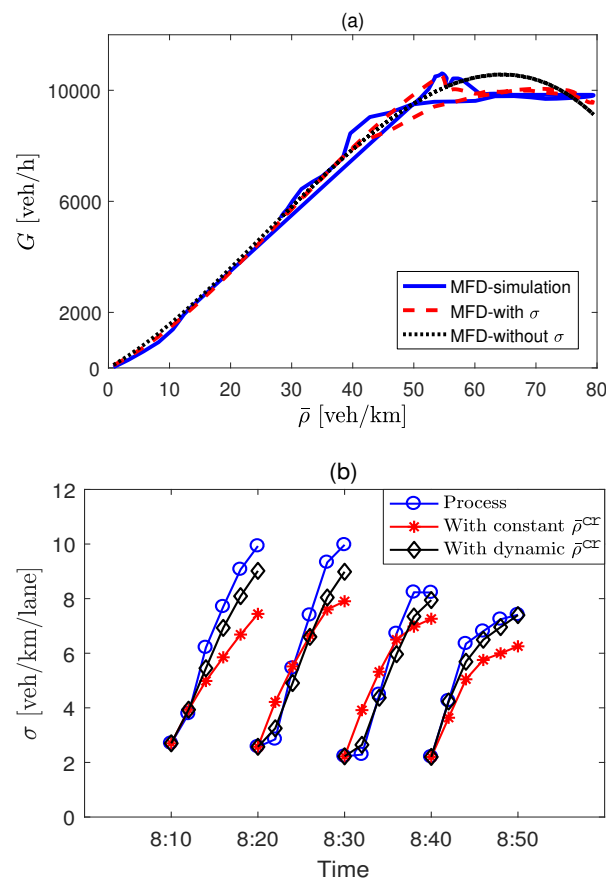


Figure 5.12: (a): The estimated MFDs in scenario 2 and 3. Note that the average density is calculated over three lanes. (b): The accuracy of the prediction of σ . Blue lines represents the evolution of σ in the process. Black lines represent the value of σ estimated by (8) which has a dynamic value of $\bar{\rho}^{\text{cr}}$. Red lines represent the values of σ predicted by the prediction model, which assumes a constant value of $\bar{\rho}^{\text{cr}}$ in the prediction horizon.

Simulation results are shown in Fig. 5.13 and Fig. 5.14, where Fig. 5.13 (b-f) show the density contour plots of the five control scenarios and Fig. 5.14 shows the queue length and ramp flow of each ramp for every control scenario. The TTS and the computation

time of each scenario are listed in Tab. 5.2, where it can be observed that the TTS of control scenario 2 (the proposed control approach) is the closest to the system optimal TTS (scenario 1). The control dynamics of both scenarios are similar where ramp metering controls are activated at the proper locations (on-ramps 3 and 5) and time (when the bottlenecks are activated). Fig. 5.14 depicts that ramp queues are mainly accumulated at on-ramps 3 and 5 with both scenarios. The proposed controller is efficient and effective because density heterogeneity is precisely modeled and predicted so that the reduction of the outflow due to the increase in vehicle accumulation is captured by the freeway MFD. Fig. 5.12 (b) shows the accuracy of the predicted heterogeneity, where blue lines represent the evolution of σ of the process model at a certain time instances, if no control action is taken. Black lines represent the values of σ estimated by (8), in which the values of $\sigma(t)$ are calculated based on dynamic values of $\bar{\rho}^{\text{cr}}$. In the prediction model of the controller, it is assumed that $\bar{\rho}^{\text{cr}}$ is constant over a control horizon, where the predictions of σ are shown with the red lines. Evidently, the model with dynamic $\bar{\rho}^{\text{cr}}$ models density heterogeneity, i.e. σ , more accurately, while the prediction model of the controller also captures the trend of increasing σ accurately. The increase of σ results in the decrease of the total outflow, thus the controller regulates the total inflow to the network to avoid the increase in σ . Furthermore, the regulated flows are distributed to the target on-ramps based on traffic state feedback, which also contributes to resolve the congestion effectively.

In contrast to the proposed controller, the controller in scenario 3 has an inferior performance, because the prediction model of the controller does not take the density heterogeneity into consideration. When breakdown occurs in the freeway, the density heterogeneity increases rapidly with the increase of the accumulation, which results in a rapid decrease of the total outflow. Since the MFD in scenario 3 does not model this effect, the controller fails to generate effective ramp metering control signals to resolve the congestion. For the controller in scenario 4, since the prediction model does not incorporate the capacity drop, the only gain of the controller comes from avoiding the off-ramp blockade. The controller underestimates the benefits of applying ramp metering. Therefore, the ramp metering rates are not sufficiently low and on time to resolve the congestion, see Fig. 5.14 where scenario 4 ramp metering is active late at on-ramp 5. For the METANET-based MPC (scenario 5), due to the mismatch between the prediction model and the process model, control signals are ineffective that lead to inferior control performance. For example, as depicted in Fig. 5.14, it can be seen that when freeway mainstream congestion is triggered at around 8:00 at on-ramp 5, there is no control action at on-ramp 5. While at around 9:00, the controller performs unnecessary actions at on-ramps 2 and 4. These ineffective control actions significantly decrease the performance of the controller.

The computation efficiency of all control scenarios are shown in Tab. 5.2. A computer with an E5-1620 processor and 16 GB RAM is used for the optimizations. The computation time of the controllers have strong correlation with the complexity of the prediction model. The MFD-based controllers (scenarios 2 and 3) have less compu-

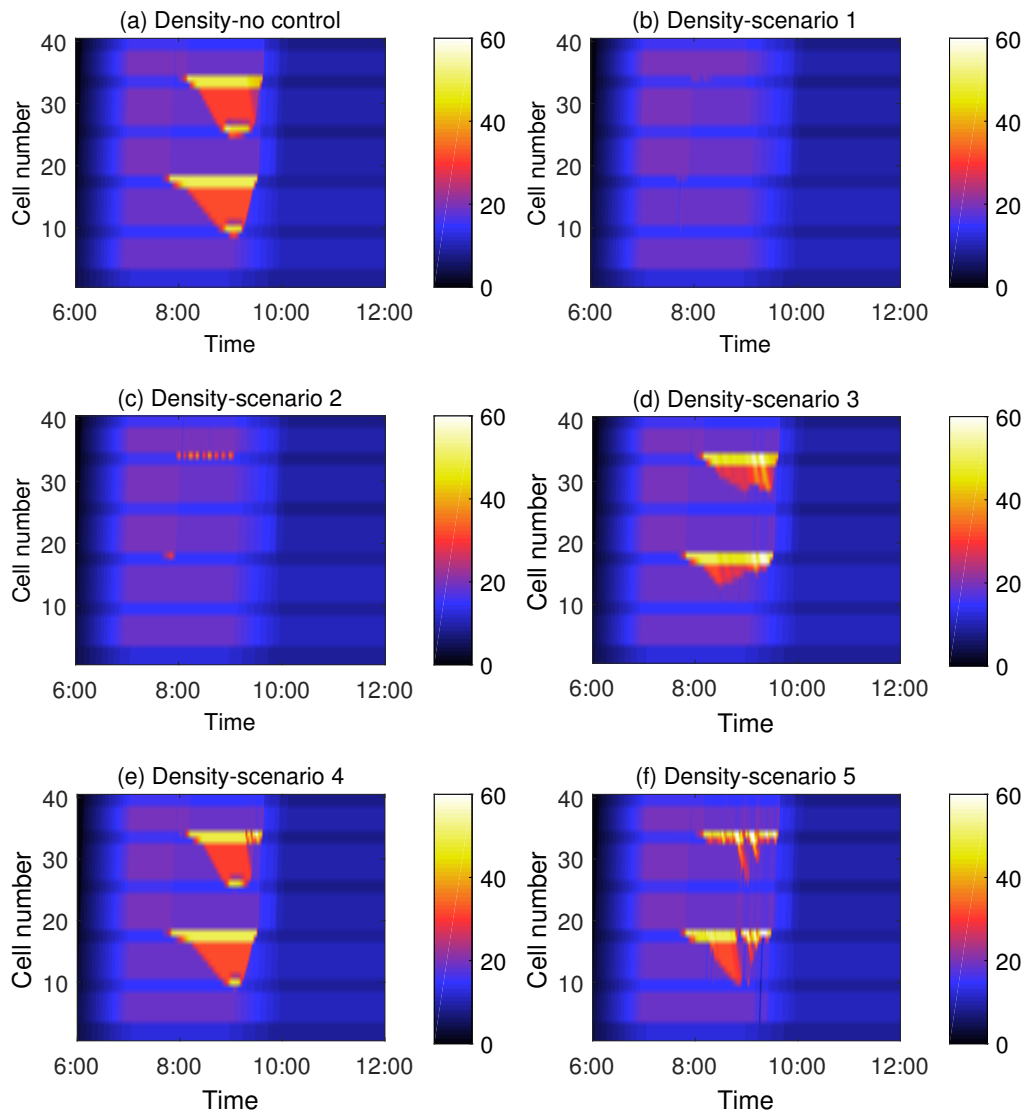


Figure 5.13: The density (veh/km/lane) contour plots of the scenario without control and all five control scenarios in case study 1. On-ramps 3 and 5 are the main bottlenecks. Evidently, the proposed MFD-based ramp metering control strategy (i.e. scenario 2) efficiently reduces traffic density at on-ramp locations. Scenario 1 represents the theoretical upper bounds of MPC performance.

tation times, the first-order model-based controllers (scenarios 1 and 4) have higher computation times, and the second-order model-based controller (scenario 5) has the highest computation time. This is of great importance for practical implementation of real-time control strategies. Note that some efficient computational algorithms, such as the feasible-direction algorithm, have been proposed to increase the computation speed of the METANET-based controllers (Kotsialos et al., 1999). In this chapter, we use the standard SQP algorithm for every scenario.

Note that in this case study, the estimated value of parameter b'_2 in (5.13), which represents the influence of capacity drop on MFD, is close to 0 because the effect of capacity drop is captured by the density heterogeneity. This is because the temporal duration of capacity drop phenomenon is limited compared to the period that the freeway is congested. Hence, to explore the influence of capacity drop on the control performance, traffic scenarios with extreme extent and number of capacity drop and moving jams should be studied. This is a future research direction.

Table 5.2: The performance of each control scenario in case study 1.

Scenario	TTS (h)	Improvement (%)	Computation time (s/control step)
Without control	4218.8	0	-
1. System optimal	3775.6	10.51	20.9
2. MFD-based	3833.9	9.12	3.1
3. MFD-based (no σ)	4170.2	1.15	2.6
4. CTM-based	4173.2	1.08	9.6
5. METANET-based	4164.0	1.30	252.9

5.4.2 Case study 2-METANET as the process model

To further explore properties and the robustness of the presented controller, the performance of the controller is tested where the METANET model is employed as the process model. The parameters of the METANET model are set to same values as the calibrated parameters values in scenario 5 of the previous section. We apply the same demand profile as in Fig. 5.11. The density contour plot of the no control scenario is shown in Fig. 5.15 (a). Three control scenarios are considered as:

1. Scenario 1. The system optimal scenario where the METANET model is used as both the process and the prediction models. This scenario demonstrates the upper bound of MPC control performance.
2. Scenario 2. The proposed control approach is applied to this scenario. The parameters in (5.13) and (5.14) are calibrated with the data from the simulation of the process model.

3. Scenario 3. The extended CTM is used as the prediction model in this scenario. The extended CTM is calibrated with the data from the simulation of the process model.

Results are presented in Fig. 5.15 and Fig. 5.16, where Fig. 5.15 (b-d) show the density contour plots of the three control scenarios and Fig. 5.16 shows the queue length and ramp flow of each ramp for every control scenario. The TTS achieved by the controllers are listed in Tab. 5.3. Fig. 5.15 shows that all three scenarios resolve the congestion in the freeway mainstream. However, for the extended CTM-based controller (scenario 3), the TTS is significantly higher than the ones achieved by the system optimal scenario and the proposed control approach (scenario 2). The differences of the control performance stem from the effectiveness of the control actions. Fig. 5.16 demonstrates that the queue lengths at the bottlenecks (on-ramps 3 and 5) are similar in scenarios 1 and 2. Whereas in scenario 3, the controller applies unnecessarily low ramp flows (e.g., after 8:00 at on-ramp 5 and after 9:00 at on-ramp 4), which results in longer ramp queues and thus larger time delays. The mismatch between the process model and the prediction model is the main reason that causes the inferior performance of scenario 3. The results of case studies 1 and 2 show that the presented control approach has a robust performance to overcome the mismatch between the prediction model and the process model in the MPC framework.

Table 5.3: The performance of each control scenario in case study 2.

Scenario	TTS (h)	Improvements (%)
Without control	4341.2	0
1. System optimal	4138.0	4.7
2. MFD-based	4156.4	4.2
3. Extended CTM-based	4619.2	-6.4

5.5 Conclusion

The chapter has investigated the properties of freeway MFD and presented a model to predict traffic dynamics based on the MFD. A main feature of the model is that the density heterogeneity is considered in the model, such that the MFD can predict the aggregated traffic states in the freeway network accurately, which benefits the MFD-based controllers. We also incorporate the capacity drop into the freeway MFD. By proposing an exponential functional form, the accuracy of the proposed MFD is validated with field traffic data.

The proposed MFD model has been incorporated in a hierarchical control structure for coordinated ramp metering on a freeway stretch. At the upper level, a model predictive control approach is developed to optimize the total inflow from on-ramps to the freeway stretch. The lower level controller distributes the optimal total inflow to each

on-ramp of the freeway based on local traffic state feedback. The proposed control approach demonstrates a performance close to the optimal MPC control in reducing the total time spent and eliminating congestion. The robustness of the controller is tested by using different traffic flow models (the extended CTM and the METANET model) as the process model and comparing with MPC controllers that are based on different prediction models. Simulation results highlight two advantages of the proposed controller, (i) the robustness to overcome the mismatch between the prediction model and the process model, and (ii) the applicability for field implementation as it requires less computation efforts than other non-linear optimal controllers.

For future research, the proposed freeway MFD model and control approach can be developed to tackle congestion control in a freeway network that consists of multiple freeway stretches. Control measures will not only be limited to ramp metering, but also integrate the variable speed limits. The influence of the capacity drop to the control performance need further investigation.

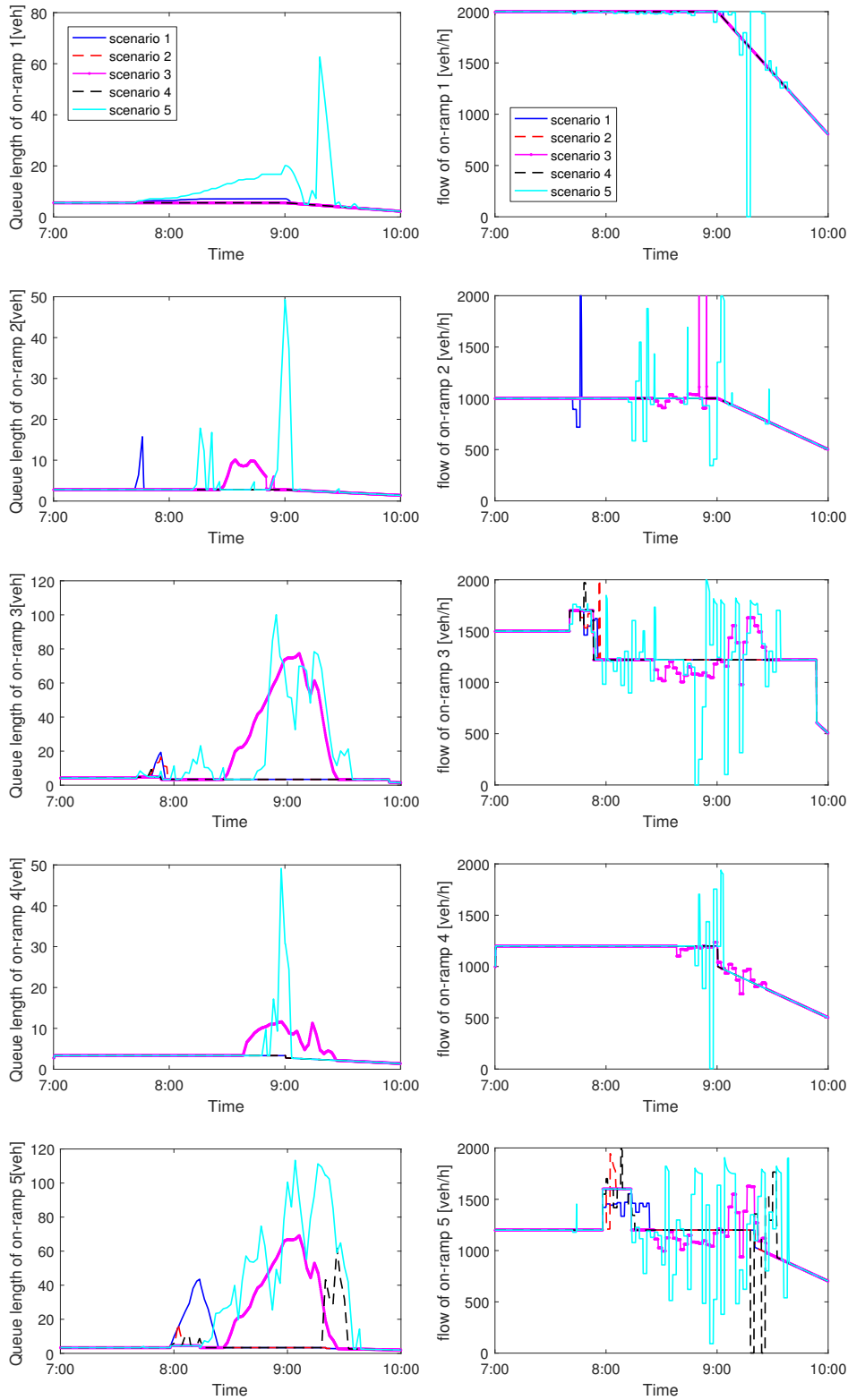


Figure 5.14: The queue lengths (on left) and ramp flows (on right) of on-ramps 1-5 of the five control scenarios in case study 1.

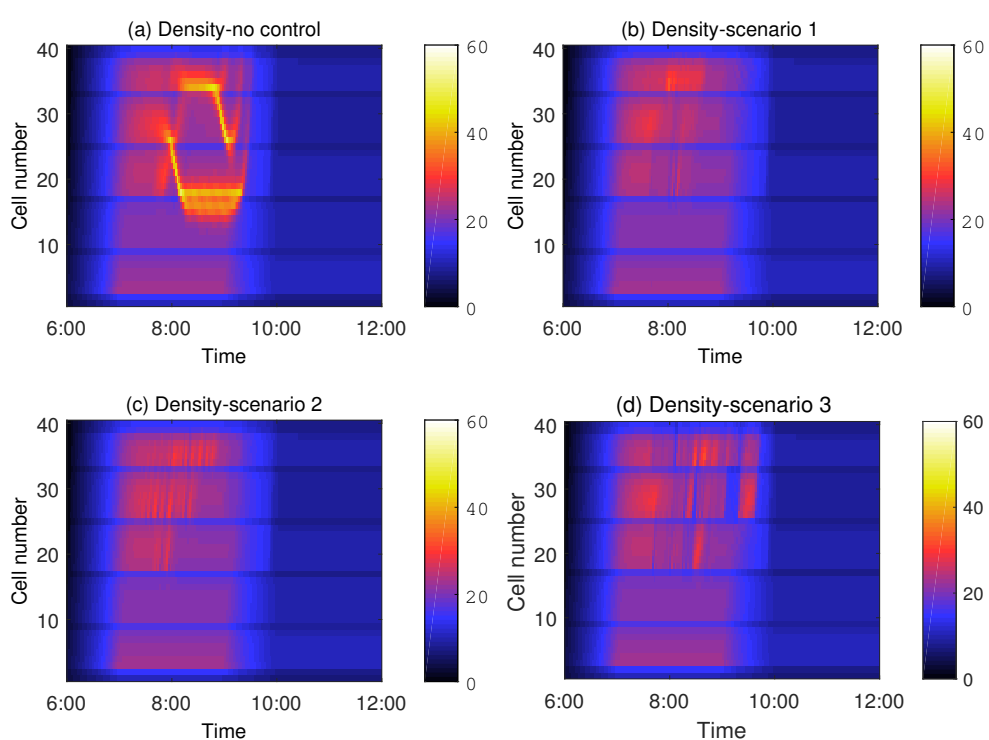


Figure 5.15: The density contour plots of the scenario without control and all three control scenarios in case study 2.

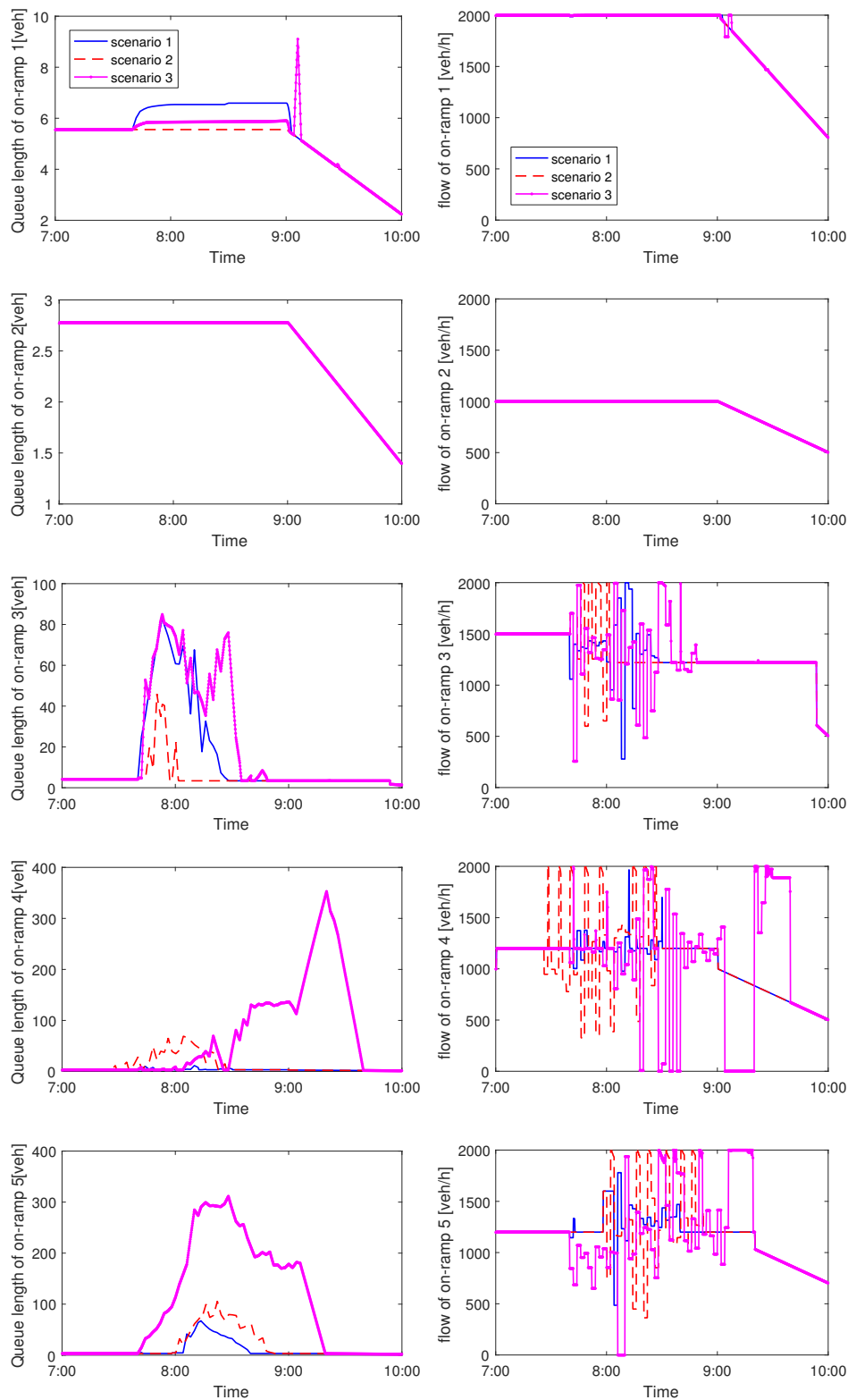


Figure 5.16: The queue lengths (on left) and ramp flows (on right) of on-ramps 1-5 of the three control scenarios in case study 2.

Chapter 6

An extended linear quadratic model predictive control approach for multi-destination urban traffic networks

This chapter extends an existing linear quadratic model predictive control (LQMPC) approach to multi-destination traffic networks, where the correct OD relations are preserved. In the literature, LQMPC approach has been presented for efficient routing and intersection signal control. The optimization problem in the LQMPC has a linear quadratic formulation that can be solved fast, which is beneficial for a real-time application. However, the existing LQMPC approach does not conserve OD relationships, and thus may send traffic to wrong destinations. This problem is tackled by a heuristic method in this chapter. We present two macroscopic models: (1) a route-specific model (non-linear) which keeps track of traffic dynamics for each OD pair, and (2) a linear model that aggregates all route traffic states, which can be embedded into the LQMPC framework. The route-specific model predicts traffic dynamics and provides information to the LQMPC before the optimization and evaluates the optimal solutions after the optimization. The information obtained from the route-specific model is formulated as constraints in the LQMPC to narrow the solution space and exclude unrealistic solutions that would lead to flows that are inconsistent with the OD relations. The extended LQMPC approach is tested in a synthetic network with multiple bottlenecks. The simulation of the LQMPC approach achieves a total time spent close to the system optimum and the computation time remains tractable.

This chapter is an edited version of the article:

Han, Y., Hegyi, A., Yuan, Y., Roncoli, C., Hoogendoorn, S. An extended linear quadratic model predictive control approach for multi-destination urban traffic networks. *IEEE Transaction on Intelligent Transportation System*, submitted for review.

6.1 Introduction

Due to growing transportation demands and the urbanization trend, traffic congestion has become a global issue that has a significant impact on our society's productivity. Expanding the current infrastructure is not always viable due to high costs. Dynamic traffic control is an effective way to make a better use of existing infrastructure. In urban traffic networks, traffic signal control strategies play an important role to improve the traffic flow conditions. Nowadays, existing traffic signal control systems in the field have different levels of coordination. The development of optimal signal control strategies that are network-wide coordinated and real-time applicable remains a challenge.

The field of urban traffic control has been studied and developed in a variety of ways during the past decades. A review of the literature related to road traffic control of the last century is given in [Papageorgiou et al. \(2003\)](#). In recent years, model predictive control (MPC) approaches have become more and more popular in road traffic control related research. MPC approaches for traffic systems predict the evolution of traffic dynamics and calculates the optimal control actions for the time period in which the relevant traffic dynamics occur. This feature enables the controller to take advantage of potentially larger future gains at a current (smaller) cost, so as to avoid myopic control actions. MPC approaches that are based on the store-and-forward model for traffic flow optimization in saturated conditions are developed in [Aboudolas et al. \(2009, 2010\)](#). Since the travel time on the roads between intersections is ignored, these approaches are less effective for under-saturated conditions. In [van den Berg et al. \(2007\)](#) the authors developed a non-linear MPC approach that is based on a detailed traffic flow model to optimize the network throughput. The complexity of the prediction model results in high computation time of the MPC problem. In [Lin et al. \(2011\)](#) the authors proposed a non-linear MPC and reformulated it into a mixed integer linear programming (MILP) problem that can be efficiently solved. In order to reduce computational complexity, this model assumes constant delay, which leads to the model inaccuracies in under-saturated traffic. Decentralized MPC strategies have been developed for urban traffic signal control in [Camponogara & De Oliveira \(2009\)](#); [De Oliveira & Camponogara \(2010\)](#). Even though decentralized MPC approaches increase the computation speed, they may result in sub-optimal control performances [Frejo & Camacho \(2012\)](#). Recently, the macroscopic fundamental diagram (MFD) has been exploited as a basis for the derivation of urban signal control approaches [Haddad et al. \(2013\)](#); [Geroliminis et al. \(2013\)](#); [Ramezani et al. \(2015\)](#). Even though the MFD provides an efficient tool for the development of traffic control strategies for urban traffic networks, its accuracy in expressing aggregated dynamics of urban traffic networks, especially in case the network is characterized by a large density heterogeneity, still needs to be validated. Due to the complexity of urban traffic system, the development of MPC approaches of traffic signal control has to make a trade-off between model accuracy and computation time.

In the aforementioned MPC approaches, fixed turn fractions at intersections are assumed in order to simplify the prediction models. As a matter of fact, there is a clear linkage between intersection control algorithms and route choice control Peeta & Ziliaskopoulos (2001). Traffic route choice behavior is an important factor that influences the arriving flows at intersections. Thus, coordinating route choice control and intersection control will further improve traffic operation efficiency. However, the combination of route choice control and intersection control increases the complexity of the model, thus results in complicated optimization problems that cannot be solved efficiently. Therefore, many studies regarding the optimization of traffic signals and route choices used heuristic approaches to solve the problem, which do not always lead to optimal results Ceylan & Bell (2004); Teklu et al. (2007). The linear quadratic model predictive control (LQMPC) approach presented in Le et al. (2013) is a successful application to efficiently optimize turn fractions and traffic signals for single destination networks. Due to the fact that the LQMPC has a discrete linear prediction model, traffic state in each discrete segment of the traffic network is aggregated and route choice behavior of traffic in each OD pair is not considered. The controller pushes out traffic flow as much as possible, regardless of traffic desired origin and destination relations. Thus, the LQMPC approach may be ineffective when applied to multi-destination traffic networks, because the traffic may end up at wrong destinations.

In this chapter, the problem that the existing LQMPC cannot preserve traffic OD relations is tackled by a heuristic method. We present two models: (1) a route-specific model (non-linear) which keeps track of traffic dynamics for each OD pair, and (2) a linear model that aggregates all route traffic states, which can be embedded into the LQMPC framework. For each OD pair, we define a set of links as the crucial links (c-links), which are the links that traffic from the OD pair has to pass to reach the destination. Before each optimization run, the route-specific model predicts the route flow of every c-link based on a non-optimizing control strategy. In this chapter, the non-optimizing control strategy we use is the integration of a simple routing strategy, which guides traffic flows at the origins to the route that has the shortest instantaneous travel time, and a simplified back-pressure algorithm for intersection signal timing plans Varaiya (2013). In principle, any other routing and signal control approach could be used as the non-optimizing control strategy. In the optimization, we set constraints to the flows of c-links and the flows of the links towards destinations, to narrow the solution space and exclude unrealistic solutions that would lead to flows that are inconsistent with the OD relations. The optimal solution is evaluated by the route-specific model after each optimization run and compared with the non-optimizing control strategy. The control signals that leads to a better performance will be implemented to the process. The evaluation process guarantees that the designed control approach never perform worse than the non-optimizing approach. The presented control approach is tested in a synthetic traffic network with multiple bottlenecks. It shows that the LQMPC achieves a total time spent close to the system optimal, while the computation time remains tractable.

The remaining of this chapter is organized as follow. Section 6.2 presents the route-specific traffic flow model. Section 6.3 and Section 6.4 present the existing LQMPC approach and the extended LQMPC approach respectively. Section 6.5 presents the design and testing results of a synthetic case. A summary and the discussion of future work conclude the chapter in Section 6.6.

6.2 Route-specific network model

In this section, we present a route-specific macroscopic model. The traffic flow model is based on a discrete model, where the time is divided into discrete time steps and the roadway is divided into discrete cells. The definition of model variables are presented as follows.

6.2.1 Definition of model variables

Our representation of traffic networks is based on the queuing model developed in [Le et al. \(2013\)](#). The model presented in this chapter extends that model in two aspects: (i) The previous model is a queuing model, which is not able to reproduce the propagation of congestion waves. In our model, we adopt the same logic as in the cell transmission model (CTM) to decide the boundary flows between cells, so as to capture the propagation of shock waves [Daganzo \(1994\)](#); [Lighthill & Whitham \(1955\)](#); [Richards \(1956\)](#). (ii) Routes are specified in each cell, to keep track of the traffic dynamics in each OD pair. The network elements and variables of the model are introduced as follows.

Cells and Links: In our model, roadway networks are divided into discrete cells and links. Cells are categorized into different types which include origin cells, destination cells, normal cells, merging cells, diverging cells, and intersection cells. The depiction of each type of cells can be found in Fig. 6.1 and Fig. 6.2. Vehicles flowing out of a cell move to downstream cells through pre-defined links. Links do not represent physical roadways, but an abstraction of the boundary that connects two consecutive cells.



Figure 6.1: The depiction of cells and links. Red arrows represent links, and black dashed lines represent the boundaries between cells. Origin cells have no incoming link and destination cells have no outgoing link. Normal cells are connected by one incoming link and one outgoing link.

Assuming that time evolves in discrete steps, $t = 1, 2, 3, \dots$, and traffic state of cell i is represented by the number of vehicles in a discrete time step t , which is denoted as

$x_i(t)$. A link always connects two cells: we define the cell upstream of link j by Γ_j^{-1} , and the cell downstream of link j by Γ_j . The number of vehicles moving from one cell to another in a time step is defined as the traffic flow of a link, denoted by $f_j(t)$.

Routes: There may be multiple routes in a network, and each route connects one origin and one destination. In our model, routes are specified in each cell, and the number of vehicles following route r in cell i is denoted as $x_i^r(t)$. Accordingly, the traffic flow on a route of a link is denoted as $f_j^r(t)$. The traffic state of a cell and a link are the aggregations of the number of vehicles of every route in the cell and the link:

$$x_i(t) = \sum_{\forall r} x_i^r(t), \quad (6.1)$$

$$f_j(t) = \sum_{\forall r} f_j^r(t). \quad (6.2)$$

Merge and diverge: The cells that have one outgoing link and multiple incoming links are called merging cells, as shown in the left picture of Fig. 6.2. The cells that have one incoming link and multiple outgoing links are called diverging cells, as shown in the right picture of Fig. 6.2. We define the outgoing links of a diverging cell as diverging links, and the incoming links of a merging cell as merging links. The proportion of the flow of a diverging link j to the total diverging flows that originate from the same diverging cell i is called turn fraction, which is denoted as $\theta_i^j(t)$. Traffic at a diverging cell may change their original routes based on the obtained information of road traffic conditions or the instructions received from the traffic control center. Thus, the variables $\theta_i^j(t)$ can be controlled by route guidance measures.

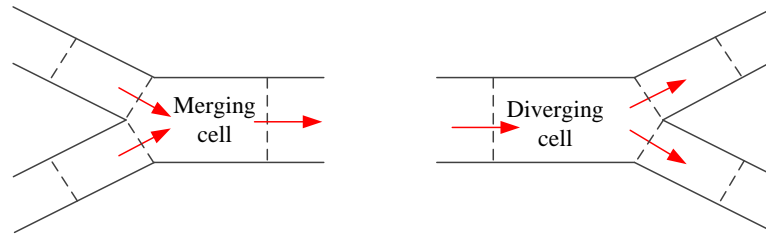


Figure 6.2: The depiction of merging and diverging cells. Red arrows represent links. The incoming links of merging cells are called merging links and the outgoing links of diverging cells are called diverging links.

Intersections: In the following, we briefly introduce the modeling of intersection flows. We define the cells directly upstream of an intersection as intersection cells. The outgoing links of intersection cells are defined as intersection links. We assume that flows of intersection links are always controlled by traffic lights. To avoid conflicting flows at an intersection, the green time of a cycle is divided into different phases such that conflicting flows do not get green at the same time. Assume that T_c represents the time duration of a cycle, then the green time duration of phase p at intersection s is $T_c \cdot u_s^p$, where u_s^p represents the green time fraction. At an intersection, the sum of the

time fraction of each phase needs to satisfy the traffic light cycle constraint,

$$\sum^p u_s^p \leq 1. \quad (6.3)$$

The set of link indices that get green at phase p is denoted as W_s^p . An intersection link may get green in different phases of a cycle, and the intersection link flow need to satisfy the green time fraction constraints,

$$f_j(t) \leq \sum^p u_s^p(t) \cdot Q_j, \quad i \in C_I, j \in W_s^p, \Gamma_j^{-1} = i \quad (6.4)$$

where C_I is the set of intersection cell indices. Q_j [veh/h] is the saturation flow of intersection link j .

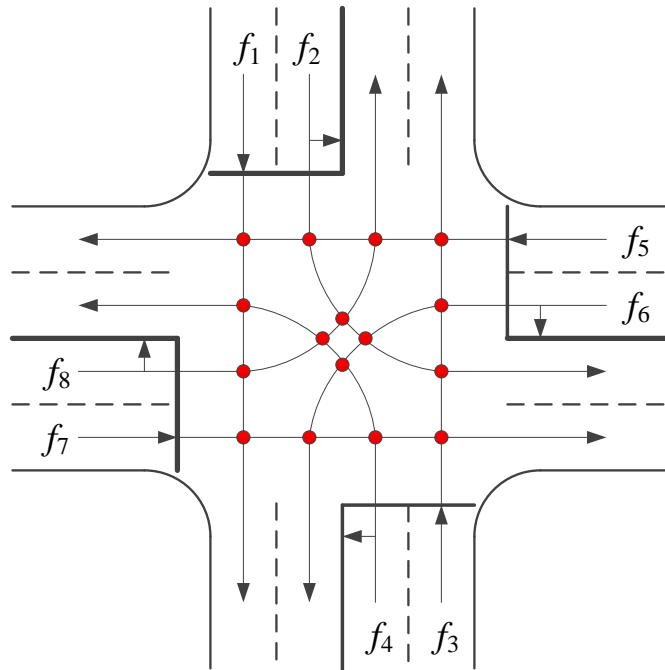


Figure 6.3: An example intersection to explain the model variables. The figure is taken from [Zhu & Ukkusuri \(2015\)](#). f_1, f_2, \dots, f_8 are the flows of intersection links. Red dots in the figure represent conflicting points of flows. The intersection is indexed as intersection 1. For simplicity, we assume that there are four phases in a cycle, and $W_1^1 = \{1, 3\}$, $W_1^2 = \{2, 4\}$, $W_1^3 = \{5, 7\}$, $W_1^4 = \{6, 8\}$. Note that there can be more phases in a cycle as long as the movements in each cycle do not conflict. u_1^1, u_1^2, u_1^3 and u_1^4 are the green time fraction of each phase so the traffic light cycle constraint is: $u_1^1 + u_1^2 + u_1^3 + u_1^4 \leq 1$. Green time fraction constraints are: $f_1 \leq u_1^1 \cdot Q_1$, $f_2 \leq u_1^2 \cdot Q_2$, $f_3 \leq u_1^1 \cdot Q_3$, $f_4 \leq u_1^2 \cdot Q_4$, $f_5 \leq u_1^3 \cdot Q_5$, $f_6 \leq u_1^4 \cdot Q_6$, $f_7 \leq u_1^3 \cdot Q_7$, $f_8 \leq u_1^4 \cdot Q_8$.

6.2.2 Updating of traffic states

The route-specific traffic state of each type of cells is updated as follows.

Origin cells:

$$x_o^r(t+1) = x_o^r(t) + (q_{o,d}(t) \cdot \alpha^r - f_j^r(t)) \cdot T, \quad (6.5)$$

$$r \in R^{o,d}, \Gamma_j^{-1} = o.$$

where o is the index of the origin cell, and (o, d) are the indices of OD pairs. $R^{o,d}$ is the set of route indices in the OD pair and $q_{o,d}(t)$ represents the demand of the OD pair at time t . T is the duration of a discrete time step. α^r is the fraction of the demand that choose route r , and it is determined by a simple logit model which is formulated as:

$$\alpha^r = \frac{e^{(-\sigma \cdot TF^r)}}{\sum_r e^{(-\sigma \cdot TF^r)}}, \quad r \in R^{o,d} \quad (6.6)$$

where, TF^r is the free flow travel time of route r . σ is a scaling parameter that describes how drivers react on a travel time difference between the alternatives. This simple logit model represents how traffic choose their routes at the origins, when they do not have any information of the traffic situation in the network Prato (2009).

Destination cells:

$$x_d^r(t+1) = x_d^r(t) + f_j^r(t), \quad \forall d \in C_D, \Gamma_j = d \quad (6.7)$$

where d is the index of destination cells and C_D is set of destination cell indices.

Normal cells and intersection cells:

$$x_i^r(t+1) = x_i^r(t) + f_j^r(t) - f_h^r(t), \quad (6.8)$$

$$\forall i \in C_N \cup C_I, \Gamma_j = \Gamma_h^{-1} = i.$$

where C_N is the set of normal cell indices.

Merging cells:

$$x_i^r(t+1) = x_i^r(t) + \sum_j^j f_j^r(t) - f_h^r(t), \quad (6.9)$$

$$\forall i \in C_M, \Gamma_j = \Gamma_h^{-1} = i,$$

where C_M is the set of merging cells.

Diverging cells:

$$x_i^r(t+1) = x_i^r(t) + f_j^r(t) - \sum_h^h f_h^r(t), \quad (6.10)$$

$$\forall i \in C_V, \Gamma_j = \Gamma_h^{-1} = i,$$

where C_V is the set of diverging cells.

The traffic flow updating procedure is presented as follows. Similar as the CTM which uses a triangular fundamental diagram, we define a sending flow, $S_i(t)$, and a receiving flow, $R_i(t)$, for cell i . $S_i(t)$ and $R_i(t)$ are formulated as,

$$S_i(t) = \min \left(\frac{x_i(t)}{L_i} \cdot v_i, c_i \right), \quad (6.11)$$

$$R_i(t) = \min \left(\beta_i \left(\frac{x_i^J}{L_i} - \frac{x_i(t)}{L_i} \right), c_i \right),$$

where v_i and c_i are the free-flow speed and capacity of cell i . β_i is the congestion wave speed of cell i . x_i^J is the maximum number of vehicles in cell i . L_i is the length of cell i . For the links whose source cell and sink cell are normal cells, the flows are determined by the minimum between the sending flow and the receiving flow,

$$f_j(t) = \min(S_i(t), R_l(t)), \quad \forall i, l \in C_N, \Gamma_j^{-1} = i, \Gamma_j = l. \quad (6.12)$$

It is assumed that there is no route choice in normal cells. Thus the flow of every route is represented as,

$$f_j^r(t) = f_j(t) \cdot \frac{x_i^r(t)}{x_i(t)}, \quad \forall i \in C_N, \Gamma_j^{-1} = i \quad (6.13)$$

For a merging cell, the sending flow of the source cell is represented by the first equation of (6.11). The receiving flow is shared by multiple merging links. If the total sending flow is higher than the receiving flow, then the receiving flow needs to be distributed to each merging link. The sending flow of the source cell of a merging link is represented as,

$$S_i(t) = \min\left(\frac{x_i(t)}{L_i} \cdot v_i, c_i\right), \quad i \in C_M \quad (6.14)$$

The merging flow updating procedure is presented as follows. Note that the receiving flow distribution process is the same as the first-order node model presented in [Tampère et al. \(2011\)](#).

1. The receiving flow is first distributed to the links, whose source cells have lower sending flow than the merging flow rate.

$$f_j(t) = S_i(t), \quad \text{if } S_i(t) \leq R_l(t) \cdot \frac{c_i}{\sum_i c_i}, \quad (6.15)$$

$$l \in C_M \Gamma_j^{-1} = i, \Gamma_j = l,$$

where, $R_l(t) \cdot \frac{c_i}{\sum_i c_i}$ is the merging flow rate of cell i . Once the receiving capacity is assigned to a link, the assigned flow is subtracted from the remaining receiving flow which is represented as $\tilde{R}_l(t)$.

2. Exclude the links which have been assigned, and calculate the remaining receiving flow. Update the merging flow rate for the remaining source cells and repeat the first step until no source cell has lower sending flow than the merging flow rate.
3. Distribute the remaining receiving flow to the remaining links, whose source cells have higher sending capacity than the merging flow rate.

$$f_j(t) = \tilde{R}_l(t) \cdot \frac{c_i}{\sum_i c_i} \quad (6.16)$$

4. The merging flow in each route is calculated by (13).

Traffic at a diverging cell may change their original routes based on the obtained information of road traffic conditions or the instructions received from the traffic control center. In our model, we assume that if there is no control action, traffic continue travelling on their original routes. The turn fraction $\theta_i^j(t)$ of each diverging link is calculated as,

$$\theta_i^j(t) = \frac{\sum_r x_i^r(t)}{x_i(t)}, \quad r : j \in J^r \quad (6.17)$$

where J^r is the set of link indices in route r . Based on $\theta_i^j(t)$, the total flow of all the links, l_1, l_2, \dots, l_n , that are leaving cell i , $F_i(t)$, is calculated as,

$$F_i(t) = \min \left(S_i(t), \frac{R_{l_1}(t)}{\theta_i^{l_1}(t)}, \frac{R_{l_2}(t)}{\theta_i^{l_2}(t)}, \dots, \frac{R_{l_n}(t)}{\theta_i^{l_n}(t)} \right), \quad \forall i \in C_V. \quad (6.18)$$

The flow of each route is calculated as,

$$f_j^r(t) = F_i(t) \cdot \frac{x_i^r(t)}{x_i(t)} \quad \forall i \in C_V, \Gamma_j^{-1} = i \quad (6.19)$$

If route guidance control is applied at diverging cells, then $\theta_i^j(t)$ is the resulting turn fraction after control. In this chapter, we assume that $\theta_i^j(t)$ equals to the optimal turn fraction obtained from the controller. Traffic at diverging cell i change their routes at time t , thus there are two traffic states for cell i , which are the state before route choice and the state after route choice, represented as $x_{i,b}^r(t)$ and $x_{i,a}^r(t)$ respectively. Based on the optimal turn fractions, we assign the vehicles in cell i to new routes based on an optimization problem, which aims to minimize the total cost of all travelers. The optimization problem is formulated as,

$$\begin{aligned} & \min \sum_r x_{i,a}^r(t) \cdot \text{TT}_i^r(t) \quad \forall i \in C_V \\ & \text{subject to} \\ & \frac{\sum_r x_{i,a}^r(t)}{x_i(t)} = \theta_i^j(t), \quad r : j \in J^r \\ & \sum_r x_{i,a}^r(t) = \sum_r x_{i,b}^r(t), \quad r : r \in R^{od} \end{aligned} \quad (6.20)$$

where $\text{TT}_i^r(t)$ is the instantaneous travel time of the part of route r , that starts from the diverging cell i . The first constraint in (6.20) ensures that the turn fraction of a diverging link j equals to θ_i^j (the optimal value obtained from the controller) after re-routing, and the second constraint guarantees that the OD relations of the traffic in cell i unchanged. For diverging links, the total flow $F_i(t)$ is calculated according to (6.18) and the flow of each route is calculated based on (6.19) by changing $x_i^r(t)$ to $x_{i,a}^r(t)$. Note that the optimization problem (6.20) represents an autonomous process of drivers' route choices, and it does not change the turn fractions.

As introduced in 6.2.1, intersection flows are determined by the following equations,

$$\begin{aligned} f_j(t) &= \min \left(S_i(t), R_l(t), \sum_s^p u_s^p(t) \cdot Q_j \right), \\ & i \in C_I, j \in W_s^p, \Gamma_j^{-1} = i, \Gamma_j = l, \end{aligned} \quad (6.21)$$

and the flow in each route is calculated by (6.13).

6.3 The classical LQMPC approach

In Le et al. (2013) the authors presented an LQMPC approach based on a queuing model for optimal routing and intersection signal control of urban networks. Later, LQMPC approaches using extended discrete LWR models as prediction models have been presented for optimal ramp metering or variable speed limit control Han et al. (2015b, 2017b). In general, any model that has a linear formulation can be applied to the LQMPC framework. In this section, we present the general LQMPC framework using a discrete LWR model as the prediction model.

In the model, the states of cells and links are aggregated numbers of vehicles and aggregated flows, which are represented by $x_i(k)$ and $f_j(k)$, where k is the index of the discrete time step in the prediction. It is worth to note that the route-specific model presented in the previous section will be used to represent the reality, and the time index is t in that model. Usually the time step size of the prediction model is larger than the process model, because in the process model, detail traffic dynamics is usually needed but in the prediction model, a trade-off between the level of predicted details and the computation time has to be made. In this chapter we use the same time step size for both the prediction model and the process model. Traffic states of each class of cells is predicted according to the following equations.

Origin cells:

$$\begin{aligned} x_o(k+1) &= x_o(k) - f_j(k) + q_o(k), \quad \Gamma_j^{-1} = o \\ q_o(k) &= \sum^o q_{o,d}(k) \end{aligned} \quad (6.22)$$

Destination cells:

$$x_d(k+1) = x_d(k) + f_j(k), \quad \Gamma_j = d \quad (6.23)$$

Normal cells and intersection cells:

$$\begin{aligned} x_i(k+1) &= x_i(k) + f_j(k) - f_h(k), \\ \forall i \in C_N, C_I, \Gamma_j &= \Gamma_h^{-1} = i. \end{aligned} \quad (6.24)$$

Merging cells:

$$\begin{aligned} x_i(k+1) &= x_i(k) + \sum^j f_j(k) - f_h(k), \\ \forall i \in C_M, \Gamma_j &= \Gamma_h^{-1} = i. \end{aligned} \quad (6.25)$$

Diverging cells:

$$\begin{aligned} x_i(k+1) &= x_i(k) + f_j^r(k) - \sum^h f_h(k), \\ \forall i \in C_V, \Gamma_j &= \Gamma_h^{-1} = i. \end{aligned} \quad (6.26)$$

The flows must satisfy the following linear constraints,

Demand constraints:

$$f_j(k) \leq \frac{x_i(k)}{L_i} \cdot v_i, \quad (6.27)$$

$$f_j(k) \leq c_i, \quad \forall i \in C_O, C_N, C_M, C_I, \Gamma_j^{-1} = i$$

$$\sum^j f_j(k) \leq \frac{x_i(k)}{L_i} \cdot v_i, \quad (6.28)$$

$$\sum^j f_j(k) \leq c_i, \quad \forall i \in C_V, \Gamma_j^{-1} = i$$

Supply constraints:

$$f_j(k) \leq \beta_i \cdot x_i^J \left(1 - \frac{x_i(k)}{x_i^J} \right), \quad (6.29)$$

$$f_j(k) \leq c_i, \quad \forall i \in C_O, C_N, C_V, \Gamma_j = i$$

$$f_j(k) \leq \beta_i \cdot x_i^J \left(1 - \frac{x_i(k)}{x_i^J} \right), \quad (6.30)$$

$$\sum^j f_j(k) \leq c_i, \quad \forall i \in C_M, \Gamma_j = i.$$

Apart from the demand and supply constraints, intersection flows need to satisfy also the constraints of (6.3) and (6.4). The objective of the LQMPC is to minimize the quadratic function of the number of vehicles in the network to approximately minimize the total travel time. The overall optimization problem is formulated as:

$$\min \sum_{k=t}^{t+K_p-1} [X(k)'UX(k) + G\bar{f}(k)]$$

subject to (6.31)

$$(3)-(4), (22)-(30)$$

$$f_j(k) \geq 0, \quad \forall k = t, t+1, \dots, t+K_p-1.$$

where $X(k)$ is the vector of traffic state of all the cells excluding destination cells in the network and $\bar{f}(k)$ is the vector of all the flows in the network. K_p is the prediction horizon. U is a matrix with all 1 elements, so the term $X(k)'UX(k)$ aims to minimize the quadratic function of the number of vehicles in the network in the prediction horizon. G is a vector of the cell lengths multiplied by a small negative value. The term $G\bar{f}(k)$ aims to maximize the flows, so as to address the so-called holding back problem. Holding back may occur for a congested state, where the traffic control measures cannot have an influence. The second term in (6.31) (which was initially suggested in Papageorgiou (1995)) mitigates this undesirable occurrence by rewarding the solutions where the flows take higher values. Note that this second term in (6.31) has a physical meaning, as it is proportional to the total travel distance (TTD). Thus, (6.31) may be

perceived as a weighted combination of TTT minimization and TTD maximization. A more detailed explanation of this objective function can be found in [Le et al. \(2013\)](#). The optimization problem can be solved by an appropriate solver (e.g., CPLEX). The outputs of the controller are optimal flows between cells.

Optimal control signals, i.e. green time fraction $u_s^p(k)$ and turn fraction $\theta_i^j(k)$, are derived from the optimal flows by the following equations:

$$\begin{aligned}\theta_i^j(k) &= \frac{f_j(k)}{\sum_{j:\Gamma_j^{-1}=i} f_j(k)}, \\ u_s^p(k) &= \frac{\sum_{j \in W_s^p} f_j(k)}{\sum_p \sum_{j \in W_s^p} f_j(k)}.\end{aligned}\tag{6.32}$$

In [Le et al. \(2013\)](#), the LQMPC approach was applied to a single destination network for optimal vehicle routing, and the network is shown as Network 1 in Fig. 6.4. In the network, route choices occur at bifurcations, which are shown as red arrows in the figure. According to (6.31), traffic at bifurcations will be assigned to the optimal route, which leads to the minimum system cost. However, when the LQMPC approach is applied to Network 2 in Fig. 6.4, the desired OD relations may not be preserved. Suppose that the whole network is in the free flow. Then the controller will assign as much traffic as possible to Destination 1 because it is closer to the origins so that the controller can achieve a lower cost by pushing out flows as quickly as possible. However, this is not realistic because traffic have their own intended destinations and the traffic that is assigned to Destination 1 may intend to go to Destination 2. In other words, if we apply the optimal turn fraction θ_i^j obtained from (6.32) to the optimization problem in (6.20) for traffic assignment, there may be no feasible solution. In addition, traffic assignment errors in the LQMPC have significant influences to the signal control performances. For example, if the LQMPC assigns traffic flows from some OD pairs to a route that belongs to a different OD pair, then the intersection links in this route may get unnecessarily high green time because traffic in reality will not use these links. This may significantly decrease the intersection control efficiency.

6.4 An extended LQMPC approach

This section extends the method of the previous section to account for multiple destination networks. The reason why previous LQMPC may not work for multi-destination networks is that the prediction model of the LQMPC does not have any information about the OD relations of traffic flows, thus traffic dynamics cannot be accurately predicted. The way we tackle this problem is to impose constraints (i.e. lower or upper bounds of flows) to the LQMPC, to reduce the freedom of the controller in such a way

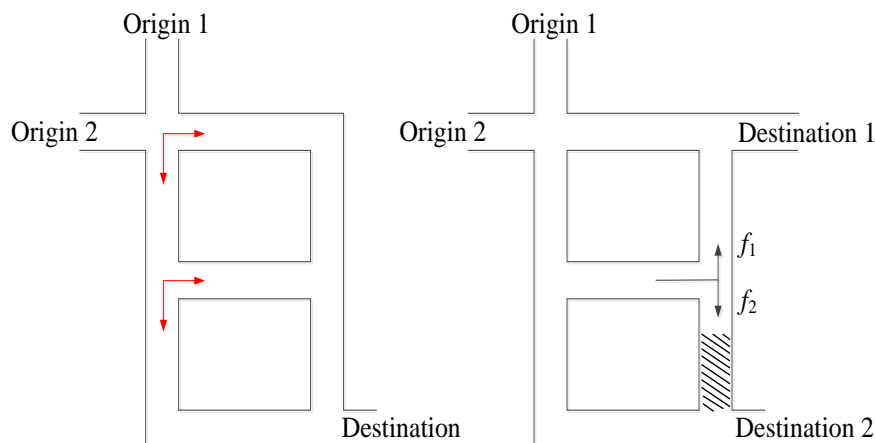


Figure 6.4: Example networks to explain the usage of the LQMPC. Network 1 is a single destination network and network 2 is a multiple destination network. Red arrows in network 1 indicate the bifurcations, where traffic can make their route choices. Grey area in network 2 represents congestion. If the receiving flow of the gray area is 0, and the turn fraction of link 2 is higher than 0, then according to (18), both f_1 and f_2 are 0.

that some unrealistic solutions, which may be optimal in the mathematical sense but not practically feasible, are excluded. The control approach is shown by a flowchart in Fig. 6.5. The extended LQMPC calculates the optimal flows based on real-time traffic states and some constraints to ensure the conservation of destinations. The constraints are set based on a heuristic method, in which the lower or upper bounds of flows are determined based on the OD relations of dynamic flow that is predicted by a forward simulation. The heuristic method intends to ensure the conservation of destinations, but cannot guarantee it. Thus, after each optimization run, the performance of the extended LQMPC is evaluated by the route-specific model. The performance of the extended LQMPC is compared with a non-optimizing control strategy, and the one that performs better is implemented into the process. This evaluation process ensures that the controller never performs worse than the non-optimizing approach.

The remaining of this section is set up as follows. Section 6.4.1 presents the forward simulation method. Section 6.4.2 - 6.4.4 introduce the heuristic approach to narrow the solution space and exclude unrealistic solutions. Section 6.4.5 explains the evaluation process after each optimization run of the extended LQMPC.

6.4.1 Forward simulation

Before each optimization run, we perform a forward simulation to predict the OD relation of dynamic flows. The forward simulation runs based on the route-specific model using a non-optimizing control strategy, which is the integration of a simple routing strategy and the back-pressure algorithm for intersection timing plans. The routing strategy guides traffic flows to the routes that have the shortest instantaneous travel time. Traffic flows are assigned at the origins based on the instantaneous travel time

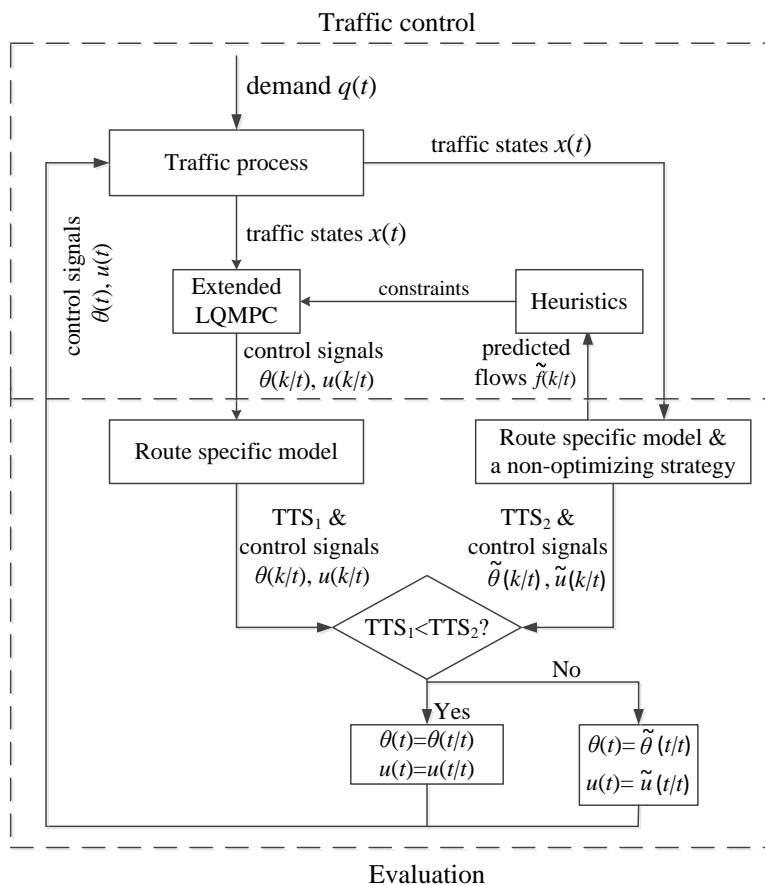


Figure 6.5: The flowchart of the control approach. There are two blocks in this flowchart, where the upper block is the traffic control part and the lower part is the control performance evaluation part.

of each route using (6.6) and (6.7), and it is assumed that there is no route choice for the traffic that is already in the network. A simplified back-pressure algorithm is used to determine the signal timing plans at intersections. For the original back-pressure algorithm, readers are referred to [Varaiya \(2013\)](#). Note that any non-optimizing control strategy can be used in the forward simulation. In this chapter, we use a simple routing strategy and a simplified back-pressure algorithm for convenience. The performance of the extended LQMPC may be related to the non-optimizing control strategy, because it serves as a lower bound of the performance of the proposed controller.

In the simplified back-pressure algorithm, there is a weight b_s^j associated with the link j at intersection s . The weight is the difference in the number of vehicles between the source cell and the sink cell.

$$b_s^j(\tau) = x_i(\tau) - x_l(\tau), \quad j \in W_s^p, \Gamma_j^{-1} = i, \Gamma_j = l \quad (6.33)$$

The back-pressure algorithm determines the active phase p of intersection s at every time slot τ based on the back pressure B of phase p , $B_s^p(\tau)$. The back pressure $B_s^p(\tau)$

is the summation of links weights in the phase, which is formulated as,

$$B_s^p(\tau) = \sum_j b_s^j(\tau). \quad j \in W_s^p \quad (6.34)$$

The phase that has the maximum pressure is active at τ . We assume that T_c is the time duration of a cycle and T_τ is the time duration of a slot time, and that T_c/T_τ is an integer. Then the green time fraction of phase p at intersection s is calculated as $\frac{n_s^p}{T_c/T_\tau}$, where n_s^p is the total number of activations of phase p in a cycle.

6.4.2 Minimum flow constraint on each OD pair

As introduced in the beginning of this section, the LQMPC may assign traffic flow to the routes that are not accessible to their desired destinations, which is not realistic and may undermine the controller performance. The reason is that the LQMPC has the freedom to assign the flows to any connected links (if the receiving flows are high enough), which may result in unrealistic flow distributions, in the sense that under the assumption that the destination are conserved, the flows cannot be realized. To this end, we restrict the freedom of the LQMPC by imposing hard constraints, to exclude some unrealistic solutions that may deteriorate the control performance. To illustrate how the hard constraints are imposed, we first introduce the concept of crucial links (c-links).

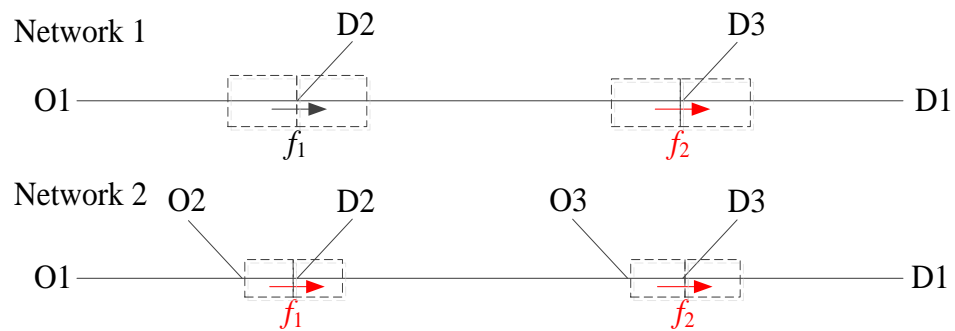


Figure 6.6: An example to show the c-links (red arrows) of an OD pair that has a single route. Dashed rectangles represent cells and arrows represent links. In Network 1, link 2 is the c-link of OD pair O1-D1. In Network 2, links 1 and 2 are the c-links of O1-D1.

For each OD pairs a set of c-links is defined, among the links that are on the routes corresponding to the given OD pair. We first give an example where there is only one route for each OD pair, shown as the networks in Fig. 6.6. In both networks, there are two bifurcations on the route that belong to the OD pair O1-D1, where the upstream one also directs to destination D2 and the downstream one also directs to destination D3. Traffic flows from O1 need to pass all the links in this route to reach the destination. Nevertheless, traffic flows may be guided to wrong destinations only at the bifurcations where the traffic from O1-D1 may be assigned to destinations D2

or D3. The same principle applies to other OD pairs of the network. In a route of an OD pair, if traffic flow of all diverging links that belong to the route comes from only one origin (e.g., the route of O1-D1 in Network 1), then only the most downstream link is a c-link because the flow that passes the most downstream diverging link has already passed all of the upstream diverging links. Therefore, only link 2 is the c-link in Network 1 and both links 1 and 2 are the c-links in network 2. The key of the extended LQMPC is to impose constraints to the flows of c-links, such that the predicted dynamic traffic flow obeys the desired OD relations or has a small bias that does not have a significant influence on the control performance.

In Network 1, the c-link constraint of OD pair O1-D1 is formulated as,

$$f_2(k) \geq \tilde{f}_2^1(k), \quad k = t, t+1, \dots, t+K_p-1. \quad (6.35)$$

where $\tilde{f}_2^1(k)$ is the flow of link 2 in route 1 (the route that belongs to O1-D1) that is predicted by the forward simulation. $\tilde{f}_2^1(k)$ is perceived as the lower bound of flow that needs to pass link 2 at time step k .

In Network 2, the c-link constraint of OD pair O1-D1 is formulated as,

$$\begin{aligned} f_1(k) &\geq \tilde{f}_1^1(k), \\ f_2(k) &\geq \tilde{f}_2^1(k), \\ k &= t, t+1, \dots, t+K_p-1. \end{aligned} \quad (6.36)$$

In Network 2, links 1 and 2 are also the c-links of OD pair O2-D1, thus links 1 and 2 are loaded with traffic flow from both O1-D1 and O2-D1. Under this circumstance, the c-link constraint is formulated as,

$$\begin{aligned} f_1(k) &\geq \tilde{f}_1^1(k) + \tilde{f}_1^2(k), \\ f_2(k) &\geq \tilde{f}_2^1(k) + \tilde{f}_2^2(k), \\ k &= t, t+1, \dots, t+K_p-1. \end{aligned} \quad (6.37)$$

where the route that belongs to O2-D1 is indexed by 2.

Now we extend the example to networks that have multiple routes in each OD pair, which are shown in Fig. 6.7. There are three routes in OD pair O1-D1, indexed as 1, 2, and 3 from top to bottom. Similar to the analysis of the single route example, link 1 (indicated by f_1) is the c-link of route 1 (i.e. traffic flow in O1-D1 that choose route 1 cannot reach the destination without passing link 1). Note that link 1 is a diverging link and other diverging links that originate from the same source cell may go to destination 3 or 4. Similarly, links 2 and 3 are the c-links of routes 2 and 3, thus links 1, 2, and 3 are the c-links of this OD pair. The c-link constraint of OD pair O1-D1 is formulated as,

$$\begin{aligned} f_1(k) + f_2(k) + f_3(k) &\geq \tilde{f}_1^1(k) + \tilde{f}_2^2(k) + \tilde{f}_3^3(k), \\ k &= t, t+1, \dots, t+K_p-1. \end{aligned} \quad (6.38)$$

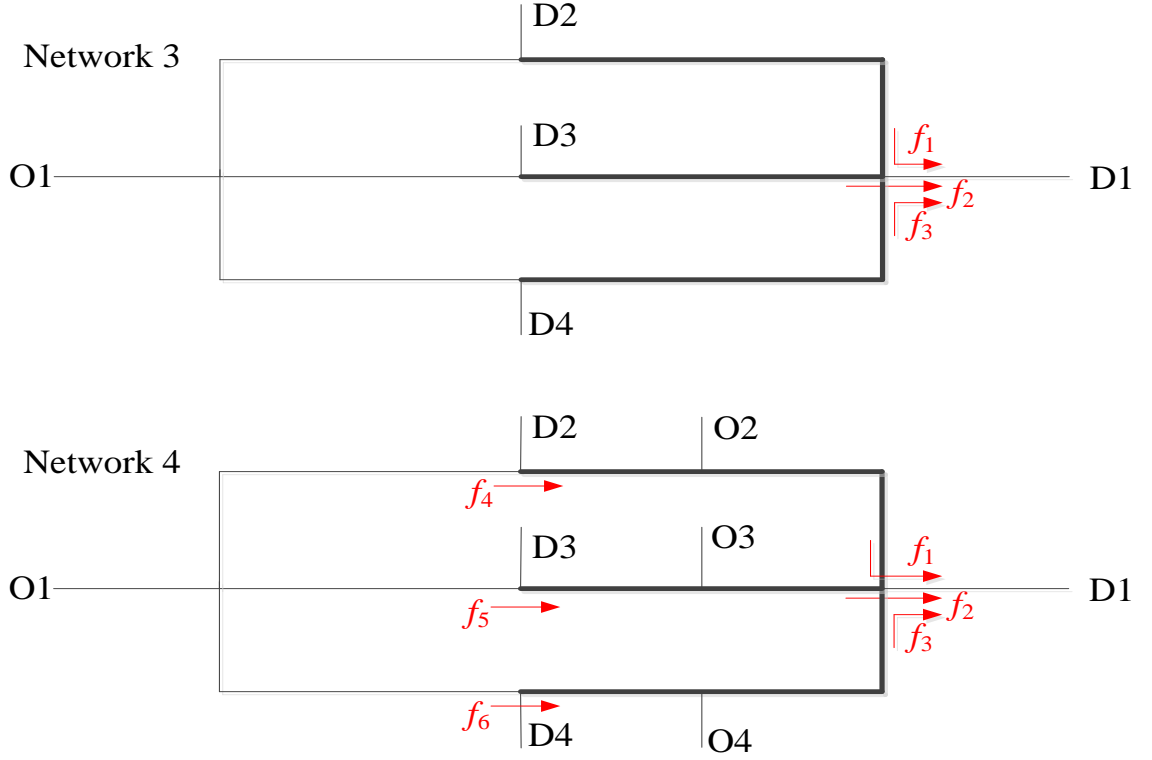


Figure 6.7: An example to show the c-links of an OD pair that has multiple routes. Bold lines represent the roadway with two directions. In Network 3, links 1, 2, 3 are the c-link of OD pair O1-D1. In Network 4, links 1, 2, 3 and 4, 5, 6 are the c-links of the OD pair.

where the right-hand side of the inequality represents the lower bound of flow that needs to pass the c-links of O1-D1. Note that such constraint does not restrict the route choice freedom of traffic flow. In Network 4, traffic flow from different origins is added to the network. Similar to the analysis of the single route example, links 1 and 4 are the c-links of route 1, links 2 and 5 are the c-links of route 2, and links 3 and 6 are the c-links of route 3. We use $J_{C(o,d)}^r$ to represent the set of c-links of route r in OD pair (o,d) . In this example, $J_{C(o,d)}^1 = \{1,4\}$, $J_{C(o,d)}^2 = \{2,5\}$, and $J_{C(o,d)}^3 = \{3,6\}$. The c-link constraint of O1-D1 is formulated as,

$$\begin{cases} f_1(k) + f_2(k) + f_3(k) \geq \tilde{f}_1^1(k) + \tilde{f}_2^2(k) + \tilde{f}_3^3(k) \\ f_1(k) + f_2(k) + f_6(k) \geq \tilde{f}_1^1(k) + \tilde{f}_2^2(k) + \tilde{f}_6^3(k) \\ \vdots \\ f_4(k) + f_5(k) + f_6(k) \geq \tilde{f}_4^1(k) + \tilde{f}_5^2(k) + \tilde{f}_6^3(k). \end{cases} \quad (6.39)$$

In each inequality, the left hand side contains one of the c-links from each route. The c-links of a route are indexed as κ^r , and $\kappa^r = 1, 2, \dots, \kappa^r$. The general formulation of

(35-36) and (38-39) is,

$$\sum_{C(o,d)}^r f_{J_{C(o,d)}^r(\mathcal{Z}^r)}(k) \geq \sum_{C(o,d)}^r f_{J_{C(o,d)}^r(\mathcal{Z}^r)}(k), \quad (6.40)$$

$$\forall r \in R^{o,d}, \forall \mathcal{Z}^r = 1, 2, \dots, \kappa^r.$$

If link j is the c-link of multiple OD pairs, i.e., $\exists r' \in R^{o',d'}$, $(o',d') \neq (o,d)$: $j = J_{C(o,d)}^r(\mathcal{Z}^r) = J_{C(o',d')}^{r'}(\mathcal{Z}^{r'})$, then the right hand side of each OD pair's inequalities that contains the route flow of link j need to be integrated, like the example shown in equation (6.37). The general formulation of (6.37) is,

$$\sum_{C(o,d)}^r f_{J_{C(o,d)}^r(\mathcal{Z}^r)}(k) \geq \sum_{o,d} \sum_{C(o,d)}^r f_{J_{C(o,d)}^r(\mathcal{Z}^r)}(k), \quad (6.41)$$

$$\forall r \in R^{o,d}, \forall \mathcal{Z}^r = 1, 2, \dots, \kappa^r.$$

To summarize, we elaborate several features of c-links. (i) C-links are the diverging links that do not direct traffic to a same destination. (ii) If the most downstream diverging link of a route has destination choice, then it is a c-link. (iii) If traffic from other origins joins the links that are between two consecutive diverging links, then the upstream diverging link is a c-link. We present a generic algorithm to search for the c-links of each route. $J_{V(o,d)}^r$ is the set of diverging links in route r . The diverging links are indexed as v^r , and the diverging links in route r are represented as $J_{V(o,d)}^{r,v^r}$. $v^r=1,2,\dots,\vartheta^r$, numbered from downstream to upstream. $\mathcal{Z}^r=1$;

$v^r=1$;

$\forall j': \Gamma_{j'}^{-1} = \Gamma_{J_{V(o,d)}^{r,v^r}}^{-1}$, if $\exists r, \exists d': j' \in J^r, r \in R^{o,d'}, d' \neq d$

$$J_{C(o,d)}^{r,\mathcal{Z}^r} = J_{V(o,d)}^{r,v^r};$$

$$v^r = v^r + 1;$$

end

for $v^r = 1: \vartheta^r$

$\forall o' \neq o$, if $\exists r \in R^{o,d'} : J_{V(o,d)}^{r,v^r} \in J^r, J_{V(o,d)}^{r,v^r+1} \notin J^r$

if $v^r+1 \leq \vartheta^r$

$$J_{C(o,d)}^{r,\mathcal{Z}^r+1} = J_{V(o,d)}^{r,v^r+1};$$

$$v^r = v^r + 1;$$

end

end

end

6.4.3 Reproducing spillback

The c-link constraint narrows the solution space of the LQMPC such that the OD relations are better preserved. However, the LQMPC with the c-link constraint cannot reproduce spillback. Spillback is the phenomenon that the congestion of a downstream link affects the possible outflow volume of the upstream link. The route-specific model presented in Section 6.2 is able to reproduce spillback. For example, if the grey area in network 2 of Fig. 6.4 is fully blocked (the receiving flow is 0), and the turn fraction of link 2 is higher than 0, then according to (6.18-6.19), the total flow of f_1 and f_2 is 0. The classical LQMPC approach introduced in the previous section is not able to reproduce the spillback phenomenon because the LQMPC can always limit the turn fraction of a link to zero (or one, or small or large enough) if one of the receiving cell is congested, in such a way that the outflow to other link is still maximal. However, such kind of traffic assignment is not always realistic. If all of the traffic who intend to pass the congested link can reach their desired destinations by changing routes to the uncongested link, then it is realistic for the LQMPC to limit the turn fraction of the congested link to 0. Nevertheless, if some traffic cannot reach their desired destinations by changing routes, then it is not realistic to avoid spillback by forcing traffic to change their routes. If the LQMPC still limits the turn fraction of the congested link to 0, there will be a discrepancy between traffic dynamics that predicted by the LQMPC and real traffic dynamics, which may have a significant influence on the control performance.

In the classical LQMPC, the supply constraint (6.30) is imposed to each individual diverging link. To reproduce spillback, the supply constraint should be imposed to the total flow of all diverging link that originate from one diverging cell, as explained in equation (6.18). To this end, we define a minimum turn fraction of a diverging link, $\theta_{\min}^{i,j}(k)$, which indicates the proportion of traffic who cannot reach their desired destination without passing link j . Since the route-specific model keeps track of the OD relations of traffic flows, $\theta_{\min}^{i,j}(k)$ can be easily obtained from the forward simulation. According to (6.18), the following constraint is added to the LQMPC to reproduce spillback.

$$\sum_{v=1}^n f_{l_v}(k) \leq \min \left(\frac{R_{l_1}(t)}{\theta_{\min}^{i,l_1}(t)}, \frac{R_{l_2}(t)}{\theta_{\min}^{i,l_2}(t)}, \dots, \frac{R_{l_n}(t)}{\theta_{\min}^{i,l_n}(t)} \right), \quad (6.42)$$

$$\forall i \in C_V, k = t, t+1, \dots, t+K_p-1.$$

6.4.4 Destination constraint

The extended LQMPC always have a tendency to assign traffic flow to attractive destinations where traffic can leave the network faster, even if the c-link constraint are imposed. To avoid that the traffic flows assigned to attractive destinations are higher than the maximum flows, we add destination constraint to balance the flows arriving at each destination.

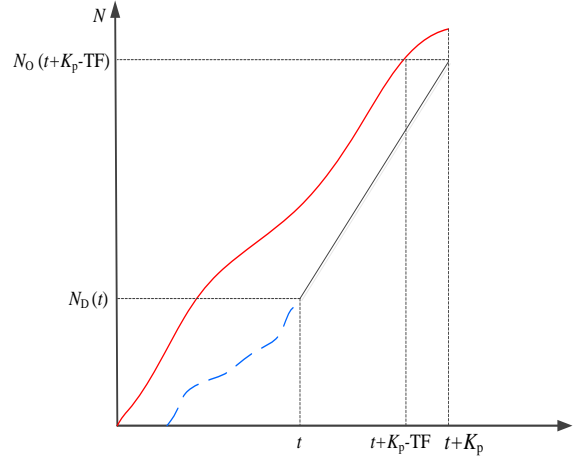


Figure 6.8: The cumulative curves at the origin and the destination of an OD pair. The red line is the cumulative curve at the origin and the dashed blue line is the cumulative curve at the destination. $N_O^{o,d}(t)$ and $N_D^{o,d}(t)$ are the cumulative number of vehicles at the origin and the destination. t is the current time that the controller runs, and $TF_{o,d}$ is the free flow travel time of the OD pair. The difference between $N_O^{o,d}(t + K_p - TF_{o,d})$ and $N_D^{o,d}(t)$ is the maximum number of vehicles that can travel to destination d from origin o from t to $t + K_p$. For destination d , the maximum number of arriving vehicles is $\sum^o \left(N_O^{o,d}(t + K_p - TF_{o,d}) - N_D^{o,d}(t) \right)$

The destination constraint is set based on the maximum possible arriving flow of each destination. As shown in Fig. 6.8, the red line represents the cumulative curve at the origin of an OD pair, and the dashed blue line represents the cumulative curve at the destination. Suppose that the controller runs at time t , then the maximum possible number of vehicles that can arrive to the destination within the prediction horizon K_p is $\sum^o \left(N_O^{o,d}(t + K_p - TF_{o,d}) - N_D^{o,d}(t) \right)$, where $TF_{o,d}$ is the free flow travel time of the OD pair. $N_O^{o,d}(t)$ and $N_D^{o,d}(t)$ are the cumulative number of vehicles at the origin and the destination. We set the following constraint to the extended LQMPC, to make the flow that arrives to each destination proportional to the maximum possible number of vehicles that can arrive to the destination within the prediction horizon.

$$\frac{\sum_{k=t}^{t+K_p-1} f_j(k)}{\sum_{k=t}^{t+K_p-1} f_{j'}(k)} = \frac{\sum^o \left(N_O^{o,d'}(t + K_p - TF_{o,d'}) - N_D^{o,d'}(t) \right)}{\sum^o \left(N_O^{o,d'}(t + K_p - TF_{o,d'}) - N_D^{o,d'}(t) \right)}, \quad (6.43)$$

$$\forall d, d' \in C_D, \Gamma_j = d, \Gamma_{j'} = d'.$$

The destination constraint balances the number of vehicles that arrive at each destination, to avoid that redundant traffic flow is assigned to attractive destinations. For example, if a bottleneck is activated, which blocks the flow that will arrive at a (unattractive) destination, then the LQMPC without the destination constraint will try to push

the flow to other destinations to maximize the throughput. The destination constraint forces the flow that arrives to each destination is proportional. Thus, under that circumstance, the LQMPC with the destination constraint tries to resolve the congestion first, because if the flow that arrives to the unattractive destination is low, the flows that arrives to any destination are low.

6.4.5 Evaluation step of the extended LQMPC

We assume that the route-specific model introduced in Section 6.2 is accurate enough to represent reality. Suppose that the route-specific model is applied in a model predictive control framework to optimize the network performance (e.g. minimize the total time spent). The optimization problem is formulated as,

$$\min \sum_r \sum_{\substack{i \\ \forall i \notin C_D}} \sum_{k=t}^{t+K_p-1} x_i^r(k) \cdot T$$

(6.44)

subject to
(1)-(19)

$$f_j(k) \geq 0, \quad \forall k = t, t+1, \dots, t+K_p-1.$$

where C_D is the set of destination cells. The above optimization problem is non-linear because the traffic dynamic represented by the route-specific model is non-linear. The original MPC can be perceived as a simplification of the non-linear optimization problem. However, the solution space of the original LQMPC is larger than the non-linear optimization, because the original LQMPC has the freedom to assign vehicles to any links that are not accessible to their desired destinations. A conceptual depiction of the solution spaces of the optimizations is shown in Fig. 6.9. The solution boundary of the non-linear optimization is shown as the black line. If the optimal solution of the original LQMPC lies in the boundary of the solution space (the red line), then it has a large bias towards the optimal solution of the non-linear optimization, which may have significant influence on the control performance.

The c-link constraint and the destination constraint introduced in the previous sections narrow the solution space of the LQMPC to exclude unrealistic solutions. In other words, the implementation of the constraints reduces the distance of the solution space boundary between the extended LQMPC and the non-linear optimization. Therefore, even if the optimal solution of the extended LQMPC lies in the boundary of the solution space (shown as the dashed blue line in Fig. 6.9), it will have less influence to the control performance.

Note that the constraints cannot guarantee the solution space of the extended LQMPC to be exactly the same as the non-linear optimization. The control performance may not be good if, for example, the optimal solution of the extended LQMPC lies in the boundary of the solution space. To avoid extreme bad performance of the extended LQMPC, we evaluate the performance of the extended LQMPC at the end of each

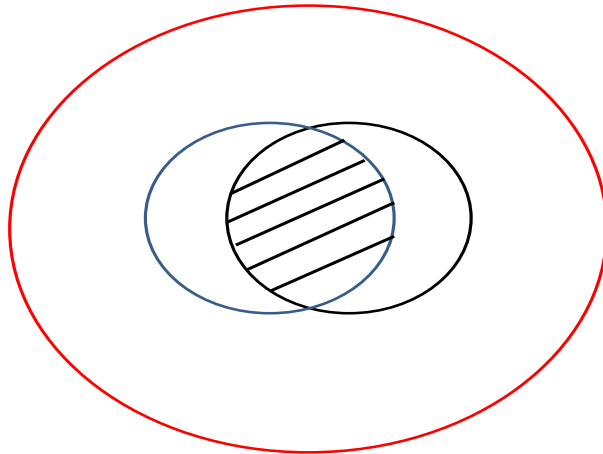


Figure 6.9: A conceptual explanation about the boundaries of the solution spaces of the original LQMPC (the red line), the extended LQMPC (the blue line), and the non-linear optimization (the black line).

optimization run. As shown in Fig. 6.5, the optimal control signals obtained from the extended LQMPC are implemented into the route-specific model and compared with the non-optimizing control strategy. If the performance of the non-optimizing strategy is better than the extended LQMPC, then the non-optimizing strategy will be used for real-time traffic control.

6.5 Case study

We test the performance of the presented control approach in a synthetic network, which is shown in Fig. 6.10. The synthetic network contains 7 OD pairs, and the demand patterns are shown in Tab. 6.1. The bold line that connects A and B represents an urban expressway and other lines represent urban arterial roads. The length of each roadway section is shown in the figure. There are two signalized intersections in the network. The movements in each intersection are shown in Fig. 6.11. The synthetic network contains two potential bottlenecks, which are shown as 'BN1' and 'BN2' in Fig. 6.10. BN1 represents an accident area, and the capacity of this area reduces to half of the free flow capacity when it is activated. BN2 represents the blockade of the downstream boundary, and the capacity of this area reduces to one-third of the free flow capacity when it is activated.

Table 6.1: The demand pattern of the synthetic case.

OD pairs	AB	AD	AF	CB	CF	EB	ED
Demand veh/h	1500	150	500	1000	500	300	150

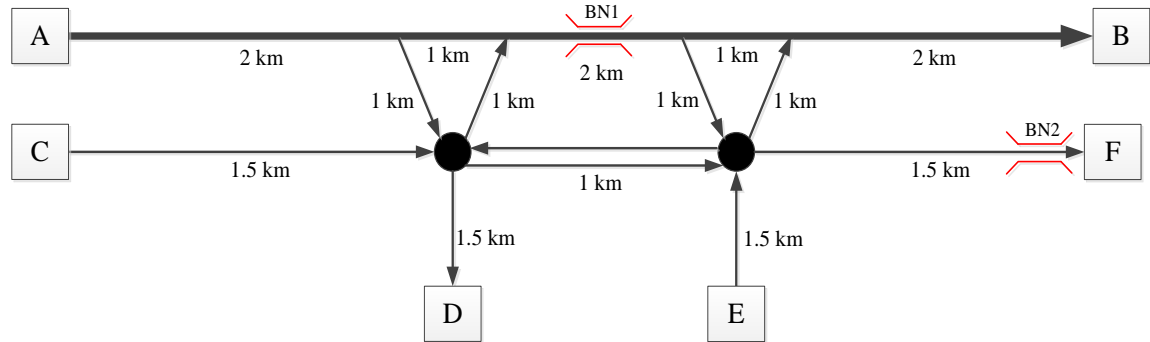


Figure 6.10: Sketch of the synthetic network. The bold line indicates freeway and other lines indicate urban arterials. 'BN1' and 'BN2' represent bottlenecks. The circles represent signalized intersections.

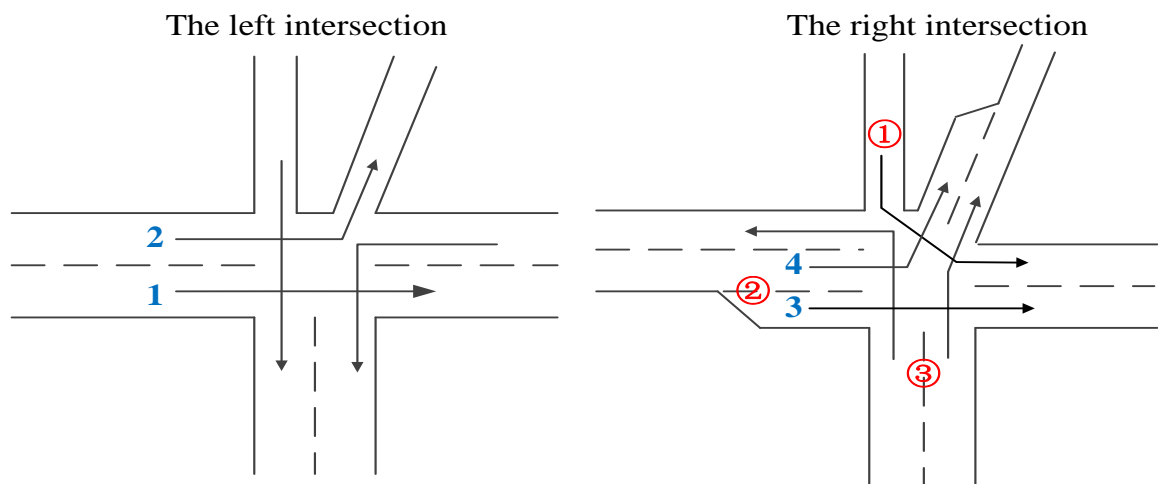


Figure 6.11: The depiction of the movements. Red circled numbers in the right intersection are the index of roadway sections at the upstream of the intersection. Blue numbers represent the index of some of the links.

The route-specific model is used as the process model to represent reality. The parameters of the process model are set up as follows. The free flow speed v and the cell length l are set to 120 (km/h) and 1 (km/cell) for the freeway, and 60 (km/h) and 0.5 (km/cell) for urban roads. The duration of one simulation step, T , is set to 30s, which satisfies the CFL condition. The capacity c is set to 2000 (veh/h/lane) for the freeway and 1800 (veh/h/lane) for urban roads. The freeway has two lanes and urban roads have only one lane. The saturation flow Q of intersection movements are set to 1800 (veh/h). The congestion wave speeds for all cells are set to -20 (km/h). The parameter σ , which is used to model traffic route choices at origins, is set to 0.05 (1/s).

Three control scenarios are tested in the synthetic case study.

1. Scenario 1 is the referential scenario which indicates that the solution is close to the system optimal. The route-specific model is used as both the prediction model and the process model. The prediction horizon K_p is set to 15 minutes and

the control horizon K_c is set to 30 seconds. The non-linear optimization problem is solved by MATLAB implementation of the SQP algorithm (fmincon).

2. Scenario 2 applies the presented extended LQMPC. K_p and K_c are set to same values as in scenario 1. In the forward simulation, the slot time τ of the simplified back-pressure algorithm is set to 1 second.
3. Scenario 3 applies the non-optimizing control strategy. The slot time τ of the simplified back-pressure algorithm is set to 1 second. Traffic is routed at the origins according to (6) based on the instantaneous travel time.

We test the performance of the three control scenarios for a simulation period of an hour. It is assumed that 'BN1' is activated during the first half an hour and 'BN2' is activated during the second half an hour. At the first half an hour, due to the capacity reduction of the freeway, controllers need to assign less traffic flow that comes from origin C to the freeway, to avoid congestion occurs at the freeway bottleneck. At the second half an hour, the demand of destination F is higher than the maximum outflow. Under such circumstance, if the congestion propagates to the intersection and blocks link 3 in Fig. 6.11, then the flow of link 4 is also blocked which results in a reduction of the outflow at destination B. Thus, controllers need to prevent the blockade of the intersection for as long as possible.

We run the simulations of the three control scenarios and the results are shown in Fig. 6.12. To investigate how OD relations are preserved by the original LQMPC, we run the simulation of another scenario (scenario 4) where both the process model and the prediction model are the macroscopic model that does not keep track of flows in each route. The TTS of four control scenarios are 322.5h, 324.7h, 544.8h, 126.2 h respectively. The outflows of destinations B, D, and F in scenario 4 (the original LQMPC) and 1 (the system optimal) are shown in Fig. 6.12 (a). It can be seen that the outflows of each destination is significantly different in scenarios 4 and 1. The original LQMPC pushes out more flows (around 2000 veh/h) from destination D (dotted pink line) and less flows (nearly 0) from destination F (dashed pink line). This is because destination D is closer to the origins so that the controller can achieve a lower cost by pushing out higher flow via destination D. However, this violates the conservation of destinations because the desired demand at destination D is very little (300 veh/h). Therefore, even though the original LQMPC achieves a much less TTS, it cannot be applied to reality because desired demand goes to wrong destinations.

Fig. 6.12 (b) shows the comparison of the outflows between scenarios 1 and 2, and (c) shows the comparison between scenarios 2 and 3. In Fig. 6.12 (c), it can be seen that at the first 60 steps (half an hour), the outflows in scenario 2 (the presented control approach) are generally higher than which in scenario 3 (the non-optimizing control approach). This is attributed to the better operation of the route guidance. Due to the fact that the capacity at BN1 is only 2000 veh/h and traffic flows of OD pairs AB and AF (the total demand of which is 2000 veh/h) have to pass BN1, extra traffic flow that

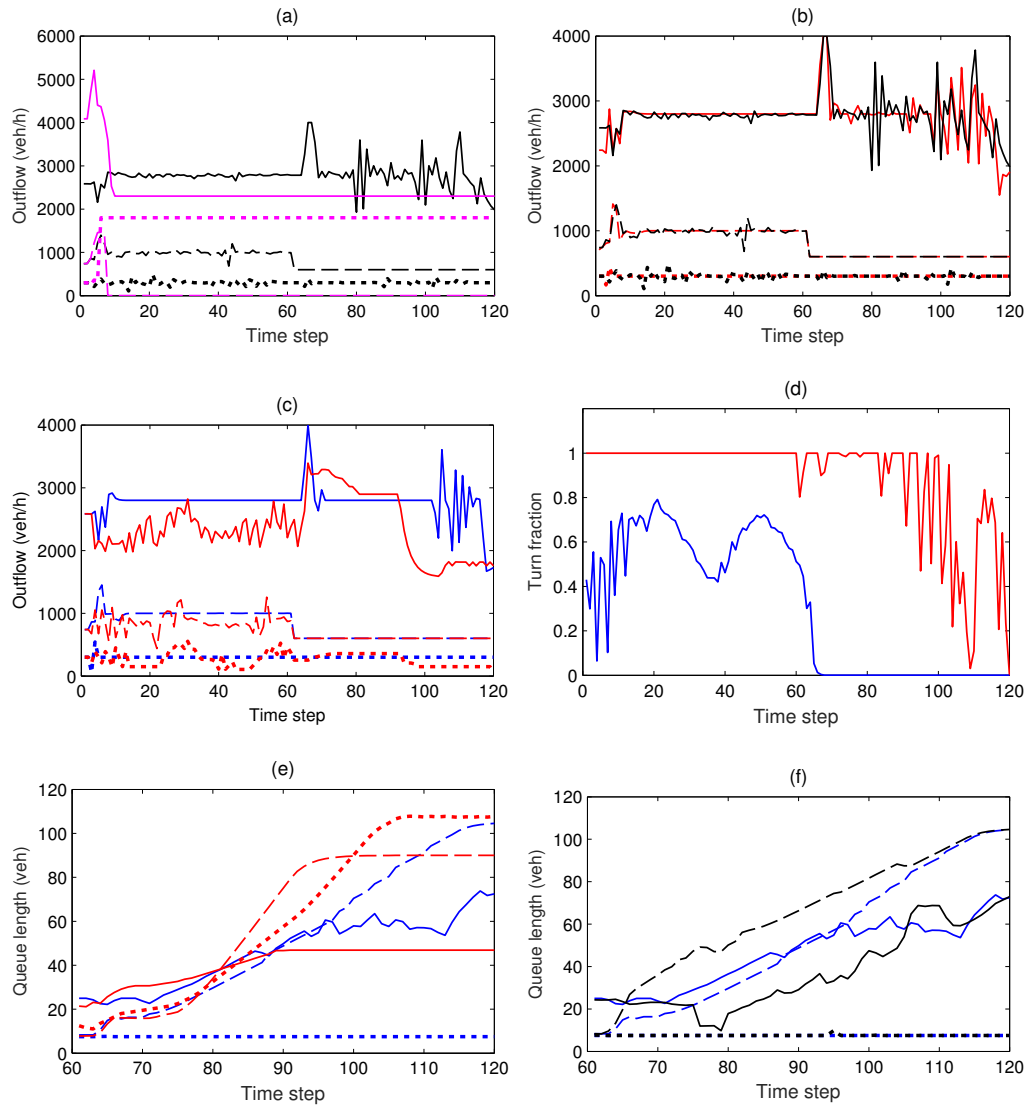


Figure 6.12: Simulation results of the case study. Line colors represent the results from different scenarios. Black: scenario 1, system optimal. Blue: scenario 2, the proposed control approach. Red: scenario 3, the non-optimizing approach. Pink: scenario 4, the original LQMPC. (a), (b), and (c) are the outflows of scenarios 4 and 1, scenarios 1 and 2, and scenarios 2 and 3. The solid lines, dashed lines, and dotted lines represent the outflows of destinations B, D, and F respectively. (d) is the turn fraction of link 1 at the left intersection in Fig. 6.11. (e) and (f) are the queue lengths of section 1, 2 and 3 in Fig. 6.11. The solid lines, dashed lines, and dotted lines represent the queue lengths of road sections 1, 2, and 3 respectively.

comes from origin C will trigger congestion on the freeway. Thus, the controller in scenario 2 assigns all of the flow from origin C to urban roads. As shown by the blue line in Fig. 6.12 (d), in the first 60 steps the turn fraction of link 1 in Fig. 6.11 is 1, which means that the controller assigns all of the flow from origin A to urban roads to avoid triggering congestion on the freeway. The non-optimizing approach assigns traffic flows based on the instantaneous travel time, and the turn fraction shows oscillated behavior. Once the freeway has a lower instantaneous travel time, the controller assigns more flows to the freeway which results in travel time increasing and turn fraction decreasing, and vice versa. Consequently, the outflow also has oscillated behavior. When the turn fraction of link 1 is high, the outflow are also high because more traffic flow avoids being delayed in the congestion that occurs on the freeway. The presented control approach keeps high outflow for the first 60 steps, which is close to the system optimal controller.

In the second 60 steps, the capacity at BN1 recovers to 4000 (veh/h) and a new bottleneck, BN2, is activated. The total demand of destination F is 1000 (veh/h), which is higher than the maximum outflow of the bottleneck, 600 (veh/h). Thus, congestion occurs at destination F and propagates to the right intersection. If the road section at the upstream of destination F is fully blocked, the flows of both link 3 and 4 in Fig. 6.11 are 0, which will reduce the outflow of destination B. Thus, the controller needs to gate the flow at road section 3 to prevent the congestion spillbacks to the intersection. Controllers of scenario 1 (system optimal) and scenario 2 (the presented control approach) have similar control patterns. As shown by the dashed lines in Fig. 6.12 (f), a queue is accumulated at road section 2 from time step 60. The gating of road section 3 generates a queue and if it spillback to the freeway, the freeway mainstream flow will reduce and the outflow of destination B will decrease. Thus, the controller should balance the queue lengths of road sections 1 and 2 to avoid the reduction of the total outflow for as long as possible. At the end of the simulation, since both sections 1 and 2 are blocked and the controller cannot get any benefit from gating, the outflow starts decreasing. For the non-optimizing control approach, as shown in Fig. 6.12 (e), the controller balances of queue lengths of sections 1, 2 and 3 because of the feature of the back-pressure algorithm. However, gating at road section 3 but releasing the flows of road sections 1 and 2 prompts the blockade of the intersection. Thus, the outflow of destination B in scenario 3 starts decreasing before time step 90, which is much earlier than which in scenarios 1 and 2.

The computer with an E5-1620 processor and 16 GB RAM is used for the optimizations. The computation time of the presented controller and the system optimal controller are 1.2 seconds per control step and 89 seconds per control step, respectively. In this case study, the presented controller achieves a performance that is close to the system optimal, while it keeps a tractable computation time.

6.6 Conclusions

In this chapter, we presented an extended LQMPC approach for multi-destination networks. The previously proposed LQMPC approach is ineffective for multi-destination networks because the prediction model cannot preserve traffic OD relations. The optimal solution of the previous LQMPC may violate the flow conservation of destinations, which may significantly influence the control performance. To overcome this shortcoming, we introduced a heuristic method to narrow the solution space of the LQMPC by constraints, so as to exclude unrealistic solutions. In principle, not all of the unrealistic solutions can be excluded by the constraints and it is still possible that the optimal solution does not conserve destinations. To this end, we evaluate the control performance at the end of each optimization run. The performance of the extended LQMPC is compared with a non-optimizing control approach, and the one that has a better performance is implemented to the process. This approach ensures that the presented control approach never has a worse performance than the non-optimizing approach.

The presented control approach is tested in a synthetic network that has multiple bottlenecks. When the bottleneck appears at the expressway, the controller effectively assigns traffic flows to the parallel arterial road, and when the bottleneck appears at an urban intersection, the controller effectively gates flow to maximize the intersection outflow. It is of particular significance that the computation time of the presented control approach is real-time tractable, which is essential for field applications. A future step is to test the presented approach by using microscopic simulation tools.

Chapter 7

Conclusions

In the previous chapters fast model predictive control strategies are developed for control measures, such as ramp metering, variable speed limits, route guidance, and intersection signal. This final chapter summarizes the main contributions, findings and conclusions of each chapter, elaborates on the practical relevance and potential field applications, and points out the directions for future research.

7.1 Main findings and conclusions

This thesis develops several traffic flow models that are accurate to reproduce traffic dynamics, and proposes several efficient model predictive control strategies for traffic control measures such as variable speed limits, ramp metering, route guidance, and intersection signals. In the following we elaborately present the main findings and conclusions of each chapter.

Chapter 2 proposes an extended discrete first-order model to reproduce the propagation of jam waves on freeways. The model has a better accuracy in reproducing the capacity drop and the propagation of jam waves than the models that have the same type, for example, the models in [Daganzo \(1994\)](#); [Roncoli et al. \(2015a\)](#); [Muralidharan \(2012\)](#), which are evidenced by simulations of a synthetic freeway stretch and a real life freeway stretch. From both the synthetic case and the real life case, qualitatively, it is found the jam wave reproduced by the proposed model propagates longer in time and space than that reproduced by the other models, which fits reality better. Quantitatively, the root-mean-square error of the proposed model is much lower than that of the other models. Besides the advancement in reproducing traffic phenomena, the proposed model keeps the advantage of the discrete first-order models in respect of computation efficiency. The linear property of the proposed model enables it to be applied to a linear optimization problem for traffic control purposes.

Chapter 3 proposes an MPC controller that has a linear quadratic formulation to resolve freeway jam waves. The proposed MPC is developed based on a new extended discrete first-order model (proposed in Chapter 2) which can reproduce capacity drop and the propagation of jam waves accurately. The proposed MPC controller is applied to a benchmark freeway stretch to test the performance in terms of computation speed and the resolution of jam wave. The second-order model METANET, which can reproduce the capacity drop phenomenon, is used as the process model to represent the reality. Two other controllers, one non-linear MPC based on the METANET model and the other linear MPC based on a first-order model ([Roncoli et al., 2015a](#)), were also tested in the benchmark problem for comparison. In the benchmark problem, the proposed controller shows a better performance in resolving the jam wave and reducing the total travel delay compared to the linear MPC. Meanwhile, the proposed controller keeps a real time feasible computation time (9 seconds per control iteration), which has a significant improvement comparing to the METANET-based MPC (6-8 minutes per control iteration). Simulation results highlight that the proposed controller has a good potential of being implemented into practice.

Chapter 4 validates the prediction model of a MPC controller of VSLs that is based on a new extended discrete LWR model. Through two synthetic cases it is shown that the model is able to reproduce the key mechanisms that are applied by existing VSL control algorithms. Simulation results show that the model can reproduce the traffic states transition of VSLs used by the SPECIALIST algorithm in [Hegyi et al. \(2008\)](#), and the gating effect applied by the mainstream metering control approach in [Carl-](#)

son et al. (2010b). Furthermore, we calibrate and validate the model with real data to demonstrate its accuracy in terms of predicting the evolution of jam waves under the VSL control. As the target for the model validation, we choose a successful case and a failed application from the SPECIALIST field test to show if the model underestimates or overestimate the effects of VSLs. The proposed model is compared with an extended METANET model which also incorporates with VSLs. The performance of two models are comparable: both models reproduce traffic states transition and the evolution of jam waves accurately. Quantitatively, both models reproduce accurate results which have a good match with real data. Flow errors of the calibration and validation are around 10%. The quantitative results are in line with other model validation works in the literature. Besides, similar quantitative errors between the prediction model and the process model are also obtained in our previous paper which used the proposed model for a MPC controller to resolve freeway jam waves, which gives confidence of the model validation results. The MPC controller has been demonstrated to be computationally more efficient. Model validation results demonstrate that the proposed MPC can make accurate prediction on the evolution of jam waves under the VSL control, which is a step towards the real life application.

In chapter 5, we first investigate the features of the freeway MFD and then present a model to predict traffic dynamics based on the MFD. A main feature of the model is that the evolution of the density heterogeneity is considered, such that the MFD can predict the aggregated traffic states in the freeway network accurately, which benefits the MFD-based controller. We also incorporate the capacity drop into the freeway MFD. By proposing an exponential functional form, the accuracy of the proposed MFD is validated with field traffic data. The proposed MFD model is incorporated in a hierarchical control structure for coordinated ramp metering on a freeway stretch. At the upper level, a model predictive control approach is developed to optimize the total inflow from on-ramps to the freeway stretch. The lower level controller distributes the optimal total inflow to each on-ramp of the freeway based on local traffic state feedback. The proposed control approach demonstrates a performance close to optimal in reducing the total time spent and eliminating congestion. The robustness of the controller is tested by using different traffic flow models (the extended CTM and the METANET model) as the process model and comparing with MPC controllers that are based on different prediction models. Simulation results highlight two advantages of the proposed controller, (i) the robustness to overcome the mismatch between the prediction model and the process model, and (ii) the proposed controller requires less computation efforts than other non-linear optimal controllers.

Chapter 6 extends an existing linear quadratic model predictive control (LQMPC) approach to multi-destination traffic networks, where correct OD relations are preserved. The original LQMPC approach was presented for efficient routing and intersection signal control (Le et al., 2013). From a simulated case it is shown that the previous LQMPC approach cannot preserve correct OD relations, so it cannot be applied to multi-destination networks because desired demand may end up at wrong destinations.

To tackle this problem, we present a heuristic method to narrow the solution space and exclude unrealistic solutions that would lead to flows that are inconsistent with the OD relations. The extended LQMPC approach is tested in a synthetic network with multiple bottlenecks. It shows that the LQMPC achieves a total time spent close to the system optimal and the computation time remains tractable.

7.2 Implications for practice

This thesis has several important practical implications and contributions, which can be categorized into the development of traffic simulation tools and the provision of real-time applicable traffic control strategies.

The extended cell-transmission model proposed in this thesis provides an efficient tool that can be used for freeway traffic simulation, short-term traffic prediction, and evaluations of traffic control performances. The original cell-transmission model has been extensively used in those aspects due to its linear formulation and simple solution method. The proposed model keeps the linear formulation and the simple solution method, whereas significantly improves the model accuracy in reproducing the propagation of freeway jam waves, as is evidenced in the model validations. More importantly, the proposed model can be embedded in model predictive controllers which have linear or linear quadratic formulations that can be efficiently solved.

This thesis proposes a freeway MFD model that can accurately predict the outflow of freeway networks. The proposed MFD model is an efficient tool to predict and evaluate the operation efficiency of freeways (in terms of total outflow or total time spent) at a network level. Compared to macroscopic simulation tools, e.g., the cell transmission model and the METANET model, the MFD model (i) has less parameters, which reduces model calibration efforts, and (ii) requires less traffic data to estimate traffic state. These features make the proposed MFD model to be a more efficient tool in predicting the aggregated traffic state of freeway networks.

This thesis proposes several model predictive control strategies that are fast enough to be implemented into the field. The proposed MPC controller of VSLs, which intends to suppress freeway jam waves, is tested successfully in a homogeneous freeway stretch in simulated cases. The proposed controller resolves jam waves in a real-time tractable computation speed (6-9s per control iteration), which is a crucial step towards field application. The prediction model of the MPC controller is validated with reasonable accuracy, which also gives confidence that the controller can work well in reality. Compared to the SPECIALIST algorithm, which has already been successfully implemented, the proposed control approach has several advantages, for example, to correctly handle disturbances such as the change of the demand, and easy to be adapted to new situations such as changes in infrastructure. These advantages may lead to a better control performance to further reduce the total time spent.

The proposed hierarchical ramp metering control strategy has two features that brings it closer for field application. First of all, it has a fast computation speed that is real-time tractable. The proposed control approach requires much less computation efforts than optimal controllers based on macroscopic models such as the cell transmission model and the METANET model, as a result of the aggregated feature of the freeway MFD model. Secondly, the proposed control approach demonstrates a robust performance to overcome the mismatch between the prediction model and the process model, which increases the confidence that the proposed approach will also accurately work in practice, where there is always a model mismatch.

This thesis proposes an extended LQMPC strategy of dynamic route guidance and intersection control. The same as other controllers that are proposed in this thesis, the extended LQMPC runs fast enough for a real time application. What's more, we propose an evaluation process after each optimization run of the controller. The performance of the extended LQMPC is compared with a non-optimizing control strategy after each optimization run, and the one that has a better control performance is implemented into the process. This evaluation procedure guarantees that the proposed control approach perform better than the non-optimizing approach, which may increase the robustness of the controller, especially under high uncertainties such as errors of demand estimation and the mismatch between the prediction model and reality.

7.3 Recommendations for future research

This thesis presents a new extended cell-transmission model to reproduce the capacity drop and the propagation of jam waves. The model is validated with reasonable accuracy to reproduce traffic dynamics in a homogeneous freeway stretch. A future step is to validate the model in a freeway network that includes on-ramps and off-ramps, and geometric changes of infrastructure. It is crucial for a model to reproduce different type of bottlenecks, e.g., on-ramp bottlenecks, lane drop bottlenecks, etc, because in reality a freeway network may not always be homogeneous. In addition, the proposed model may be extended to multi-class and multi-lane to account for different classes of vehicles and the lateral movements. This may increase the accuracy of the model and makes it possible to integrate more control measures such as lane changing control, which may further increase the control performance. What's more, the proposed model considers only self-driving cars. In recent years, autonomous vehicle technology have matured sufficiently in testing on public roads. Since automatic vehicles may have a significant impact on traffic flow characteristics, it will be interesting to extend the proposed model accounting for different market penetration rate of automatic vehicles.

The proposed MPC strategy of VSLs may be extended in several aspects. First of all, safety constraint needs to be implemented into the controller. For safety, normally the VSL difference between the neighboring cells should be lower than a certain value. This is challenging because the controller does not determine the optimal VSL rates

directly, but calculate it as the ratio of the optimal flow and the density based on the fundamental relation. Implementing such kind of constraint results in a non-linear optimization problem, which may have a significant influence to the computation speed. This problem may be tackled by developing a hierarchical control approach, where the upper level controller determines the optimal flows and the lower level controller determines the VSL rates that satisfy with the safety constraints. Secondly, the proposed controller assumes that every driver follows the suggested speed limit value, which may not be true in reality. Thus, a future step is to explore drivers compliance behavior and take it into account in the optimization. Last but not least, the proposed controller is tested in a homogeneous freeway stretch. A future step is to test the control approach for integrated control measures (such as ramp metering, route guidance, etc.) in large-scale networks.

This thesis proposes a hierarchical ramp metering control approach which is developed based on a freeway MFD model. The influence of capacity drop on the modeling of freeway MFD is investigated. While it has been shown that incorporating the capacity drop improves the accuracy of the MFD, the influence of the capacity drop to the control performance needs further investigation. This requires traffic scenarios with extreme extent and number of capacity drop and moving jams. This proposed control approach assumes infinite storage space of on-ramps, which is not realistic. A future step is to impose constraint to the controller, such that ramp queues satisfy with available on-ramp storage spaces. Another future research direction is that the proposed freeway MFD model and control approach can be developed to tackle congestion control in a freeway network that consists of multiple freeway stretches. Control measures will not only be limited to ramp metering, but also integrate the variable speed limits.

The LQMPC approach proposed in this thesis is tested by a route-specific macroscopic model. Although the model can correctly handle OD relations and it is able to reproduce many important traffic phenomena such as spillback and the propagation of shock wave, it has not been validated with real traffic data and it has a high similarity to the prediction model in the LQMPC. Thus, a future step is to test the proposed LQMPC approach using validated models or microscopic simulations that have enough difference with the prediction model of the LQMPC.

In general, this thesis focuses on developing real time traffic control strategies. To implement a control strategy into practice, other works such as traffic demand prediction and traffic state estimation also need to be done. These research directions will also be explored in the future.

Bibliography

- Aboudolas, K., M. Papageorgiou, E. Kosmatopoulos (2009) Store-and-forward based methods for the signal control problem in large-scale congested urban road networks, *Transportation Research Part C: Emerging Technologies*, 17(2), pp. 163–174. 118
- Aboudolas, K., M. Papageorgiou, A. Kouvelas, E. Kosmatopoulos (2010) A rolling-horizon quadratic-programming approach to the signal control problem in large-scale congested urban road networks, *Transportation Research Part C: Emerging Technologies*, 18(5), pp. 680–694. 118
- Ampountolas, K., N. Zheng, N. Geroliminis (2017) Macroscopic modelling and robust control of bi-modal multi-region urban road networks, *Transportation Research Part B: Methodological*. 90
- Aw, A., M. Rasle (2000) Resurrection of “second order” models of traffic flow, *SIAM journal on applied mathematics*, 60(3), pp. 916–938. 16
- Bellemans, T., B. De Schutter, B. De Moor (2006) Model predictive control for ramp metering of motorway traffic: A case study, *Control Engineering Practice*, 14(7), pp. 757–767. 2
- Burger, M., M. Van Den Berg, A. Hegyi, B. De Schutter, J. Hellendoorn (2013) Considerations for model-based traffic control, *Transportation Research Part C: Emerging Technologies*, 35, pp. 1–19. 9
- Camacho, E. F., C. Bordons (2012) *Model predictive control in the process industry*, Springer Science & Business Media. ISBN: 978-1-4471-3010-9. 44, 103
- Camponogara, E., L. B. De Oliveira (2009) Distributed optimization for model predictive control of linear-dynamic networks, *IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans*, 39(6), pp. 1331–1338. 2, 118
- Carlson, R. C., I. Papamichail, M. Papageorgiou (2011) Local feedback-based mainstream traffic flow control on motorways using variable speed limits, *IEEE Transactions on Intelligent Transportation Systems*, 12(4), pp. 1261–1276. 34

- Carlson, R. C., I. Papamichail, M. Papageorgiou, A. Messmer (2010a) Optimal motorway traffic flow control involving variable speed limits and ramp metering, *Transportation Science*, 44(2), pp. 238–253. vii, ix, 5, 6, 7, 8, 35, 42, 60, 66, 68, 70, 73
- Carlson, R. C., I. Papamichail, M. Papageorgiou, A. Messmer (2010b) Optimal motorway traffic flow control involving variable speed limits and ramp metering, *Transportation Science*, 44(2), pp. 238–253. 9, 11, 146
- Cassidy, M., S. Ahn (2005) Driver turn-taking behavior in congested freeway merges, *Transportation Research Record: Journal of the Transportation Research Board*, (1934), pp. 140–147. 62
- Cassidy, M., K. Jang, C. Daganzo (2011) Macroscopic fundamental diagrams for freeway networks: theory and observation, *Transportation Research Record: Journal of the Transportation Research Board*, (2260), pp. 8–15. 90
- Cassidy, M. J., R. L. Bertini (1999) Some traffic features at freeway bottlenecks, *Transportation Research Part B: Methodological*, 33(1), pp. 25–42. 34
- Cassidy, M. J., J. Rudjanakanoknad (2005) Increasing the capacity of an isolated merge by metering its on-ramp, *Transportation Research Part B: Methodological*, 39(10), pp. 896–913. 62
- Ceylan, H., M. G. Bell (2004) Traffic signal timing optimisation based on genetic algorithm approach, including drivers routing, *Transportation Research Part B: Methodological*, 38(4), pp. 329–342. 8, 119
- Chen, D., S. Ahn, A. Hegyi (2014a) Variable speed limit control for steady and oscillatory queues at fixed freeway bottlenecks, *Transportation Research Part B: Methodological*, 70, pp. 340–358. 6, 60
- Chen, D., S. Ahn, J. Laval, Z. Zheng (2014b) On the periodicity of traffic oscillations and capacity drop: the role of driver characteristics, *Transportation research part B: methodological*, 59, pp. 117–136. 62
- Chung, K., J. Rudjanakanoknad, M. J. Cassidy (2007) Relation between traffic density and capacity drop at three freeway bottlenecks, *Transportation Research Part B: Methodological*, 41(1), pp. 82–95. 34
- Cremer, M. (1979) *Der Verkehrsfluss auf Schnellstrassen*, Springer Verlag. 5, 60
- 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. 10, 16, 36, 62, 73, 91, 120, 146
- Daganzo, C. F. (1995) Requiem for second-order fluid approximations of traffic flow, *Transportation Research Part B: Methodological*, 29(4), pp. 277–286. 16

- Daganzo, C. F. (2007) Urban gridlock: macroscopic modeling and mitigation approaches, *Transportation Research Part B: Methodological*, 41(1), pp. 49–62. 11, 90
- Daganzo, C. F. (2011) On the macroscopic stability of freeway traffic, *Transportation Research Part B: Methodological*, 45(5), pp. 782–788. 90
- De Oliveira, L. B., E. Camponogara (2010) Multi-agent model predictive control of signaling split in urban traffic networks, *Transportation Research Part C: Emerging Technologies*, 18(1), pp. 120–139. 118
- Diakaki, C., M. Papageorgiou, K. Aboudolas (2002) A multivariable regulator approach to traffic-responsive network-wide signal control, *Control Engineering Practice*, 10(2), pp. 183–195. 7
- Duret, A., S. Ahn, C. Buisson (2012) Lane flow distribution on a three-lane freeway: General features and the effects of traffic controls, *Transportation research part C: emerging technologies*, 24, pp. 157–167. 5, 60
- Frejo, J. R. D. (2015) *Model predictive control for freeway traffic networks*, Ph.D. thesis, Universidad de Sevilla. 3
- Frejo, J. R. D., E. F. Camacho (2012) Global versus local mpc algorithms in freeway traffic control with ramp metering and variable speed limits, *IEEE Transactions on intelligent transportation systems*, 13(4), pp. 1556–1565. 118
- Gayah, V. V., C. F. Daganzo (2011) Clockwise hysteresis loops in the macroscopic fundamental diagram: an effect of network instability, *Transportation Research Part B: Methodological*, 45(4), pp. 643–655. 90
- Gazis, D., R. Potts (1963) The oversaturated intersection, *Proceedings of the 2nd International Symposium of Traffic Theory*. 36
- Geroliminis, N., C. F. Daganzo (2008) Existence of urban-scale macroscopic fundamental diagrams: Some experimental findings, *Transportation Research Part B: Methodological*, 42(9), pp. 759–770. 11, 90, 94
- Geroliminis, N., J. Haddad, M. Ramezani (2013) Optimal perimeter control for two urban regions with macroscopic fundamental diagrams: A model predictive approach, *IEEE Transactions on Intelligent Transportation Systems*, 14(1), pp. 348–359. 11, 90, 118
- Geroliminis, N., J. Sun (2011a) Hysteresis phenomena of a macroscopic fundamental diagram in freeway networks, *Transportation Research Part A: Policy and Practice*, 45(9), pp. 966–979. 90, 91, 96, 98
- Geroliminis, N., J. Sun (2011b) Properties of a well-defined macroscopic fundamental diagram for urban traffic, *Transportation Research Part B: Methodological*, 45(3), pp. 605–617. 90

- Godunov, S. K. (1959) A difference method for numerical calculation of discontinuous solutions of the equations of hydrodynamics, *Matematicheskii Sbornik*, 89(3), pp. 271–306. 62
- Gomes, G., R. Horowitz (2006) Optimal freeway ramp metering using the asymmetric cell transmission model, *Transportation Research Part C: Emerging Technologies*, 14(4), pp. 244–262. 4, 10, 16, 19, 44, 91
- Haddad, J. (2015) Robust constrained control of uncertain macroscopic fundamental diagram networks, *Transportation Research Procedia*, 7, pp. 669–688. 90
- Haddad, J., M. Ramezani, N. Geroliminis (2013) Cooperative traffic control of a mixed network with two urban regions and a freeway, *Transportation Research Part B: Methodological*, 54, pp. 17–36. 11, 90, 103, 118
- Haddad, J., A. Shraiber (2014) Robust perimeter control design for an urban region, *Transportation Research Part B: Methodological*, 68, pp. 315–332. 90
- Hadiuzzaman, M., T. Z. Qiu (2013) Cell transmission model based variable speed limit control for freeways, *Canadian Journal of Civil Engineering*, 40(1), pp. 46–56. 44, 61
- Han, Y., D. Chen, S. Ahn, A. Hegyi (2015a) Analysis of driver response and traffic evolution under variable speed limit control, *Transportation Research Record: Journal of the Transportation Research Board*, 2490, pp. 1–10. 60
- Han, Y., A. Hegyi, Y. Yuan, S. Hoogendoorn (2017a) Validation of an extended discrete first-order model with variable speed limits, *Transportation research part C: emerging technologies*, 83, pp. 1–17. 11
- Han, Y., A. Hegyi, Y. Yuan, S. Hoogendoorn, M. Papageorgiou, C. Roncoli (2017b) Resolving freeway jam waves by discrete first-order model-based predictive control of variable speed limits, *Transportation Research Part C: Emerging Technologies*, 77, pp. 405–420. 61, 64, 65, 67, 126
- Han, Y., Y. Yuan, A. Hegyi, S. Hoogendoorn (2016a) A new extension of discrete first-order model to reproduce the propagation of jam wave, in: *Transportation Research Board 95th Annual Meeting*, 16-3482. 39, 41
- Han, Y., Y. Yuan, A. Hegyi, S. P. Hoogendoorn (2015b) Linear quadratic MPC for integrated route guidance and ramp metering, in: *Intelligent Transportation Systems (ITSC), 2015 IEEE 18th International Conference on*, IEEE, pp. 1150–1155. 4, 16, 18, 19, 22, 26, 38, 40, 91, 126
- Han, Y., Y. Yuan, A. Hegyi, S. P. Hoogendoorn (2016b) New extended discrete first-order model to reproduce propagation of jam waves, *Transportation Research Record: Journal of the Transportation Research Board*, (2560), pp. 108–118. 63, 98, 106

- Hegyi, A. (2004) *Model predictive control for integrating traffic control measures*, PhD dissertation, TRAIL Thesis series. Delft University of Technology. ISBN: 90-5584-053-X. vii, 2, 9
- Hegyi, A., B. De Schutter, H. Hellendoorn (2005a) Model predictive control for optimal coordination of ramp metering and variable speed limits, *Transportation Research Part C: Emerging Technologies*, 13(3), pp. 185–209. 4, 8, 9, 11, 13, 35, 42, 49, 62, 70, 73, 74, 91, 103
- Hegyi, A., B. De Schutter, J. Hellendoorn (2005b) Optimal coordination of variable speed limits to suppress shock waves, *Intelligent Transportation Systems, IEEE Transactions on*, 6(1), pp. 102–112. 6, 35, 50, 60
- Hegyi, A., S. Hoogendoorn (2010) Dynamic speed limit control to resolve shock waves on freeways-field test results of the specialist algorithm, in: *Intelligent Transportation Systems (ITSC), 2010 13th International IEEE Conference on*, IEEE, pp. 519–524. 2, 6, 35, 47, 60, 67, 70
- Hegyi, A., S. Hoogendoorn, M. Schreuder, H. Stoelhorst, F. Viti (2008) Specialist: A dynamic speed limit control algorithm based on shock wave theory, in: *Intelligent Transportation Systems, 2008. ITSC 2008. 11th International IEEE Conference on*, IEEE, pp. 827–832. 5, 6, 17, 34, 35, 47, 60, 66, 67, 146
- Hoogendoorn, S., R. Landman, J. van Kooten, M. Schreuder (2013) Integrated network management amsterdam: Control approach and test results, in: *16th International IEEE Conference on Intelligent Transportation Systems (ITSC 2013)*, IEEE, pp. 474–479. 2
- Ji, Y., W. Daamen, S. Hoogendoorn, S. Hoogendoorn-Lanser, X. Qian (2010) Investigating the shape of the macroscopic fundamental diagram using simulation data, *Transportation Research Record: Journal of the Transportation Research Board*, (2161), pp. 40–48. 90
- Kerner, B. S., H. Rehborn (1996) Experimental features and characteristics of traffic jams, *Physical Review E*, 53(2), p. R1297. 5, 17, 18, 34
- Keyvan-Ekbatani, M., A. Kouvelas, I. Papamichail, M. Papageorgiou (2012) Exploiting the fundamental diagram of urban networks for feedback-based gating, *Transportation Research Part B: Methodological*, 46(10), pp. 1393–1403. 90
- Keyvan-Ekbatani, M., M. Yildirimoglu, N. Geroliminis, M. Papageorgiou (2015) Multiple concentric gating traffic control in large-scale urban networks, *IEEE Transactions on Intelligent Transportation Systems*, 16(4), pp. 2141–2154. 90
- Knoop, V., S. Hoogendoorn (2013) Empirics of a generalized macroscopic fundamental diagram for urban freeways, *Transportation Research Record: Journal of the Transportation Research Board*, (2391), pp. 133–141. 90, 96

- Kotsialos, A., M. Papageorgiou (2004) Nonlinear optimal control applied to coordinated ramp metering, *IEEE Transactions on control systems technology*, 12(6), pp. 920–933. 4, 91
- Kotsialos, A., M. Papageorgiou, C. Diakaki, Y. Pavlis, F. Middelham (2002a) Traffic flow modeling of large-scale motorway networks using the macroscopic modeling tool metanet, *IEEE Transactions on Intelligent Transportation Systems*, 3(4), pp. 282–292. 10, 16, 35, 60
- Kotsialos, A., M. Papageorgiou, M. Mangeas, H. Haj-Salem (2002b) Coordinated and integrated control of motorway networks via non-linear optimal control, *Transportation Research Part C: Emerging Technologies*, 10(1), pp. 65–84. 2
- Kotsialos, A., M. Papageorgiou, A. Messmer (1999) Optimal coordinated and integrated motorway network traffic control, in: *14th International Symposium on Transportation and Traffic Theory*. 2, 35, 49, 111
- Kouvelas, A., M. Saeedmanesh, N. Geroliminis (2017) Enhancing model-based feedback perimeter control with data-driven online adaptive optimization, *Transportation Research Part B: Methodological*, 96, pp. 26–45. 90
- Laval, J. A., F. Castrillón (2015) Stochastic approximations for the macroscopic fundamental diagram of urban networks, *Transportation Research Part B: Methodological*, 81, pp. 904–916. 90
- Laval, J. A., C. F. Daganzo (2006) Lane-changing in traffic streams, *Transportation Research Part B: Methodological*, 40(3), pp. 251–264. 62
- Laval, J. A., L. Leclercq (2010) A mechanism to describe the formation and propagation of stop-and-go waves in congested freeway traffic, *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 368(1928), pp. 4519–4541. 17
- Le, T., H. L. Vu, Y. Nazarathy, Q. B. Vo, S. Hoogendoorn (2013) Linear-quadratic model predictive control for urban traffic networks, *Transportation Research Part C: Emerging Technologies*, 36, pp. 498–512. 8, 11, 46, 119, 120, 126, 128, 147
- Lebacque, J. (2003) Two-phase bounded-acceleration traffic flow model: analytical solutions and applications, *Transportation Research Record: Journal of the Transportation Research Board*, (1852), pp. 220–230. 38
- Lebacque, J.-P. (1996) The godunov scheme and what it means for first order traffic flow models, in: *International symposium on transportation and traffic theory*, pp. 647–677. 10, 16, 37, 61
- Leclercq, L., N. Chiabaut, B. Trinquier (2014) Macroscopic fundamental diagrams: A cross-comparison of estimation methods, *Transportation Research Part B: Methodological*, 62, pp. 1–12. 90

- Leclercq, L., V. L. Knoop, F. Marczak, S. P. Hoogendoorn (2016) Capacity drops at merges: New analytical investigations, *Transportation Research Part C: Emerging Technologies*, 62, pp. 171–181. 17, 62
- Leclercq, L., J. A. Laval, N. Chiabaut (2011) Capacity drops at merges: An endogenous model, *Procedia-Social and Behavioral Sciences*, 17, pp. 12–26. 17
- LeVeque, R. J. (1992) *Numerical methods for conservation laws*, Springer Science & Business Media. 67
- Lighthill, M. J., G. B. Whitham (1955) On kinematic waves. **II. A theory of traffic flow on long crowded roads**, in: *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, vol. 229, The Royal Society, pp. 317–345. 10, 16, 19, 36, 60, 62, 120
- Lin, S., B. De Schutter, Y. Xi, H. Hellendoorn (2011) Fast model predictive control for urban road networks via milp, *IEEE Transactions on Intelligent Transportation Systems*, 12(3), pp. 846–856. 118
- Lo, H. (1999) A dynamic traffic assignment formulation that encapsulates the cell-transmission model, in: *14th International Symposium on Transportation and Traffic Theory*. 36
- Lo, H. K., E. Chang, Y. C. Chan (2001) Dynamic network traffic control, *Transportation Research Part A: Policy and Practice*, 35(8), pp. 721–744. 2
- Mahajan, N., A. Hegyi, V. De Weg, G. Sterk, S. P. Hoogendoorn (2015) Integrated variable speed limit and ramp metering control against jam waves—**A COSCAL v2 Based Approach**, in: *Intelligent Transportation Systems (ITSC), 2015 IEEE 18th International Conference on*, IEEE, pp. 1156–1162. 6, 35, 60
- Masher, D. P., D. Ross, P. Wong, P. Tuan, H. Zeidler, S. Petracek (1975) Guidelines for design and operation of ramp control systems. 4
- Mazlounian, A., N. Geroliminis, D. Helbing (2010) The spatial variability of vehicle densities as determinant of urban network capacity, *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 368(1928), pp. 4627–4647. 90, 96
- Muralidharan, A. (2012) *Tools for modeling and control of freeway networks*, Ph.D. thesis. 38, 44, 146
- Muralidharan, A., G. Dervisoglu, R. Horowitz (2009) Freeway traffic flow simulation using the link node cell transmission model, in: *American Control Conference, 2009. ACC'09.*, IEEE, pp. 2916–2921. 16, 17, 19
- Muralidharan, A., R. Horowitz (2012) Optimal control of freeway networks based on the link node cell transmission model, in: *American Control Conference (ACC), 2012*, IEEE, pp. 5769–5774. 27

- Muralidharan, A., R. Horowitz (2015) Computationally efficient model predictive control of freeway networks, *Transportation Research Part C: Emerging Technologies*, 2, 7, 9, 27, 36, 44, 61
- Papageorgiou, M. (1990) Dynamic modeling, assignment, and route guidance in traffic networks, *Transportation Research Part B: Methodological*, 24(6), pp. 471–495. 2
- Papageorgiou, M. (1995) An integrated control approach for traffic corridors, *Transportation Research Part C: Emerging Technologies*, 3(1), pp. 19–30. 36, 46, 127
- Papageorgiou, M., C. Diakaki, V. Dinopoulou, A. Kotsialos, Y. Wang (2003) Review of road traffic control strategies, *Proceedings of the IEEE*, 91(12), pp. 2043–2067. 118
- Papageorgiou, M., H. Hadj-Salem, J.-M. Blosseville (1991) ALINEA: A local feedback control law for on-ramp metering, *Transportation Research Record*, 1320(1), pp. 58–67. 91
- Papageorgiou, M., H. Hadj-Salem, F. Middelham (1997) Alinea local ramp metering: Summary of field results, *Transportation Research Record: Journal of the Transportation Research Board*, (1603), pp. 90–98. 2
- Papageorgiou, M., E. Kosmatopoulos, I. Papamichail (2008) Effects of variable speed limits on motorway traffic flow, *Transportation Research Record: Journal of the Transportation Research Board*, (2047), pp. 37–48. 6, 42, 44, 61, 68
- Papageorgiou, M., A. Kotsialos (2000) Freeway ramp metering: An overview, in: *Intelligent Transportation Systems, 2000. Proceedings. 2000 IEEE*, IEEE, pp. 228–239. 3, 4, 91
- Papamichail, I., A. Kotsialos, I. Margonis, M. Papageorgiou (2010) Coordinated ramp metering for freeway networks—a model-predictive hierarchical control approach, *Transportation Research Part C: Emerging Technologies*, 18(3), pp. 311–331. 4, 91
- Peeta, S., A. K. Ziliaskopoulos (2001) Foundations of dynamic traffic assignment: The past, the present and the future, *Networks and Spatial Economics*, 1(3), pp. 233–265. 7, 119
- Prato, C. G. (2009) Route choice modeling: past, present and future research directions, *Journal of choice modelling*, 2(1), pp. 65–100. 123
- Ramezani, M., J. Haddad, N. Geroliminis (2015) Dynamics of heterogeneity in urban networks: aggregated traffic modeling and hierarchical control, *Transportation Research Part B: Methodological*, 74, pp. 1–19. 11, 90, 96, 100, 118
- Richards, P. I. (1956) Shock waves on the highway, *Operations research*, 4(1), pp. 42–51. 10, 16, 19, 36, 60, 62, 120

- Robertson, D. I., R. D. Bretherton (1991) Optimizing networks of traffic signals in real time—the scoot method, *IEEE Transactions on vehicular technology*, 40(1), pp. 11–15. 7
- Roncoli, C., M. Papageorgiou, I. Papamichail (2015a) Traffic flow optimisation in presence of vehicle automation and communication systems—part i: A first-order multi-lane model for motorway traffic, *Transportation Research Part C: Emerging Technologies*, 57, pp. 241–259. viii, 16, 17, 19, 27, 38, 39, 50, 65, 146
- Roncoli, C., M. Papageorgiou, I. Papamichail (2015b) Traffic flow optimisation in presence of vehicle automation and communication systems—part ii: Optimal control for multi-lane motorways, *Transportation Research Part C: Emerging Technologies*, 57, pp. 260–275. 2, 6, 7, 9, 13, 27, 35, 36, 44, 61
- Saberi, M., H. Mahmassani (2012) Exploring properties of networkwide flow-density relations in a freeway network, *Transportation Research Record: Journal of the Transportation Research Board*, (2315), pp. 153–163. 90
- Schönhof, M., D. Helbing (2007) Empirical features of congested traffic states and their implications for traffic modeling, *Transportation Science*, 41(2), pp. 135–166. 5, 17, 18, 34
- Schreiter, T., E.-S. Smits, H. Van Lint, S. Hoogendoorn (2010) The cell-transmission model with capacity drop, in: *Proceedings of the 11th International Congress of the Research School Transportation, Infrastructure and Logistics (TRAIL)*. 17, 38
- Sims, A. G., K. W. Dobinson (1980) The sydney coordinated adaptive traffic (scat) system philosophy and benefits, *IEEE Transactions on vehicular technology*, 29(2), pp. 130–137. 7
- Smith, M., M. Ghali (1990) The dynamics of traffic assignment and traffic control: A theoretical study, *Transportation Research Part B: Methodological*, 24(6), pp. 409–422. 2
- Smulders, S. (1990) Control of freeway traffic flow by variable speed signs, *Transportation Research*, 24B, pp. 111–132. 5, 60
- Soriguera, F., I. Martínez, M. Sala, M. Menéndez (2017) Effects of low speed limits on freeway traffic flow, *Transportation Research Part C: Emerging Technologies*, 77, pp. 257–274. 5, 60
- Srivastava, A., N. Geroliminis (2013) Empirical observations of capacity drop in freeway merges with ramp control and integration in a first-order model, *Transportation Research Part C: Emerging Technologies*, 30, pp. 161–177. 17, 34, 38
- Srivastava, A., W.-L. Jin, J. P. Lebacque (2014) A modified cell transmission model for signalized intersections, in: *Transportation Research Board 93rd Annual Meeting*, 14-4713. 38

- Tampère, C. M., R. Corthout, D. Cattrysse, L. H. Immers (2011) A generic class of first order node models for dynamic macroscopic simulation of traffic flows, *Transportation Research Part B: Methodological*, 45(1), pp. 289–309. 124
- Teklu, F., A. Sumalee, D. Watling (2007) A genetic algorithm approach for optimizing traffic control signals considering routing, *Computer-Aided Civil and Infrastructure Engineering*, 22(1), pp. 31–43. 8, 119
- Tettamanti, T., I. Varga (2010) Distributed traffic control system based on model predictive control, *Periodica Polytechnica. Civil Engineering*, 54(1), p. 3. 2
- van de Weg, G. S., A. Hegyi, S. P. Hoogendoorn, B. D. Schutter (2015) Efficient model predictive control for variable speed limits by optimizing parameterized control schemes, in: *Intelligent Transportation Systems (ITSC), 2015 IEEE 18th International Conference on*, IEEE, pp. 1137–1142. 2, 35
- van den Berg, M., A. Hegyi, B. De Schutter, J. Hellendoorn (2007) Integrated traffic control for mixed urban and freeway networks: A model predictive control approach, *European journal of transport and infrastructure research*, 7(3), pp. 223–250. 118
- Van Katwijk, R. T. (2008) Multi-agent look-ahead traffic-adaptive control, *PhD thesis, Delft University of Technology*. 7
- van Wageningen-Kessels, F., H. Van Lint, K. Vuik, S. Hoogendoorn (2015) Genealogy of traffic flow models, *EURO Journal on Transportation and Logistics*, 4(4), pp. 445–473. 16
- Varaiya, P. (2013) Max pressure control of a network of signalized intersections, *Transportation Research Part C: Emerging Technologies*, 36, pp. 177–195. 7, 119, 130
- Wang, J. J., A. D. May (1973) Computer model for optimal freeway on-ramp control, *Highway Research Record*, (469). 4
- Webster, F., B. Cobbe (1963) Traffic signals, *Traffic Engineering Practice*, pp. 117–121. 7
- Wilson, R. E. (2008) From inductance loops to vehicle trajectories, in: *Proc. Symp. Fundam. Diagram*, pp. 134–143. 20
- Yildirimoglu, M., M. Ramezani, N. Geroliminis (2015) Equilibrium analysis and route guidance in large-scale networks with mfd dynamics, *Transportation Research Procedia*, 9, pp. 185–204. 90
- Yuan, K., V. L. Knoop, S. P. Hoogendoorn (2015) Capacity drop: A relation between the speed in congestion and the queue discharge rate, in: *Transportation Research Board 94th Annual Meeting*. 27, 38, 63, 100

- Zackor, H. (1972) Beurteilung verkehrsabhängiger geschwindigkeitsbeschränkungen auf autobahnen, *Strassenbau und Strassenverkehrstechnik*, 128, pp. 1–61. 5, 60
- Zegeye, S. K., B. De Schutter, J. Hellendoorn, E. A. Breunese, A. Hegyi (2012) A predictive traffic controller for sustainable mobility using parameterized control policies, *Intelligent Transportation Systems, IEEE Transactions on*, 13(3), pp. 1420–1429. 2, 35
- Zhu, F., S. V. Ukkusuri (2015) A linear programming formulation for autonomous intersection control within a dynamic traffic assignment and connected vehicle environment, *Transportation Research Part C: Emerging Technologies*, 55, pp. 363–378. xii, 8, 122
- Ziliaskopoulos, A. K. (2000) A linear programming model for the single destination system optimum dynamic traffic assignment problem, *Transportation science*, 34(1), pp. 37–49. 10, 16, 36

Summary

Traffic congestion has become a global issue that has a significant impact on our society's productivity. Its negative effects not only lie in the travel delays and unsafe conditions that it brings to road users, but also many aspects of our lives such as the air we all breathe. Construction and traffic management are typical alternatives for traffic researchers and practitioners to reduce congestion. Traffic management, which intends to make a better use of existing infrastructure, is more economical and environmentally friendly and becoming an increasingly preferred option. Dynamic traffic control proves to be efficient in the management of network traffic flows. This thesis focuses on the development of dynamic traffic control strategies to reduce congestion.

Advanced dynamic traffic control strategies using model predictive control (MPC) approaches can considerably reduce traffic congestion. MPC for traffic systems utilizes a traffic model to predict traffic states evolutions based on the current states of the system, and determines the optimal control actions that result in the optimum value of an objective function. This feature enables the controller to take advantage of potentially larger future gains at a current (smaller) cost, so as to avoid myopic control actions.

Although model predictive control approaches for road traffic have been extensively investigated in the literature, none of them has been implemented into the field. Generally, traffic model predictive control approaches face with two main problems towards field application: (i) high computation complexity, (ii) low prediction accuracy, i.e., a high mismatch between the predicted traffic dynamics and real traffic dynamics. This thesis aims to develop model predictive control approaches that are accurate in predicting traffic dynamics and fast in determining optimal control signals. To achieve this goal, this thesis focuses on developing more accurate models, and developing faster MPC-based control approaches.

This thesis develops different traffic flow models to describe different traffic phenomena. An extended discrete first-order model is proposed to reproduce the propagation of jam waves on freeways. The proposed model keeps the linear formulation and the simple solution method of the original discrete first-order model, whereas significantly improves the model accuracy in reproducing the propagation of freeway jam waves, as is evidenced in model validation. More importantly, the proposed model can be embedded in model predictive controllers which have a linear or linear quadratic formulation that can be efficiently solved.

Furthermore, a model that is based on the freeway macroscopic fundamental diagram (MFD) is proposed to predict the aggregated traffic state (weighted average density and weighted average flow) of freeway networks. A main feature of the model is that the evolution of the density heterogeneity is considered, such that the MFD can predict the aggregated traffic state in the freeway network accurately, which benefits the MFD-based controller. We also incorporate the capacity drop into the freeway MFD. By proposing an exponential functional form, the accuracy of the proposed MFD is validated with field traffic data.

This thesis proposes several model predictive control strategies that are fast enough to be implemented in the field. Some MPC control strategies are developed based on the proposed traffic models. For example, this thesis develops an MPC controller of VSLs that is based on the extended discrete first-order model to suppress freeway jam waves. The proposed controller is tested successfully in a homogeneous freeway stretch in simulated cases and demonstrates a better performance in resolving freeway jam waves compared to a linear MPC and a faster computation speed compared to a classical non-linear MPC.

The proposed MFD model is incorporated in a hierarchical control structure for coordinated ramp metering on a freeway stretch. At the upper level, a model predictive control approach is developed to optimize the total inflow from on-ramps to the freeway stretch. The lower level controller distributes the optimal total inflow to each on-ramp of the freeway based on local traffic state feedback. Simulation outcomes of several numerical experiments highlight that the MFD-based hierarchical controller (i) is better able to overcome the modeling mismatch between the prediction model and the plant (process model) in the MPC framework and (ii) requires less computation effort than other nonlinear controllers.

This thesis proposes an extended linear quadratic MPC approach of integrated routing and urban intersection control for multi-destination networks. The extended linear quadratic MPC uses a heuristic approach to narrow the solution space and preserve traffic OD relations, which overcomes the drawback of the original linear quadratic MPC approach that desired OD relations may end up at wrong destinations. In simulated cases, the proposed control approach demonstrates a good performance in reducing the TTS, and keeps the computation time tractable.

To sum up, this thesis enhances MPC approaches for dynamic traffic control in terms of prediction accuracy and computation speed. The proposed MPC approaches have good potential to be applied to the field. In the future, the proposed control approach will be extended such that different type of vehicles (both self-driving cars and autonomous cars) are considered, since autonomous vehicles will appear on the road in the near future.

Samenvatting

Verkeersopstoppingen zijn een wereldwijd probleem geworden dat een significante impact heeft op de productiviteit van onze samenleving. De negatieve effecten zitten niet alleen in de reisvertragingen en onveilige situaties die het voor weggebruikers oplevert, maar ook in de lucht die we allen inademen en de algemene kwaliteit van ons leven. Constructie en verkeersmanagement zijn typische mogelijkheden voor verkeersonderzoekers en practici om congestie te verminderen. Verkeersmanagement, dat beoogt om bestaande infrastructuur beter te benutten, is goedkoper en milieuvriendelijker en verdient steeds meer de voorkeur. Dynamische verkeersregeling bewijst efficiënt te zijn in het beheersen van verkeersstromen in het netwerk. Dit proefschrift focust op de ontwikkeling van strategieën voor dynamische verkeersregeling om congestie te verminderen.

Geavanceerde strategieën voor dynamisch verkeersregeling die gebruik maken van modelgebaseerde voorspellende regeling (MVR) kunnen verkeersopstoppingen behoorlijk verminderen. MVR voor verkeerssystemen gebruikt een verkeersmodel om de evolutie van verkeers toestanden te voorspellen op basis van de huidige toestanden van het systeem, en bepaalt de optimale regelacties die leiden tot de optimale waarde van een doelfunctie. Deze eigenschap laat de regelmatig gebruikmaken van mogelijk grote winsten in de toekomst tegen (lage) kosten in het heden, waardoor kortzichtige regelacties worden vermeden.

Hoewel modelgebaseerde voorspellende regelingen voor wegverkeer uitgebreid zijn onderzocht in literatuur, is geen daarvan in de praktijk toegepast. In het algemeen hebben verkeersmodelgebaseerde voorspellende regelingen twee hoofdproblemen voor praktijktoepassingen: (i) hoge complexiteit van de berekening, (ii) lage nauwkeurigheid van de voorspelling, d.w.z., een groot verschil tussen de voorspelde verkeersdynamiek en de werkelijke verkeersdynamiek. Dit proefschrift beoogt modelgebaseerde voorspellende regelingen te ontwikkelen die nauwkeurig zijn in het voorspellen van de verkeersdynamiek en snel in het bepalen van de optimale regelsignalen. Om dit te bereiken, concentreert dit proefschrift zich op het ontwikkelen van nauwkeurigere modellen en snellere MVR-regelmethoden.

Dit proefschrift ontwikkelt verschillende verkeersstroommodellen om verschillende verkeersfenomenen te beschrijven. Een uitgebreid discreet eersteordemodel wordt voorgesteld om de propagatie van filegolven op snelwegen te reproduceren. Het voorgestelde model behoudt de lineaire formulering en eenvoudige oplossingsmethode van het originele eersteordemodel, terwijl het de modelnauwkeurigheid in het reproduceren van filegolven op snelwegen significant verbetert, zoals blijkt uit de modelvalidatie. Nog belangrijker is dat het voorgestelde model kan worden opgenomen in modelgebaseerde voorspellende regelaars met een eenvoudig oplosbare lineaire of lineair-kwadratische formulering.

Daarnaast wordt een model voorgesteld dat is gebaseerd op het macroscopisch fundamenteel diagram (MFD) voor snelwegen om de geaggregeerde verkeerstoestand (gewogen gemiddelde dichtheid en gewogen gemiddelde intensiteit) van snelwegnetwerken te voorspellen. Een hoofdeigenschap van dit model is dat de evolutie van de heterogeniteit van de dichtheid wordt beschouwd, zodat het MFD de geaggregeerde verkeerstoestand in het snelwegnetwerk nauwkeurig kan voorspellen, waar de MFD-gebaseerde regelaar van profiteert. We nemen ook een capaciteitsval op in het MFD voor snelwegen. Door een exponentiele functieform te nemen is de nauwkeurigheid van het voorgestelde MFD gevalideerd met gemeten verkeersgegevens.

Dit proefschrift stelt verscheidene modelgebaseerde voorspellende regelstrategieën voor die snel genoeg zijn om in de praktijk te worden toegepast. Een aantal MVR-regelstrategieën is ontwikkeld op basis van de voorgestelde verkeersmodellen. Dit proefschrift ontwikkelt bijvoorbeeld een MVR-regelaar van variabele snelheidslimieten gebaseerd op het uitgebreide eerste-orde model om filegolven op snelwegen te bestrijden. De voorgestelde regelaar is succesvol getest op een homogeen stuk snelweg in simulatiescenario's. Hij laat betere prestaties zien m.b.t. het oplossen van filegolven dan een lineaire MVR en heeft een hogere rekensnelheid dan een klassieke niet-lineaire MVR.

Het voorgestelde MVR-model is opgenomen in een hiërarchische regelstructuur voor gecoördineerde toeritdosering op een stuk snelweg. Voor het bovenste niveau is een modelgebaseerde voorspellende regelaar ontwikkeld die de totale instroom van de snelwegtoeritten optimaliseert. De regelaar op het onderste niveau verdeelt de optimale totaalinstroom over de snelwegtoeritten op basis van feedback over de lokale verkeerstoestand. Simulatieresultaten van diverse numerieke experimenten tonen dat de MFD-gebaseerde hiërarchische regelaar (i) beter omgaat met het verschil tussen het voorspellingsmodel en het procesmodel in het MVR-raamwerk en (ii) minder rekenkracht vereist dan andere niet-lineaire regelaars.

Dit proefschrift stelt voor een uitgebreide lineair-kwadratische MVR-aanpak van geventueerde routing en kruispuntregeling voor netwerken met meerdere bestemmingen. De uitgebreide lineair-kwadratische MVR gebruikt een heuristische benadering om de oplossingsruimte te verkleinen en behoudt HB-verkeersrelaties, waardoor wordt vermeden het probleem van de oorspronkelijke lineair-kwadratische MVR-aanpak dat gewenste HB-relaties op verkeerde bestemmingen terechtkomen. In gesimuleerde casussen laat de voorgestelde regelaanpak goede prestaties zien m.b.t. het reduceren van de totale reistijd, terwijl de rekestijd werkbaar blijft.

Samenvattend verbetert dit proefschrift MVR-aanpakken voor dynamische verkeersregeling in termen van nauwkeurigheid en rekestijd. De voorgestelde MVR-aanpakken hebben goede potentie om in de praktijk te worden toegepast. In de toekomst zal de voorgestelde regelaanpak worden uitgebreid zodat verschillende typen voertuigen (zowel zelfrijdend als autonome autos) worden meegenomen, aangezien autonome voertuigen in de nabije toekomst op de weg zullen verschijnen.

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.

Han, Y., *Fast Model Predictive Control Approaches for Road Traffic Control*, T2017/13, December 2017, TRAIL Thesis Series, the Netherlands

Wang, P., *Train Trajectory Optimization Methods for Energy-Efficient Railway Operations*, T2017/12, December 2017, TRAIL Thesis Series, the Netherlands

Weg, G.S. van de, *Efficient Algorithms for Network-wide Road Traffic Control*, T2017/11, October 2017, TRAIL Thesis Series, the Netherlands

He, D., *Energy Saving for Belt Conveyors by Speed Control*, T2017/10, July 2017, TRAIL Thesis Series, the Netherlands

Beinovi, N., *Integrated Capacity Assessment and Timetabling Models for Dense Railway Networks*, T2017/9, July 2017, TRAIL Thesis Series, the Netherlands

Chen, G., *Surface Wear Reduction of Bulk Solids Handling Equipment Using Bionic Design*, T2017/8, June 2017, TRAIL Thesis Series, the Netherlands

Kurapati, S., *Situation Awareness for Socio Technical Systems: A simulation gaming study in intermodal transport operations*, T2017/7, June 2017, TRAIL Thesis Series, the Netherlands

Jamshidnejad, A., *Efficient Predictive Model-Based and Fuzzy Control for Green Urban Mobility*, T2017/6, June 2017, TRAIL Thesis Series, the Netherlands

Araghi, Y., *Consumer Heterogeneity, Transport and the Environment*, T2017/5, May 2017, TRAIL Thesis Series, the Netherlands

Kasraian Moghaddam, D., *Transport Networks, Land Use and Travel Behaviour: A long term investigation*, T2017/4, May 2017, TRAIL Thesis Series, the Netherlands

Smits, E.-S., *Strategic Network Modelling for Passenger Transport Pricing*, T2017/3, May 2017, TRAIL Thesis Series, the Netherlands

Tasseron, G., *Bottom-Up Information Provision in Urban Parking: An in-depth analysis of impacts on parking dynamics*, T2017/2, March 2017, TRAIL Thesis Series, the Netherlands

Halim, R.A., *Strategic Modeling of Global Container Transport Networks: Exploring the future of port-hinterland and maritime container transport networks*, T2017/1, March 2017, TRAIL Thesis Series, the Netherlands

Olde Keizer, M.C.A., *Condition-Based Maintenance for Complex Systems: Coordinating maintenance and logistics planning for the process industries*, T2016/26, December 2016, TRAIL Thesis Series, the Netherlands

Zheng, H., *Coordination of Waterborn AGVs*, T2016/25, December 2016, TRAIL Thesis Series, the Netherlands

Yuan, K., *Capacity Drop on Freeways: Traffic dynamics, theory and Modeling*, T2016/24, December 2016, TRAIL Thesis Series, the Netherlands

Li, S., *Coordinated Planning of Inland Vessels for Large Seaports*, T2016/23, December 2016, TRAIL Thesis Series, the Netherlands

Berg, M. van den, *The Influence of Herding on Departure Choice in Case of Evacuation: Design and analysis of a serious gaming experimental set-up*, T2016/22, December 2016, TRAIL Thesis Series, the Netherlands

Luo, R., *Multi-Agent Control of urban Transportation Networks and of Hybrid Systems with Limited Information Sharing*, T2016/21, November 2016, TRAIL Thesis Series, the Netherlands

Campanella, M., *Microscopic Modelling of Walking Behavior*, T2016/20, November 2016, TRAIL Thesis Series, the Netherlands

Horst, M. van der, *Coordination in Hinterland Chains: An institutional analysis of port-related transport*, T2016/19, November 2016, TRAIL Thesis Series, the Netherlands

Beukenkamp, W., *Securing Safety: Resilience time as a hidden critical factor*, T2016/18, October 2016, TRAIL Thesis Series, the Netherlands

Mingardo, G., *Articles on Parking Policy*, T2016/17, October 2016, TRAIL Thesis Series, the Netherlands

Duives, D.C., *Analysis and Modelling of Pedestrian Movement Dynamics at Large-scale Events*, T2016/16, October 2016, TRAIL Thesis Series, the Netherlands

Wan Ahmad, W.N.K., *Contextual Factors of Sustainable Supply Chain Management Practices in the Oil and Gas Industry*, T2016/15, September 2016, TRAIL Thesis Series, the Netherlands

Liu, X., *Prediction of Belt Conveyor Idler Performance*, T2016/14, September 2016, TRAIL Thesis Series, the Netherlands

Gaast, J.P. van der, *Stochastic Models for Order Picking Systems*, T2016/13, September 2016, TRAIL Thesis Series, the Netherlands

Wagenaar, J.C., *Practice Oriented Algorithmic Disruption Management in Passenger Railways*, T2016/12, September 2016, TRAIL Thesis Series, the Netherlands

About the author

Curriculum Vitae

Yu Han was born in Shandong, China, on June 12, 1989. He obtained his bachelor degree in traffic engineering from Harbin Institute of Technology in 2011. Afterwards he continued master study in Beijing Jiaotong University. In 2013, he obtained his master degree in system engineering.

In September 2013, he started his PhD research with Professor Serge Hoogendoorn. His research interests include road traffic control and traffic flow modelling. Now he is working as a research engineer in Didichuxing company, China. He is involved in the projects of implementing traffic signal control and ramp metering control strategies to multiple cities in China.