

**Road Incidents and Network Dynamics  
Effects on driving behaviour and traffic congestion**

Victor L. Knoop

Cover illustration: [petervandorst.nl](http://petervandorst.nl)

# **Road Incidents and Network Dynamics Effects on driving behaviour and traffic congestion**

## **Proefschrift**

ter verkrijging van de graad van doctor  
aan de Technische Universiteit Delft,  
op gezag van de Rector Magnificus prof.dr.ir. J.T. Fokkema,  
voorzitter van het College van Promoties,  
in het openbaar te verdedigen op woensdag 2 december 2009 om 10:00 uur  
door Victor Lambert KNOOP  
doctorandus in de natuurkunde  
geboren te Arnhem.

Dit proefschrift is goedgekeurd door de promotoren:

Prof. dr. H.J. van Zuylen

Prof. dr. ir. S.P. Hoogendoorn

Samenstelling promotiecommissie:

Rector Magnificus

Prof. dr. H.J. van Zuylen

Prof. dr. ir. S.P. Hoogendoorn

Prof. dr. M.G.H. Bell

Prof. dr. ir. B. De Schutter

Prof. ir. L.H. Immers

Prof. dr. ir. W.H.K. Lam

Prof. dr. ir. J.A.A.M. Stoop

voorzitter

Technische Universiteit Delft, promotor

Technische Universiteit Delft, promotor

Imperial College London

Technische Universiteit Delft

Katholieke Universiteit Leuven

Hong Kong Polytechnical University

Lund University

### **TRAIL Thesis Series T2009/13, TRAIL Research School**

P.O. Box 5017

2600 GA Delft

The Netherlands

T: +31 (0) 15 278 6046

F: +31 (0) 15 278 4333

E: info@rsTRAIL.nl

ISBN 978-90-5584-124-0

Keywords: incidents, traffic flow theory

Copyright © 2009 by Victor L. Knoop

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

Printed in the Netherlands

Πάντα ῥεῖ (Everything flows)  
– Ἡράκλιτος ὁ Ἐφέσιος (Heraclitus of Ephesus)



# Preface

It was in 2004 that I had the first contact with Henk van Zuylen. He first answered an e-mail by me, encouraging me to come and talk about a position as PhD student. He wrote me that it could be very fruitful to start in traffic theory with a background in physics.

Henk, in the talk we then had you made me enthusiastic about working in the Transport and Planning group in Delft. The only (minor) problem was that there was no project which could finance my PhD research. There was funding for one year and in that one year I had to apply for funding for the rest of my PhD. The confidence you showed made me decide to take that offer.

Financing was an issue afterwards. However, you always were confident that there would be funding. This is also the place to thank the organisations and grants without which this research would not have been possible. That is the Transport Research Centre Delft, the Dutch research program Next Generation Infrastructures and the research program Transumo-ATMA. The very attractive helicopter flights were financed by the research program Tracing Congestion Dynamics – with Innovative Data to a Better Theory (sponsored by the Dutch Foundation of Scientific Research MaGW-NWO).

This thesis was written to finish my PhD research. The chapters are often written as individual papers or articles and send to conferences or journals. The comments by anonymous reviewers were very helpful in improving the paper or article, and therefore also in improving this thesis.

Henk, during my PhD research you often made remarks about my work. They often addressed fundamental parts of the research. Your comments often made the research I did much better, for which I would like to thank you here.

I also would like to thank my daily supervisor Serge Hoogendoorn here. Serge, I really liked working with you. Discussions with you were always interesting. The ideas coming from our discussions always gave me plenty of opportunities to continue the research. You always showed much interest in my work. Even when I came with some new ideas to you, you thought for just a moment and then immediately you were able to discuss the ideas in detail with me. Other PhD students usually suffer every now and then from motivational problems during their research. I never did, which might be due to your inspiration, Serge. I hope we will continue this corporation for some more time!

Also Michael Bell from Imperial College London has supervised a part of this research when I visited Imperial College. Mike, the way you welcomed me during my stay in

London was superb. You were always willing to explain or discuss the work, a point you emphasised by calling supervising PhD students “the fun part of the job”. Your commitment made the stay in London a very interesting period where I learned a lot.

Not only the supervisors inspired me during my research. The city of London was, for me, not very attractive to live in. However, the colleagues at Imperial College made it a nice stay. My roommates in “318”, as well as the other transportation scientists: thanks a lot for a great time!

I also have to thank the colleagues at the Delft University of Technology. The diversity of the group was one of the reasons to move my research from “hardcore physics” to traffic flow. I appreciated the discussions, sometimes about traffic, sometimes about other things. It always was possible to have a nice chat – or to play table tennis... The people I thank here are not only the researchers: the supporting staff also contributed to a large extent to the atmosphere in the group.

Some colleagues I would like to mention explicitly, like Maaïke Snelder. Maaïke, you work at the Delft University of Technology for just one day per week (did you make it?). Nevertheless, you are one of the colleagues I cooperated with most. The idea for chapter 8 came during a break at TRB. Did we work closely together because our research topics were closely connected or have the topics grown towards each other because we worked closely together? Anyway, I liked, and like, working with you and you became a friend as well as a colleague. Winnie Daamen also became a friend as well as a colleague. Winnie, I always enjoyed popping by your office for a chat. Sometimes we discussed topics in transportation planning. It was also nice to exchange day-to-day worries and successes. Saskia Ossen was the most special colleague for me. Saskia, it was you who first told me how nice it was to work in the Transport & Planning group. At that moment, I never expected that I would find such a nice partner in work. It was always nice to walk in your office and have a chat – and very often a laugh! Of course also thanks to my roommates which I distracted a lot (computer breaks come really often), and thanks to Femke for answering many questions about L<sup>A</sup>T<sub>E</sub>X.

I would also express my thanks to the two persons assisting me in the defence, my friend Vincent Lonij and my sister Saskia. Vincent, it has been really nice to keep so closely in touch after you have moved to the USA. Sometimes it was a pity that we could not meet too often. At the other hand, I enjoyed the holidays we had very, very much. I also would like to thank you for coming over to Europe and assist in my defence: I appreciate! With you I discussed the theoretical parts of the traffic flow theory. My other assistant during the defence, my sister Saskia, gave me input about the practical state of the traffic condition. She commented that I never experienced congestion myself. Thanks for sharing your experience, Saskia, and thanks to offer to assist in the defence.

Of course I have to thank my parents. It is said sometimes that the best thing one can give to a child is a good education. I think that this thesis is a proof you have allowed me this education. Moreover, I felt your love and care. Please know that I appreciate the education and I do really love you too!

Finally, this is possibly the best place to apologise to everyone that I did not solve the

traffic jams. You at least remembered the field I was working in. In the thesis you can read what I have been doing exactly. However, I will continue solving the traffic jams...

– Victor Knoop



# Contents

<b>Preface</b>	<b>vii</b>
<b>1 Introduction</b>	<b>1</b>
1.1 Motivation . . . . .	1
1.2 Context . . . . .	2
1.3 Research objectives and questions . . . . .	3
1.3.1 Incident location . . . . .	4
1.3.2 Network effects . . . . .	4
1.4 Thesis setup . . . . .	6
1.5 Main contributions . . . . .	6
1.5.1 Scientific contributions . . . . .	6
1.5.2 Practical contributions . . . . .	7
<b>2 Change in traffic processes due to an incident</b>	<b>11</b>
2.1 Incident and direct consequences . . . . .	11
2.2 Traffic situation . . . . .	13
2.3 Feedback to travellers and travellers' choices . . . . .	13
2.4 Network policy and robustness . . . . .	14
<b>3 Microscopic traffic behaviour around incidents</b>	<b>15</b>
3.1 Introduction . . . . .	15
3.2 Theory and behavioural hypotheses . . . . .	17
3.3 Experimental set up . . . . .	18
3.3.1 Need for trajectory data . . . . .	19
3.3.2 Description of incidents . . . . .	19
3.3.3 Description of the data . . . . .	21
3.4 Data analysis approach . . . . .	22
3.4.1 Microscopic flow characteristics . . . . .	22
3.4.2 Calibrating a car-following model . . . . .	24
3.4.3 Macroscopic flow characteristics . . . . .	27
3.5 Results . . . . .	28
3.5.1 Microscopic properties . . . . .	29
3.5.2 Macroscopic properties . . . . .	31
3.6 Conclusions and discussion . . . . .	34

<b>4</b>	<b>Motorway queue discharge rate reductions at incidents sites</b>	<b>37</b>
4.1	Introduction . . . . .	37
4.2	Literature review . . . . .	39
4.3	Data collection set-up . . . . .	41
4.3.1	Data type . . . . .	41
4.3.2	Incident selection . . . . .	42
4.4	Data analysis . . . . .	43
4.4.1	Queue discharge rates in incident conditions . . . . .	43
4.4.2	Queue discharge rate in normal conditions . . . . .	44
4.4.3	Comparing efficiency . . . . .	45
4.5	Resulting queue discharge rates . . . . .	47
4.6	Application: incident management . . . . .	50
4.6.1	Case set-up . . . . .	50
4.6.2	Results of the incident management study . . . . .	51
4.7	Conclusions and discussion . . . . .	53
<b>5</b>	<b>Influence of incidents on route choice</b>	<b>55</b>
5.1	Introduction . . . . .	55
5.2	Previous studies . . . . .	56
5.2.1	Theory . . . . .	56
5.2.2	Practice . . . . .	57
5.3	Methodology and data selection . . . . .	58
5.3.1	Possible alternative routes . . . . .	58
5.3.2	Data selection . . . . .	60
5.3.3	Indicators for route choice change . . . . .	60
5.4	Incident description . . . . .	63
5.5	Observed route choice changes . . . . .	67
5.5.1	Minor incidents . . . . .	67
5.5.2	Major incidents . . . . .	67
5.6	Conclusions and future work . . . . .	74
<b>6</b>	<b>Stochastic incident duration – impacts on delay</b>	<b>77</b>
6.1	Introduction . . . . .	78
6.2	State of the Art . . . . .	78
6.2.1	Incident duration prediction . . . . .	78
6.2.2	Delays due to incidents . . . . .	79
6.3	Theory for mathematical formulation of queue lengths . . . . .	80
6.3.1	Fundamental Diagram . . . . .	80
6.3.2	Shock Waves and Their Speeds . . . . .	81
6.3.3	Delays . . . . .	82
6.4	Elementary Road Layouts . . . . .	82
6.4.1	Road layout for considered cases . . . . .	82
6.5	Applying the model . . . . .	84
6.5.1	Scenario 1: no influence of junctions . . . . .	84
6.5.2	Scenario 2: incident upstream of a junction . . . . .	89

6.5.3	Scenario 3: queues longer than the distance to the junction . . . . .	94
6.6	Other configurations . . . . .	96
6.7	Implications . . . . .	99
6.7.1	Mathematical consequences . . . . .	100
6.7.2	Real-life incident data . . . . .	101
6.8	Conclusions . . . . .	103
<b>7</b>	<b>Consequences of spillback for blockades in a road network</b>	<b>105</b>
7.1	Introduction . . . . .	106
7.2	State of the Art . . . . .	107
7.3	Research approach . . . . .	109
7.3.1	Traffic Flow Modelling and Spillback Modelling . . . . .	110
7.3.2	Route Choice Modelling . . . . .	110
7.4	Case study description . . . . .	114
7.5	Case study results . . . . .	115
7.6	Conclusions and further research . . . . .	119
<b>8</b>	<b>Link-level vulnerability indicators for real-world networks</b>	<b>121</b>
8.1	Introduction . . . . .	121
8.2	Literature-overview of methodologies to find vulnerable links . . . . .	123
8.3	Calculation of vulnerability indicators . . . . .	123
8.3.1	Selection criteria . . . . .	123
8.3.2	Assignment . . . . .	125
8.4	Analysis . . . . .	127
8.4.1	Redundancy of criteria . . . . .	127
8.4.2	Predictive value of criteria for simulation result . . . . .	128
8.4.3	Multi-linear fit of criteria . . . . .	128
8.5	Networks . . . . .	129
8.6	Results . . . . .	130
8.6.1	Simple Network . . . . .	130
8.6.2	Delft network . . . . .	130
8.6.3	Rotterdam network . . . . .	132
8.7	Conclusions and recommendations . . . . .	137
<b>9</b>	<b>Risk-averse traffic assignment using a dynamic simulator</b>	<b>139</b>
9.1	Introduction . . . . .	139
9.2	Literature review . . . . .	141
9.3	Methodology . . . . .	143
9.3.1	Mathematical modelling . . . . .	143
9.3.2	Solution algorithm . . . . .	147
9.3.3	Reducing computation time . . . . .	149
9.3.4	Interpretation for vulnerability . . . . .	151
9.4	Case study . . . . .	151
9.4.1	Scenarios . . . . .	151
9.4.2	Network and OD-matrix . . . . .	152

9.4.3	Simulation model . . . . .	155
9.5	Results . . . . .	157
9.6	Conclusions and possible application . . . . .	160
<b>10</b>	<b>Synthesis and conclusions</b>	<b>163</b>
10.1	Aim and set-up of the chapter . . . . .	163
10.2	Findings . . . . .	163
10.2.1	Traffic properties at incident location . . . . .	163
10.2.2	Network effects . . . . .	164
10.3	Synthesis and conclusions . . . . .	165
10.3.1	Research methodology . . . . .	165
10.3.2	Incident location . . . . .	166
10.3.3	Network effects . . . . .	166
10.4	Implications for policy makers and road authorities . . . . .	167
10.4.1	Operations . . . . .	167
10.4.2	Planning . . . . .	168
10.5	Future research . . . . .	169
10.5.1	Methodological advancements . . . . .	169
10.5.2	Modelling advancements . . . . .	171
	<b>Bibliography</b>	<b>173</b>
<b>A</b>	<b>Processing traffic data collected by remote sensing</b>	<b>185</b>
A.1	Introduction . . . . .	185
A.2	Literature Review . . . . .	187
A.2.1	Microscopic Traffic Data . . . . .	187
A.2.2	Collecting and Processing Video Data . . . . .	188
A.3	Transformations . . . . .	189
A.3.1	Lateral Cross-Section . . . . .	190
A.3.2	Longitudinal Cross-Section . . . . .	192
A.4	Vehicle Detection . . . . .	193
A.4.1	Threshold Values . . . . .	194
A.4.2	Smoothing . . . . .	198
A.5	Conclusions . . . . .	198
<b>B</b>	<b>Statistics</b>	<b>201</b>
	<b>Summary</b>	<b>203</b>
	<b>Samenvatting</b>	<b>205</b>
	<b>Curriculum vitae</b>	<b>209</b>

# List of figures

1.1	A macroscopic fundamental diagram . . . . .	3
1.2	The layout of this thesis . . . . .	9
2.1	Model of traffic processes that play a role in incident situations . . . . .	12
3.1	Proposed mechanism relating attention level to macroscopic flow characteristics . . . . .	17
3.2	The incident location near Apeldoorn (right=east) . . . . .	20
3.3	Incident location near Gorinchem (right=east) . . . . .	21
3.4	The accumulation of error . . . . .	25
3.5	The headway distribution for Apeldoorn westbound left lane. . . . .	29
3.6	The distribution of reaction times . . . . .	30
3.7	The speed profile along the road as a function of the distance along the road . . . . .	32
3.8	The queue discharge rates for the different locations . . . . .	33
4.1	The fitted fundamental diagram for one incident . . . . .	45
4.2	The probability that an incident causes a queue and the correction factor . . . . .	46
4.3	The fluctuation of the capacity flow for one incident . . . . .	47
4.4	The capacity factor for the three types of incidents which block the roadway and for the rubbernecking . . . . .	48
4.5	The capacity factors found for different incidents . . . . .	49
4.6	The length of a queue for a fictitious incident . . . . .	51
4.7	The delays for different scenarios . . . . .	52
5.1	The case study area with the incident locations (numbers) and points for route decisions (letters) . . . . .	58
5.2	The traffic situation at 23 June 2008 at 16.00 hours (accident 1) . . . . .	64
5.3	The traffic situation at 22 August 2008 at 15.45 hours (accident 2) . . . . .	64
5.4	The traffic situation at 12 December 2007 at 7.30 hours (accident 3) . . . . .	65
5.5	The traffic situation at 22 February 2008 at 7.30 hours (accident 4) . . . . .	66
5.6	The traffic situation at 22 September 2008 at 13.30 hours (accident 5) . . . . .	66
5.7	The resulting route choice for accident 1 . . . . .	68
5.8	The resulting route choice for accident 2 . . . . .	69
5.9	The resulting route choice for accident 3 . . . . .	70
5.10	The resulting route choice for accident 4 . . . . .	72
5.11	The resulting route choice for accident 5 . . . . .	73

6.1	Possible road layouts and accident locations with one node . . . . .	83
6.2	The road layout used – incident locations indicated by a shaded arrow . .	83
6.3	Traffic states for a temporal bottleneck where the queue does not spill back to an upstream link . . . . .	85
6.4	The traffic states on the density-flow diagram . . . . .	86
6.5	The delay as function of $r$ and $\Delta T$ for a long link . . . . .	88
6.6	Traffic states for a temporal bottleneck on link 1 . . . . .	90
6.7	The traffic states on the density-flow diagram . . . . .	90
6.8	The contribution of the total delay that is encountered in G . . . . .	92
6.9	Traffic states when the queue is longer than $x^{\text{inc}}$ . . . . .	95
6.10	The traffic states in case of spillback on the density-flow diagram . . . . .	95
6.11	The delay caused by the incident, $x^{\text{inc}} = 6\text{km}$ . . . . .	96
6.12	The contribution of the total delay that is encountered in G . . . . .	97
6.13	Configuration 21B . . . . .	97
6.14	The fundamental diagram and the traffic states for configuration 21B . . .	98
6.15	Traffic states and shockwaves for configuration 21B . . . . .	98
6.16	The duration of incidents . . . . .	102
7.1	Outline of the research approach . . . . .	109
7.2	Congestion dynamics in a space-time plane. From left to right: (a) re- alistic queue development, including spillback, (b) implementation of a non-spillback model . . . . .	111
7.3	The route choice module . . . . .	112
7.4	The case study area, the ring road around Rotterdam, left to right around 25 km . . . . .	115
7.5	Comparison of the impacts of link closure in scenarios between spillback simulations and non-spillback simulations . . . . .	116
7.6	The impact of closure a link. The shading of a link indicates its impact: delays are large when there is a closure on a dark link; delays are minor if there is a closure on a light link. . . . .	117
7.7	Relative decrease of delay when providing information compared to fixed paths; cumulative distribution for blocking a randomly chosen link . . . . .	118
8.1	Simple test network Source Li (2008) . . . . .	129
8.2	Network of Delft, the Netherlands . . . . .	131
8.3	Network of Rotterdam, the Netherlands . . . . .	131
8.4	The scatterplot of the results of the full calculation compared with the calcalated criteria . . . . .	133
8.5	Scatterplot of the ranks of the full calculation and the criteria . . . . .	134
8.6	The number of links that should be evaluated after pre-selecting using all criteria . . . . .	135
8.7	Links that are difficult to pre-select with the selection criteria . . . . .	136
8.8	The goodness of fit of the multi-linear models . . . . .	136

---

9.1	The fraction of risk sensitive truck drivers in category 3 plotted against the difference in range - from Bogers and Van Zuylen (2004) . . . . .	141
9.2	Schemes of different solution algorithms . . . . .	148
9.3	The network used in the case study . . . . .	152
9.4	The traffic demand . . . . .	154
9.5	The convergence of assumed link blocking probabilities in composition 7; the other links have a probability of approximately 0. . . . .	157
9.6	The fraction of risk-averse travellers taking the motorway (link 4) over time in the different scenarios of table 9.3 . . . . .	158
9.7	Relative traffic volumes over each of the links in period 5 . . . . .	158
9.8	The anticipated blocking probability $f$ for the risk-averse user class in composition 2 – the width of the line indicates the anticipated blocking probability . . . . .	159
A.1	A three-dimensional box of pictures with the lateral (black) and longitudinal (gray) cross-section . . . . .	190
A.2	A lateral cross section, size 1.25 min x 50 m . . . . .	191
A.3	A longitudinal cross-section . . . . .	192
A.4	The intensity of the image over space . . . . .	195
A.5	The trajectories found on top of the image of the longitudinal cross-section	197
A.6	Trajectories for several vehicles . . . . .	198



# List of tables

1.1	The different analysis levels and spatial scopes . . . . .	4
3.1	The used indicators . . . . .	19
3.2	Properties of the incidents for which data are collected. . . . .	21
3.3	The number of vehicles used for computing average speeds and the number of leader-follower pairs used for determining headways and calibrating parameters of a car-following model . . . . .	22
3.4	The median headways and standard deviation . . . . .	29
3.5	The reaction times . . . . .	31
3.6	Capacity values for different locations . . . . .	34
4.1	Overview of the remaining capacities . . . . .	40
4.2	Road stretches used for analysis . . . . .	42
4.3	Number of incidents used . . . . .	43
4.4	The used variables . . . . .	45
4.5	The resulting capacities . . . . .	49
4.6	The effectiveness of shortening the actions in Incident Management . . . . .	52
5.1	The alternative routes . . . . .	59
5.2	The measured flows . . . . .	60
5.3	The used indicators and their meaning . . . . .	61
5.4	The incidents used in this study . . . . .	63
5.5	Summary of the results . . . . .	74
6.1	The properties of the road . . . . .	84
6.2	The symbols used . . . . .	85
6.3	Properties of the incident duration . . . . .	101
6.4	The delays in vehicle hours split up per source . . . . .	102
7.1	Reported values of travel time unreliability for car trips – source Van Amelsfort (2009) . . . . .	106
7.2	The fit parameters for the relationship between the non-spillback performance and the spillback performance . . . . .	115
8.1	List of symbols used . . . . .	124
8.2	Network characteristics . . . . .	130
8.3	The correlation coefficients . . . . .	132

9.1	The symbols used . . . . .	144
9.2	Properties of the links . . . . .	153
9.3	The distribution of the travellers over the users classes for each scenario .	155
9.4	The fraction of all travellers traversing link 4 . . . . .	159
10.1	The different levels of analysis and spatial scopes . . . . .	164

# Notation

The following variables are used throughout the thesis

symbol	meaning
$v$	Speed of a car
$v_f$	Free flow speed
$k$	Density
$q$	Flow
$u$	Average speed
$c$	Sensitivity parameter in the car-following model
$\mathcal{C}$	Set of parameters in the car-following model
$\varepsilon$	Error between the model and the reality
$s$	Distance headway between two cars
$S$	sensitivity of the parameter estimate
$w$	Weight
$C$	Capacity
$r$	Capacity reduction
$T^{\text{junct}}$	Time congestion reaches a junction
$\Delta T$	Duration of an incident
$t^{\text{end}}$	End time of congestion
$x^{\text{end}}$	Most upstream position of the (tail of a) queue
$x^{\text{inc}}$	Position of the incident
$k_2$	Density in free flow conditions
$k_1$	Density in congested conditions
$k_c$	Critical density
$k_j$	Jam density
$\omega$	Shock wave speed
$Q$	Demand
$dl$	Delay for one vehicle at one time instant $t$
$dD$	Delay for all vehicles at one time instant $t$
$D$	Total delay for all road users
$l$	Number of lanes
$\phi_{\text{non-incident}}$	The queue discharge rate in normal conditions

---

symbol	meaning
$F$	Fraction of queue discharge rate that remains available
$R$	Reduction of the queue discharge rate
$\eta$	Efficiency of the road use in incident situations
$q_{\text{queue}}$	The rate at which drivers drive into a traffic jam
$N$	Number of delayed vehicles
$A$	Average delay per traveller
$\mathcal{R}$	Sum of the correlation coefficients
$I$	Indicator for link vulnerability
$\mathcal{I}$	Set of criteria indicating link vulnerability
$\mathcal{K}$	subset of criteria indicating vulnerability
$rD$	Relative delay
$N$	Cumulative vehicles
$L$	Length of a link
$\Psi$	Split fraction towards each of the links (multi-dimensional)
$\psi$	Split fraction to one direction
$\chi$	Fraction of travellers changing route in incident situation
$\pi$	route choice
$G$	Network
ss	boolean variable indicating whether spillback is included or not
$f$	Anticipated link failure probability
$U$	User class
$t$	Time, discretised in time slices numbered by the integer $t$
$\Delta t$	The duration of one time slice
$z$	Path number
$\mathbf{f}_U$	Vector of link failure probabilities for user class $C$
$h_{U,t}$	Route choice for user class $U$ : fraction of travellers in class $U$ taking route $z$
$a_{iz}$	Link incidence matrix (1 if link $i$ is on path $z$ , 0 otherwise)
$\mathbf{tt}_{ijt}^{\text{link}}$	Travel time in time slice $t$ on link $i$ under scenario $j$
$\mathbf{tt}_{jzt}^{\text{path}}$	Travel time in time slice $t$ on path $z$ under scenario $j$
$\tau_{j,U}$	Travel time of a user in class $U$ under scenario $j$
$\langle \mathbf{tt}_{it} \rangle_U$	The anticipated travel time in time slice $t$ on link $i$ for users in class $U$
$\vartheta$	Constant indicating the risk-averseness of a user class
$n$	Iteration number of the algorithm
$m$	The reciprocal frequency [iteration numbers] for computing the Total Cost of all scenarios
$w$	Number of scenarios assessed when iteration number is not a multiple of $m$
$D$	Operator for the distance between 2 vectors
th	Threshold value for converging algorithm
$p_U$	Probability assigned to the blocking scenarios for the risk-averse users ( $0 < p_U < 1$ )

---

---

The following notation is used to show properties of the variable  $a$ :

symbol	meaning
$\langle a \rangle$	Expectation value of $a$
$\bar{a}$	Mean value of $a$
$\tilde{a}$	Approximation of $a$
$a^*$	Optimal $a$

---



# Chapter 1

## Introduction

### 1.1 Motivation

Traffic incidents are an important cause of congestion and the total cost of traffic congestion is high. It is estimated that annually 44 million hours are lost due congestion in the Netherlands, which, translated into monetary costs, is equivalent to a loss of 700 million euros (Dutch Road Authority, 2007a). About 25% if this is caused by incidents (Dutch Road Authority, 2007a).

In general, traffic becomes congested when the demand is larger than the supply, i.e. there are more travellers than the road can cope with. In general, two types of congestion can be distinguished: recurrent congestion and non-recurrent congestion. Travellers dislike non-recurrent and therefore unexpected congestion, even more than recurrent congestion. Travellers can attach a value to a short travel time (value of time, VOT), but also to the reliability of travel times (value of reliability, VOR). These values can be compared with each other, as is for instance done by Noland et al. (1998) or Bates et al. (2001). Generally, the VOT (in euros per minute travel time) and the VOR (in euros per standard deviation in travel time) are similar, although different studies show different values, e.g. Eliasson (2004). Differences are explicitly discussed by Börjesson (2007). Some studies De Jong et al. (2007) separate the value of reliability on the one hand from the value of departing early and the value of arriving late on the other hand. These argue that there is a disutility of the variation of travel time which is independent of the disutility of being late or departing early. Other studies like Bates et al. (2001) find that either the value of reliability is significant or the value of arriving early or departing late is significant. How this changes travel behaviour, is shown by Li (2009).

Travel time and delay variability can be caused by stochastic fluctuations in demand or capacity. The second possibility is for instance shown by Tu et al. (2008) who describe how traffic flow influences the probability on breakdown. This probability increases if the flow increases, but there is no fixed “critical” demand under which level the flow will not break down and above which level it will certainly break down. This means that there are

traffic demand levels for which sometimes traffic breaks down and sometimes not, which causes a stochastic element in the travel times.

Another possibility for non-recurrent congestion is an event which causes a sharp increase in the demand or an incident which causes a decrease in road capacity. Kwon et al. (2006) analyse the total traffic congestion and find that around 25% of the delay in their study area in the US is caused by incidents. So incidents cause a large part of the delay and drivers dislike this delay the most because it is unexpected and, therefore, they may be late for a meeting. This thesis discusses the traffic flow at incident sites which are the origin of congestion, and how this influences many drivers in the whole network.

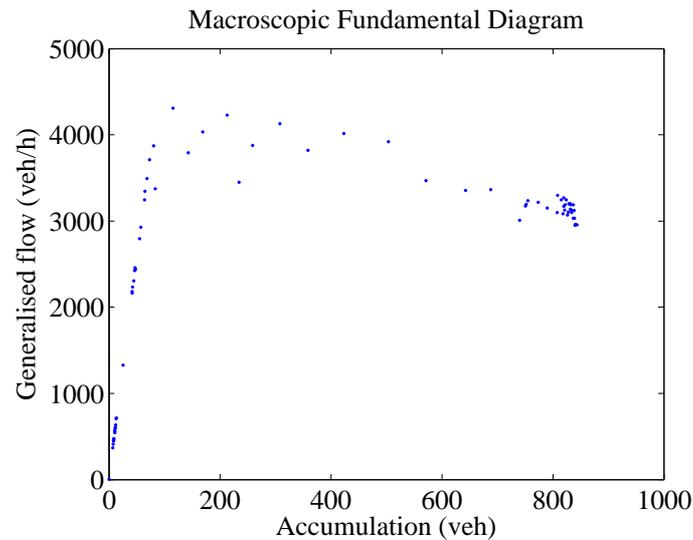
Traffic network congestion caused by incidents is also scientifically challenging. Much effort has been put into describing traffic flows and travel times over the years. It is not known yet how drivers change their operational driving behaviour at an incident site and how incidents influence the traffic situation in the whole network. In recent years, advanced and detailed models to describe the individual traffic behaviour of drivers have been proposed and tested Ossen (2008). However, this has never been analysed under incident conditions. Therefore, neither the resulting traffic conditions nor the delays can be predicted. However, in order to take appropriate measures of Dynamic Traffic Management like dynamic traffic guidance, it is essential to be able to forecast the traffic situation with different measures. This does not only hold for the traffic at the incident site, but also for the propagation of the queues through the network. In sum, it is necessary to know what happens at an incident site to take appropriate actions to reduce delays. It also is required to know what the road users' reaction to delays is in terms of route choice: will people change their route if one lane of the motorway is blocked? These aspects will be discussed in this thesis.

The duration of an incident is an important aspect in this forecast. There have been many studies (e.g, Knibbe et al. (2006), Wang et al. (2005) and Zhang et al. (2007a)) which also examined this subject, and these are discussed in section 6.2.1. This thesis will therefore not develop new incident duration prediction models.

## 1.2 Context

This research has been carried out as part of the research program of “The Next Generation Infrastructures Foundation” (Next Generation Infrastructures, 2009). Its goal is “to improve the way our vital infrastructures are planned, designed, operated and maintained.” This is a multidisciplinary program where different types of networks are studied, for instance ICT networks, electricity networks and road networks.

The performance of a road network under load differs fundamentally from many other types of networks, like internet networks (IP-network). In fact, if the load (traffic demand, in veh/h or bit/h) increases, the performance increases up to a certain load level, usually indicated by a critical density. Figure 1.1 shows the relationship between the load of the network and its performance. If the load exceeds the critical load, the cars in the



**Figure 1.1: A macroscopic fundamental diagram**

network block each other and the network performance decreases (Geroliminis and Daganzo, 2008). This is different, for instance, in an IP-network, where performance will not increase if capacity is reached, but there is no “blocking back” or gridlock either. So if the load increases, the rate of data exchange stays at the highest possible level until all tasks have been performed.

One of the projects in this program aims to improve the robustness of road traffic network operations. Road networks are designed for normal conditions. In this thesis robustness of a road network is defined as the way the network copes with a short-term, strong decrease in capacity or a short-term strong increase of demand. The first situation is discussed in this thesis: an incident on the road which decreases the capacity. Incidents are unpredictable short-term (typically a few hours) events blocking a part of the road, like car-crashes which we will call “accidents”. The thesis of Li (2009), which discusses reliability-based network design, is also part of the same project.

### 1.3 Research objectives and questions

This thesis aims to provide insights into the relevant traffic processes during an incident, as well as to describe these processes mathematically. It also aims to model the traffic situations which result from the incidents.

To reach the main research objective, the following research question has to be answered: “How does the traffic situation change due to a road incident?” This question can be unraveled in several subquestions which are stated below. There are two dimensions of the main problem, the aggregation dimension and the spatial dimension. The first dimension is the aggregation level, which gives a distinction between processes at the level of the individual driver, the microscopic level, and processes at a more aggregated

**Table 1.1: The different analysis levels and spatial scopes**

Level \ Spatial scale	microscopic	macroscopic
Incident location	I	II
Network	III	IV

level of traffic flows, the macroscopic level. The second dimension is the location. The processes at the incident location can be studied or the consequences for the traffic in the whole network (see table 1.1).

### 1.3.1 Incident location

The first research questions discuss what happens at the location of the incident.

#### **Research question 1: How do drivers change their driving behaviour when passing by an incident location?**

This is a broad research question which aims to describe qualitatively the car following behaviour of the drivers as well as the choice of the free speed in case of free-flow driving conditions. This question is discussed in chapter 3.

The formation of queues is largely determined by the capacity at an incident location, which is a macroscopic quantity. The change in capacity is caused by a decrease in the number of available lanes and a change in (microscopic) driving behaviour. Mitigating actions to increase the maximum flow can be only taken if this behavioral change is known.

Due to the limited number of lanes that is available and changes in driving behaviour, the capacity of the road at an incident location decreases. Therefore, the following question is raised:

#### **Research question 2: What is the capacity of a motorway at the position of an incident location and which are the influencing factors?**

Research question 2 requires less detailed data than the first research question. These data are abundantly available so that capacity effects can be studied for many situations and locations. This question is discussed in chapter 4.

### 1.3.2 Network effects

With the first two research questions, the traffic characteristics on the incident site are analysed. The other research questions examine the consequences at other places in the network. The resulting traffic conditions depend largely on the route choice that travellers make, which leads to questioning how people change their route choice behaviour.

**Research question 3: Do travellers change their intended route when they are faced with unexpected congestion due to an incident?**

This question is discussed in chapter 5.

In principle, once the capacity and the route choice are known, the traffic conditions and the delay for general traffic conditions resulting after an incident can be derived. This requires more knowledge about the dynamics of the congestion and about the delay of the congestion.

**Research question 4: How do the queues caused by an temporal bottleneck propagate through a traffic network and how can the delay be computed for general traffic conditions and road layouts?**

This question is discussed in chapter 6.

The answer to that question will provide insights into the magnitude of the delays and the typical key parameters determining the delay. One of the parameters is the location where the incident takes place. This raises the following question:

**Research question 5: At which positions in a network does an incident cause the greatest delay?**

This question is discussed in chapter 7.

These locations determine the vulnerability of the network and these locations are called vulnerable. Note that this definition for vulnerability does not contain an element of how often incidents happen at these places. Throughout the thesis, the probability on an incident is taken as an external factor. The most vulnerable links will be sought by simulation. However, as stated by Li (2008), much computation time is required to compute the delay of an incident at all places. Moreover it does not provide insight into the risk factors. That makes one wonder whether it is possible to estimate the reduction in network performance caused by an incident at a specific location based on traffic conditions which can be derived from recurrent conditions, such as for example traffic volume or road capacity? In short, this is formulated as follows:

**Research question 6: How can the vulnerability of links be predicted based on properties of the flow in non-incident conditions?**

This question is discussed in chapter 8.

The research questions about vulnerable links assumes that people do not include the risk of a blockage in their route choice. Travellers could expect the possibility of an incident when making the decision on their route. In particular, risk-averse travellers would anticipate the worst possible situation. Therefore, the following question will be discussed:

**Research question 7: What routes are taken by travellers who are risk-averse?**

This question is discussed in chapter 9.

## 1.4 Thesis setup

Figure 1.2 shows how the different chapters in this thesis are connected. The figure shows by arrows a suggested reading order indicated by arrows if one prefers to focus on either the macroscopic or microscopic level. A dotted arrow means that there is a connection between these chapters from one aggregation level to another. Additionally, it should be stated that after each of the chapters 3 to 9, the reader can turn to the conclusions.

In chapter 2 the context of the research is shown. Chapters 3 to 9 then all answer one of the research questions mentioned above. All these chapters are self-contained, including an introduction and conclusions, and can therefore be read as individual articles.

## 1.5 Main contributions

This thesis describes the scientific work that has been carried out to find a solution for a practical problem (“How does the traffic situation change due to a road incident?”). In the next two sections the contributions are discussed: first the scientific contributions, then the practical contributions.

### 1.5.1 Scientific contributions

#### **Methodological contribution**

In this thesis a dynamic traffic simulation program is developed to simulate traffic operations. It can be switched between two states: with spillback and without spillback. “Spillback” is the phenomenon when a queue grows longer than the link length and also vehicles which do not need to pass the bottleneck location are delayed. The only difference is the presence of spillback; all other features of traffic operations, including the way traffic jams grow and dissolve is the same.

The thesis also shows a new method to identify vulnerable links. It is shown that in this process spillback or blocking back effects play a crucial role. Because this is shown to be important, it is concluded that proper modelling of the queue length is important. It also means that it is important to choose an appropriate network structure to avoid unnecessary spillback delays. This thesis also shows the (limited) value of link-based indicators to identify vulnerable links (chapter 8).

A new method for recording traffic operations on video is applied to study the driving behaviour at incident locations. With a helicopter the observation team could fly to the incident location and observe traffic from a high position. This made it possible to model the driving behaviour at the incident location.

Bogers and Van Zuylen (2004) show that most road users are risk-averse in their route choice. Up to now, risk-averse dynamic traffic assignment has not been combined with

dynamic traffic simulation models which have a realistic queuing model. This thesis shows how risk-averse routing can be combined with a dynamic traffic simulator.

### **Theoretical contribution**

First of all, there is a contribution in the field of driving behaviour. Incidents cause delays in several ways. First, the roadway is partially blocked which reduces the capacity. Secondly, the capacity of the remaining lanes decreases due to changed driving behaviour. One of the main contributions of this thesis is that for the first time the driving models are fitted to measured trajectories of vehicles at the moment they pass the incident. Also, the road capacity at the incident site is studied and described.

Secondly, the influence of the duration is quantified analytically. The total delay that an incident causes depends on the duration of the blockage. This thesis shows analytically how the total delay depends on the variability of the incident duration. It even shows the consequences when the queue grows and the tail of the queue reaches an intersection and blocks it.

A third contribution is that route choice during and after incidents is described. The traffic situation also depends on the routes travellers take. Chapter 5 describes how travellers deviate from their intended route if that route is severely congested, but that minor delays do not have an influence on the chosen routes.

Finally, this thesis makes a contribution on risk-averse route choice. It shows how the link loads and therefore the vulnerable links change if traffic is assumed to make risk-averse choices for their routes. The main motorways are then used less and the traffic diverts to the underlying road network.

## **1.5.2 Practical contributions**

One of the aspects of engineering work is that it aims to solve practical problems. In this case, the practical problems are the traffic jams that many face on a regular basis. In fact, the problem has been analysed in detail and the critical points for improving the situation in practice have been presented. These points are summarised below.

### **On-line stage**

By studying a large amount of incidents, the maximum flow values that can be obtained at accident sites are determined. These can be input for on-line (or ex-ante) traffic predictions, which can help to provide correct and predictive information about the queue length to travellers. Earlier research by Khattak et al. (2003) showed that this type of information, if correct, is highly valued by travellers.

Alternatively, this information could be used to advise travellers in taking an appropriate alternative route which reduces the delay. For giving the best possible routing advice, traffic managers are assisted by computer systems. These can better predict the impact of possible control actions (like route advice or lane closure) if they are provided with more accurate information about the capacity. The thesis shows that this information can change the traffic situation

**Planning stage**

First of all, people drive differently when passing an incident site. Therefore, they keep a larger headway to their leaders, causing the capacity of the road considerably decreases. A decreased capacity yields an imbalance between demand and supply and can cause traffic jams and delays. The capacity does not only decrease in the direction of the incident, but also in the opposite direction of the motorway. This effect is called “rubbernecking”. A possible solution is to shield the incident location with opaque screens.

So-called vulnerable links can be identified based on the methods proposed in this thesis. These are links which, by their position in the network, are valuable for the traffic and if they are blocked the traffic would encounter long delays. Note that travellers aiming for the blocked route may be delayed, but also travellers aiming for different routes can be delayed since the queue formed might spill back onto their route. If the vulnerable parts are known, the road authority could focus their efforts on solving traffic jams on those links. For instance, they could try to reduce the number of incidents there (e.g., by changing the road layout or lowering the speed limit) or improve the emergency response time, for instance by relocating emergency vehicles or improving incident detection. Another possibility is to build buffer space where cars can queue without hindering other cars.

**Policy**

This thesis shows that the duration of an incident plays a crucial role in the total travel delay. The influence of duration is for instance more important than the remaining capacity. It is therefore recommended to put great effort into actions which reduce the incident duration. An example thereof is a public campaign that in case there are no personal injuries, drivers are requested to move their damaged cars to a parking place, instead of finding the responsible driver at the roadway. Another example is to locate more recovery vehicles at strategic locations to remove wrecks which cannot move anymore.

Furthermore, the thesis also shows that spillback is important, i.e. the delay caused by a queue which grows and blocks more upstream links. This can for instance be reduced by separating traffic flows, for instance by a hard barrier to separate local traffic from through traffic. Also, the use of the emergency lane as a buffer for cars wanting to leave the motorway could improve the flow for through traffic if the off ramp is congested.

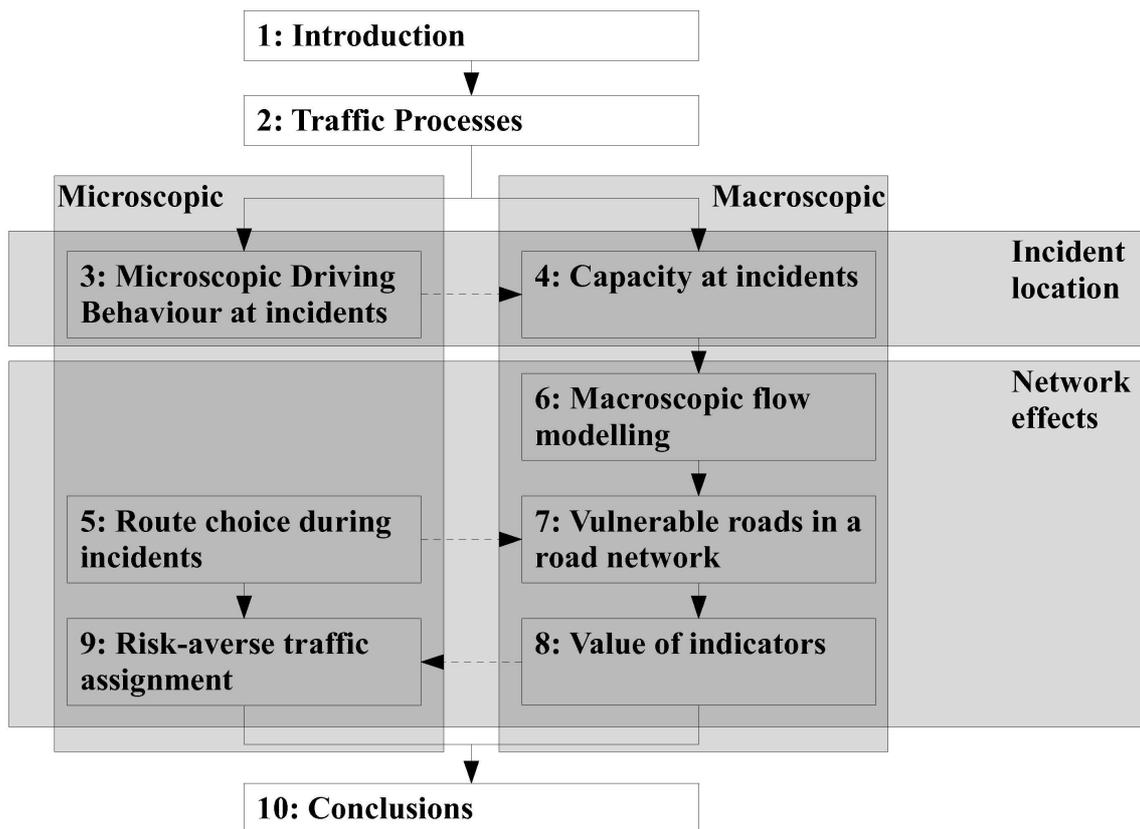


Figure 1.2: The layout of this thesis



# Chapter 2

## Change in traffic processes due to an incident

---

This chapter discusses the traffic processes that play a role when there is an incident on the road. Figure 2.1 shows these processes and their relationship graphically. The boxes are the elements and the arrows indicate causal relationships. In this chapter, the boxes and the arrows in the figure will be explained.

---

### 2.1 Incident and direct consequences

This study only considers the traffic situation after an incident has occurred. Therefore, it will neither discuss the processes which lead to an incident, nor the influence of factors which increase the probability of an incident. The incident is considered as an exogenous process to the scheme in figure 2.1.

The incident changes the traffic operations at the incident site in two ways. First of all, the incident reduces the number of lanes that is available to the traffic. But travellers will also change their driving behaviour due to the incident. They will possibly look at the incident or drive more carefully because there are workers on the roadway. This distraction changes the headways of the drivers. How driving behaviour changes is studied in this thesis. Chapter 3 shows the drivers' reaction to an incident and how this incident changes their behaviour. This chapter thus describes the traffic behaviour at an individual, microscopic level. In this chapter, video data on driving behaviour near incidents are analysed. Chapter 4 shows the results of this changed behaviour on a macroscopic level. It shows the capacities at incident sites, obtained by analysing loop detector data.

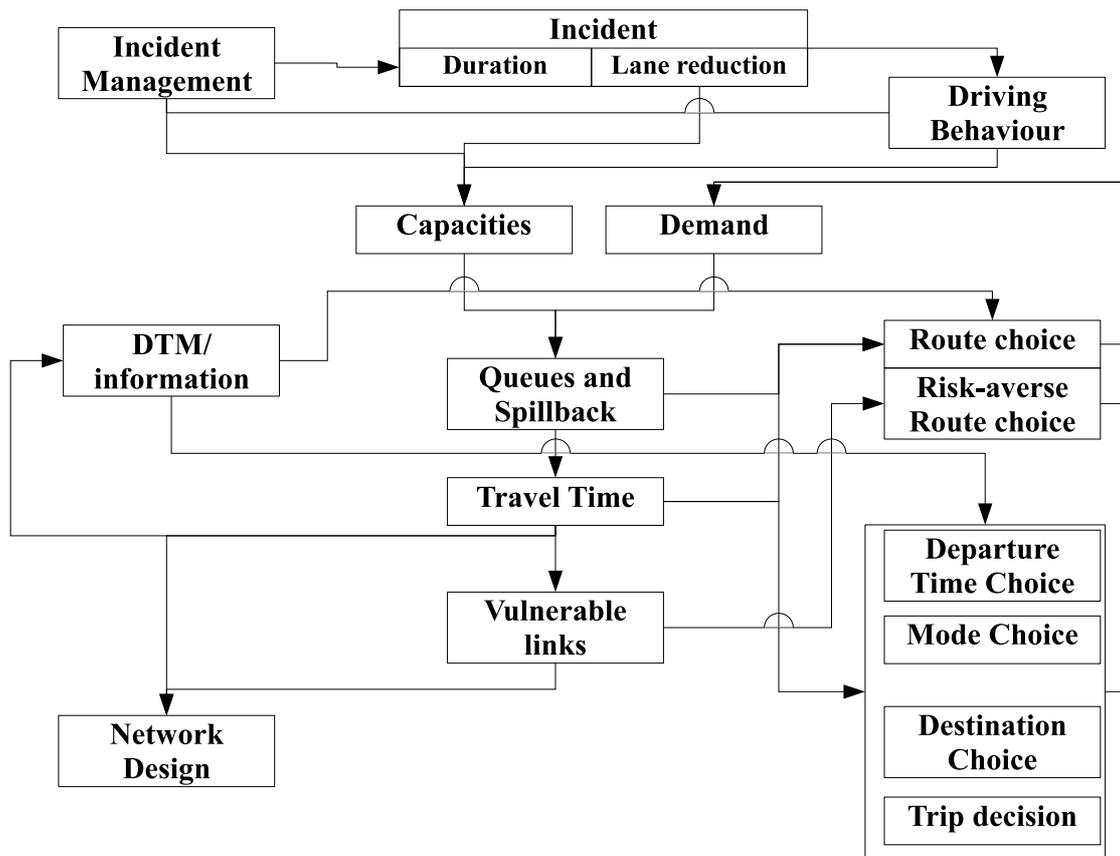


Figure 2.1: Model of traffic processes that play a role in incident situations

## 2.2 Traffic situation

If the capacity of a link is lower than the traffic demand, i.e. the amount of travellers wanting to pass that link per unit time, congestion will set in. The delay will not take place at the link where the capacity is exceeded, but the head of the queue will be at that point, the bottleneck. The vehicles that cannot pass the bottleneck have to wait in a queue upstream of the bottleneck. A macroscopic, analytical model for a description of the traffic situation is presented in chapter 6.

Insufficient capacity leads to queues which grow in the network. One of the typical aspects of a network, discussed in detail in chapters 6 and 7, is that the queue can grow so long that the link with the bottleneck is overfilled. This way, upstream links can become blocked and travellers who do not need to pass the bottleneck can also be delayed. The influence of this spillback effect, as well as the influence of location is analysed by a simulation model in which spillback can be switched on and off. This model is presented and discussed in chapter 7.

## 2.3 Feedback to travellers and travellers' choices

A queue on a particular stretch of road will lead to an increase in travel time on a route which incorporates that stretch, making it less attractive to travellers. Travellers are informed about the traffic situation, for instance on the radio broadcasting or through navigation devices. Using that information, they can change their intended route. They might also change their route based on the traffic situation they are actually facing. Chapter 5 analyses the percentage of travellers who actually change their route when faced by an incident leading to delays. However, this thesis does not discuss in detail all the possibilities of informing travellers. This is for instance studied by Sodnik et al. (2008) or Dicke-Ogenia et al. (2008). It also does not analyse how travellers respond to such information. This problem is addressed in the PhD thesis of Bogers (2009). In this thesis, the route choice change is derived from changes in flows to the incident route and to alternative routes. The resulting route choice is considered to be an effect of the combination of traffic situation and information. The route choice change is used as input in the chapters on queuing, chapters 7 and 8.

During the process of choosing routes in advance, travellers might even consider the possibilities of an incident even before it has actually happened. This risk-averse behaviour can also be included in the traffic assignment. How this can be done in a dynamic traffic simulation program is shown in chapter 9.

Apart from the change in route choice, travellers can also make larger changes in their travel plan. If the delay is large, they can choose to change their departure time. They can depart earlier to still be on time. If possibility is too unattractive, they may accept being late and depart after most of the queue has dissolved. In that way, they are late but they do not have a very large travel time.

One of the other options is to reschedule the activities on a planned trip with multiple destinations. In that way, travellers can avoid the congested road stretches and thus avoid the delay caused by the incident. Alternatively, travellers can choose to take another form of transportation, for instance using the train instead of the car. A final option is not to travel at all due to the extra delay caused by the incident. However, this thesis only describes behaviour on the operational and tactical level and therefore these changes are not discussed in this thesis.

The process above is the main loop that feeds information back to travellers. In short: there is a capacity reduction which causes queues and leads to delays. This increased travel time may stimulate travellers to take another route, which reduces the inflow. Consequently the queues decrease compared to an unchanged demand.

## **2.4 Network policy and robustness**

For the strategy of a road authority it is important to find the road sections where incidents cause the largest delays. This knowledge can be used, for instance, to locate the emergency vehicles to reduce incident duration. On the other hand, the network can also be designed in a risk-averse way. That is, a policy maker can choose to design a network in such a way that the consequences of incidents are small. This is one of the possible policy goals, besides for instance total travel time in a normal situation, and environmental issues. The process of network design lies outside the scope of this thesis. However, the knowledge which links are vulnerable provides input for the network design. Chapters 7 and 8 indicate how the most vulnerable links can be found. Chapter 7 shows a “brute force” method which checks by simulation the vulnerability of each of the links. Chapter 8 evaluates indicators to pre-select possible vulnerable links.

Another aspect for network operators is incident management. The paragraph above has already mentioned how information on vulnerable links could help in choosing the location of emergency vehicles. On the other hand, there are many measures which can shorten the incident duration. The analytical formulation of the queuing model in chapter 6 focusses specifically on incident duration.

## Chapter 3

# Microscopic traffic behaviour around incidents

---

The delays caused by incidents are partly caused by blockage or closure of lanes, but also by the change in driving behaviour in the remaining lanes. This chapter analyses traffic flow conditions near an incident both microscopically. A theory is proposed to describe drivers' behaviour, which is tested using real-life traffic data of individual vehicles, collected using a helicopter. A bimodal headway distribution is observed, centred around two mean values, 2 seconds and 4 seconds. To understand the underlying mechanisms a car-following model is fitted to the drivers' behaviour. The model parameters show that the reaction time is much higher than usual. The measurements are restricted to vehicle movement, and no psychological factors are actually measured. Using a model-based analysis, we derive that the incident distracts the drivers. The consequence is that the queue discharge rate for the unblocked lanes is 30% lower than the usual queue discharge rate per lane. This chapter is a edited version of: Knoop, V. L., Van Zuylen, H. J. and Hoogendoorn, S. P. (2009) Microscopic Traffic Behaviour near Accidents. In Lam, W. H. K., Wong, S. C. and Lo, H. K. (Eds.) *Proceedings of the 18th International Symposium of Transportation and Traffic Theory*. Springer, New York.

---

### 3.1 Introduction

The Highway Capacity Manual (Transportation Research Board, 2000) provides estimates for capacities around incidents. Qin and Smith (2001) carry out a more detailed analysis, but their analysis is macroscopic, based on macroscopic data. Incidents are also modelled in micro-simulation packages. Using several of these software tools, Sinha et al. (2007) find a capacity reduction of about 50% if one of three lanes is blocked. The empirical reduction found by Knoop et al. (2008b) is around 65%.

Given the observed changes in macroscopic flow characteristics, we know that drivers change their behaviour when driving along an incident site. However, to gain more insight into these behavioural changes, empirical microscopic traffic data is required. To the best of our knowledge, microscopic measurements of the behaviour of single vehicles around incidents have not been collected and analysed in detail until now. The change in individual driving behaviour causes a change in the macroscopic traffic flow characteristics of the road, such as the capacity. It is still a scientific challenge to describe traffic flow along incidents, especially at the level of individual drivers. This contribution takes on this challenge. For the first time, a theory is proposed for microscopic driving behaviour around incidents which is tested with real-life data. This conceptual model is proposed and then tested using data of individual vehicles in real life.

To this end, we have collected empirical trajectory data using a digital camera mounted under a helicopter. At two incident locations, the traffic operations have been captured on high resolution video from the helicopter which flew high enough not influence the traffic operations. From the video, we derived trajectories, as described in appendix A of this thesis.

We have calibrated a car-following model for these incident locations. At first, data is analysed at the level of individual vehicles and drivers. We show considerable differences in individual driving behaviour between driving under normal conditions and driving along an incident site. We show how these changes translate into changes in macroscopic properties which are also found in the data.

The behaviour of the traffic on the carriageway in both directions has been analysed. Particularly the behaviour on the carriageway which is not blocked is of interest from a behavioural point of view. Since there is no physical obstruction that causes a change in traffic operations, the changes are solely due to a changed driving behaviour. Analysing the effects on this carriageway quantifies the effects of “rubbernecking”. Note that this chapter describes the differences in vehicle movements between normal conditions and conditions near an accident site. The chapter will not investigate the psychological causes for these differences. It will indicate some likely causes, but further research will be necessary for the psychological explanation.

Although drivers might act differently at another incident site, the change in driving found in these two cases appears to have generic features and may have a certain generic validity for application to a different location. The qualitative changes are likely to be the same, although quantitatively, effects may vary (for instance on the incident type, the presence of emergency services etc.). More importantly, it shows the microscopic mechanisms that cause the flow to change and thus which parameters to tune to calibrate a model for incident situations.

The remainder of the chapter is set up as follows. The next section poses the hypotheses that are tested in the chapter, section 3.3 describes the experimental setup. Section 3.4 and 3.5 are closely related and describe respectively the data processing and the results. If the reader prefers, these sections can be read together, paragraph by paragraph. Section 3.6 summarises the chapter and gives the concluding remarks.

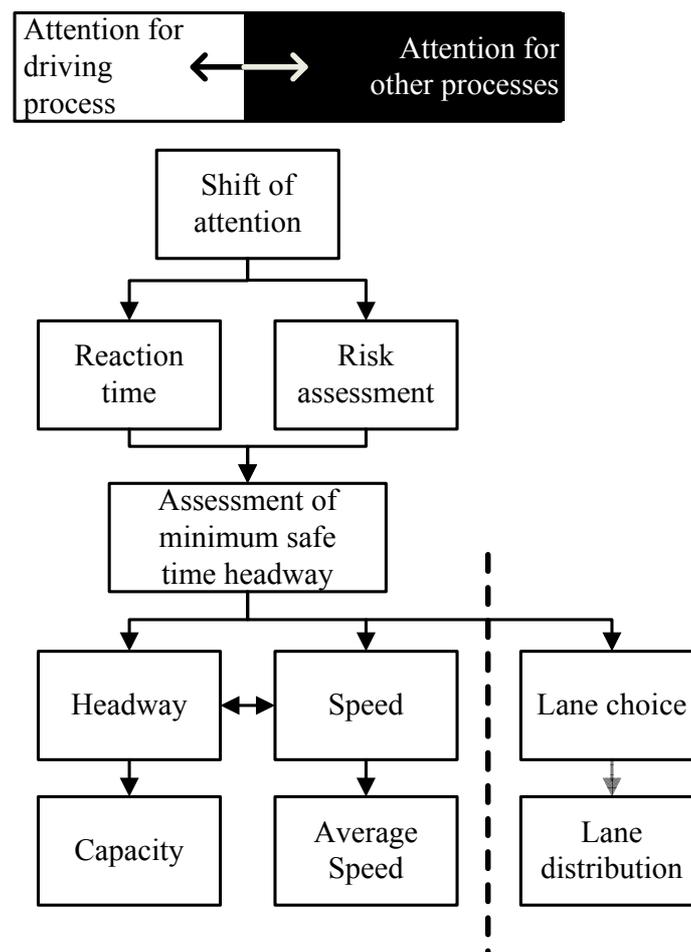


Figure 3.1: Proposed mechanism relating attention level to macroscopic flow characteristics

## 3.2 Theory and behavioural hypotheses

This contribution proposes a theory predicting how driving behaviour around incidents is likely to differ from normal conditions. In our theory, drivers shift their attention towards the incident when passing by the incident site. The degree of attention shift differs per driver, with some drivers only focussing on the traffic whereas others focus mainly on the incident.

This attention shift changes the driving behaviour at incident sites in a way depicted in figure 3.1. We propose that the drivers are distracted and due to the limited attentional resources, the attention for the driving task reduces. Two changes now take place. Firstly, the reaction time increases. Secondly, the risk of an accident which involves themselves increases. The drivers therefore choose to have a larger headway. In practice, drivers could obtain a larger headway by reducing speed, in which case an equal space headway becomes a larger time headway or alternatively by changing lanes. These driver characteristics change the macroscopic properties of the flow. An increased minimum

headway causes a reduced capacity, a reduced speed causes a reduced average speed and a lane change will change the distribution of vehicles across the available lanes. Under congested flow conditions all lanes are occupied and a lane change does not increase the headway. Since the measurements presented here are taken in congestion, this part of the model will not be discussed.

The theory and the other parts of the model are tested using four hypotheses about the driving behaviour stated below:

1. Drivers choose a different headway in a bottleneck caused by an incident compared to a bottleneck in normal traffic and the mean headway is larger.
2. Drivers have a different reaction time at an incident site compared to normal conditions and the mean reaction time is larger.
3. People reduce speed in view of an incident.
4. The queue discharge rate at the incident site is lower than the outflow capacity under normal conditions for the same roadway geometry.

Note that a “bottleneck” is defined as the place where the demand exceeds the capacity, and therefore the traffic conditions upstream are in congested, in the bottleneck there is capacity and downstream of the bottleneck traffic is in a free-flow condition or also at capacity. Acceptance of the hypotheses shows a change in driving behaviour. The underlying psychological process, the shift of attention, can not be proven by this type of measurements. However, if the driving behaviour changes considerably, this is a good indication that the proposed model holds.

Hypotheses 1 and 2 are based on individual driving behaviour, whereas hypotheses 3 and 4 are macroscopic properties. The summary of the data that have been extracted is given below in table 3.1 where for each of the hypotheses, the indicator used is stated. The indicators have been compared with the same indicators for normal traffic. For one indicator at one site, we had to use a literature value for the reference situation stating an average value for the Netherlands as well as the spread. For all other indicators, a comparable indicator is found, although they are sometimes derived from data at another, comparable location with a comparable driver population. This was needed since the traffic conditions at the incident sites differ from the conditions at the same locations in non-incident situations, and therefore, at same sites no comparable data for non-incident situations could be derived. Most importantly, the locations of reference situations have the same traffic conditions.

### **3.3 Experimental set up**

This section discusses the required measurements for this study. The first part discusses the necessity of detailed trajectory information in order to gain insight into the macroscopic and microscopic traffic flow operations near incidents. The second part introduces the observations that are used for this chapter and in the third part the properties of the resulting data are given.

**Table 3.1: The used indicators**

	Hypothesis	Indicator
1	Headways	Headway distribution
2	Reaction time distribution	Results of a fit of a car-following model
3	Traffic flow speed	Average speed profile and interval bounds
4	Outflow capacity	Outflow capacities and spread

### 3.3.1 Need for trajectory data

There are many ways to measure traffic behaviour. Recently, Schaap et al. (2008) showed how drivers react on an unexpected event in a driving simulator. They showed how an unexpected action of a leading vehicle influences the driving over longer time. We performed measurements of the effects of incident in real-world. Loop detectors are the most common way to gather traffic data, but these will not provide sufficiently detailed information about the driving behaviour dynamics around incidents, since they only provide local, cross-sectional, information. Spatial information, or, in fact, trajectory information, is needed, because this will yield information of the behaviour upstream of the incident and at the incident location itself. This allows for instance to observe speed adaptation and car-following behaviour, including estimation of the reaction times.

For microscopic behaviour, it is essential to follow a single vehicle and its leaders over a certain time period. Only from the exact place of the vehicle and the leaders, one can derive the stimuli that are possibly input for the driver to accelerate. To derive the acceleration, one needs a time series of accurate position measurements.

For the capacity estimation passing times are sufficient. However, data from one loop detector will not be sufficient, since one needs to know whether there is a queue waiting; this information cannot be derived from one local measurement. So we need information which has both a temporal and a spacial extend, also for capacity measurements. Trajectory data has this information.

### 3.3.2 Description of incidents

The following approach was taken to get the trajectory data of an incident (Hoogendoorn et al., 2003). A digital photo camera was attached to a helicopter. The helicopter stayed approximately at the same position, above the incident. The camera could move in all directions to compensate for the helicopter movements. Digital photographs were taken at a rate of 15.1 images per second and saved to a hard disc drive. The size of the pictures is 1392 x 1040 pixels. The altitude of the helicopter is around 400 meter and the length of the long size of the image is also around 400 meter. This implies that one pixel equals about 30x30 cm on the road.

The observation team waited at the Traffic Management Centre in the centre of the Netherlands until an incident was reported nearby, after which it flew with the helicopter to



**Figure 3.2: The incident location near Apeldoorn (right=east)**

the incident location. From the moment of arrival, traffic operations for both directions have been recorded. For traffic in the non-incident direction, the incident was visible but it formed no physical obstruction. The video shows the so-called “rubbernecking effect” (i.e. people watching the incident at the other side of the guardrail). The altitude of the helicopter was large enough not to influence the traffic operations.

The remainder of the chapter presents the data for two incidents. All properties are summarised in table 3.2 below. At the first incident a van rolled over at 6 June 2007, at around 9:15 am, near the city of Apeldoorn at the motorway A1 in the Netherlands. It ended in the median strip, the unpaved area between the two carriageways of the motorway. The road has two carriageways in each direction and no gradient. Congestion occurred in both directions with the heads of both queues at the location of the incident which means that the incident formed a bottleneck. For the eastbound direction, the emergency vehicles blocked one lane which was therefore unavailable for the traffic. For the other, westbound, direction, the delay was only caused by rubbernecking. Since there was a tunnel around 100 meters west of the blocking of the lane, the traffic operations there were invisible (see figure 3.2). The crosses mark the blocked lanes, the arrows the available lanes.

The lanes used changed over time for traffic in the eastbound direction. Sometimes, the emergency lane was used, sometimes, the right lane was used, and sometimes, they were both used. From the video we have observed that some drivers choose to avoid the emergency lane without an apparent reason. There is no indication that the conditions for taking the emergency lane changed over time. Due to this intermittent use of both lanes, taking the flow values of each of the lanes separately would yield nonsensical results (since often no-one uses the lane, although there is a queue waiting). The sum of the flows yields the capacity for the case at hand, incorporating both the headway choice and the lane choice of drivers.

The second incident happened at 6 June 2007, at around 11am, at the westbound carriageway of the two-lane motorway A15. In this incident, several trucks and passenger cars were involved which blocked one lane of the two-lane motorway as well as the shoulder lane. From time to time, the police stopped the traffic for a while to recover a car from the incident. The incident happened several hundreds of meters downstream of the motorway

**Table 3.2: Properties of the incidents for which data are collected.**

Nearby city	Apeldoorn	Gorinchem
Road	A1	A15
Date	6 June 2007	6 June 2007
Time	9:39 – 10:15	11:59 – 12:56
Weather	Clear	Clear
Type	Motorway	Motorway
Number of lanes	2 x (2 + shoulder)	2 x (2 + shoulder)
Gradient	none	none
Weather	Clear	Clear
Lanes used by traffic:		
eastbound	right & emergency	right & left
westbound	left & right	left
Jammed	eastbound & westbound	westbound
imaged area	400 m	400 m
Remarks	just east of tunnel	just east of merging

**Figure 3.3: Incident location near Gorinchem (right=east)**

junction Gorinchem (see figure 3.3); there is no gradient. Traffic had to merge twice in several hundreds of meters: the regular merging of traffic from the two freeways and the merging from two to one lane. The traffic demand in the eastbound direction was not sufficiently high to cause a traffic jam.

### 3.3.3 Description of the data

The video was taken from a helicopter. This means that the image area is not stable. Due to this movement of the recorded area we could not use the full length of the recorded road image. Around 200-300 meters of the road remains that overlapped all recorded image and the full length of the available image is used to track the vehicles. This means that typically, a vehicle appears on 200 different images before it reaches the end of the analysed stretch.

The operations of several hundreds of vehicles were recorded on video. Table 3.3 states how many vehicles were used in the further analysis. It also shows there are differences in the number of vehicles used for further analysis depending on the type of analysis.

**Table 3.3: The number of vehicles used for computing average speeds and the number of leader-follower pairs used for determining headways and calibrating parameters of a car-following model**

	Eastbd right ln	Eastbd left ln	Westbd right ln	Westbd left ln
Apeldoorn	644 headways, left + right		76 car-following 366 headways	123 car-following 331 headways
Gorinchem	378 speed	90 speed	-	123 car-following 402 headways

For car-following analysis, for instance, one needs the complete trajectories of a leader-follower pair, and therefore the pairs which change over time (lane changing manouevres) are discarded for this analysis. However, for the headways one only needs to know the headway of a leader-following pair at one moment in time. This excludes less leader-follower pairs.

Due to the resolution of the images, the trajectories have to be smoothed before they can be analysed (Thiemann et al., 2008). To this end, the same filtering is applied as described by Ossen et al. (2006) and Toledo et al. (2007) where the position of the vehicle at a certain moment in time is replaced by a point calculated by a local regression using weighted linear least squares and a polynomial model. The weight is given by a tricube function of distance in time, with a typical decrease of 1 second. The positions at times which differ more than 2 seconds from time  $t$  are not considered at all.

### 3.4 Data analysis approach

Both the behaviour of the drivers (microscopic) and the characteristics of the flow (macroscopic) have been studied. At the microscopic level, the chosen time headway at one location and the dynamic car-following characteristics are studied, which is described in section 3.4.1. Section 3.4.2 provides the details of the process of fitting car-following models. It also discusses explicitly the robustness thereof. Section 3.4.3 discusses the way the macroscopic properties of traffic are processed. The results are stated in section 3.5 which ordered the same way as this section, which therefore can be read together with this section.

#### 3.4.1 Microscopic flow characteristics

The basis for all changes in traffic flow is the change in human behaviour when driving. Differences in capacities can be derived from differences in headways of individual drivers, which are likely to be caused by differences in dynamical car-following behaviour. This section presents how insights at this microscopic level can be gained. It is divided in

three subsections, of which the first discusses the headways. The results are presented in the first subsection of section 3.5.1. The second subsection discusses how a car-following model can be calibrated. The third subsection finally discusses the robustness of the parameter values found in the calibrating process. Results of the calibrating process are written in the second subsection of 3.5.2.

### **Headways**

The trajectory data of the vehicles contains, by interpolation, the moment a vehicle passes a predetermined point, which will be called a “virtual detector”. For all vehicles for which the trajectories are reconstructed, passing times are collected. In case a leader-follower combination is tracked, the difference of the passing times gives the headway. This headway could be analysed at any point along the trajectory. The most interesting point was chosen, which is the point that forms the bottleneck, found by the moment that cars start accelerating again. Note that this point is a fixed point and does not vary over time.

We have analysed the distribution of headways (in the congested directions) in order to check whether the headways around an incident are larger compared to normal driving. Apart from the average value, the distribution of the headways is interesting.

The headway distribution only is relevant if a bottleneck is present in the monitored area and people are bound to a lane. This is the case for the following three lanes:

- Apeldoorn westbound right lane
- Apeldoorn westbound left lane
- Gorinchem westbound left lane

It is only relevant to analyse headways collected at a bottleneck itself. For instance upstream of a bottleneck every vehicle could be in car-following mode, since there is congestion, but nevertheless the average headway is determined downstream at the bottleneck. We did not have access to remote sensing data of individual drivers’ behaviour at a comparable bottleneck or, in fact, at any bottleneck. However, the median headway for traffic flowing out of a queue is the inverse of the median queue discharge rate, which is the maximum number of vehicles per unit time flowing out of the queue.

The queue discharge rate can be determined from loop detector data if both free flow and congested conditions occur. The road at both incident sites is equipped with loop detectors. At the Gorinchem site, there is sometimes congestion from downstream which spills back. In contrast, at the Apeldoorn site there is no congestion at all. Therefore, we have to use another study that gives the queue discharge rate and the spread thereof for Dutch motorways. There is no reason why the queue discharge rate at the Apeldoorn site would differ from those at other Dutch motorways.

At the Apeldoorn site in the eastbound direction, drivers do not drive consequently within one lane. People drive at the right lane, at the emergency lane or sometimes even between these lanes. Directly after they have passed the incident location, they change lanes. Consequently, a leader-follower couple could not be identified for a long time and the

follower is likely to keep adjusting his headway to changing leaders. An estimate of the car-following model on this fluctuating behaviour would give unreliable outcomes, so we have not used data of Apeldoorn eastbound. For the Gorinchem eastbound direction there was not enough traffic to make a car-following analysis useful. Only a small fraction of the travellers was in car-following mode.

### 3.4.2 Calibrating a car-following model

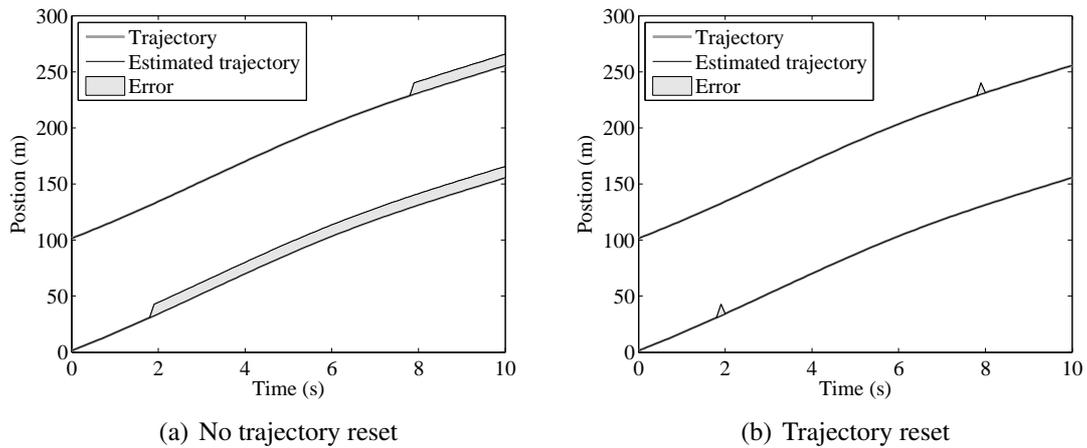
The longitudinal driving behaviour at the location of the bottleneck determines the headway distribution and capacity, given that there are no lane changes. This driving behaviour can be described by a car-following model. To quantify the driving processes, a car-following model is fitted to the vehicles passing an incident location. To find the differences with normal driving, the parameters describing the operational driving are compared with the same parameters for driving under similar conditions without an incident. This section will explain the choice of the model and the fitting process.

The goal of fitting a model is that the fitted model parameters can be compared to the fitted parameters under normal, non-incident conditions. However, it is impossible to monitor the same drivers at the same site under the same conditions without an incident. The best option is to compare the parameters with parameters from a similar drivers' population at a similar site under similar driving conditions. However, it is impossible to collect comparable data from the incident sites since they are usually uncongested. Therefore, we compared the model parameters with parameters obtained from an observation at other, but comparable sites in terms of road geometry, driving conditions and drivers' population, which is the Everdingen and Waalhaven site as presented in Ossen (2008).

For describing the longitudinal driving behaviour around the incident location, a model with both a car-following part and a free driving part is required, since upstream of the incident traffic is congested and downstream of the incident it is not congested. Furthermore, it is expected that the reaction time of drivers plays an important role. Therefore, the car-following model we choose incorporates the reaction time. Apart from these considerations, it is best to choose a model which fits well on Dutch drivers. Combining the requirement with the fit results from Ossen (2008), we chose for the car-following model proposed by Tampère (2004).

The model parameters will be compared with model parameters of the same model for driving under normal conditions. Therefore, it is not required that the Tampère model is the best fitting model. However, it is required that the model fits reasonably well on a vehicle. This is tested for each vehicle. In fact, for each vehicle it is tested whether the fitted parameters are sensitive enough. If not, the parameter is not considered in the further analysis. It has to be checked whether not too much data is ignored. A detailed description of the process of testing is given in the next section.

In the Tampère model the driver has a desired distance  $s^*(v)$ , which is linearly dependent on its own speed  $v$ . The acceleration  $\dot{v}$  is the minimum of two terms: (1) A free driving term and (2) A car-following term. In this car-following term, the desired acceleration is



**Figure 3.4: The accumulation of error**

in fact a linear combination of two stimuli. One stimulus is the speed difference with the predecessor,  $\Delta v$ . The other stimulus in the same term is the difference between the actual distance with the predecessor,  $s$ , and the desired distance,  $s^*(v)$ .

The acceleration is at maximum a constant times the difference between the desired or free speed,  $v^*$ , and the actual speed. A reaction time is taken into account by delaying the response to the stimuli by  $\tau_r$ . In mathematical terms, this model is expressed as follows:

$$\begin{aligned}
 s^*(v) &= c_4 + c_5 v \\
 \dot{v}(t + \tau_r) &= \min \left\{ \begin{array}{l} c_1 \Delta v(t) + c_2 (s(t) - s^*(v)) \\ c_3 (v^* - v(t)) \end{array} \right\} \quad (3.1)
 \end{aligned}$$

This paragraph explains the way the fitting process is implemented, which is similar to the one described by Ossen et al. (2006). The goal of the optimisation was to minimize an error function  $\varepsilon$ , which consisted of the error in predicted location and the error in the predicted speed. For each time step the next position and speed were calculated based on the measured trajectories of the follower and the leader up to that moment. If one would, in contrast, choose to predict the complete trajectory based on the trajectory of the leader and only the initial position and speed of the leader, an error at the beginning of the trajectory would accumulate and thus the errors at the beginning of the trajectory would get a larger weight. If the measured trajectory is used, like in our case, an error at the beginning does not accumulate and therefore is not more important than an error at the end of a trajectory. This is shown graphically in figure 3.4

The error is a function of trajectories of a leader-follower pair, which is fixed, a fixed reaction time and a set of 5 parameters. Using a constrained optimisation algorithm with one initial point for each of the criteria, we find the minimum error in this 5 dimensional space for each leader-follower pair. The parameters have been bound to a feasible range which differed per variable, and a typical value has to be put in for the initial computation step.

This has been done for different reaction times in a certain range. For each reaction time an optimal parameters set  $c_1 - c_5$  has been found and a corresponding remaining error. The reaction time with the minimum error has been selected as best and consequently the corresponding parameters  $c_1 - c_5$  as well as the corresponding reaction time are considered to be the best parameters to describe the follower's behaviour with the Tampère model.

### Sensitivity of the fitting parameters

Even in the best fit, not all parameters of each vehicle could reliably be estimated. Therefore, equation 3.2 below shows how to calculate the sensitivity  $S$  of the error for each of the 6 parameters. Note that in this equation,  $\varepsilon^n$  is the error function for the  $n^{\text{th}}$  vehicle and  $c_i^n$  is any of the 6 parameters  $c_1 - c_5$  or the reaction time; these six parameters together are called  $\mathcal{C}^n$  and the optimal value of a parameter or a parameter set is indicated with an asterisk. The optimal set of parameters for the  $n^{\text{th}}$  vehicle,  $\mathcal{C}^{n*}$  is found if the derivative to each of the variables is 0. How much the error fluctuates by a small change in the parameter is expressed in the second derivative. However, this shows the variation of the error when varying the parameter by one unit. Since the scales of the units differ per parameter, we can normalise this to a dimensionless parameter scale by multiplying the second derivative by the square of the value of the parameter. In order to obtain a scale without units, we divide the sensitivity by the error value for the best parameter set, which gives the relative sensitivity of the parameters.

$$S_{c_i}^n = \frac{(c_i^{n*})^2}{\varepsilon^n(\mathcal{C}^{n*})} \frac{\partial^2 \varepsilon^n}{\partial c_i^2}(\mathcal{C}^{n*}) \quad (3.2)$$

The sensitivity shows how reliably a parameter can be estimated; in fact, a sensitivity is comparable with the inverse of a variance. We account for the sensitivity in the further analysis of the parameter values by introducing a weight factor  $w$  which increases for increasing sensitivity, as proposed by Hoogendoorn and Van Lint (2007). This weight is for instance used to calculate a weighted average of a parameter over all vehicles. Additionally, parameters which could not be estimated reliably enough are neglected. If  $S$  is 1, the variation of the error is as large as the error when the parameter is twice its optimal value. If the estimation of a parameter is less sensitive than that, it is not considered at all. These requirements are captured in the following weight function.

$$w_{c_i}^n = \begin{cases} \log S_{c_i}^n & S_{c_i}^n > 1 \\ 0 & S_{c_i}^n \leq 1 \end{cases} \quad (3.3)$$

For approximately 40% of the vehicles all six parameters can be reliably calibrated, i.e. for approximately 40% of the vehicles all six calibrated parameters have a sensitivity larger than 1. For the other vehicles it differs which parameters are sensitive; for 1 vehicle none of the parameters could not be reliably calibrated, for the others only some of the parameters could be calibrated. The calibration of  $c_4$  and  $c_5$  was the most reliable. 70% of all parameters could be reliably estimated.

We have chosen not to fit the desired speed. The sensitivity of the trajectory for the value of  $v^*$ , is low and therefore no reliable estimate for  $v^*$  can be made; the only requirement for a plausible fit is that  $v^*$  is not too low (Ossen, 2008). We therefore have fixed  $v^*$  at a value near the speed limit, 30 m/s.

### 3.4.3 Macroscopic flow characteristics

This section discusses the analyses for the macroscopic flow variables. The results of these analyses can be found in section 3.5.2.

#### Average speed

A change in microscopic driving behaviour possibly shows a lower speed. For each point along the road, the average speed of the vehicles passing at that point has been calculated. We have analysed the profile of the average speed along the road for both directions and have compared the right and the left lane. It is expected that in the left lane, the speed difference is larger (it is closer to the incident) and the acceleration is higher, because in the left lane are no trucks (which accelerate slower) and the left lane is probably occupied by the more aggressive drivers. We performed this analysis for the two locations where two lanes are available: Apeldoorn westbound and Gorinchem eastbound.

#### Queue discharge rate

The headways give an indication of the reciprocal value of the queue discharge rate. However, in determining the headways we only consider the vehicles that could be tracked over the whole stretch. To obtain a more reliable estimate of the queue discharge rate and the fluctuations thereof, for all passing vehicles the passing time is recorded. This is done using the video data; also the counts of the loop detectors can be used, but they do not distinguish between cars and trucks. From these passing times we calculated flows by aggregating them over time and, if relevant, over the carriageway.

The actual queue discharge rate is stochastic while its distribution depends on external factors (such as weather, road geometry) and individual characteristics of the drivers. Usually, it is characterised by the median value: the queue discharge rate that is exceeded in half of the cases. For one measurement site, the external conditions are fixed. The advantage of using the median over the mean is that it is less dependent on a long “tail” of the distribution. Moreover, the inverse of the median flow rate is the median headway, and this does not hold for the means.

The variability of the queue discharge rate indicates the extent of the inter-driver differences. Not all aggregation intervals are suitable to estimate the queue discharge rate. For instance, there should be a queue of cars waiting to pass and the flow should be uninterrupted which is not always the case at the incident location near Gorinchem. If there is no continuous queue discharge in some aggregation interval, this interval should be ignored in estimating the queue discharge rate. Therefore, the amount of data that should be ignored increases with the aggregation time. For this reason, we have taken a relatively short aggregation time of 30 seconds. Using shorter intervals increases the spread of the measurements, but will not change the median value.

Using 30 seconds intervals, there are 35-40 aggregation intervals for the location near Apeldoorn (the number of usable intervals depends on the lane and direction) and 45 for the location near Gorinchem. To obtain a single value for the flow, all passing vehicles have been converted to passenger car equivalents. Hence, the passing of a truck is counted as 1.5 passenger car as prescribed by both the Highway Capacity Manual (Transportation Research Board, 2000) and its Dutch equivalent (Dutch Road Authority, 2007c). The flows have been converted in this way to passenger car units per hour lane, pcu/lane-h.

Only locations with an bottleneck are useful for determining the outflow capacity. These are:

- Apeldoorn eastbound - carriageway
- Apeldoorn westbound - right lane
- Apeldoorn westbound - left lane
- Apeldoorn westbound - carriageway
- Gorinchem westbound – left lane

The first value is for the amount of traffic that passes the incident near Apeldoorn, the eastbound direction. Adding the flows of the emergency and the right lane (used intermittently) gives information about the realised queue outflow discharge rate if these two lanes are available. In this way, we have not obtained a capacity value for each lane separately. However, the value we have got for the carriageway is a result of drivers' behaviour to pass the incident location. The resulting flow values have been divided by 2 to get an outflow capacity per lane.

Traffic in the non-incident direction used both lanes continuously. Therefore, we could calculate a capacity value for each of the 2 lanes, as well as the average outflow capacity (distribution) per lane. The fifth capacity value is the capacity of the remaining lane passing the incident at the incident location near Gorinchem.

Let us finally remark that the queue discharge rate is lower than free flow capacity, the maximum number of passenger car equivalents per unit of time that can pass a cross-section of a road in free flow, which is usually obtained before congestion sets in. This phenomenon is called the capacity drop. It is described extensively in literature and estimations for the reduction vary, but are typically around 10% (Cassidy and Bertini, 1999; Chung et al., 2007; Dijkster et al., 1997; Hall and Agyemang-Duah, 1991). In the analyses we will only use the queue discharge rate.

### 3.5 Results

The data have been analysed in the way described in the last section. This section presents the results and is divided into two subsections. The first one presents the results on the level of individual cars, the microscopic flow characteristics, and the second one presents the findings on an aggregated level, the macroscopic flow characteristics.

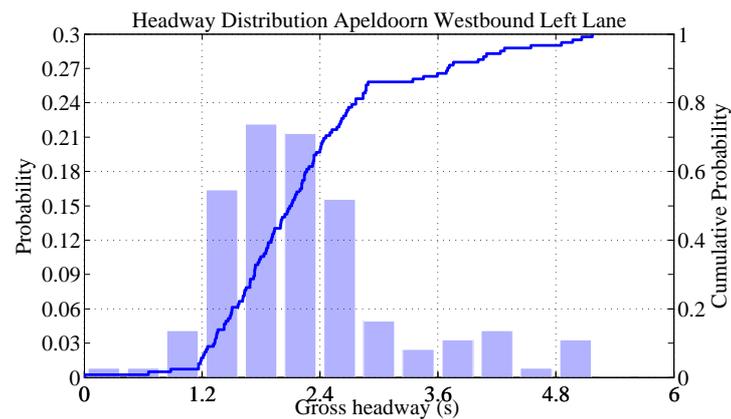


Figure 3.5: The headway distribution for Apeldoorn westbound left lane.

Table 3.4: The median headways and standard deviation

	Apeldoorn westbound	Apeldoorn westbound	Gorinchem westbound
Lane	Right	Left	Left
Median headway	3.2 s	2.1 s	3.7 s
Standard dev.	2.7 s	0.9 s	2.4 s
Normal	1.9 s	1.9 s	2.0 s

### 3.5.1 Microscopic properties

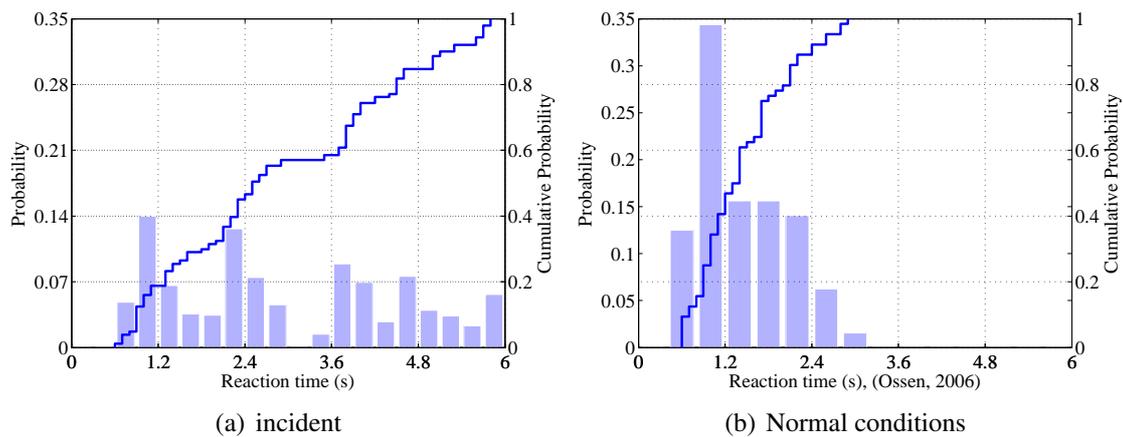
Let us first consider the microscopic characteristics determined from the helicopter data.

#### Headways

Analysing the flows at the level of detail of individual vehicles, we could obtain the headway distribution on a cross-section. As an example, consider the distribution of the headways on a cross-section at left lane of Apeldoorn westbound, shown in figure 3.5. Both the cumulative curve and the histogram are shown, which respectively show what part of the headways is lower than a certain value, and the fraction of vehicles having a headway according to the bin size. The majority has a headway of around 2 seconds, derived from the highest bar, while another part has a headway at around 4 seconds. Similar distributions are found for the other locations; the data of all locations are summarised in table 3.4.

Both sites are no bottlenecks under normal conditions, so an alternative method was used to derive the queue discharge rate and from that the minimum headways in a bottleneck. An intensive research studying 2 years of data showed that the median queue discharge rate of Dutch motorways is around 1875 veh/lane-h (Heikoop et al., 2007) which implies a median queue discharge headway for normal traffic of 1.9 seconds in the Netherlands, which is used as reference for Apeldoorn site.

Also the Gorinchem site is no bottleneck under normal conditions, but there are queues spilling back to the site. Therefore, both the congested branch and the free flow branch of the fundamental diagram could be fitted on measurements from loop detector data using



**Figure 3.6: The distribution of reaction times**

the functions proposed by Wu (2002). The intersection of these two branches gives a queue discharge rate of 1760 veh/lane-h, which implies a median queue outflow headway of 2.0 s.

The significance of the differences in a distribution can be shown using a Kolmogorov-Smirnov test. There is a significant difference if for any value of the variable the difference between the two cumulative distributions is larger than the Kolmogorov distance (Chakravarti et al., 1967). In this case, however, only one level of the cumulative distribution for the reference situation is known, namely that 50% of the headways is lower than 1.9 respectively 2.0 seconds for Apeldoorn or Gorinchem. These points are a single point at each of the respective distribution functions. The distance of this point of the reference distribution function to the distribution function of the headways at the respective incident site is now calculated. This distance is larger than the Kolmogorov distance for a significance level of 5%, which means that the two distributions are significantly different. This test was repeated for all lanes at the incident sites, which showed that the headways at each site differ significantly from the headways under normal driving. Therefore, hypothesis 1 on page 18 is accepted.

### Reaction times

Figure 3.6 shows the distribution of the fitted reaction times over the different drivers, in a graph similar to figure 3.5; the values are summarised in table 3.5. Note that the reaction time  $\tau_r$  should not be interpreted as the time one physically needs to process an input.  $\tau_r$  can be interpreted as the time a driver takes before an input has effect on his driving operations. For the reference situation in figure 3.6b, reaction times, presented in Ossen et al. (2006), are shown without a weighting factor, but for the incident measurements an extra weight is used for fitted parameters which are more reliable according to equation 3.3. In fact, the line in figure 3.6a shows which part of the weight of the fitted reaction time estimates is lower than a value and the bars show how much weight of the reaction time estimates is located in each of the bins.

The reaction times are distributed similar to the headways. Similarly, there are two values

**Table 3.5: The reaction times**

	Apeldoorn westbnd	Apeldoorn westbnd	Gorinchem eastbnd	Netherlands nrml
Lane	Right	Left	Right	Average
Mean	3.9 s	2.9 s	3.8 s	1.3 s
Stdev.	1.4 s	1.6 s	1.3 s	1.0 s

around which the reaction times centres, 2 seconds and 5 seconds, as can be seen for instance for the left lane of Apeldoorn westbound (figure 3.6a). A Kolmogorov-Smirnov test is performed on each of the incident reaction time distributions and it shows that these distributions all differ significantly at a significance level of 5% from the reaction times at a non-incident location. In fact, the reaction time at the incident sites is time is much larger and therefore hypothesis 2 on page 3.2 is accepted.

### 3.5.2 Macroscopic properties

Let us now discuss the main macroscopic properties of the flow during the incident and compare it with normal circumstances.

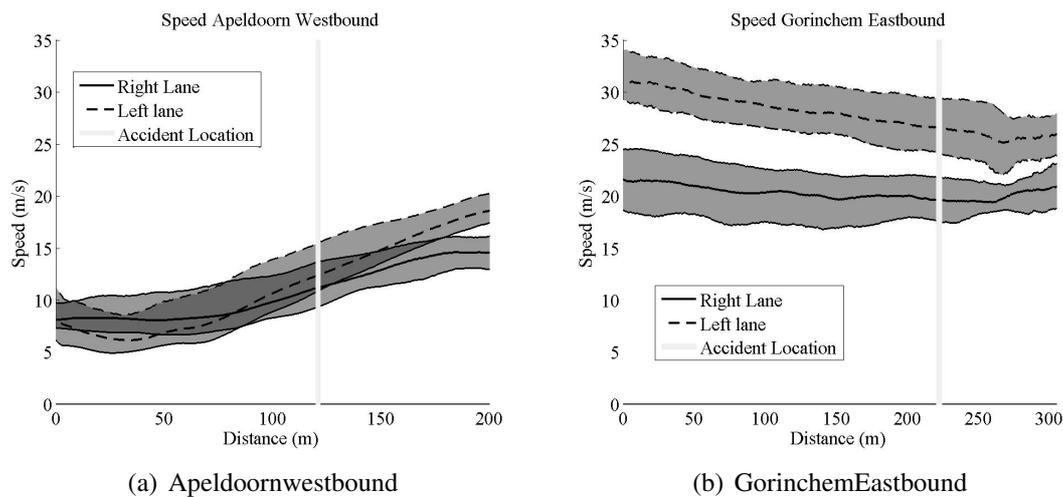
#### Average speed

The speed profile for Apeldoorn and Gorinchem is shown in figure 3.7. The bold line is the median speed (of the cars) for all positions along the road. It also shows the interval bounds (60 % of the speeds lies within the gray area).

At the Apeldoorn site the average speed drops when drivers approach the incident site. After people have passed a certain point, they start accelerating again. This point appears to be the point providing the best view on the incident location. At this point, the average speed in the right lane is actually higher than in the left lane; in fact, the minimum speeds are 7.8 m/s for the right lane and 6.2 m/s in the left lane. However, the vehicles in the left lane accelerate faster and they have a speed of 18.9 m/s when they reach the point at which they drive out of the picture, whereas the average speed in right lane is just 14.6 m/s at this point. In order to test the difference in the speed reduction in the right and left lane, we used *t*-test, which yields a *P* value of  $5 \cdot 10^{-6}$  indicating a significant difference.

In Gorinchem dataset there is no queue in the non-incident, eastbound direction. The speeds (figure 3.7b) show that people nevertheless reduce their speed. Drivers have passed the most eastbound wreck at 220 meters. The minimum of the average speed lies at a point a little downstream, at around 250 meters. This is possibly caused by looking at the traffic jam (including the merging) at the other side. In addition, there is less incentive to accelerate quickly. Namely, contrary to the Apeldoorn site, the drivers do not drive in congestion. Therefore, they might not have been aware for a second that their speed has been reduced.

Only the start of the acceleration lies within the area that is captured on video. At this location, we therefore focus on the speed reduction. The average speed in the left lane



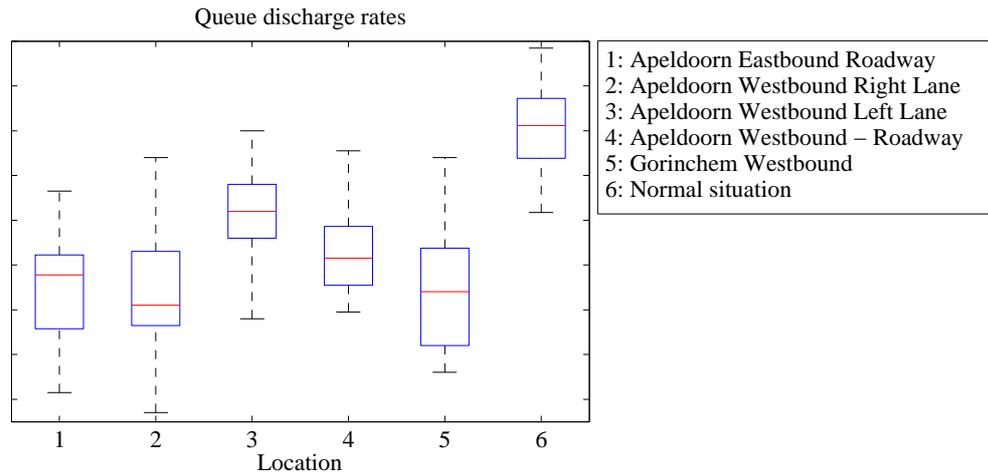
**Figure 3.7: The speed profile along the road as a function of the distance along the road**

drops from 31.1 m/s to 24.8 m/s, which is a difference of 6.3 m/s. In contrast, the average speed in the right lane reduces by just 2.6 m/s, from 22.0 m/s to 19.4 m/s. A t-test (for the explanation of the statistical tests and values: see appendix B) has shown that this speed reduction is significantly larger in the left lane than in the right lane with a  $P$  value of  $1 \cdot 10^{-4}$ . This difference in free speed can partially be explained by the European law which requires the drivers, under non-congested conditions, to keep the right lane when possible, and to overtake at the left lane. Therefore, since the traffic is still under free flow conditions, the right lane has a lower speed than the left lane. Additionally, 25% of the vehicles in the right lane are trucks with a lower speed limit, whereas there are no trucks in the left lane. To confirm the idea that trucks influence the magnitude of the speed reduction, we analysed the difference in speed between trucks and passenger cars in the right lane, we have found that trucks keep a more constant speed. Passenger cars enter the area at higher speed than trucks and decelerate to a lower speed than the trucks.

So, at the location of the incident, the average speed of the vehicles drops, which confirms hypothesis 3 on page 18. Moreover, the average speed in the left lane decreases more than the average speed in the right lane which can be explained in three ways. Firstly, drivers in the left lane are probably more distracted since they are closer to the incident. Secondly, there could be a psychological difference between drivers in the right and the left lane. More aggressive drivers are more likely to drive in the left lane. Thirdly, drivers of less powerful cars and trucks, having a slower acceleration, probably stay in the right lane.

### Queue discharge rates

Figure 3.8 is a boxplot of the queue discharge flows. The middle line indicates the median flow value; the box is made of horizontal lines at the 25% and 75% percentile. The whiskers (at the outside) give the total range of the values. In addition, the queue discharge rates of the incident locations are stated in table 3.6. As normal queue discharge rates we use the values as found in section 3.5.1, namely 1875 veh/h-lane for the Apeldoorn site and 1760 veh/h-lane for the Gorinchem site.



**Figure 3.8: The queue discharge rates for the different locations**

As expected, the queue discharge flow values for the locations with an incident drop to values under the usual ones based on the remaining lanes (table 3.6). The reference queue discharge rates are in vehicles/lane-h, rather than pcu/lane-h. We need to be sure that we do not wrongly indicate the values as (significantly) different due to a wrongly chosen pcu-value. Hence, we need to be sure not to overestimate the truck share in the reference values, and therefore we have conservatively processed the value as was the truck share 0%. This means that, by converting the values to pcu-values, the reference estimate is certainly not too high. Z-tests have shown that for each of the 5 locations the queue discharge rate is significantly different (5% significance), from the normal queue discharge rates. The stated expected capacity is the reference free flow capacity for the number of lanes that is open. So, for example, in the case of Gorinchem, the table 3.6 states the reference capacity of one lane. Thus, the reduction of capacity cannot be explained by the reduction of number of lanes. Since we compared outflow values with the outflow capacities, it is also not the capacity drop. Thus, hypothesis 4 on page 18 could be accepted.

In the first case, Apeldoorn eastbound, one of the two remaining lanes is the emergency lane. This has two consequences for the estimation of the capacity reduction. First, the reference capacity is probably lower than the capacity for two full-width lanes. Secondly, as described in section 3.2, the use of both lanes at this location varies in time: in some periods the shoulder lane is used, in some periods the right lane is used and in some periods, both of them are used. If they were using both lanes continuously, the flow would probably have been higher. However, the data represents the drivers' behaviour around the incident, including their lane choice.

This reduction may be caused by “rubbernecking”, the fact that people are distracted by watching what has happened. Another possibility is that people drive extra carefully since there are people working at the roadway. Compared to the normal queue discharge rate, also stated in table 3.6, the queue discharge rate reduces by 25-40%. Other publications as by Qin and Smith (2001), or the Highway Capacity Manual (Transportation Research Board, 2000) show lower reductions, even when compared to the higher free flow capa-

**Table 3.6: Capacity values for different locations**

Location	Lanes	Median pcu/lane-h	St dev pcu/lane-h	Expected without rubber- necking pcu/lane-h	Percent of queue discharge rate remaining
Apeldoorn eastbound carriageway	2	1170	239	1875	59%
Apeldoorn westbound right lane	1	1020	249	1875	58%
Apeldoorn westbound left lane	1	1440	195	1875	76%
Apeldoorn westbound carriageway	2	1230	163	1875	66%
Gorinchem westbound	1	1080	326	1760	61%

city. The confidence bounds for the values in these references are unclear, and therefore it is impossible to tell whether the difference is significant. A possible explanation for a larger reduction found in this study is the high queue discharge rate under normal conditions in the Netherlands. In case the outflow under incident conditions is the same in the Netherlands as in other countries, there is a larger reduction of capacity.

It is remarkable that the resulting maximum flow rates at the incident near Apeldoorn are about the same for the eastbound and the westbound direction, whereas in the eastbound direction one lane is blocked. Drivers will use the emergency lane from time to time. Blocking of lanes and looking at the incident causes the same reduction of flow at the carriageway with the incident as the rubbernecking on the carriageway for the opposite direction.

### 3.6 Conclusions and discussion

This contribution proposes a new conceptual model for driving in near an incident. The model states that a shift of the driver's attention leads to a different traffic characteristics. To prove the model, microscopic data of traffic operations at two incident sites were used. We accepted the following hypotheses:

1. Drivers choose a different headway in a bottleneck caused by an incident compared to a bottleneck in normal traffic and the mean headway is larger.
2. Drivers have a different reaction time at an incident site compared to normal conditions and the mean reaction time is larger.
3. People reduce speed in view of an incident.
4. The queue discharge rate at the incident site is lower than the outflow capacity under normal conditions for the same roadway geometry.

Although it does not prove the theory, the findings support the theory that drivers shift their attention towards the incident when passing by the incident site. We found evidence implying that there may be two groups of drivers, those who have a slightly changed driving behaviour and a smaller group with a considerably changed driving behaviour, shown by a large headway and reaction time.

These behavioural changes lead to a reduction of the queue discharge rate. At the incident site, the queue discharge rate per lane is 60-75% of normal queue discharge rate per lane. This holds for both directions, the incident direction and the incident direction. This reduction is additional to the reduction of the queue outflow rate due to the reduction of the number of available lanes.

For practical applications a simulation model for traffic flows around incidents would be useful. Every simulation model, both microscopic and macroscopic, needs calibration and validation. It is difficult to calibrate a simulation model for an incident situation, since there are only few incident situations, which are all different. Although the findings presented here are based on observations of many drivers, drivers at another incident might drive differently.

This chapter only showed the changes in vehicles movement. To know the psychological causes (distraction, increase safety), future research should investigate the psychological effects in a more controlled situation, as for instance a driving simulator. Other future work includes the lane choice behaviour around accidents.



## Chapter 4

# Motorway queue discharge rate reductions at incidents sites

---

The last chapter described how the individual driving behaviour changes due to an incident. Consequently, the capacity and the queue discharge rate change. The traffic situation changes only significantly if there is congestion, so a detailed knowledge of the queue discharge rate is important. It can improve for instance the traffic prediction and thereby improve delay information or routing advice. Therefore, this study determines the queue discharge rate for many incident locations during an incident situation and these are compared with the queue discharge rate at the same location in normal conditions. Ninety incidents meet the requirements to apply the proposed methodology. It is found in case a driving lane is blocked, the queue discharge rate for each available lane is reduced by 50%. In case the driving lanes are open but there is a distraction of an incident at the emergency lane or on the roadway for the opposite direction, the queue discharge rate is reduced by 30%. This paper is an edited version of: Knoop, V. L., Hoogendoorn, S. P. and Adams, K. Incidents at Motorways: Capacity Reductions and the Effects of Incident Management. *European Journal of Transportation and Infrastructure Research*. In press.

---

### 4.1 Introduction

The amount of delay due to a traffic incident depends on the demand and the supply on the road. In fact, if the supply (the remaining capacity) reduces to a level lower than the demand, it causes a traffic jam. Therefore, to compute the delays it is essential to know the capacity of the road at the location of an incident. Obviously, one reason for a reduced capacity at the incident site is the reduction of the number of available lanes. We hypothesize that, in addition, the capacity reduces because the remaining lanes are used less efficient because people drive differently.

This contribution investigates the queue discharge rate (the maximum throughput when drivers drive out of congestion) at the position of an incident when a jam has occurred rather than the free flow capacity. We consider both the incident-direction and the non-incident direction (i.e., the direction in which no physical obstruction occurs). Due to limited data availability, we restrict the study to freeways with 3 lanes in each direction. This has the advantage that the road layout is the same for all measurements and the results can be directly compared.

Incidents are categorized in two groups. One of these is an accident situation where at least one of the lanes normally available for traffic – such lanes will be called “driving lanes” – is blocked. The other the situation where a car has broken down and is stopped at the shoulder lane of the freeway. With shoulder lane we mean a paved lane at the right of the driving lanes which is only meant for emergencies and which is not one of the driving lanes. In the Netherlands, there is usually no shoulder lane at the left of the driving lanes in case there are less than 4 driving lanes. In both situations, a change in driving behaviour can reduce the capacity of the remaining lanes. This study will show the magnitude of this behavioural effect.

The objective of this chapter is to describe the maximum flows at incidents for traffic in both the incident-direction (with a partially blocked roadway) and in the non-incident direction. We expect a lower capacity at that point due to distraction also at the other side of the guardrail, an effect which is called “rubbernecking” as consequence of the microscopic changes described in section 3.5.1. To this end, data from a large set of incidents is used. Ninety incidents out of this set were suitable to use for the computation of the capacity reduction. These incident capacities, for both directions, have never been studied on this scale before. There are handbook values, but those often give just a number for the incident capacities. This chapter adds the methodology that is used. Another contribution of the chapter is that it adds a distribution of the capacities which is possible because of the large number of incident capacities analysed. Section 4.2 will show the available literature on incident capacity; there, the gap is explained in more detail.

The capacity values are used to quantify the effects of shortening the incident handling time. The chapter only states how large the effects of a time reduction in a specific phase are which can be used to assess the impact of specific measures and specific measures, are not named or assessed within this chapter because there are many measures. However, because the findings are general, the findings proposed here can for instance be used in assessing the effectiveness of *new* measures which do not exist yet. For each specific measure, the costs will vary. In this way, the method presented can be used to get an insight of the benefits, which then have to be weighted against the costs.

The remainder of the chapter is set up as follows. In section 4.2, an overview is given of the research that has been carried out on the incident capacity reductions over the past years, both in the US and in the Netherlands, since the values can differ for different countries. It also indicates how this study fits into the existing literature. Then, we discuss which data are used for analyzing flows. Section 4.4 discusses how the data were processed. The actual reductions of queue discharge rates in case of incidents are then

shown in section 4.5. Then, an example is presented which shows the practical relevance of this research. Finally, section 4.7 gives the discussion and the concluding remarks.

## 4.2 Literature review

Goolsby (1971) published an early overview of incident reductions due to incidents. The findings of his work and all other work mentioned in this section are stated in table 4.1. In nearly 40 years the traffic operations are likely to have changed. Apart from that, he determined the reduction compared to a reference capacity which he assumed the same for all locations.

Handbook values are useful for practitioners. However, often a description of the origin of the values lacks. This is also the case for an early publication of the Transportation Research Board by Blumentritt et al. (1981). Nowadays, the Highway Capacity Manual (Transportation Research Board, 2000) is the most referred source. Both values can be found in table 4.1.

The capacity values at incident sites might be dependent on the country since the capacities also vary per country. For the Netherlands there is also a handbook comparable to the Highway Capacity Manual, *Nieuwe Ontwerprichtlijn Autosnelwegen* (especially for motorways), (Dutch Road Authority, 2007c). It does not state a capacity for incident sites separately. However, there is an entry on the capacity for partially closed lanes (used mainly for road works), which is given in table 4.1. In general, the road capacity in a work zone is also lower (e.g., Kim et al. (2001); Dixon et al. (1996)). That might lead to a conclusion that work zones and incident situations are comparable. However, Al-Kaisy and Hall (2001), and Heaslip et al. (2008) argue that driver familiarity also is important in the capacity of the work zone. Clearly, drivers can get used to a work zone, but drivers are never familiar to a incident situation. Another difference between an incident situation and a work zone is that there are proper markings for the work zones which increase the capacity which lack at incident sites.

Recently, two research projects are carried out in the Netherlands to determine the capacity in incident situations. Schrijver et al. (2006) use the Highway Capacity Manual as starting point. They modify these values based on experts' opinions. One of the changes is that they state that the queue discharge rate is 80% of the free flow capacity. The phenomenon that the queue discharge rate is lower than the free flow capacity is well known and discussed for instance by Hall and Agyemang-Duah (1991), Dijkster et al. (1997), Cassidy and Bertini (1999) and Chung et al. (2007). Additionally, Schrijver et al. (2006) distinguish between different phases in the incident. The main distinction between the phases is the presence of emergency services (e.g. police, ambulance). They assume that the presence of workers on the roadway halves the remaining capacity.

The other project is reported by Van Toorenburg and Nijenhuis (2007). They use traffic data to check the values used in the report by Schrijver et al. (2006). They find a considerable difference for the situation in which there is an accident at the shoulder lane: 45%

**Table 4.1: Overview of the remaining capacities**

Type of blocking	Shoulder	1 out of 3 blocked	2 out of 3
Goolsby (1971)	0.67	0.50	0.21
Blumentritt et al. (1981)	0.84	0.53	0.22
Transportation Research Board (2000)	0.83	0.49	0.17
Schrijver et al. (2006)	0.77	0.35 <sup>1</sup>	0.17
Dutch Road Authority (2007c)	-	0.36	0.17
Smith et al. (2003)	-	0.37	0.27
Van Toorenborg and Nijenhuis (2007)	0.45	0.38 <sup>1</sup>	-

<sup>1</sup> There is no distinction for different lane closures: the value for an accident is given.

of the capacity is used, whereas Schrijver et al. (2006) assume 77% is still available. In their report, Van Toorenborg and Nijenhuis (2007) do not distinguish between different phases in the incident or different numbers of lanes closed.

Regarding methodology, Smith et al. (2003) describe best how capacity values are found, with a more detailed description available (Qin and Smith, 2001). Contrary to Goolsby (1971) and Blumentritt et al. (1981), Smith et al. (2003) describe the capacity as stochastic variable. In their research, macroscopic data of loop detectors is used to determine the maximum flow out of a queue. Thereby, their study describes the queue discharge rate, like all other quoted studies. To the best of our knowledge there has been no research to the maximum possible flow in free flow state around an incident location. Our research also shows the queue discharge rate values. Section 4.3 will describe the methodology we applied.

Not much is known about the “rubbernecking effects”, i.e. the reduction of capacity in the non-incident direction. The capacity reductions are only caused by a change of driving behaviour. Sinha et al. (2007) use several microscopic simulation tools to predict the reduction of capacity. They use the developers’ default values for the capacity reductions or rubbernecking effects. Chapter 3 presented the driving behaviour and mentioned the capacity, but for only one accident site a queue discharge rate for a rubbernecking queue was found. The rubbernecking queue occurred on a freeway with 2 lanes per direction, whereas this chapter is restricted to freeways with 3 lanes per direction. The conclusion from chapter 3 was that there is a considerable change in driving behaviour around incidents compared to normal data. Therefore, we also included the rubbernecking queues in this research.

This contribution fills several gaps in our current knowledge. First of all, it is an extensive data analysis, using data of over 55,000 incidents. Even though there is a selection on the number of incidents, the remaining number of incidents (90) is still more than described in any other study. This increases the existing knowledge on capacity, but the large number makes it also possible to show the variations in the capacity. Secondly, rather than giving the numbers in the handbook values, the chapter provides the methodology to perform this analysis. The contribution thereof is that it explicitly discusses how a bias, as often seen in existing methods, can be overcome. Thirdly, it compares the reduction of the capacity

in the Netherlands, where the non-incident capacity is high, with the reduction in the USA where the non-incident capacity is lower. That shows whether the queue discharge rate at an incident are a fixed fraction of the queue discharge rate in non-incident situations. Fourthly, it provides insights into rubbernecking effects observed at traffic in the non-incident direction which are not studied on this scale before.

Additionally, this chapter also gives an example application for the use of these capacities, namely the assessment of possible incident management measures which reduce the incident handling time. Furthermore, it is essential to know the capacity to provide a good state estimation of future traffic conditions and these predictions form the basis for route advice as consequence of the incident.

### 4.3 Data collection set-up

The first part of this section describes which traffic data on during incidents are needed and which are available. The second part describes which data were used and which are the exact selection criteria. It also points out why the capacity reductions found here can only be applied in case of a traffic jam.

#### 4.3.1 Data type

In this contribution, we just discuss the queue discharge rate which differs from the free flow capacity (which is discussed in detail in section 4.4.2). The queue discharge rate is the flow at the location of a bottleneck. Bottlenecks are characterized by the condition that upstream traffic is in a congested state, and downstream traffic is in a free flowing state. The traffic demand at the bottleneck is therefore higher than the capacity of the road. To be able to find these bottlenecks, one needs to know the traffic state, which can be derived from the average speed at the road or alternatively, the occupancy along a road stretch.

The traffic states can be found by traffic flow data, which are often logged automatically, but reliable data on incidents lacks often since time and location have to be put into a database manually. Also information on the available lanes is required. It is necessary to combine the data on incidents with the data on the traffic flow properties, which makes a good registration of starting time, end time, and location essential. Also the number of lanes which are available are needed which is not recorded in the incident database. However, dedicated Variable Message Signs (VMS) in the Netherlands show every 500 meters for each lane whether it is opened or closed. The settings of these Variable Message Signs are logged automatically and therefore there are no manual errors in the logs. From these logs the number of available lanes can be retrieved, which indicate the space and time of accidents.

**Table 4.2: Road stretches used for analysis**

Road number	Direction	from km	to km
A1	Eastbound	0	29
A1	Westbound	29	0
A2	Southbound	37	94
A2	Northbound	94	37
A4	Northbound	49	21

### 4.3.2 Incident selection

This contribution focusses on the capacity reduction on freeways with three lanes per direction. A large part of the Dutch freeway network is equipped with double loop detectors. At every 500 meter interval, there is a double loop detector. The aggregate data of the loop detectors is stored. This shows the average speed and the flow, both aggregated over 1 minute.

Five roadway stretches in the Netherlands were selected based on the available data and on incident frequencies. Table 4.2 presents these stretches.

We divide the incidents in two categories: (1) an accident and (2) a car break down at the emergency lane. The requirements for selecting the event are different for these two categories. We want that the incident causes a traffic jam and provides enough measurements. We require the following:

- Upstream of the incident, the average speed is lower than 70 km/h (44 mph) and downstream of the incident, the average speed is over 70 km/h (44 mph). This means that the incident causes congestion and we only measure queue discharge rates.
- An accident needs to have a duration of at least 30 minutes and one or more lanes need to be closed for at least 30 minutes. This ensures there are enough data points for a good estimate.
- A broken down car needs to be at the same spot for at least 15 minutes. This ensures there are enough data points for a good estimate. It would be better to take even longer (30 minutes) but a car that broke down usually does not stay there for such a long time.

The first requirement says that only incidents which cause congestion are considered. Without congestion, there are no extra delays and the delays form the main incentive for this research. The other two requirements imposes restrictions on the duration of the incident. If the duration is too short, there are too few measurements and the capacity can not be determined. There might be a few capacity measurements, but for reliable

**Table 4.3: Number of incidents used**

	Number used	Fraction passed bottleneck criterion
Shoulder lane	20	9%
1 lane closed	21	43%
2 lanes closed	20	65%
Rubbernecking	29	7%

information the confidence bounds are needed. The spread of the capacities can only be derived from a series of measurements.

Only incidents which are a bottleneck (i.e., cause a traffic jam, as explained in section 4.3.1) are included. These are now filtered based on the incident type (accident or car break down) and location (selected stretches). Table 4.3 shows the numbers of incidents which were suitable based on all criteria, including the requirement that the incident forms an bottleneck (the first requirement). The second column shows the fraction of incidents that fulfilled this requirement.

For rubbernecking the same criteria apply. Table 4.3 shows that the larger the disruption is, the more likely it is that an incident causes a traffic jam. For example, a rubbernecking queue occurred in just 7% of the incidents, but in 65% of the incident that blocked 2 out of the 3 lanes a queue occurred. Although it is noteworthy, this is not surprising since the capacity reduction in case of a physical blocking is much larger.

Note furthermore that incidents included in the analysis for rubbernecking are not necessarily a subset of the cases for which the incident was included for further analysis. For instance it is possible that there is no queue in the incident direction and hence the incident is discarded for the analysis of the capacity in the incident direction. However, the same incident could create a traffic jam in the non-incident direction (if the traffic flow is high enough) and so be included for the rubbernecking analysis.

## 4.4 Data analysis

To determine the reduction of the queue discharge rate at incident locations it is required to have good estimates of the queue discharge rate during the incident situation and in the normal situation. This section explains how the queue discharge rates are determined in case of an bottleneck during an incident situation, in section 4.4.1, and in normal conditions without an incident, in section 4.4.2. Section 4.4.3 presents how the queue discharge rates are compared and a road efficiency is computed.

### 4.4.1 Queue discharge rates in incident conditions

In order to determine the queue discharge rate, the incidents are selected in such a way that they form a bottleneck during the time of the incident. This means that the head of the

queue is located at the incident site. Traffic flows out at the maximum flow rate possible for that location at that moment. The queue discharge rate at the incident site, the outflow out of the queue, can thus be derived from the counts at the downstream detector. This gives an average and median queue discharge rate as well as the interval bounds around it.

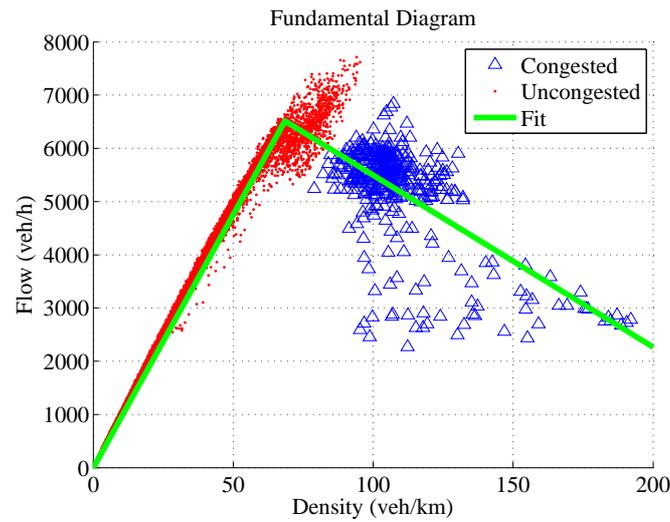
We use data aggregated over 1 minute, which is a relatively short time. Brilon et al. (2005) show the stochastic nature of the capacity measurement and discuss the consequences of a large or short aggregation time. The short time available for measurements, the duration of an incident, requires to make a choice on fluctuating measurements aggregated over a short time or a few more stable measurements aggregated over a longer time. Taking a short interval means that the spread of the measurements is larger, but more important, it will provide sufficient measurements to have an indication about the reliability of the value. Although the median value might slightly change due to stochastic effects, there is no bias of choosing a short aggregation interval. Therefore, we choose a short aggregation time of 1 minute.

#### 4.4.2 Queue discharge rate in normal conditions

The queue discharge rate is site-specific, like normal capacities. Apart from the site-specific influences, there are likely to be behavioural influences, which can be best described if one computes the relative queue discharge rate compared to the normal queue discharge rate. This section explains how the normal queue discharge rate is computed, even for sites which are not a bottleneck in normal conditions.

There are many methods to determine the free flow capacity in the bottleneck, such as the Product Limit Method introduced by Kaplan and Meier (1958). Other methods, like the one proposed by Brilon et al. (2005), can compute the capacity at any point, but give the free flow capacity. This is generally based on analyzing the maximum flow. The location of the incident is in general not a bottleneck for non-incident situations. Therefore, it is not possible to use the same method as during an incident (described in the last paragraph) and another method has to be used. To find the queue discharge rate, we use a fit of an reverse-lambda shaped fundamental diagram (Koshi et al., 1981) in which the intersection of the fit of the free flow branch and the congested branch is taken as queue discharge rate – see figure 4.1. Even though the incident locations are not a bottleneck in normal conditions, there is congestion from time to time due to growing queues caused by downstream bottlenecks. Therefore also at the incident locations there are points in the fundamental diagram in the congested branch.

Previously collected data were obtained for the periods of 10 days before and 10 days after the incident. If the average speed was under 70 km/h (44 mph), the traffic state was classified as congested traffic; if the average speed exceeded 70 km/h, the traffic state was classified as uncongested. For both traffic states (both branches of the fundamental diagram), we made a linear fit in the density-flow diagram. The queue discharge rate is found at the point where these two lines intersect. Note that the free capacity can be



**Figure 4.1: The fitted fundamental diagram for one incident**

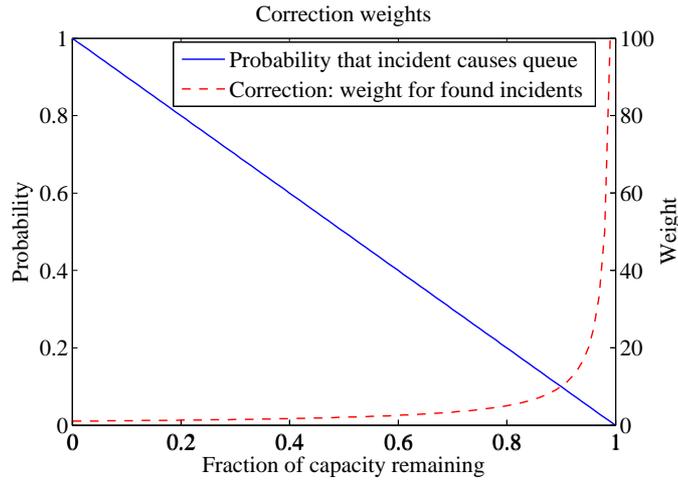
**Table 4.4: The used variables**

variable	unit	meaning
$l_{incident}$	-	The number of lanes available for traffic in one direction during the incident
$l_{non-incident}$	-	The number of lanes available for traffic in one direction in normal conditions
$\phi_{incident}$	$veh/h$	The queue discharge rate during the incident
$\phi_{non-incident}$	$veh/h$	The queue discharge rate in normal conditions
$F$	-	The fraction of queue discharge rate that remains available
$R$	-	The reduction of the queue discharge rate
$\eta$	-	The efficiency of the road use in incident situations

found on the free flow branch which is usually higher than the intersection point, see also for instance Hall and Agyemang-Duah (1991), Dijkster et al. (1997), Cassidy and Bertini (1999) or Chung et al. (2007); this is called the capacity drop. This can also be seen in figure 4.1 which shows measurements in free flow with flow well above the queue discharge rate.

### 4.4.3 Comparing efficiency

This section uses the variables as used and shown in table 4.4.  $\phi$  is the total hourly queue discharge rate for all lanes. Under normal conditions the queue discharge rate is  $\phi_{non-incident}$ , and during an incident  $\phi_{incident}$ . The quotient of the two queue discharge



**Figure 4.2: The probability that an incident causes a queue and the correction factor**

rates is the fraction of the capacity that remains,  $F$ :

$$F = \frac{\phi_{incident}}{\phi_{non-incident}} \quad (4.1)$$

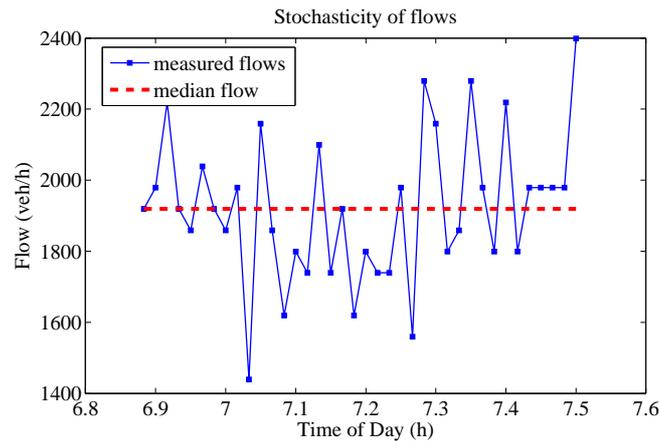
The reduction of the queue discharge rate is a combined effect of the reduction of the number of lanes and the less efficient use of the remaining lanes. We therefore expressed the efficiency  $\eta$  of the use of the remaining lanes by dividing the capacity factor by the fraction of the roadway that is available:

$$\eta = \frac{F}{\frac{l_{incident}}{l_{non-incident}}} = \frac{\phi_{incident}}{\frac{l_{incident}}{l_{non-incident}} \phi_{non-incident}} \quad (4.2)$$

In this formula  $l$  is the number of lanes that is available. The use of the formula is best explained by an example, which is a (fictitious) situation where the capacity reduces to 40% if 1 out of 3 lanes is closed. The remaining 2 lanes could have provided 67% of the original capacity with unchanged traffic behaviour, but the actual flow is only 40% of the original capacity. So, the road is used at  $\frac{40}{67} = 60\%$  of the original efficiency. Note that this is an example and the real outcomes of the study can be found in section 4.5

In the methodology applied in this study (and in any other) is a possible bias towards higher capacity reductions (capacity reduction  $R$  is  $1 - F$ ). Incident locations with a higher capacity reduction are a bottleneck at lower demand levels. This means that incidents with large capacity reductions are more likely to be incorporated in the study, which biases the results. Assuming a uniform demand distribution over the time (e.g., day), the probability to find an incident blocking the road is proportional to the capacity reduction  $R$ . Using the relationship  $R = 1 - F$ , we find the probability that an incident with fraction  $F$  remaining causes congestion:

$$P(Q > FC) = R = 1 - F \quad (4.3)$$



**Figure 4.3: The fluctuation of the capacity flow for one incident**

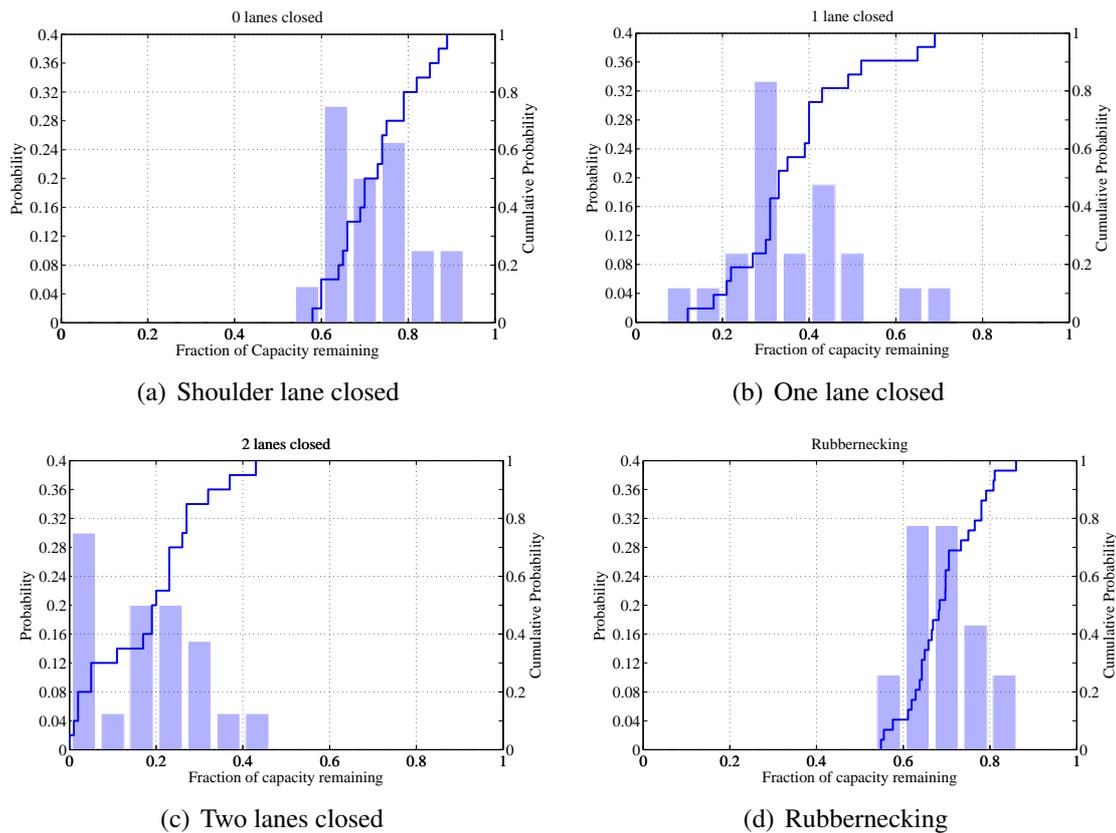
This is shown in figure 4.2. It can now also be computed which weight one should give to an incident reduction. In order to get an unbiased set, this weight should be inversely proportional to the probability of finding an incident with that reduction. The weight is also shown in figure 4.2.

One could check how biased the set of incidents is. To this end, the incidents are first grouped in bins with similar reductions. Then, one can divide the found number of incidents by the probability for each bin of capacity reduction. However, this might give unbalanced results since one incident can get a very high weighting factor and this way the distribution might depend heavily on one or two incidents that are included. This is not desirable since the inclusion of incidents is a stochastic process. Therefore, we choose to check the corrected distribution, ignoring some incidental cases, and to present the more homogeneous set of uncorrected (but possibly biased) capacity reductions. Note that this bias is only a serious concern if the capacity reduction varies much. In the limit where no variation of the capacity reduction exists, there is no bias at all.

## 4.5 Resulting queue discharge rates

Figure 4.3 shows for one incident the fluctuation of the flow over time. In this particular incident, one out of the three lanes was closed due to the incident. Consequently, two lanes remain opened. The median flow is indicated with a dotted line. The fluctuation is larger because we aggregated over short time intervals of 1 minute. For the example given here, the median is 1919 vehicles/h and a standard deviation is 209 vehicles/h.

For the same incident location, figure 4.1 shows the construction of the reference capacity. The free part and the congested part of the diagram are separated based on a 70 km/h threshold. This value is based on experience for the critical speed in the Netherlands. The figure shows congested and non-congested measurements, as well as linear fits thereof. Figure 4.1 shows that this value agrees with the critical speed, the speed at the point

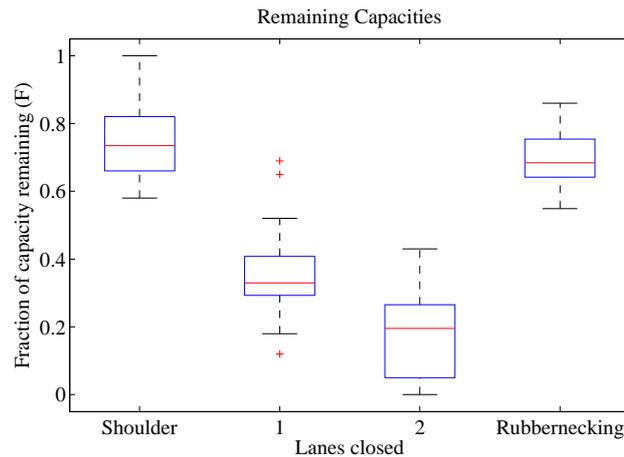


**Figure 4.4: The capacity factor for the three types of incidents which block the roadway and for the rubbernecking**

where the flow is maximum. The flow value at the point where the fits of both branches cross is the queue discharge rate. For this incident the reference queue discharge rate is 6500 vehicles/h over three lanes. Since the queue discharge rate is 1919 vehicles/h (see last paragraph), the resulting capacity factor  $F$  for this incident now is  $\frac{1919}{6500} = 0.30$

For all incident situations we calculated the capacity factor  $F$  as mentioned in equation 4.1, which were grouped by category, being the shoulder lane blocked, 1 of the 3 lanes blocked, 2 of the 3 lanes blocked, and rubbernecking. In figure 4.4 the full distributions of the capacity factors per group can be found. The figure shows a spread of distributions even within one group. Furthermore, it is remarkable that the group where the shoulder lane is closed and all driving lanes are available, still shows a considerable capacity reduction. The same holds for the case of rubbernecking. Figure 4.5 combines the values of the capacity factors. In the figure, the median value is indicated by the middle line, and the 25th and 75th percentile values are indicated by the edges of the box. The line ends show the range of the capacity reductions. Some incidents cause a reduction outside the normal range; these reductions are indicated with a cross in figure 4.5. Due to a lack of detailed information, we could not analyse in detail why these situations were different.

Table 4.5 shows the results in numbers. The table states the mean values of the capacity factors and the standard deviation. It also states how efficient the remaining lanes are



**Figure 4.5: The capacity factors found for different incidents**

**Table 4.5: The resulting capacities**

Lanes blocked	Shoulder	1 out of 3	2 out of 3	0 (rubbernecking)
Capacity Factor $F$				
Mean	0.72	0.36	0.18	0.69
Standard deviation	0.09	0.14	0.12	0.08
Efficiency of Lane Use $\eta$	0.72	0.54	0.54	0.69

used. This is calculated by dividing the capacity factor by the fraction of the number of lanes of the roadway that is available, see equation 4.2.

There were 29 rubbernecking queues that met all criteria and could be analysed using the method described in section 4.4.3. For these queues, the fraction of the queue discharge rate that remains,  $F$ , equals the efficiency rate  $\eta$  in formula 4.2 since no lanes are closed. This efficiency value (0.69 as shown in table 4.5) is higher than stated earlier in a microscopic analysis in chapter 3. However, that earlier work calculated the quotient of the queue discharge rate during an incident and the *free* capacity. This value can be higher because queue discharge rate is lower than the free capacity (the capacity drop) and thus the queue discharge rate expressed as fraction of the changing reference value is higher. At the other hand, this reduction is much more than predicted by microscopic traffic simulators using default settings (Sinha et al., 2007). Another interpretation for the efficiency of 69% is that the queue discharge rate at the incident site is (1-69%=) 31% lower than in normal conditions.

When one or two of the driving lanes are blocked, the efficiency of the remaining lanes reduces to 54%. Note that the efficiency factor indicates how much the driving behaviour changes compared to normal driving. That this number is the same for the case one driving lane is blocked and for the case two lanes are blocked means that both situations lead to the same behavioural effects. In general, the efficiency factor  $\eta$  is higher in case there is no disruption in one of the lanes normally available for traffic, so in case there

is a blocking on the shoulder lane or on the carriageway in the opposite direction. We conclude that if the driving lanes are disturbed, this affects the driving behaviour more.

A check as explained in 4.4.3 shows the effect of a possible bias. For the cases with 1 or 2 lanes blocked, this bias turned out to be negligible. For the cases of rubbernecking and a blocking of the hard shoulder, the bias in efficiency is around 0.05.

## 4.6 Application: incident management

The capacity reductions found in this study are relevant for various purposes. One of them is explained here where it is analysed what the effect of quicker actions of incident management would be qualitatively. First, the case is explained in section 4.6.1. Section 4.6.2 then gives the results of the case study.

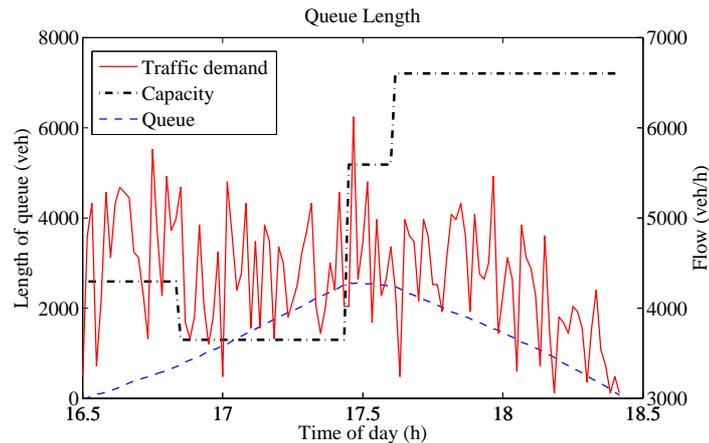
### 4.6.1 Case set-up

This section discusses the possible effectiveness of shortening the clearing of an incident. It describes how it can be found how the queue length and the delay change if these processes are sped up. In this analysis we use the capacity factors found from the data analysis.

Usually, the emergency services start working on the roadway (on the lanes for through traffic) but move the wrecks as soon as possible to the shoulder lane. According to the Dutch study on incident management by the Dutch Road Authority (2007b), we distinguish three phases during the incident: (1) the time until the first emergency services arrive, (2) the time that the emergency services work on the roadway and (3) the time that the emergency services work at the shoulder lane. For this case study, it is supposed that during the first phase, one of the three lanes is blocked. When the emergency services arrive, a second lane is blocked to have a safer working space. In phase 3 only the shoulder lane is blocked.

The queue discharge rates which come out of the capacity analysis can now be used to determine the maximum flow during each of these phases. The incident management study not only determines the phases, but also gives the average time for each of these phases. To determine the delay, the demand for a non-incident day and the capacities during the incident are put into a traffic simulator. This equals a situation where people would not change their route because of the incident. The simulator now predicts the queue length and delay if the outflow was blocked for a while by the incident. Note that only the flow values for a non-incident situation can be used, otherwise flows do not indicate the demand, but are limited at the queue discharge rate. We use a vertical queuing model to compute the delay. That means that vehicles will encounter delay at the moment they pass the incident location.

We analysed the differences in queue length and total delay if the time in different phases reduces by 2 or 4 minutes. We derive two measures from the simulations: the duration



**Figure 4.6: The length of a queue for a fictitious incident**

of the queue and the total delay. It will show the effect of shortening a period when all other periods remain the same. It can be expected that the sensitivity is largest for the shortening of the period with the lowest capacity. In that phase, the effective inflow of vehicles into the queue, and so the queue growth, is the largest.

In the analysis we quantify how much the queue duration and the total delay of the queue can be reduced. We analyse the effectiveness of a time gain for two different, typical (but fictitious) incidents in the peak period.

## 4.6.2 Results of the incident management study

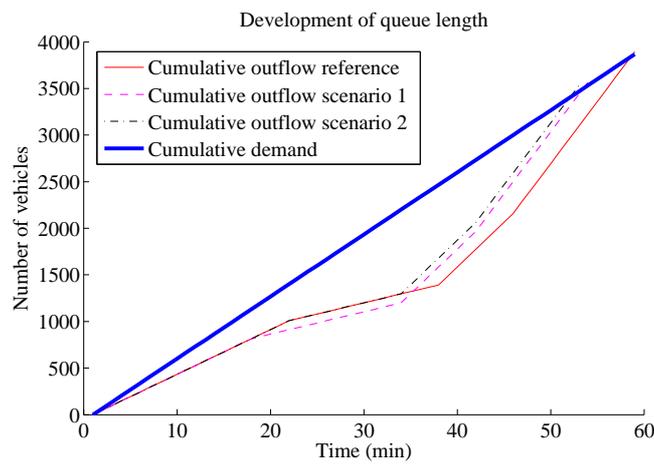
Figure 4.6 graphically shows the method used to compute the indicators for one set of durations per time phase. The changing capacity is shown, as well as the demand, which fluctuates at a high frequency. The queue length is calculated as the cumulative sum of the difference between the demand and the capacity and is also shown in the figure. From this derivation, the indicators queue duration and delay can be calculated.

Table 4.6 shows the effects of the shortening of actions. We simulated the traffic with a 2 and 4 minute shortening for each of the phases. The total effects are divided by the decrease of shortening of the phase in order to get comparable sensitivity numbers for both 2 and 4 minutes shortening. For both times, comparable results in reduction of delay and queue length are per minute shortening are found.

An error that can easily be made is that one expects a large sensitivity for the earliest phases because all following vehicles profit from the total shortening. However, this reasoning does not hold. It is best explained graphically (figure 4.7). In the example we show results for a case with a constant demand of 4000 veh/h and a reference capacity of 8000 veh/h. For the capacity reductions at each of the phases we use the values of the according link configuration (as explained in section 4.6.1) stated in table 4.5. For the duration we used the results of the study of the Dutch Road Authority (2007b). The figure shows the cumu-

**Table 4.6: The effectiveness of shortening the actions in Incident Management**

	Decrease of queue duration (min queue length/ min phase reduction)	Decrease of delay (vehicle hours/ min phase reduction)
$\Delta t = 2$ min		
Arriving	2.2	70
Service on roadway	2.8	101
Service on shoulder lane	0.6	21
$\Delta t = 4$ min		
Arriving	2.5	77
Service on roadway	3.2	106
Service on shoulder lane	0.85	30

**Figure 4.7: The delays for different scenarios**

lative departure and arrival curve at the (vertical) bottleneck. The area between the two curves is the total delay. Therefore it holds that the larger that area is, the larger the delay is. We plot the same curve for three situations: the reference situation using the actual lengths of a period, a situation with a shorter first phase (scenario 1) and a situation with a shorter second phase (scenario 2). The second phase is the phase in which emergency services are working on the freeway and the capacity is reduced most. The figure shows that the delays are most reduced if phase 2 is shortened (scenario 2). One of the reasons for this is that in both scenarios the end of phase two is at the same moment. This way it can be reasoned that the shortening the phase with the largest capacity reduction is most effective.

This is also shown in the numbers: a shortening of the period with the largest capacity reduction, the period of emergency workers at the roadway, has the most effect. If the emergency services block 2 lanes of the roadway one minute less, the duration of the queue reduces with about 3 minutes (and 100 vehicle hours of delay). For the reduction of the time until the emergency services arrive, these gains are lower, but still considerable.

## 4.7 Conclusions and discussion

In this chapter the maximum outflow out of a jam which is caused by an incident is analysed by studying 90 traffic jams. The most important finding is that the capacity per lane reduces significantly due to a change in driving behaviour. The size of this reduction depends on the incident type. If one of the driving lanes is blocked, the remaining lanes are used 46% less efficient, which yields an “efficiency factor” of 54%. To compute the resulting queue discharge rate, one has to do the following. First, one has to take the proportional part of the road that is available (e.g., 33% if 2 out of 3 lanes are blocked). To compute the capacity reduction one takes the proportionally factor (33%) and multiplies this by the efficiency factor (54%). This is how much of the normal queue discharge rate remains (18%). Note that this is the reduction compared to the normal queue discharge rate; the reduction compared to the free flow capacity is even larger.

If there is an incident at the shoulder lane, the efficiency reduces by 28%. This is only due to a change in driving behaviour since all the lanes are open. A similar efficiency drop (31%) is found in case there is an incident at the other side of the guardrail, the “rubbernecking” effect.

Scientifically, the value of the chapter is threefold. First of all, it shows the size of the effect of changed driving behaviour at the incident site. The road efficiency is several tens of percents lower than in normal conditions, which means that the drivers changed driving behaviour plays an important role. Secondly, it is the first time that it is shown how much capacity decreases in the non-incident direction because of an incident at the other side of the guardrail. This effect is solely due to a change in driving behaviour. For both directions, the chapter shows the bandwidth of the road efficiencies for different types of configurations. Thirdly, for scientific purposes, it is important to have the methodology of the research described, rather than just a value in a handbook. The chapter provides also

insight in the used methodology and the possible flaws in it, being a possible bias due to a selective inclusion of capacities.

Practitioners can use this paper in the following ways. In the first place, the found capacity during an incident can be used by the road authority. The capacity is one of the most important road characteristics for traffic engineers. For instance rerouting of traffic in a situation after an incident will be based on the capacity of the blocked road. In case there is no other alternative, travellers can be informed about the delay.

Knowing the large influence of the distraction, it is worthwhile for practitioners to try to reduce the distraction and therefore increase the capacity. In the longer term, the road authority can experiment with increasing the road capacity. One of the possibilities would be to place screens in the middle of the road to prevent travellers in the non-incident direction of looking at the incident site. This could decrease the distraction and thus increase the road efficiency and capacity.

Another possibility is the quicker handling at the incident site. The impact of such a strategy is assessed quantitatively in the case study described the chapter. Delay can be avoided most by reducing the servicing time on the roadway and the time until the emergency services reach the incident site. These findings can provide the basis on a decision to invest in a quicker emergency response system: the benefits can now be shown and can, for individual measures, be compared with the costs.

Finally, the efficiency values found in this study for the Netherlands can be placed in a international context. Some values can be compared directly with values in literature for other countries. Comparing these shows that the reductions are similar for different countries, although the capacity is usually lower in other countries. Therefore, it is likely that the other relative reduction values which are not studied elsewhere, like the rubbernecking, are also valid in other countries, which is a value of this study for practitioners.

# Chapter 5

## Influence of incidents on route choice

---

Chapter 4 showed that the capacity of the road decreases considerably at an incident location. That will cause queues and extra delays. This chapter investigates how many people change their route when faced with this unexpected congestion caused by an incident. To this end, traffic data from days with serious incidents are analysed in this contribution. The flows on the routes past the incident and on alternative routes are compared with the same values on days without this extra delay. It is found that for major incidents up to 50% of the travellers take another route. This is important for instance when providing route information or suggestions on alternative routes. It is furthermore essential if one wants to compute vulnerable links in a network and is therefore an input to chapter 7. This chapter is a revised version of: Knoop, V.L., Hoogendoorn, S.P. and Van Zuylen, H.J. (2009) Route Choice Under Exceptional Traffic Conditions, in: *International Conference on Evacuation Management*, 23-25 September 2009, The Hague, the Netherlands.

---

### 5.1 Introduction

An incident influences both the traffic supply and the traffic demand. The total delay depends on the capacity and demand, not only because of the changing number of travellers that queue to pass the bottleneck itself, but a long queue also can cause delays for people travelling to destinations upstream of the bottleneck. If the queue grows longer than the distance to the closest (upstream) offramp, travellers who are not passing the incident location are delayed. Delays for these travellers are called spillback delays. Chapter 7 will show that spillback delays are an essential part of the total delay caused by an incident. In that chapter, the route choice in incident conditions is an input.

In this chapter, this route choice is studied. We do this by measuring traffic flows and derive the route choice change. Other studies usually use a simulator set-up or questionnaires which give the route choices made by a sample group of travellers and sometimes even give stated preferences instead of revealed preferences. There is, to the the best of

our knowledge, currently no study in the literature describing the actual change in route choice due to an incident by using the data of *all* vehicles. This might be due to the computational complexity of combining measurements of traffic monitoring devices for large amount of vehicles with an accurate lane closure accident database. This study fills that gap. The contribution of this chapter is that we find that up to 50 % of the travellers change their route during incident conditions. It shows furthermore that more people change their route if there is an accident on the intended route than in case there is a queue of the same length without an accident. It even shows a hysteresis effect: the travellers' reaction is delayed compared to the traffic situation.

In the next section, previous studies discussing route choice in incident conditions are briefly described, as well as the way these studies relate to work presented in this chapter. It then continues by explaining the methodology in-depth in section 5.3. That section also gives a description of the type of data that has been used. Section 5.4 describes the five incident situations which have been studied. The results, in terms of route choice, are given in section 5.5. That section also discusses differences found between the five incidents. Section 5.6 concludes the contribution and states some ideas of future research.

## 5.2 Previous studies

Many articles describe the influence of information on for instance awareness of alternative routes or the routes that are considered. Chorus et al. (2006) provides an excellent overview of different studies carried out. We will refer to their article for a comprehensive overview. Here the most important ones are listed, which are divided into two different categories: studies describing a theoretical framework and studies presenting data on drivers' preference.

### 5.2.1 Theory

For our study, it is particularly important to which extent the travellers deviate from their intended routes when facing unexpected queues. A sensitivity analysis of this deviation percentage can be found in chapter 7 of the thesis of Li (2008). It states that for an optimal network performance, there exists an optimal percentage of travellers that adapt their route en-route. Using simulation, Li shows that the impact of the fraction of people changing routes is large.

A theoretical explanation of what the influence of route information could be is given for instance by De Palma and Picard (2005). Using a mathematical game-theoretical framework, they show possible advantages of giving route information, assuming a certain compliance rate.

## 5.2.2 Practice

This section describes studies which describe the personal choice (stated or revealed in practice) of travellers.

### Stated preference

Koo and Yim (1998) study the behaviour of individual travellers in practice and analyse how they adapt their behaviour to traffic information. The study is restricted to 1052 participants that have filled out a questionnaire on their behaviour and it is limited to one specific type of incidents, namely the larger incidents giving a delay of more than 30 minutes, but not completely blocking the motorway. They find that even if travellers are informed about the traffic situation, 70% of them still keep stick to their original plan of departure time and route choice. A similar methodology with a questionnaire is applied by Jou et al. (2005). Amongst others, they conclude that while the travel times are within a certain band width, the travellers do not change their routes.

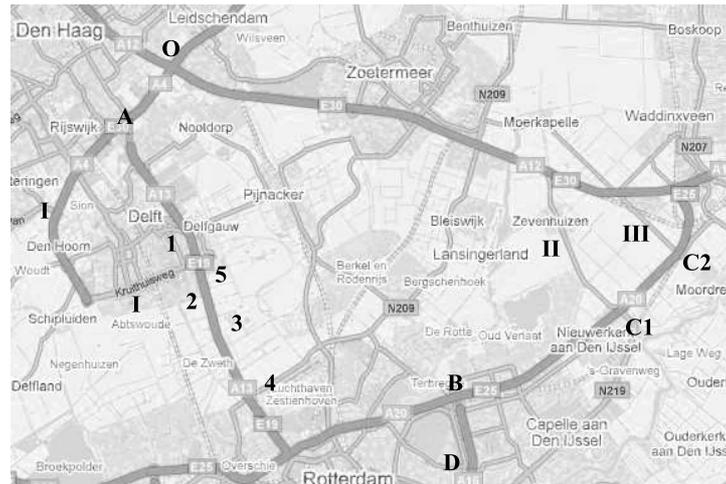
Some studies describe how people choose their route and how Advanced Traveller Information Services (ATIS) influences them using a simulated route choice environment, for instance Chen and Mahmassani (2004) and Bogers et al. (2004). The last paper discusses how people weight the influence of on-line information compared to their previous experiences. The authors conclude that people learn from bad experiences and that travellers only rely on information they consider to be correct. That is, as soon as they find out that the information is incorrect, their reaction is not always in line with the provided information.

### Revealed preference

Muller et al. (2005) discuss the type of data that can be used to measure traffic data in congested network. One of the examples shown is how loop detector data can show what alternative route people take when an incident has happened.

Kraan et al. (1999) show the influence of Variable Message Signs showing the queue length. They find that “each additional kilometer queue length displayed leads to a reduction of the proportion of drivers that select that route between 0.8 and 1.6 percent.” These are percentages of the total flow which possibly consist of travellers which have a destination upstream of the point where the two alternative routes come together. Therefore, the fraction of travellers changing their route might be larger then this 0.8% to 1.6% percent, conclude the authors. Their findings are based on the (stochastic) change of recurrent congestion. They do not study the route choice changes in incident situations in particular. Whether queues caused by incidents have a different effect will be studied in this contribution.

The study presented here is inspired by findings of Kraaijeveld (2008) who uses detector data and an incident database. He analyses five cases, in which he finds one case in which travellers deviate significantly. His conclusion is that, for the specific situation considered, a serious incident makes around 7% of the drivers deviate, whereas in the other cases the disruption was too small to cause a change. The final conclusion is based on one observation on one day, and the statistical variations in the normal route choice are not given.



**Figure 5.1: The case study area with the incident locations (numbers) and points for route decisions (letters)**

This contribution also studies revealed preference in real-life, like Kraan et al. (1999) and Kraaijeveld (2008). Rather than Kraan, we study the situations with an incident. We use a similar methodology, but a larger set of incidents. Furthermore, we choose another location where it is possible to analyse the flows in more detail.

## 5.3 Methodology and data selection

This section discusses how the route choice change is determined and how the incidents are selected that are used in this study.

### 5.3.1 Possible alternative routes

The motorway A13, east of Delft, a town in the Netherlands (figure 5.1), is the main corridor for traffic with an origin in The Hague or more north to Rotterdam and other destinations more south. We consider the network of the motorway A13 and its alternative routes. In the study we will focus on traffic from “O” to “D” or vice versa. To get from “O” to “D”, one has to pass the motorway junction indicated with a “B” in figure 5.1. The quickest route between these motorway junctions next to the letters “O” and “B” in figure 5.1 is along the A13. This trip is 24 kilometers long and would take 17 minutes in free flow.

In a database of incidents we looked for incidents on the A13 near Delft. This location is particularly interesting because it provides two alternative routes that can be followed, depending on incident severity and incident location.

The remainder of this section describes the alternative routes. There are two routes, one with a large overlap with the old route and one with a small overlap. Some properties can

**Table 5.1: The alternative routes**

Alternative	I	II	III
Exit at	A	C1	C2
Extra free flow time	9 min	2 min	4 min
Extra distance	6 km	5 km	7 km
Distance on non-motorway	4 km	7 km	3 km

be found in table 5.1. Alternative route I is described for traffic from “O” to “D”, whereas the alternative routes II and III are described for traffic from “D” to “O”. This is because for the considered accidents (see section 5.4), these turn out to be the relevant directions for these routes. However, they can be driven in both directions.

Route I is a good alternative for travellers if the main delay takes place on the motorway between Delft and The Hague (“Den Haag” on figure 5.1). It uses an alternative motorway, the A4 motorway at the west side of Delft (route I). Traffic can return to the original route by a connecting highway just south of Delft. This is a short detour of 6 km. Coming from the north, the A4 motorway entrance is convenient, namely continue on the A4. In fact, to turn to the A13, one has to take the exit. The location of this junction is indicated by an “A” in figure 5.1.

For larger queues there is a detour completely avoiding the A13 between Delft and The Hague. This is only possible for people wanting to go from Rotterdam or origins more south towards destinations which are as north as The Hague or more north (see figure 5.1), or the other direction. However, traffic does not pass towns in between, such as for instance Delft and is therefore unsuitable for traffic with these destinations. The alternative route is here described for traffic coming from the south at the eastern part of the ring road around Rotterdam. When reaching the northern branch of the ring road of Rotterdam (at the point marked with a “B” in the map), traffic can divide from the original route, the A20 westbound towards the A13, and take the A20 eastbound and then travel westbound on the A12 instead. It is not possible to turn directly from the A20 to the A12 westbound at the motorway intersection. Therefore, traffic has to cross on a non-motorway to the A12, which is possible at “C1” (route II in table 5.1) or “C2” (route III). The route turning at C2 is the quicker of the two, but has a larger detour compared to the main route (7 km) including 3 kilometers on non-motorways.

Smaller disruptions will not make people change their route towards route II or III. These routes are namely more congestion prone than alternative I in peak hours, which means there is an extra risk on delays. Furthermore, the point at which travellers need to take that route decision is further upstream, which means that they will take route II or III only if the traffic conditions on the originally intended route are really bad. For smaller disruptions route I is a sensible alternative.

Both at the northbound decision point (for accident 1 and 2, “A” on figure 5.1) and at the southbound decision point (accidents 3 to 5, “B” on figure 5.1) a variable message sign

**Table 5.2: The measured flows**

Day \ Route	Main	Alternative	Total
Reference	$q_{main}^{ref}$	$q_{alt}^{ref}$	$q_{total}^{normal}$
Accident	$q_{main}^{acc}$	$q_{alt}^{acc}$	$q_{total}^{acc}$

(VMS) is present which could suggest an alternative route. It was not possible to find the messages shown at the VMS, which might include a route advice.

### 5.3.2 Data selection

For this study we use incident data from december 2007 to september 2008. The database of incidents mentions the location of the incident, the moment it happens<sup>1</sup> as well as the time the emergency vehicles leave the incident location. From the database we select incidents on the A13 motorway if there are wrecks or emergency workers at the roadway on working days and if there is a significant queue. The last requirement means that most of the considered incidents are within the peak period.

All motorways within the study area are equipped with double loop detectors every 500 meters. To analyse the route choice, flow data obtained from these double loops is used. At the detectors, speeds and counts are recorded and then aggregated over one minute. These one-minute data are stored and can be accessed for each detector individually. Flows on the main route and the alternative route are detected for the days of the incident and the reference days. The symbols we will use for these flows are given in table 5.2.

Unfortunately, it was impossible to track back if route advice messages were given to the drivers by Variable Message Signs or radio broadcasted traffic information. On the radio messages in the Netherlands usually the reason of the delay is given, so it is likely that travellers are informed in case the road is completely closed.

### 5.3.3 Indicators for route choice change

To find the amount of traffic that reacts on these flows, several indicators are computed. These are summarised in table 5.3. In the remainder of this section, each of the indicators is discussed in more detail.

The first step to get insight in the route choice process is computing the split fractions (i.e., the quotient of the flow to the main direction and the total flow) at the decision points. In traffic there are many stochastic processes, like for instance demand. To be independent for these stochastic changes in demand, we choose to find the route choice as a relative

<sup>1</sup>this might deviate a few minutes since it is filled out by hand

**Table 5.3: The used indicators and their meaning**

Symbol	Meaning
$\psi$	Split fraction, the relative amount of traffic to the direction of the incident
$\chi$	Fraction of traffic changing routes in incident situation

number which expresses the number of travellers going in the main direction compared to the total demand. Mathematically, this can be expressed as follows:

$$\psi = \frac{q_{main}}{q_{total}} \quad (5.1)$$

$\psi$  is the split fraction as function of time, which will be calculated for the day of the accident and for several reference days. For each time interval, this gives a several reference measurements and a value which possibly differs from this reference value. For each time there is a series of values for the reference days and one measurement for the incident day. Using a t-test it is verified whether  $\psi$  significantly differs from the values in the reference days. This is repeated for each time interval and this will show in which time interval  $\psi$  differs.

This split fraction qualitatively show whether people change their routes. However, the difference of the two split fractions will not tell which fraction of the travellers actually take another route. We will assume that the change in split fraction comes only from people that in normal conditions would travel to the main direction. Then, the relative amount of people rerouting can be expressed by:

$$\chi = \frac{\psi^{normal} - \psi^{acc}}{\psi^{normal}} = 1 - \frac{\psi^{acc}}{\psi^{normal}} \quad (5.2)$$

While this might be intuitively right, we will also derive equation 5.2 in equations 5.3 to 5.8. To this end, we will introduce a hypothetical situation as it would have been if there was no incident. In this situation the total flow equals the total flow of the incident situation but the split fraction equals the split fraction of the non-incident situation. The variables for this situation are indicated by an asterisk. So, mathematically we have:

$$q_{total}^* = q_{total}^{acc} \quad (5.3)$$

$$\frac{q_{main}^*}{q_{total}^*} = \psi^* = \psi^{normal} = \frac{q_{main}^{normal}}{q_{total}^{normal}} \quad (5.4)$$

Now we compute which fraction of the travellers past the incident changes their route,  $\chi$ . The total number of vehicles changing is the difference between the flow to the incident direction if the split fraction remained the same,  $q_{main}^*$ , the *actual* flow to the incident direction,  $q_{main}^{acc}$ . To get the relative fraction, this has to be divided by the number of vehicles that would have passed the incident location,  $q_{main}^*$ . In mathematical form, that

is:

$$\chi = \frac{q_{total}^* - q_{main}^{acc}}{q_{main}^*} = 1 - \frac{q_{main}^{acc}}{q_{main}^*} \quad (5.5)$$

By definition (equations 5.1, 5.3 and 5.4) we have

$$q_{main}^{acc} = \psi^{acc} q_{total}^{acc} \quad (5.6)$$

$$q_{main}^* = \psi^{normal} q_{total}^{acc} \quad (5.7)$$

Substituting these values for  $q_{main}^*$  and  $q_{main}^{acc}$  into equation 5.5, we obtain:

$$\chi = 1 - \frac{\psi^{normal} q_{total}^{acc}}{\psi^{acc} q_{total}^{acc}} = 1 - \frac{\psi^{normal}}{\psi^{acc}} \quad (5.8)$$

Which is the same as equation 5.2.

By using split fractions rather than flows, we remove the effect of fluctuating demand. Also, if the congestion spills back onto the link where the decision has to be made, the flow on that particular link is reduced and traffic to both directions is hindered. Only for a few minutes there will be a queue on some lanes of a multilane motorway. Although there can be temporarily an effect that one lane to the alternative route is not congested. This situation, however, will not exist very long since congestion grows further upstream and blocks drivers in both directions. This means that traffic in both directions is equally influenced and the split fraction remains the same.

The changes in route choice will be analysed in combination with differences in instantaneous queue length and instantaneous travel time between the main route and the alternative routes. To this end, loop detector data are collected. For routes II and III, the queue length and the travel times can be constructed. Only a small fraction of route I is equipped with loop detectors, so the speeds are unknown. However, from experience it is known that route I is usually not very congested.

The road is split up into sections in such a way that the detectors are exactly halfway each section with a length of approximately 500 meters each. It is assumed that the average speed is constant over the section and that that speed equals the average speed on the detector. A section is assumed to be congested if the average speed is below 70 km/h. By summing the lengths of the congested sections, the (instantaneous) queue length is constructed. We will compare the queue length to the fraction of travellers that take another route. We choose for the instantaneous queue length, rather than the travel time that will be experienced by the travellers since this is the information that is given to the road users by VMS signs, or radio broadcast. By VMS signs or navigation systems at best the instantaneous travel times are communicated. We therefore also construct the instantaneous travel times based on the section distances and average speeds measured by the detectors.

The route choice is first analysed as function of time. A t-test is used to test differences in split fraction at the same time between the day with an accident and comparable days. However, this will not reveal whether people change their route due to an incident directly

**Table 5.4: The incidents used in this study**

Nr.	Direction	Date	From	To	Complete blocking
1	Southbound	23 June 2008	13h15	16h45	no
2	Southbound	22 August 2008	16h50	18h45	no
3	Northbound	12 December 2007	6h25	8h05	no
4	Northbound	22 February 2008	6h20	9h00	no
5	Northbound	22 September 2008	12h35	14h20	yes

or due to the queues caused by the accident. We also analyse this. We compare the route choice for similar queue lengths on days with an accident and without. If for the same queue length difference (or travel time difference) the route choices differ, travellers do not only react on the congestion, but also on the occurrence of the accident itself. To analyse this, we will make a scatter plot of the split fraction and the queue length. Similarly, a scatter plot of the split fraction and the instantaneous travel time will show whether people react only on travel time differences or also on the fact that an accident has occurred.

All measured data is aggregated over one minute. Because this is a short time, the data fluctuate much. In order to see a trend in the graphical representations, we smoothed the data using a moving average filter. In particular, the filter replaces each data point with a weighted average of which the weight factor depends on the distance to the considered time step. Data which is collected at times which differ more than 15 minutes from the considered time is not considered at all and therefore get a weight factor of 0.

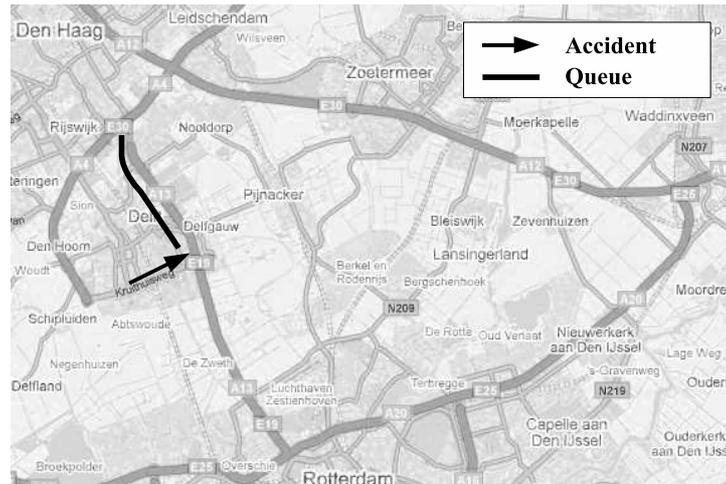
Another effect causing differences in the flows is the weather which can influence the number of trips for specific origins and destinations. The number of trips can also depend on the day of the week. Obviously, a different number of trips on several OD-pairs can influence the split fractions. Therefore, as reference for the split fractions, we use the split fractions on the same day of the week on days with comparable weather conditions.

## 5.4 Incident description

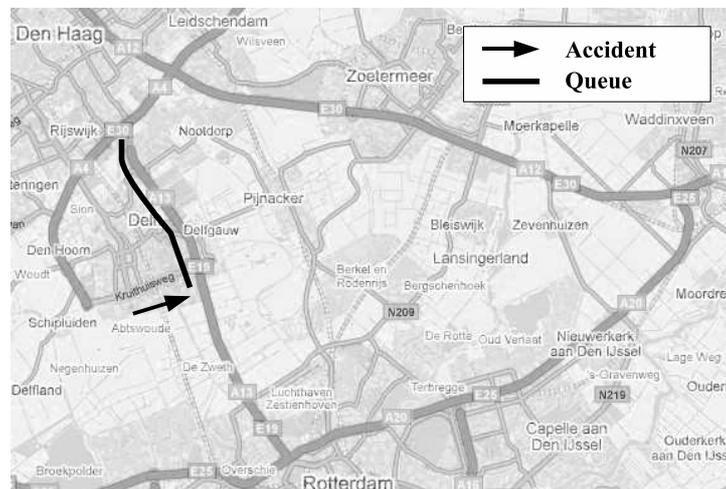
This section describes five incidents which meet the criteria posed in section 5.3. An overview of the incidents is given in table 5.4.

The first incident is a car accident which took place at Monday 23 June 2008 at 12.30 hours, blocking one lane of the A13 southbound during the afternoon peak hour. It was a clear day, so as comparison we use 9, 16 and 30 June 2008, all Mondays with clear weather. The accident caused extra congestion during the peak hour on the A13 southbound. The resulting traffic conditions are depicted in figure 5.2.

The second incident is similar to the first incident. At Friday 22 August 2008 a car crashed at around 16.50 hours during the afternoon peak hour. The location of the incident is



**Figure 5.2: The traffic situation at 23 June 2008 at 16.00 hours (accident 1)**

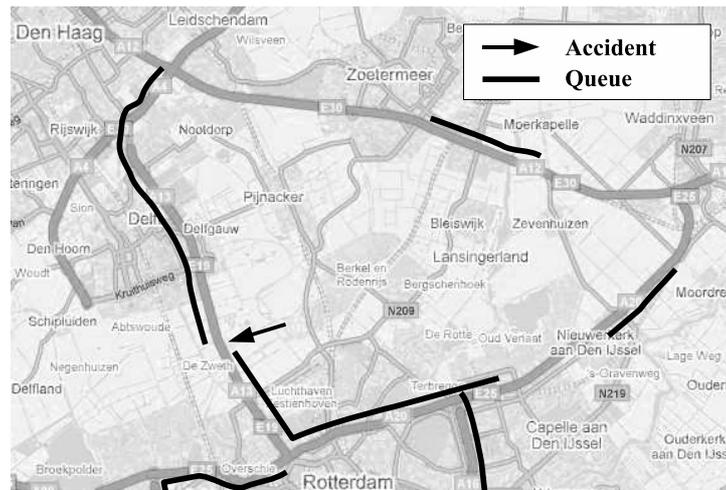


**Figure 5.3: The traffic situation at 22 August 2008 at 15.45 hours (accident 2)**

marked with a 2 on the map in figure 5.1. Also in this case, traffic at the A13 southbound was delayed during the afternoon peak hour. The resulting traffic conditions are shown in figure 5.3. The traffic conditions on this rainy Friday are compared with other Fridays with similar amounts of rain, being 1 August, 5 September, 12 September and 3 October 2008.

For both incidents, the jams are not more than a few kilometers. Since the road is not completely blocked, we expect people not to take detour II or III (as explained in section 5.3.1). For both these incidents we therefore consider the amount of traffic taking alternative route I.

Accidents 3 to 5 cause a larger disruption of the traffic flow for traffic and happen in the northbound direction. The third incident takes place at 12 December 2007 at the A13 in northbound direction. At the accident location, indicated with a 3 on the map in figure



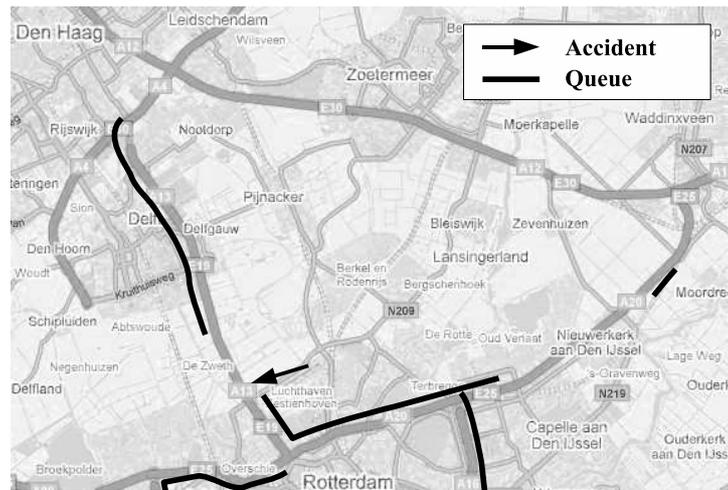
**Figure 5.4: The traffic situation at 12 December 2007 at 7.30 hours (accident 3)**

5.1, two out of the three lanes are closed in the morning peak from 6.25 am to 8.05 am. This resulted in long queues: 5 km on the A13 and then a spillback queue on the A20 westbound of more than 7.5 km, which means that it spills back further than the motorway junction indicated with a “B” in figure 5.1.

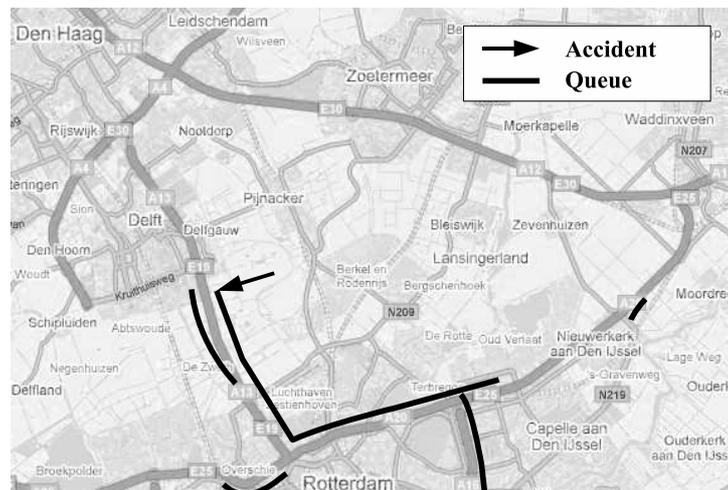
The resulting traffic jams are depicted in figure 5.5. The weather was clear at 12 December 2007, the temperature was a few degrees Celsius and there was no precipitation. The Wednesdays between 19 December and Wednesday 2 January were not considered to give good reference on the normal traffic state because many people in the Netherlands might take off around Christmas and New Year. Therefore, other days in November and January are taken as reference (not in early December because the weather was different). Only 14 November 2007 and 16 January 2008 were suitable.

At the fourth day we consider, Friday 22 February 2008, an incident happened at the beginning of the morning peak hour at around 6.30 hours in the northbound direction at the location indicated with a 4 in figure 5.1. Later a second incident happened in the southbound direction at about the same position. When the accident was removed at around 8.45 hours in the morning, a new accident happened in the northbound direction at the position indicated with a 5 in figure 5.1. In spite of both accidents, the road was never completely closed; however, the northbound traffic was seriously hindered by the congestion. The traffic situation at 7.30 hours in the morning is shown in figure 5.5. The weather at 22 February 2008 was cloudy with some rain and therefore also cloudy and rainy Fridays are used as reference: 8 February, 15 February and 7 March 2008.

The fifth situation we analyse is an incident happening at the A13 in northbound direction (see “5” in figure 5.1). In the tail of the queue, a second, larger, incident took place. To have enough safety during the emergency work, the police closed down the road completely. The results are long queues in northbound direction, spilling back on the A20 and even on the A16, as depicted in figure 5.6. The incident took place at Monday 22 September 2008. As reference we take other sunny Mondays in September and October:



**Figure 5.5: The traffic situation at 22 February 2008 at 7.30 hours (accident 4)**



**Figure 5.6: The traffic situation at 22 September 2008 at 13.30 hours (accident 5)**

15 September, and 6 and 13 October 2008.

For case 3 to 5, the traffic volume taking detour I is not analysed, since the main delay takes place south of the part for which route I provides an alternative (see section 5.3.1). Therefore, we only analyse the use of the alternative routes II and III during this incident for travellers coming from the south (“D” in figure 5.1).

For accident 4 and 5, the larger accidents, historical news messages could be found traced on a Dutch website (nu.nl). In both cases it was stated that traffic had been advised to take the alternative route II or III. That route advice could lead to an increased part of the travellers taking an alternative route.

## 5.5 Observed route choice changes

This section presents the actual route choices and split fractions which come out of the data analysis. In the first part of this section, the raw numbers of travellers and their route choice are given. At the end of the section the results are summarised in a table.

For all 5 accidents the same analysis is carried out. The result is shown in figure 5.7 to 5.11. It shows the split fraction  $\psi$  (equation 5.1) and the fraction of traffic that takes another route  $\chi$  (equation 5.2). For accident 1 and 2, the split fraction at point A is indicated in figure 5.7 and figure 5.8. For accidents 3 to 5, the traffic to alternative routes II and III is considered and therefore the decision at point B is analysed. First, the two smaller incidents (1 and 2) are discussed and then the 3 larger incidents (3 to 5).

### 5.5.1 Minor incidents

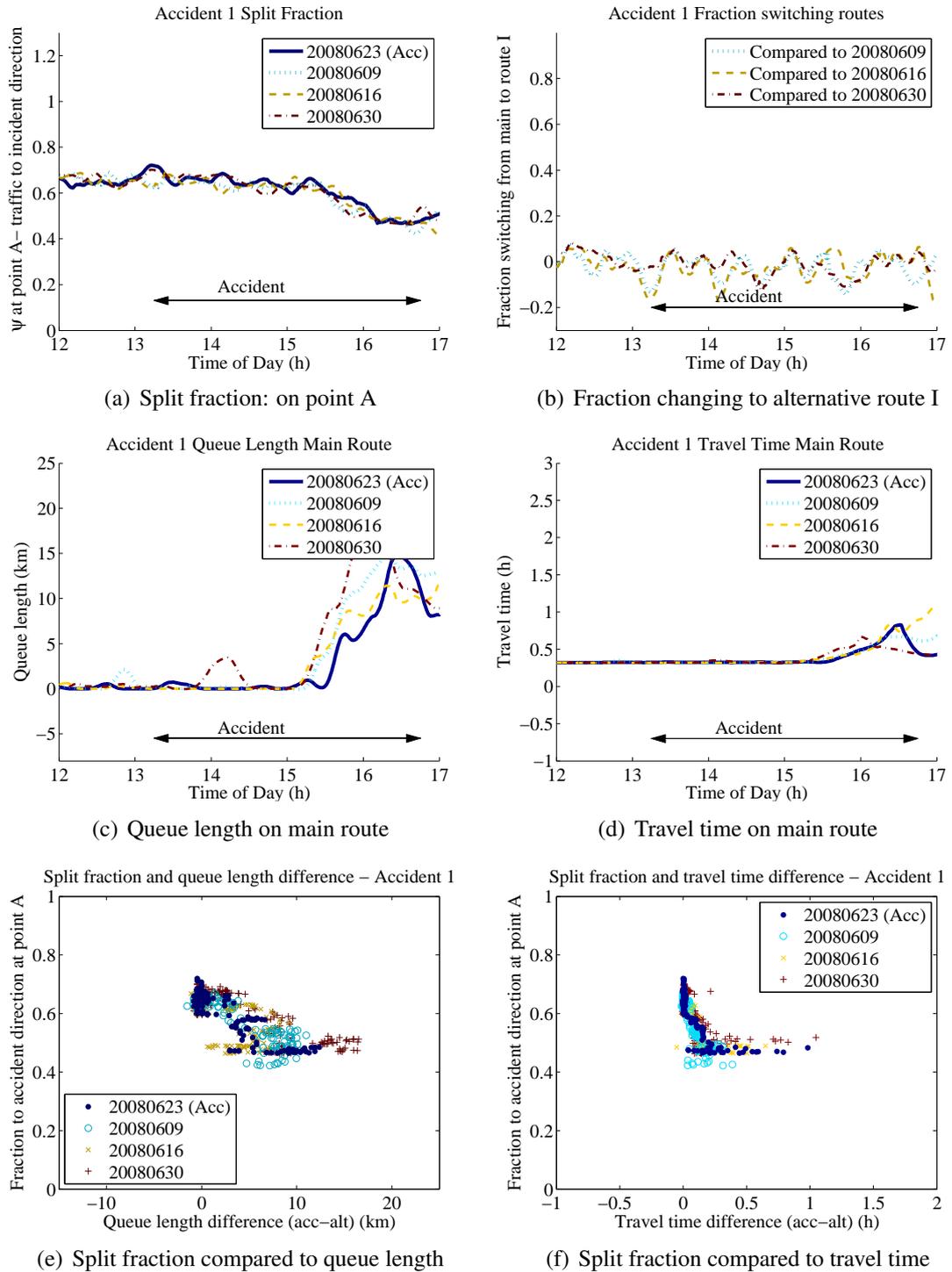
The different lines on figure 5.7a indicate the split fraction over the different days. The figures show that the split fraction at the day of the incident does not differ compared to the route choice on the other days. This is also seen in figure 5.7b, where  $\chi$  is plotted. The fraction of travellers changing route is close to zero. Figure 5.7c and 5.7d show that this accident does not cause much extra delay. In fact, the southbound direction is always busy which means that people are used to drive in congestion in that part of the road. The main difference is that the queue now is further upstream and an alternative route could have been taken.

In case of accident 2 more or less the same phenomena are observed as in accident 1. Although there is an incident, the delays are minor compared to the normal congestion: for some of the reference days, the congestion is even (much) more. The normal congestion, however, is caused downstream, on the Rotterdam ring road. This is all included in the queue length and travel time functions plotted in figure 5.8c and 5.8d. Now, the congestion already starts a bit more upstream and then there is a part free flow driving, south of the point where alternative route I joins the main route. The first part of the queue could be avoided by taking route I. In total, it would be (a bit) shorter to change routes. In this situation, no changes in route choice are observed (figure 5.8a and 5.8b).

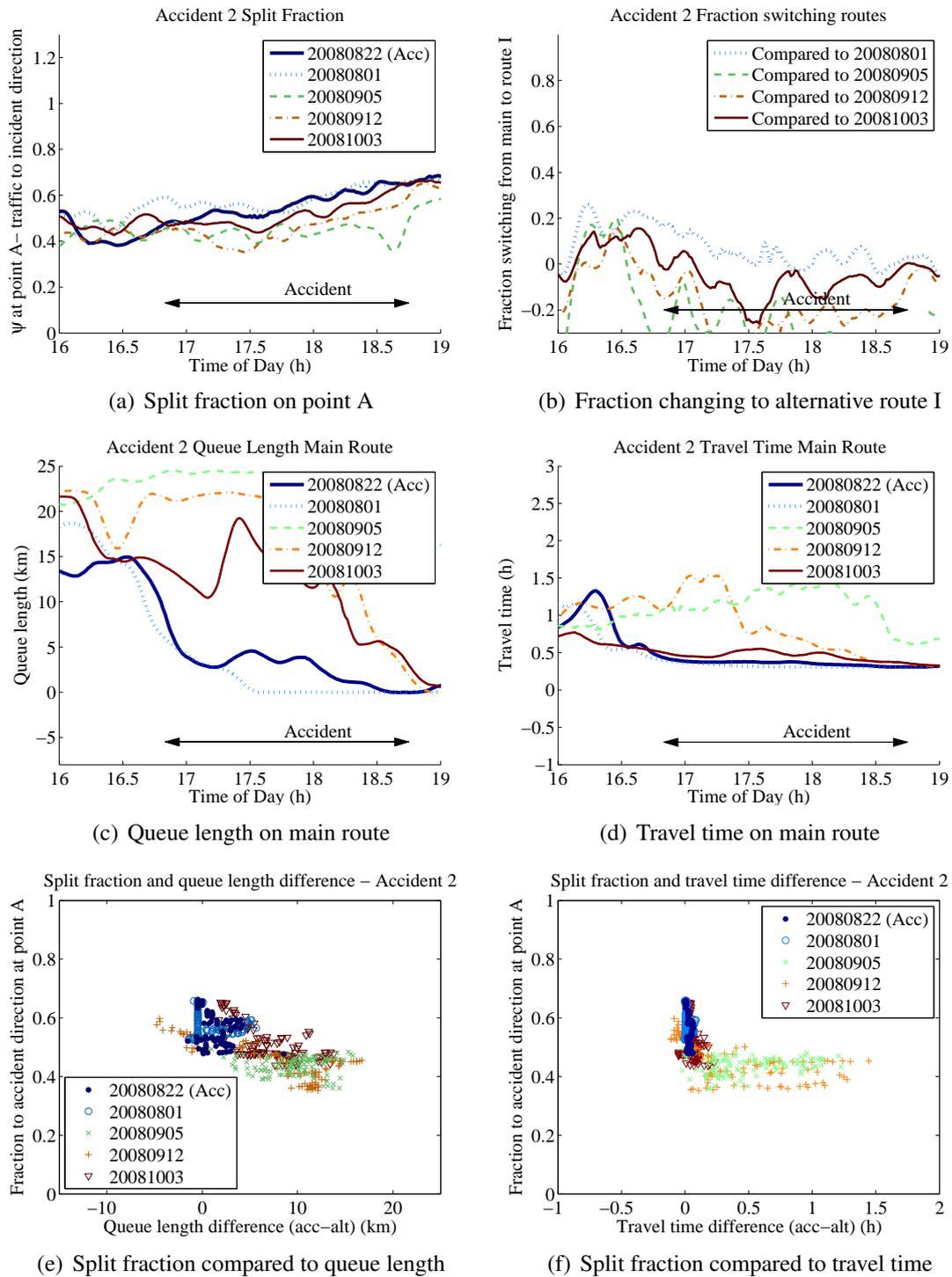
### 5.5.2 Major incidents

For accident 3 to 5, we analyse the split fraction at point B, indicating how much traffic will deviate to routes II and III. For accident 3 the split fraction considerably changes, as shown in figure 5.9a. This shows that between 7h30 and 8h45 the traffic arriving at the motorway junction takes an alternative route. A t-test shows that the difference is statistically significant (p-value < 0.01) between 7h25 and 8h10.

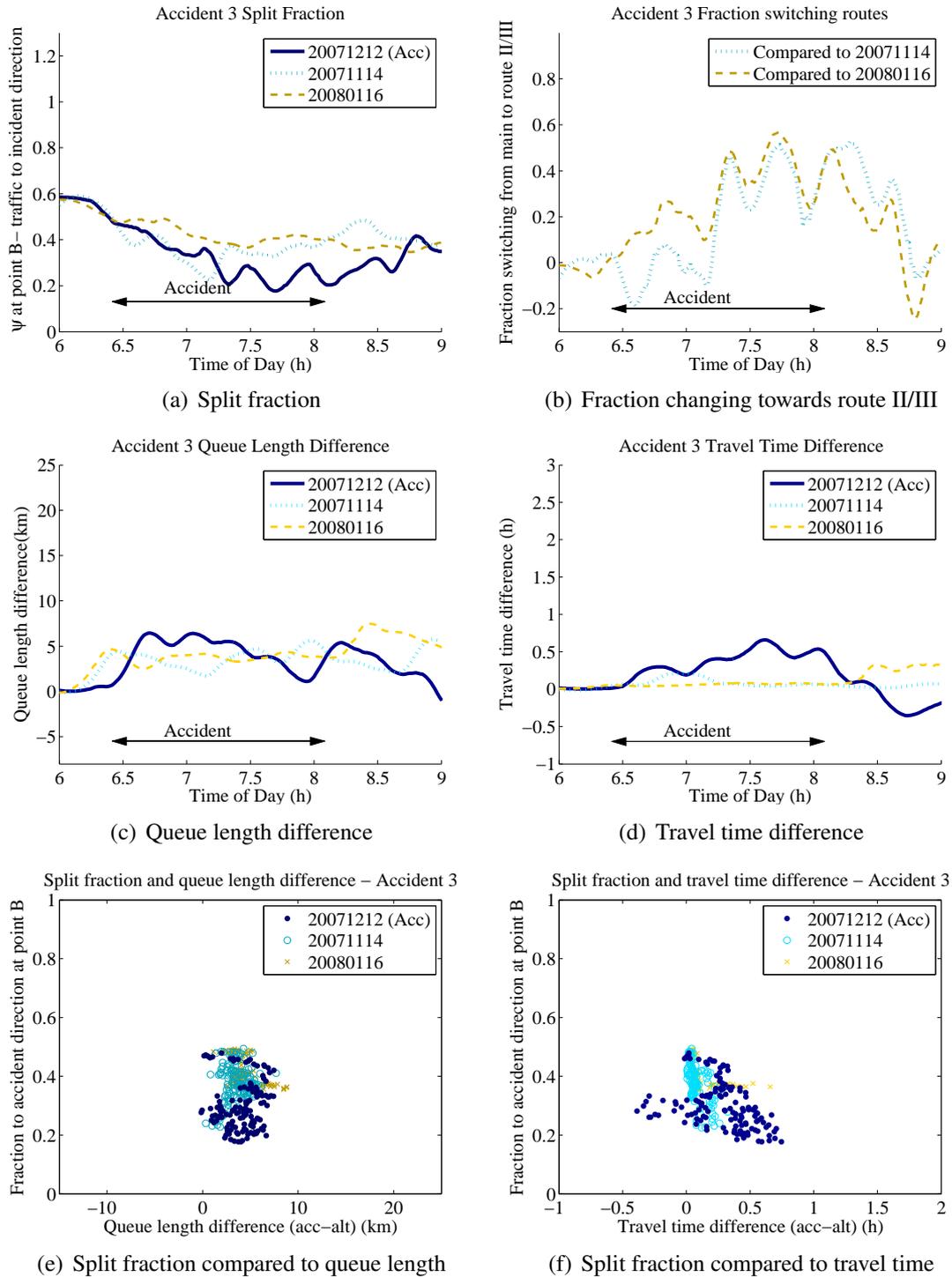
Using equation 5.2, we calculate the fraction of travellers taking an alternative route. Figure 5.9b shows that up to 30% of the travellers changes their route. It also shows the



**Figure 5.7: The resulting route choice for accident 1**



**Figure 5.8: The resulting route choice for accident 2**



**Figure 5.9: The resulting route choice for accident 3**

travel time difference at one moment, the instantaneous travel time at the alternative route II/III is over 30 minutes shorter (figure 5.9d). The extra flow to the alternative route causes congestion of several kilometers on route II/III. However, the amount of congestion on the normal, main route is much more, as can be seen in figure 5.9c. Figure 5.9e shows traffic changes more to alternative routes with similar queue lengths. In figure 5.9f it can be seen that there is a larger travel time difference than normal which could be the cause for travellers to take an alternative route.

At accident 4 the situation is similar to accident 3. The split fraction to the original route strongly reduces as can be seen in figure 5.10a. The fraction of people switching routes, computed by equation 5.2, is shown in figure 5.10b. This percentage increases up to values above 50%. The queue length on the main route is up to 12 kilometers in length, whereas the queue length on the alternative route never exceeds several kilometers; the queue length difference is plotted in figure 5.10c. Alternative route III is up to 45 minutes shorter, whereas in normal conditions they have the same travel time (figure 5.10d). Figure 5.10e shows that there are longer queue lengths at the normal route and less travellers taking the normal route.

Figure 5.10 showed that the travellers will not immediately deviate after the incident has happened. Also, they will not change back to the original route once the incident has been cleared. This delayed reaction causes a hysteresis loop in Figure 5.10f.

For accident 5 the split fraction differs considerably, as depicted in figure 5.11a. From the travellers normally taking the A13, 40% (see figure 5.11b) takes another route. This fraction is less than for accident 4, although the main road now is blocked. From 13.00 hours to 15.00 hours there is a significantly higher fraction of travellers turn towards the route II and II, although the incident itself has been cleared at 14.20h hours.

The extra volume to the A20 east is around 1000 vehicles per hour. This extra volume is supported at the motorway, but not at the underlying roads at the connection points C1 and C2 (see figure 5.1 and 5.11). Therefore, extra congestion sets in on the motorway just upstream of C1, the off ramp to the road connecting both motorways of the alternative route. This extra congestion is around 7 kilometers. This can be an explanation why the fraction of travellers choosing another route is lower than at accident 4. Furthermore, it is interesting to see that at the end of the incident, the queue on the main route rapidly decreases, whereas the queue on the alternative route takes longer to solve since it also includes a secondary road with a low capacity. Therefore, shortly after the incident has been cleared, the congestion on the alternative route is longer than on the main route.

The travel time fluctuates in the same way as the queue length as seen in figure 5.11. During the incident the alternative route is up to 30 minutes quicker. However, after the incident has been cleared, the travel time on the alternative route is around 15 minutes longer than on the main route. An explanation for the fact that the alternative route is more congested than the main route where the incident took place could be that people still did not get an update of the travel advice to take an alternative route.

Figure 5.11e and f clearly show that the incident is an extra incentive for drivers to take an alternative route. For the same queue length difference or travel time difference, a larger

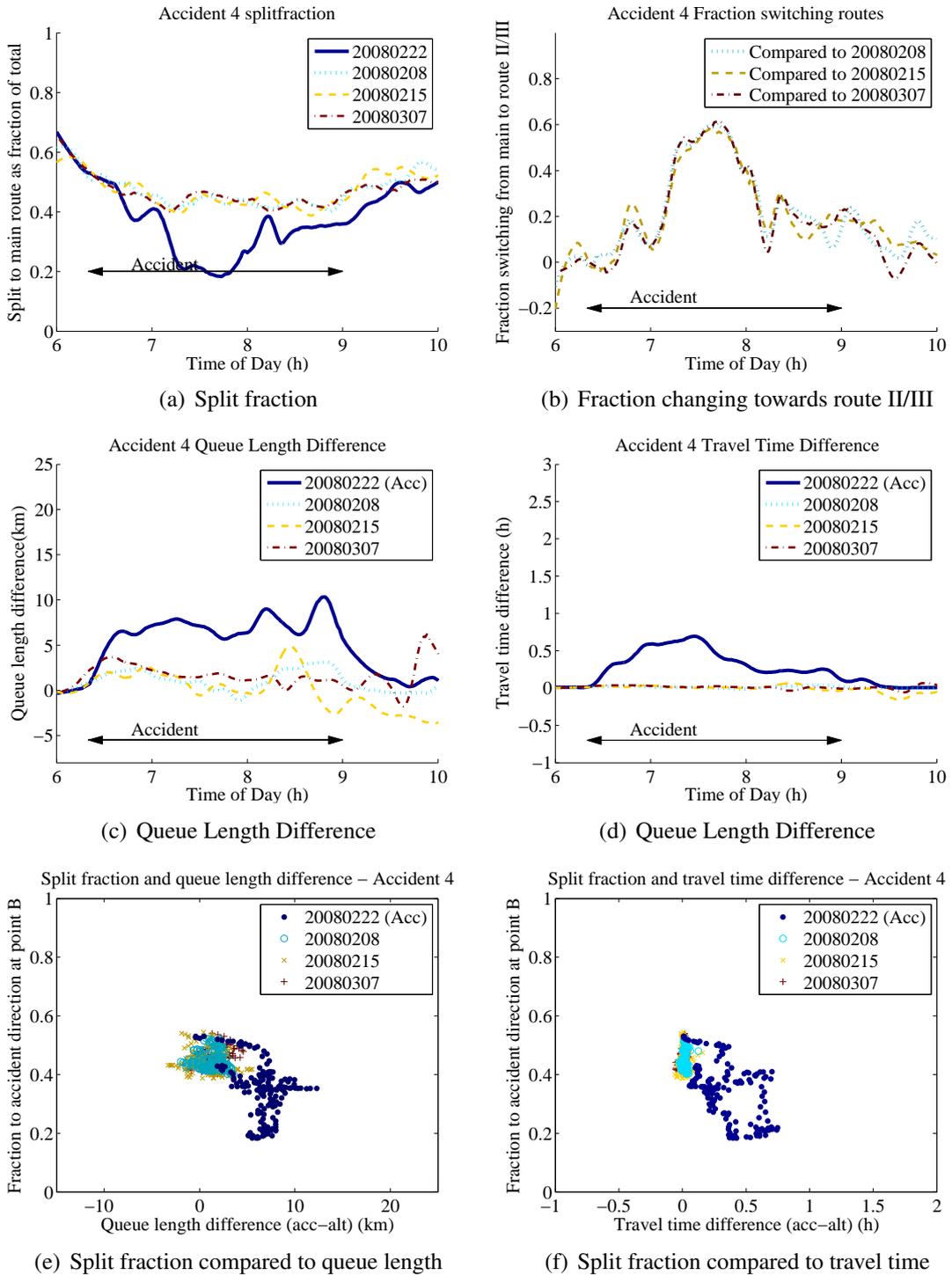
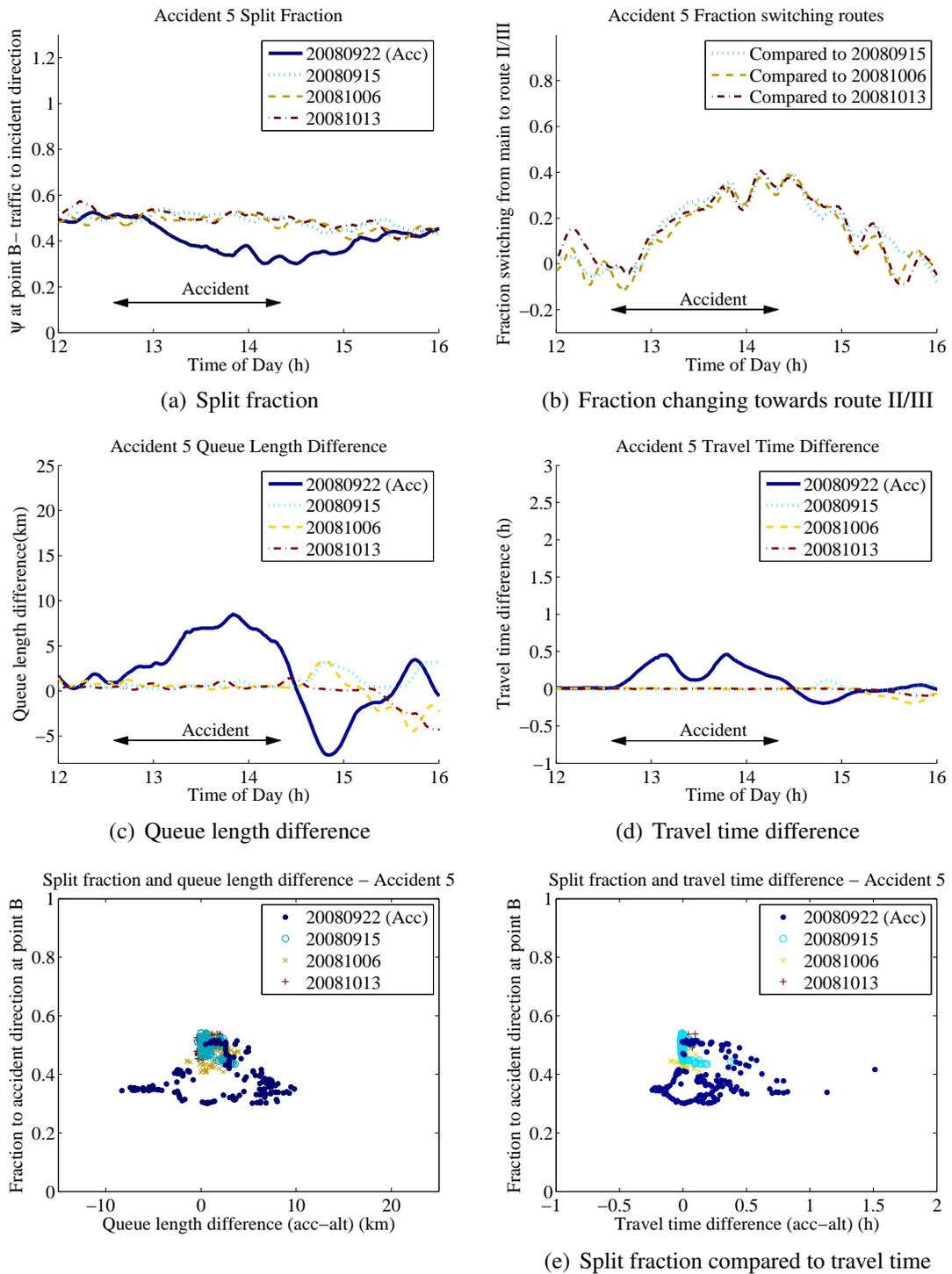


Figure 5.10: The resulting route choice for accident 4



**Figure 5.11: The resulting route choice for accident 5**

percentage than normal takes an alternative route. Also at this incident, a hysteresis loop caused by the delayed reaction of drivers on the traffic conditions is visible in figure 5.11 e and f.

Table 5.5 summarises the results. It shows a clear difference in route choice between ac-

**Table 5.5: Summary of the results**

Number	1	2	3	4	5
Peak	Evening	Evening	Morning	Morning	No
Extra delays	Minor	Minor	Major	Major	Major
Road completely blocked	No	No	No	No	Yes
Percentage changing routes	0	0	30%	>50%	40%

idents 1 and 2 at one hand and accidents 3 to 5 at the other hand. In the first two cases the queue is shorter, but still the alternative route is shorter than the original, congested route. The nature of accidents 3 to 5 is different. Those accidents reduce the capacity considerably; for accident 5 the road is even completely blocked. This means that the traffic is moving very slowly over more than 10 km. Therefore, one of the reasons for traffic to deviate from their intended route could be the length of the queue they are facing. If it is long, then the travellers will change their route, but when the queue is not longer than several kilometers, they will not change routes even if the alternative route has become slightly faster than the original.

## 5.6 Conclusions and future work

The research shows to which extent travellers change their route if the intended route is blocked due to an incident. Hereto, the actual route choice behaviour is studied for five incident cases. It is found that the severity of the capacity reduction incident and therefore the severity of the delays play an important role in the decision to deviate from the intended route.

A considerable percentage of the travellers will take an alternative route when faced with an incident on their intended route, even causing congestion on the alternative route. The percentage found changing their routes could be more than 50% even though it is not yet the last possibility to get off the motorway. This means that traffic which is assumed to take the original route still has possibilities to avoid the queues by taking another alternative route. Also, traffic having a destination before (i.e. upstream of) the bottleneck is now included in the volume of traffic that might consider to change their route. Both these effects mean that the fraction that changes its route during an incident as found in this study is a lower bound for this value.

If it were possible to track the individual vehicles over two different routes, there would have been no need to make the assumption that all traffic needs to pass the bottleneck. This could be achieved for instance by data from licence plate cameras or track vehicles by GSM signal. One would expect that in that case the fraction of deviating traffic is even higher than the values found here. This could be an interesting approach for the future, but that requires a different type of data which was not (yet) widely available at the time of this research.

It is also found that travellers actually avoid a route passing an incident location. The split fractions towards the normal route is lower in case of an incident compared to a situation with the same travel time difference but without an incident. Finally, it is shown that travellers have a delayed reaction on the traffic situation.

This delayed reaction can be due to the delayed information on which they react. It would be therefore be interesting to analyse the amount of information that is given to the drivers. For instance, Dynamic Route Information Panels could be used to provide information. In addition, nowadays, more and more information about the traffic state is presented in-car using a navigation device with Traffic Message Channel (TMC) options or – in the Netherlands – a “Live Traffic” function which gives information on the traffic speeds in the network based on the speeds of other cars equipped with the same navigation devices. In this study it was impossible to track back which information was communicated to the drivers, and at which moment. This would be a valuable addition for future research. Moreover, it would be interesting to see how the values of traffic adapting their route change over time as the penetration rate of dynamic navigation systems increases. Finally, an interesting new development is traffic information which predicts the situation for the future, rather than one which presents the current traffic state. An system which could reliably predict the traffic situation is potentially very valuable for road users since they will be advised the best possible route for their intended trip.



## Chapter 6

# Stochastic incident duration – impacts on delay

---

The duration of incidents is a stochastic variable. This chapter analyses the consequences of this stochastic nature of the duration in terms of average delay. It uses shockwave theory to describe emerging traffic states in space and time. As opposed to a point queue model, the head and the tail of the queue are separately modeled and in this way the spatial extent of the queue is described properly. Using the traffic states, the delay is analytically calculated for different basic network elements.

The chapter distinguishes between three scenarios: (1) an incident happens on a continuous road stretch; (2) an incident happens upstream of a junction; a queue forms upstream of the incident and capacity of the downstream links is insufficient to handle the queue discharge rate; (3) an incident happens downstream of a junction and the tail of the queue crosses the junction.

We derive a formula for the total delay. Because the delay is a non-linear function of the duration, the expected delay is not equal to the delay of the incident with the expected duration. In the scenarios without spillback (the first two scenarios), the delay is proportional to the square of the blocking duration. The expected delay is expressed as a function of the variance of the blocking duration. Also, the variance of the average delay per involved traveler is expressed as function of the variance of the delay. In case spillback occurs, the delay grows faster than proportional to the duration squared. This chapter is an edited version of Knoop, V.L., Hoogendoorn, S.P. and Van Zuylen, H.J. (2010) Stochastic Incident Duration - Impacts on Delay, in: *Proceedings of the 89th Annual Meeting of the Transportation Research Board, Washington D.C.*, 10-14 January 2010, Washington D.C. Accepted for publication.

---

## 6.1 Introduction

According to Bates et al. (2001) road users dislike unexpected delays even more than an expected delay. Models have been proposed to compute the delay caused by an incident. However, they often do not include the spatial extension of a queue or the stochasticity in incident duration or remaining capacity at the incident location.

It is important to compute the cost of the delay of an incident accurately. Besides other effects, the costs of delays play a role in deciding upon investment for instance in safety measures and incident management. This chapter will show that it is important to include network characteristics in computing delays. For each of the three general network layouts, an equation is formulated to compute the total delay caused by an incident with a variable duration. For this computation, only the average duration and the variability are needed; the shape of the distribution function is not needed. Also, the variability of the individual delay can be expressed as function of (only) the average duration and the variation of the duration.

This contribution is as follows: first, a review is given of the literature on incident duration models and models that predict the delay caused by an incident. Section 6.3 describes the theory of queuing and explains the shockwave theory that we will use in this contribution. It also gives the formulas to compute the delay. The theory is applied to three test cases, that are a basic set of network elements, that are described in section 6.4. Section 6.5 shows the traffic states which are predicted by shockwave theory and computes the delays for deterministic durations and capacity reductions. Section 6.6 shows how the case studies can be used for general road layouts. Section 6.7 deals with the stochasticity and discusses how uncertainties are processed. It also shows numerically the impact of the variation of the duration of the incident on the delay. The concluding remarks can be found in section 6.8.

## 6.2 State of the Art

This section discusses the literature on two subjects. In section 6.2.1 the duration of incidents is discussed and how the variability of this duration can be reduced by classifying the incidents. Section 6.2.2 gives an overview of the literature on calculating the delay of incidents.

### 6.2.1 Incident duration prediction

There are several models to estimate the incident duration. It nevertheless remains difficult to get an accurate prediction of the incident duration at the moment an incident starts. This section gives an overview of the methods that are reported in literature to predict the incident duration.

A simple approach would be to take some (independent) observable variables and fit a linear-regression model on the duration of the incident. This is done, for instance, by Khattak et al. (1995) and Garib et al. (1997). For the Dutch motorways a linear regression model is fitted by Knibbe et al. (2006). These models show that fitting a model gives an indication of the incident duration, but the indications are not very accurate.

Wang et al. (2005) predict the incident duration using two different models. Based on four input parameters (report mechanism, vehicle type involved, time of day, location), they predict the time a vehicle that breaks down remains at the same place using a fuzzy logic model and an artificial neural network. They conclude that the artificial neural network performs better with a root mean square error of 20 minutes and an  $R^2$  (see appendix B) of 0.4. The models “had difficulties in predicting the outliers”, by which they mean that they did not find a cause for the cases in which the duration was very long. They conclude that the errors might be due to imperfect or insufficient information about the incident type.

Boyles et al. (2007) apply Bayesian techniques. They use many variables (44) that characterize an incident. Their paper focusses on how characteristics of the incidents become available over time and how the incident prediction gets more accurate using more information. They use a naïve Bayesian model to classify the incidents in three duration groups: less than 30 minutes, 30-60 minutes and more than 60 minutes; the final decision to which group the incident belongs is made by a probability maximization for each of the groups. The chapter shows a (small) increase of incidents which were classified in the right group, compared to a linear regression model. When information was added on the time at which information became available, the improvement increased.

Zhang et al. (2007b) propose to combine the linear regression model of Knibbe et al. (2006) with a classification tree. Compared to a normal linear regression model, they find a better prediction model. Zhang et al. (2007b) explicitly mention that in 10% of the cases, the incident duration is much longer than the model predicts. Therefore, they neglect these cases. However, particularly these cases cause the largest delays. Therefore, in calculating the duration we will keep the incident duration variable. This contribution will show that the variability of the incident duration causes a large part of the total delay.

## 6.2.2 Delays due to incidents

Fu and Rilett (1997) discuss the influence of a stochastic incident duration on the delay. They conclude that using a mean value for the duration leads to an error in the delay. They show how one could use the probability density function to calculate the delay. They also show how this can be used on-line, during the incident, applying Bayesian theory. In their analysis they assume that the traffic jam does not occupy any space (vertical queuing model, Vickrey (1969)).

Olmstead (1999) discusses the delay on the road as a result of an incident. He derives an equation for the delay for all travellers,  $D$ , in case of a deterministic duration  $\Delta T$  (and a vertical queue):

$$D(\Delta T) = \frac{1}{2} \frac{\Delta T^2 (C - rC) (Q - rC)}{C - Q} \quad (6.1)$$

The equation is presented using the symbols described in table 6.2:  $C$  is the capacity,  $r$  the capacity reduction, and  $Q$  the demand. They all are kept constant throughout the incident duration  $\Delta T$ , thereby assuming a uniform distribution of the arrivals.

Using this formula, he shows that the expectation value of an incident with a stochastic duration is larger than the delay of an incident with the expectation value of the duration. He proposes a additional term, linear with the variation of the duration,  $\text{Var}(\Delta T)$ , to compute the expectation value (indicated with angle brackets) of the total delay of an incident with a *stochastic* duration:

$$\langle D \rangle = D(\langle \Delta T \rangle) + \frac{1}{2} \frac{(C - rC)(Q - rC)}{C - Q} \text{Var}(\Delta T) \quad (6.2)$$

The article mentions explicitly that this only holds for this vertical queuing model.

Li et al. (2006) show that the average delay per (delayed) traveler can be calculated using the average duration, but that variation of the delay per delayed traveler has the same problem as the total delay. They present the following equation to compute the variation of delay per delayed traveler  $A$ :

$$\text{Var}(A) = \frac{(Q - C)^2 \text{Var}(\Delta T)}{3Q^2} + \frac{(Q - C) \langle \Delta T \rangle}{12Q^2} \quad (6.3)$$

Note that this is the variation of the delay for the travelers that are delayed. The equation does not provide the number of travellers that encounter this delay.

Both articles assume vertical queues (or point queues). In this chapter the spatial dimension is also considered. We will show how these equations change if the queues have a spacial extent.

## 6.3 Theory for mathematical formulation of queue lengths

This section explains the queuing theory used in this chapter. The first subsections discusses the flow-density relationship that is used. The second subsection presents a mathematical expression for the speed of the boundary between phases. The third subsection presents a way to compute the delay from a space-time state diagram.

### 6.3.1 Fundamental Diagram

The traffic flow modeling using shock wave theory assumes there are distinguished traffic states. A traffic state is uniquely characterized by the flow,  $q$  and the density  $k$ . A “fundamental diagram” poses the relationship between the density  $k$  and the flow  $q$ . We use a triangular fundamental diagram. This is a relatively simple fundamental diagram,

but traffic can be modeled well using this shape, as shown by Daganzo (1997)

$$\begin{aligned} q &= v_f k & k < k_c \\ q &= -(k - k_j) \frac{C}{k_j - k_c} & k > k_c \end{aligned} \quad (6.4)$$

Two branches are separated; traffic states with a density lower than the critical density  $k_c$  non-congested, traffic states with a density larger than  $k_c$  are congested. Furthermore,  $k_j$  is the jam density and  $v_f$  the free flow speed. For an example of the shape, see figure 6.4.

For the sequel of this contribution, it is useful to rewrite equation 6.4 from  $q(k)$  to  $k(q)$ . For one flow, there are two different densities possible, a free flow density and congested density. We will refer to these densities as  $k_2$  and  $k_1$  respectively. The relationship expressed in equation 6.4 now becomes

$$\begin{aligned} k_2 &= \frac{q}{v_f} = \frac{q}{\frac{k_c}{v_f}} & \text{free flow} \\ k_1 &= k_j + \frac{k_c - k_j}{C} q & \text{congested} \end{aligned} \quad (6.5)$$

### 6.3.2 Shock Waves and Their Speeds

Suppose there are two traffic states, A and B, with properties  $(q_A, k_A)$  and  $(q_B, k_B)$  respectively. In shock wave theory, it is assumed that traffic states are homogeneous. The separation between two traffic states will move in time. This moving boundary is, by definition, called a “shock”. The boundary between the two states propagates, by definition, with speed  $\omega_{AB}$ . The speed  $\omega_{AB}$  can be derived from the law of conservation of vehicles and the assumption that the flow depends on the density (see for instance (Helbing, 1997)). It can be expressed as follows:

$$\omega = \frac{q_2 - q_1}{k_2 - k_1} \quad (6.6)$$

The rate at which drivers drive into a traffic jam,  $q_{\text{queue}}$  can be computed by multiplying the traffic density  $k$  with the speed at which they approach the traffic jam ( $v - \omega$ ):

$$q_{\text{queue}} = k(v - \omega) \quad (6.7)$$

This equation can also be used to compute the flow out of a traffic jam. The total number of travelers that drive into a traffic jam,  $N$ , now is the integral over time of equation 6.7:

$$N = \int q_{\text{queue}} dt = \int k(v - \omega) dt \quad (6.8)$$

### 6.3.3 Delays

A consequence of choosing a triangular fundamental diagram is that travelers only encounter delays when driving in a congested traffic state. At every state at the free branch of the fundamental diagram, the speed equals the free flow speed  $v_f$ .

We now compute the delay for a driver in traffic state A, who travels with speed  $v_A$  instead of the free speed  $v_f$ . We compute the delay he encounters in an infinitesimal small time period  $dt$  assuming that in other time periods he travels without delay. The infinitesimal delay  $dl$  in  $dt$  is:

$$dl = \frac{v_f - v_A}{v_f} dt \quad (6.9)$$

In total, there are  $N_A(t)$  drivers in traffic state A at moment  $t$ . The total delay at the infinitesimal time  $dt$  caused by a traffic state now is the product of the number of drivers in that times the delay for a driver.

$$dD = N_A dl = N_A \frac{v_f - v_A}{v_f} dt \quad (6.10)$$

To find the total delay  $D$  for all drivers, the partial delays at  $dt$  (equation 6.10) have to be integrated over time:

$$D = \int N_A(t) \frac{v_f - v_A}{v_f} dt \quad (6.11)$$

## 6.4 Elementary Road Layouts

The equations in the previous section describe how traffic state boundaries propagate and delay can be calculated. However, in a traffic network, the traffic flow dynamics are influenced by nodes. This section lists the node layouts which are possible and mentions the effect on the queue dynamics. Many node configurations are possible. For all link configurations, the incident can happen on an inflow link or an outflow link. The configurations and the incident locations are depicted in figure 6.1, where the number of inflow links and the number of outflow links is restricted to two. Also other node configurations are possible, with multiple input or output links. However, no more than one of the outflow links will restrict the flow over the node, being either a link with an incident or without an incident. Both these possibilities are in the configurations listed here and therefore these configurations cover the basis of all possibilities in road networks.

In the sequel of the contribution, we will discuss three elementary configurations in detail and perform calculations on these. Other configurations listed here are basically a combination of the elements presented in detail, which will be shown in section 6.6.

### 6.4.1 Road layout for considered cases

The road layouts considered in detail are shown in figure 6.2. The location of the incidents

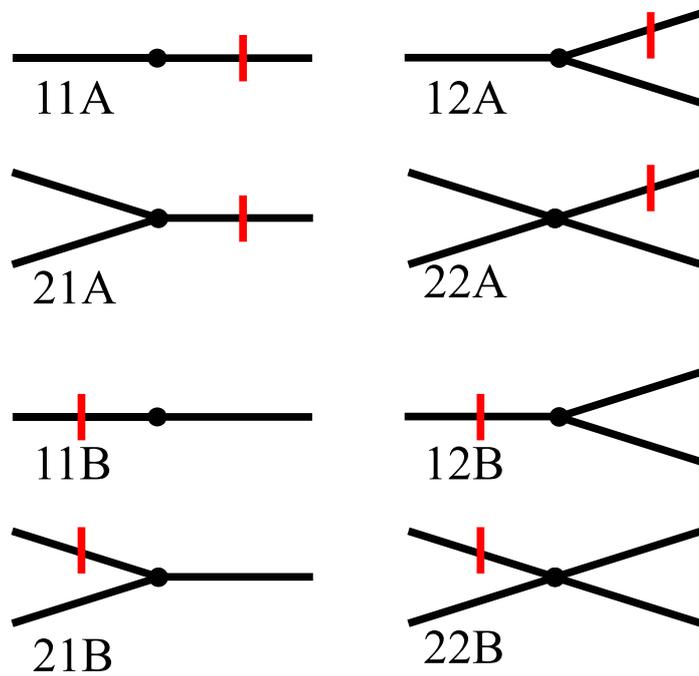


Figure 6.1: Possible road layouts and accident locations with one node

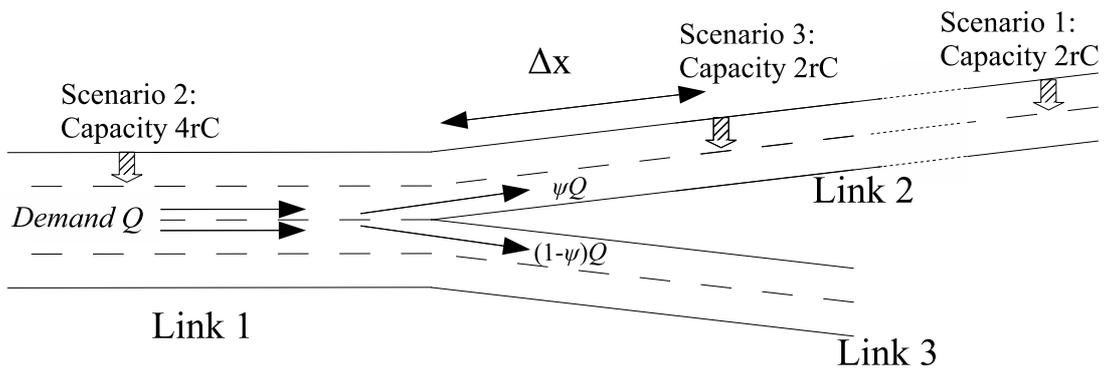


Figure 6.2: The road layout used – incident locations indicated by a shaded arrow

**Table 6.1: The properties of the road**

Property	symbol	value
Capacity	$C$	2200 <i>veh/h/lane</i>
Critical density	$k_c$	25 <i>veh/km/lane</i>
Jam density	$k_j$	150 <i>veh/km/lane</i>
Demand	$Q$	5800 <i>veh/h</i>
Fraction of traffic to link 2	$\psi$	60%

is indicated figure 6.2 with a shaded arrow. In case 1, discussed in section 6.5.1, there is a straight road with a capacity reduction. The next two cases use the network where traffic from 1 link have to split into 2 links. In case 2, discussed in section 6.5.2, a capacity is temporally reduced on the inflow link. After the incident has been cleared, the initial queue dissolves, but the new demand (being the queue discharge rate) is higher than the downstream capacity at the junction. This causes a secondary queue. The third case, discussed in section 6.5.3, has a temporal capacity reduction on one of the outflow links and the resulting queue grows further than the junction.

For the sake of understanding of the concepts, the road properties are specified according to typical values. However, the mathematical formulae are derived for a general case where the capacities of the links are included as model parameters.

The road layout we chose for illustration purposes is a four lane freeway which splits into two times two lanes. Properties of the road are based on (2007c), the Dutch equivalent of the Highway Capacity Manual (2000). The table 6.1 presents these values are presented, as well as the demand for the network.

The capacity at the location of the incident is considered as stochastic variable  $r$ .  $r$  is the fraction of the free capacity that is available during the incident, and varies from 0 to 1.

## 6.5 Applying the model

This section describes the traffic states that are present for three different cases. Table 6.2 shows the symbols that are used throughout the section and at other places in the chapter.

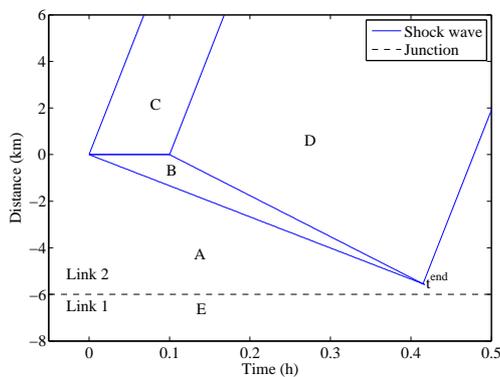
### 6.5.1 Scenario 1: no influence of junctions

#### Traffic states

A typical pattern of traffic states is shown in Figure 6.3. For case in the figure, we choose  $r = 0$  and  $\Delta T = 0.1h$  and we assume the length of the link is long enough that the tail of the queue will not reach the end of the link. We choose the point of the accident to be at

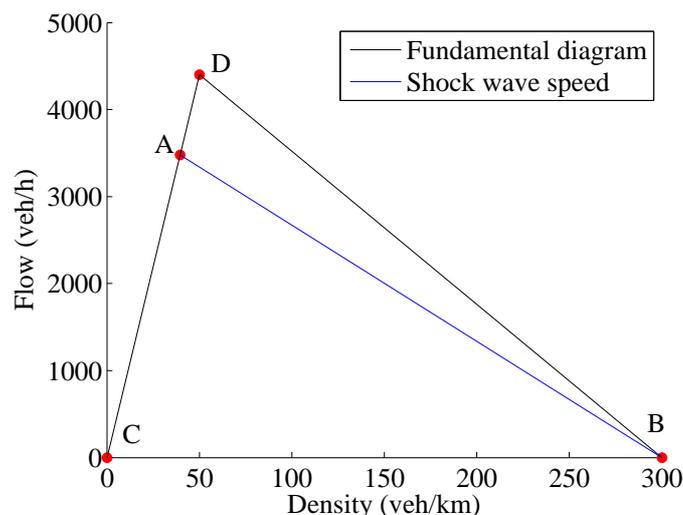
**Table 6.2: The symbols used**

Symbol	Meaning
1, 2, 3	Links
$L$	Length of a queue
$Q$	Demand ( <i>veh/h</i> )
$i$	link number
$\Psi_i$	Fraction of traffic turning to link $i$
$\psi$	Split fraction to link 2
$C_i$	Capacity on link $i$
$r$	reduction factor for capacity due to incident
$q$	Traffic flow
$k_2(q)$	Density of traffic, non-congested regime
$k_1(q)$	Density of traffic, congested regime
$\Delta T$	Duration of the incident
$\omega$	Shock wave speed
$x$	Position at the road
$t$	time
$t^{\text{end}}$	Time when the congestion at the downstream link ends (without intersection)
$x^{\text{end}}$	Position at which the congestion at the downstream link ends (without intersection)
$x^{\text{inc}}$	Distance between the incident and the junction
$A$	Average delay per delayed traveler
$N$	Number of travelers delayed
$q_{\text{queue}}$	The rate at which drivers drive into a traffic jam



State	Lanes	Flow	Congested
A	2	$\psi Q$	Free
B	2	$2rC$	Congested
C	2	$2rC$	Free
D	2	$2C$	Free
E	4	$Q$	Free

**Figure 6.3: Traffic states for a temporal bottleneck where the queue does not spill back to an upstream link**



**Figure 6.4: The traffic states on the density-flow diagram**

$x = 0$  and at  $t = 0$ . During a time  $\Delta T$  the road is blocked at  $x = 0$ . The traffic states and their characteristics are also found in figure 6.3.

Lines mark the separation of different traffic flow regions. The demand at link 2 is  $\psi Q$ , which is the traffic flow in state A. The incident reduces the traffic flow to  $rC_2$ . Downstream of incident there is a free flow traffic state C with flow  $rC_2$ . Upstream of the incident, an area with congestion builds up (B). The flow there also is  $rC_2$ . After the incident has cleared, the outflow of the traffic jam is  $C_2$ , in a free flow state (D). The traffic states are also indicated on the density-flow in Figure 6.4.

Using equation 6.6, we can compute the speed at which the boundary between traffic state A and B travels backwards,  $\omega_{AB}$ :

$$\omega_{AB} = \frac{Q\psi - rC_2}{k_2(d\psi) - k_1(rC_2)} \quad (6.12)$$

Similarly, the boundary between traffic state B and D travels backwards with speed  $\omega_{BD}$ :

$$\omega_{BD} = \frac{Q\psi - C_2}{k_2(Q\psi) - k_1(C_2)} \quad (6.13)$$

This shock wave travels upstream with a larger absolute speed than the shock wave at the upstream end of the congested area. The travel position of the wave fronts in time and space are:

$$x_1(t) = \omega_{AB}t \quad (6.14)$$

$$x_2(t) = \omega_{BD}(t - \Delta T) \quad (6.15)$$

If the link is long enough, the congestion is solved when  $x_1$  equals  $x_2$  (at moment  $t^{\text{end}}$ )

$$\begin{aligned}\omega_{AB}t^{\text{end}} &= \omega_{BD}(t^{\text{end}} - \Delta T) \\ t^{\text{end}} &= \frac{\omega_{BD}\Delta T}{\omega_{BD} - \omega_{AB}}\end{aligned}\quad (6.16)$$

Substituting  $t^{\text{end}}$  in equation 6.14 then gives the position on the road where the two waves meet:

$$x^{\text{end}} = \frac{\omega_{AB}\omega_{BD}\Delta T}{\omega_{BD} - \omega_{AB}};\quad (6.17)$$

During this scenario, we assume that congestion does not reach the junction upstream of the incident location. Having a value for  $x^{\text{end}}$ , this now can be quantified:

$$\begin{aligned}x^{\text{inc}} &> x^{\text{end}} \\ x^{\text{inc}} &> \frac{\omega_{AB}\omega_{BD}\Delta T}{\omega_{BD} - \omega_{AB}}\end{aligned}\quad (6.18)$$

### Delay

To compute the delay, one first needs to know the number of vehicles in the traffic jam. This can be computed from the length of the queue in meters,  $L(t)$ , and the density in the congested area,  $k_1(rC_2)$ . In this case, the length of the queue  $L(t)$  is:

$$L(t) = x_2(t) - x_1(t)\quad (6.19)$$

The number of vehicles in this area,  $N_B$  is computed as:

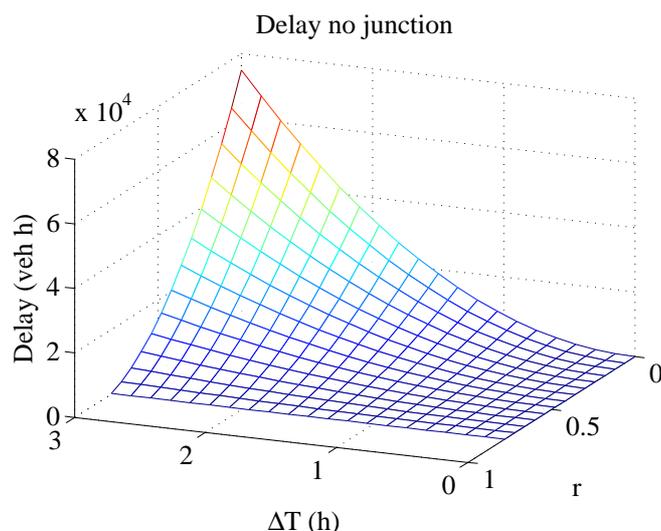
$$\begin{aligned}lN_B &= L_B(t)k_1(rC_2) \\ N_B &= (x_2(t) - x_1(t))k_1(rC_2)\end{aligned}\quad (6.20)$$

For the simplicity of the computations, we split the time during which there is congestion ( $0 < t \leq t^{\text{end}}$ ) into two parts. We distinguish a first part in which congestion builds up ( $0 < t \leq \Delta T$ ), and a second part the congestion solves ( $\Delta T < t \leq t^{\text{end}}$ ). The respective queue lengths are:

$$\begin{aligned}N_B &= \omega_{AB}t & 0 < t \leq \Delta T \\ N_B &= \omega_{AB}\Delta T - (\omega_{AB} - \omega_{BD})(t - \Delta T) & \Delta T < t < t^{\text{end}}\end{aligned}\quad (6.21)$$

If we split equation 6.11 for the two time periods (queue growing and queue solving), we obtain the following:

$$D = \int_0^{\Delta T} N_B(t) \frac{v_f - v_B}{v_f} dt + \int_{\Delta T}^{t^{\text{end}}} N_B(t) \frac{v_f - v_B}{v_f} dt\quad (6.22)$$



**Figure 6.5:** The delay as function of  $r$  and  $\Delta T$  for a long link

Substituting the value for  $N_B$  (equation 6.21) into this equation, gives the following

$$\begin{aligned}
 &= \frac{1}{2} \frac{\omega_{AB}(v_f - v_B)\Delta T^2}{v_f} \\
 &\quad + \frac{1}{2} \frac{(-\omega_{AB} + \omega_2)(v_f - v_1) \left( (t^{\text{end}})^2 - \Delta T^2 \right)}{v_f} \\
 &\quad + \frac{(\omega_{AB}\Delta T + (\omega_{AB} - \omega_{BD})\Delta T)(v_f - v_B)(t^{\text{end}} - \Delta T)}{v_f} \quad (6.23)
 \end{aligned}$$

Note, by the way, that the total delay is not equal to area in  $B$  since it depends on the speeds in  $B$ , see equation 6.10.

All variables in equation 6.23 can be substituted into basic variables, using the relationships stated in equations 6.12, 6.13 and 6.5. This yields:

$$D = \frac{1}{2} \frac{C_2 \Delta T^2 (r - 1)(r C_2 - \psi Q)}{C_2 - \psi Q} \quad (6.24)$$

Note that this is the same function as Olmstead (1999) finds, equation (6.2). However, the assumptions we made are less restricting than his assumptions. Contrary to Olmstead, the traffic jam has a non-zero length. Vehicles enter at the traffic jam when they reach the tail of the queue, which in practice is upstream of the bottleneck location. Furthermore, the traffic jam dissolves from the head of the queue.

As expected (Olmstead, 1999), the delay is proportional to  $\Delta T^2$ . It also is quadratic in  $r$ , but not proportional to the fraction of the capacity that remains,  $(1 - r)^2$ . In fact, it is proportional to  $(r - 1)(r - \frac{\psi Q}{C_2})$ . However, when the total capacity is used,  $\psi Q = C_2$ , then  $\psi Q - r C_2 \propto 1 - r$  and therefore the delay is then proportional to  $(1 - r)^2$ .

For the test case given in section 6.4.1, the shape of the function can be seen in figure 6.5. For this figure, we assumed that no spillback occurs at all, or in other words,  $x^{\text{inc}}$  is larger than the queue length spillback.

### Average delay

The number of drivers encountering delay can be computed using equation 6.8. The number of drivers that encounter delay is:

$$N = \int_0^{t^{\text{end}}} q_{\text{queue}} dt \quad (6.25)$$

Substituting  $q_{\text{queue}}$  (equation 6.7) and  $t^{\text{end}}$  (equation 6.16) now gives:

$$\begin{aligned} N &= \int_0^{t^{\text{end}}} k_A(v_A - \omega_{AB}) dt \\ &= k_A(v_A - \omega_{AB}) t^{\text{end}} \\ &= k_A(v_A - \omega_{AB}) \frac{\omega_{BD} \Delta T}{\omega_{BD} - \omega_{AB}} \end{aligned} \quad (6.26)$$

Note that this number of drivers is proportional to  $\Delta T$ . Because the total delay is proportional to  $\Delta T^2$ , the average delay per delayed driver is proportional to  $\Delta T$ :

$$\begin{aligned} A &= \frac{D}{N} = \frac{1}{2} \frac{C_2 \Delta T^2 (r-1)(rC_2 - \psi Q)}{C_2 - \psi Q} \bigg/ k_A(v_A - \omega_{AB}) \frac{\omega_{BD} \Delta T}{\omega_{BD} - \omega_{AB}} t^{\text{end}} \\ &= \Delta T \frac{1}{2} \frac{C_2 (r-1)(rC_2 - \psi Q)}{C_2 - \psi Q} \bigg/ k_A(v_A - \omega_{AB}) \frac{\omega_{BD}}{\omega_{BD} - \omega_{AB}} t^{\text{end}} \end{aligned} \quad (6.27)$$

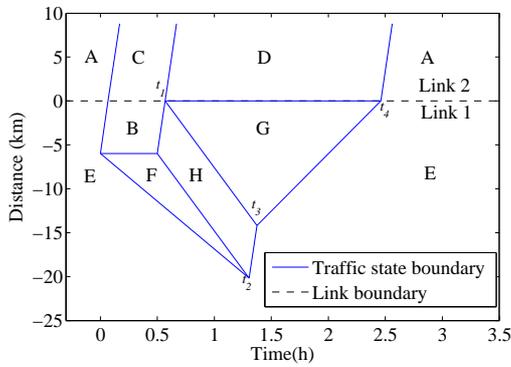
## 6.5.2 Scenario 2: incident upstream of a junction

When an incident happens upstream of a junction, the junction can have an influence on the number of delayed travellers since also travellers to the other direction might be delayed. First, we discuss the traffic states that occur in such a case. Then, the total delay is determined, followed by an analysis which part of the delay is caused by the junction. Finally, the average delay per traveller will be computed.

### Traffic states

Suppose that an accident happens upstream of the junction (on link 1) which reduces the capacity temporarily to  $rC_1$ . Figure 6.6 shows a pattern of a traffic situation that one would typically find (in this case,  $r = 0$  and  $\Delta T = 0.5$  and the incident takes place 6 km upstream of the bottleneck).

This paragraph explains the traffic states as seen in Figure 6.6; the states are also indicated in figure 6.7. Areas A and E are the states in the non-incident situation, for links 2 and



State	Lanes	Flow	Congested
A	2	$\psi Q$	Free
B	4	$4rC$	Free
C	2	$4rC\psi$	Free
D	2	$2C$	Free
E	4	$Q$	Free
F	4	$4rC$	Congested
G	4	$\frac{2C}{\psi}$	Congested
H	4	$4C$	Free

Figure 6.6: Traffic states for a temporal bottleneck on link 1

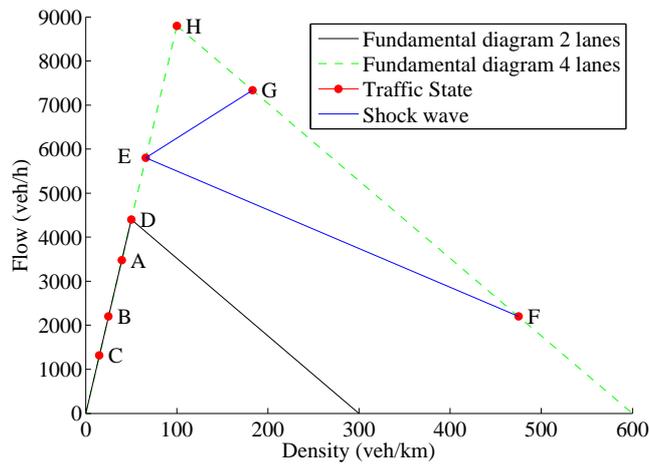


Figure 6.7: The traffic states on the density-flow diagram

1 respectively. There is an incident upstream of the bottleneck. Upstream of the incident a queue builds up (area F), and during the incident the outflow is lower (area B). This lower outflow reduces the demand to link 2 from  $t_0$  to  $t_1$ . After the capacity is restored, the traffic flows out of the traffic jam at the capacity of link 1 (area H). This flow is larger than the original demand. If the capacity of any of the downstream links is lower than the new demand to that link (the split fraction remains the same), a new area of congestion arises (area G). In this case, where  $\frac{\psi}{C_2} > \frac{1-\psi}{C_3}$  link 2 forms a bottleneck in case  $\psi C_1 > C_2$ . Since the demand is high, the flow on the downstream link equals capacity, and on link 1 the flow is maximised by the outflow to the downstream links. In an equation, we could write

$$q_G = \min \left\{ \min \left( \frac{C_i}{\Psi_i} \right), C_1 \right\} \quad (6.28)$$

In the second case (equation 6.28, the demand is lower than the possible flows on the downstream links and the outflow of link 1 is not restricted by any of the downstream links. In that case, the problem reduces to the situation described in section 6.5.1. If the outflow is restricted, the traffic states as introduced in this section are applicable. Therefore, the traffic states as shown in figure 6.6 are only applicable if

$$C_1 > \min \left( \frac{C_i}{\Psi_i} \right) \quad (6.29)$$

### Total delay

Only in areas F and G (see figure 6.6) travelers encounter delay. Using shockwave theory one can compute the total delay. Again, equation 6.11 has to be used to compute the total delay:

$$D = \int N_A(t) \frac{v_f - v_A}{v_f} dt \quad (6.30)$$

This formula is written for the general traffic state A ( $N_A$ ,  $v_A$ ) where this, of course, should be used for areas F and G (i.e., use  $N_F$ ,  $N_G$ ,  $v_F$ ,  $v_G$ ).

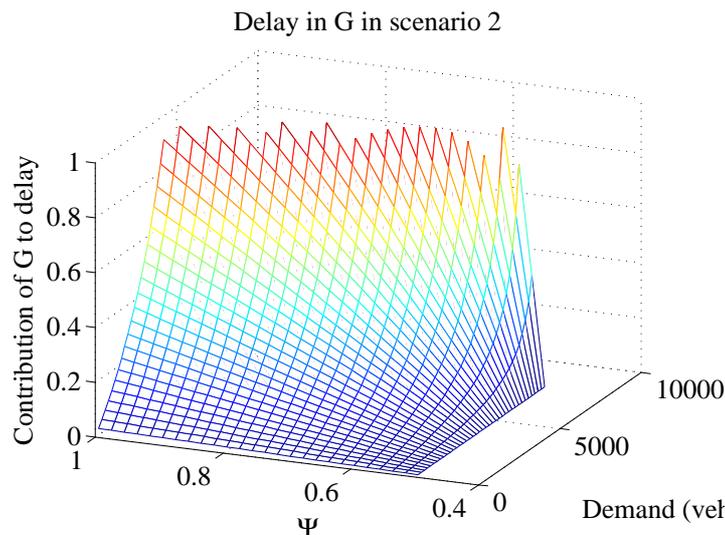
In this formula, the number of vehicles can be substituted by the length of the queue times the density:

$$N(t) = kL(t) \quad (6.31)$$

$k$  can be derived using equation 6.5, whereas  $L$ , the distance between the boundaries of the traffic states can be computed using equation 6.6. This substitutions can be performed manually, or by a computer program. The result of substituting all variables in equation 6.31 gives the total delay in a closed equation:

$$D = \frac{1}{2} \frac{\Delta T^2 (r^2 \psi C_1^2 - C_1 \psi r Q - r C_1 C_2 + C_2 Q)}{C_2 - \psi Q} \quad (6.32)$$

Note that, like in the case without the influence of a junction, the delay is proportional to  $\Delta T^2$ .



**Figure 6.8: The contribution of the total delay that is encountered in G**

### Delay caused by junction

We will now show the impact the delay caused by an insufficient downstream capacity. In particular, we show the part of the delay that is encountered in G as function of  $\psi$ . We use the test case (section 6.4.1) as example. We can assume that  $\psi$  is larger than or equal to 0.5. Since the situation is symmetrical, this assumption is not restrictive because either  $\psi$  or  $(1 - \psi)$  will be. Now we find that the delay in G is determined by the queue at moment  $t_2$  and the flow in G (see equation 6.28).

The initial network can only be congestion free with the following condition:

$$\psi_i Q < C_i \quad (6.33)$$

for all outflow links  $i$ . In our situation where link 2 forms the bottleneck (by convention of numbering)  $\psi$  varies between 0.5 and  $\frac{C_2}{Q}$ . For all demands between 0 and  $C_1$  and for all appropriate split fractions, we plot the fraction of delay encountered in G in figure 6.8.

If the outflow rate ( $C_2/\psi$ ) equals the inflow rate, the queue length does not change, and thus the queue will not solve at all. This means that all vehicles that will pass, encounter delay. In the limit that  $\psi = \frac{C_2}{Q}$ , the queue will never solve. Therefore, all vehicles will encounter the delay and all other delays are negligible.

The queue length can also be described. Figure 6.7 shows that the length of the traffic jam does not decrease as  $\frac{C_2}{\psi}$  approaches the demand  $Q$ , the gradient of the line E – G decreases to 0. This means that the shock wave speed that defines the tail of the traffic jam approaches to zero, and the queue length remains the same.

### Mean delay

Let's finally consider the mean delay per driver encountering congestion. Only in areas F and G (see figure 6.6) there is a speed reduction. The propagation speed of the state

boundaries BH and EH (figure 6.6),  $\omega_{BH}$  and  $\omega_{EH}$  equals the free flow speed  $v_f$  and therefore the speed of the vehicles in area H. That means that the vehicles that enter area G via boundary HG are the same vehicles as exit area F. Therefore, the total number of delayed vehicles,  $N$ , is the sum of the vehicles encountering delay in area F ( $N_F$ ) and the number of vehicles that enter area F via boundary EG ( $N_{EG}$ ):

$$N = N_F + N_{EG} \quad (6.34)$$

Furthermore, we know

$$t_3 - t_1 = t_2 - \Delta T \quad (6.35)$$

For the number of drivers that face congestion in area F, the same holds as for the number of drivers facing congestion in scenario 1 (equation 6.26):

$$\begin{aligned} N_F &= \int_0^{t^{\text{end}}} k_E(v_E - \omega_{EF})dt \\ &= k_E(v_E - \omega_{EF})t^{\text{end}} \\ &= k_E(v_E - \omega_{EF})\frac{\omega_{FH}\Delta T}{\omega_{FH} - \omega_{EF}} \end{aligned} \quad (6.36)$$

To compute  $N_{EG}$  we use equation 6.8:

$$N_{EG} = \int k_E(v_E - \omega_{EG})dt \quad (6.37)$$

$$= (t_4 - t_3)k_E(v_E - \omega_{EG}) \quad (6.38)$$

We will now show that  $(t_4 - t_3)$  is proportional to the duration  $\Delta T$  to show that  $N_{EG}$  is proportional to  $\Delta T$ .

Since FH is parallel to HG and BH is parallel to EH, we can write

$$\begin{aligned} t_3 &= t_1 + t_2 - \text{duration} \\ &= t_1 - \frac{\omega_{EF} * \Delta T}{\omega_{EF} - \omega_{FH}} \end{aligned} \quad (6.39)$$

Therefore, the duration of the growing of area G,  $t_3 - t_1$  is proportional to the duration  $\Delta T$ . This means that the maximum length of queuing area G is

$$\begin{aligned} x_3 &= \omega_{GH} * (t_3 - t_1) \\ &= -\omega_{GH} \frac{\omega_{EF} * \Delta T}{\omega_{EF} - \omega_{FH}} \end{aligned} \quad (6.40)$$

This length decreases with  $\omega_{EG}$  per unit time, so

$$t_4 - t_3 = x_3 / \omega_{EG} \quad (6.41)$$

Substituting equation 6.40 into equation 6.41 gives

$$t_4 - t_3 = \frac{-\omega_{GH} \omega_{EF} * \Delta T}{\omega_{EG} \omega_{EF} - \omega_{FH}} \quad (6.42)$$

This shows that  $t_4 - t_3$  is proportional to the duration and therefore  $N_{EG}$  is proportional to the duration.

We can now compute the total number of delayed vehicles, substituting equations 6.36, 6.38 and 6.42 in equation 6.34

$$\begin{aligned} N &= N_F + N_{EG} \\ &= \frac{k_E(v_E - \omega_{EF})\omega_{FH}\Delta T}{\omega_{FH} - \omega_{EF}} + k_E(v_E - \omega_{EG})\frac{-\omega_{GH} \omega_{EF} * \Delta T}{\omega_{EG} \omega_{EF} - \omega_{FH}} \end{aligned} \quad (6.43)$$

Note that the number of travelers that encounters delay is proportional to the duration  $\Delta T$ . Therefore, the average delay per traveler  $A$ , the total delay (equation 6.32) divided by the number of delayed travellers, is also proportional to the duration:

$$\begin{aligned} A &= D/N \\ &= \Delta T \left( \frac{1}{2} \frac{(r^2\psi C_1^2 - C_1\psi rQ - rC_1C_2 + C_{12}Q)}{C_2 - \psi Q} \right) / \dots \\ &\quad \left( \frac{k_E(v_E - \omega_{EF})\omega_{FH}}{\omega_{FH} - \omega_{EF}} + k_E(v_E - \omega_{EG})\frac{-\omega_{GH} \omega_{EF} * \Delta T}{\omega_{EG} \omega_{EF} - \omega_{FH}} \right) \end{aligned} \quad (6.44)$$

### 6.5.3 Scenario 3: queues longer than the distance to the junction

When an incident happens *downstream* of a junction, on link 2, the traffic states that occur depend on the length of the queue. If the queue is shorter than the distance to the upstream intersection and there is no restricting junction downstream, the traffic follows the pattern described in section 6.5.1. However, when it reaches the upstream junction, different traffic states occur. This section will first discuss the traffic states and the delays, and subsequently the influence of the junction.

#### Traffic states and delays

A typical pattern is given in Figure 6.9. For this particular graph, we use  $r = 0.3$ ,  $x^{\text{inc}} = 6\text{km}$ , and  $\Delta T = 1\text{h}$ .

The congestion that the incident causes, spills back to the more upstream link. From the moment that congestion reaches the junction, the outflow of that link is reduced. In particular, fraction  $\psi$  that would like to turn to link 2 is reduced to  $rC_2$ . Consequently, the flow on link 1 is  $\frac{rC_2}{\psi}$ .

After the congestion on link 2 has solved (at  $t_1$ ) there is still congestion on link 1 if the turn fraction is not 50% – similar to the congestion described in section 6.5.2). The demand to links 2 and 3 equals the capacity of the link 1, rather than the demand. In the example case, the new demand from link 1 is  $4 * 2200\text{veh}/\text{h} = 8800\text{veh}/\text{h}$ . Link 2

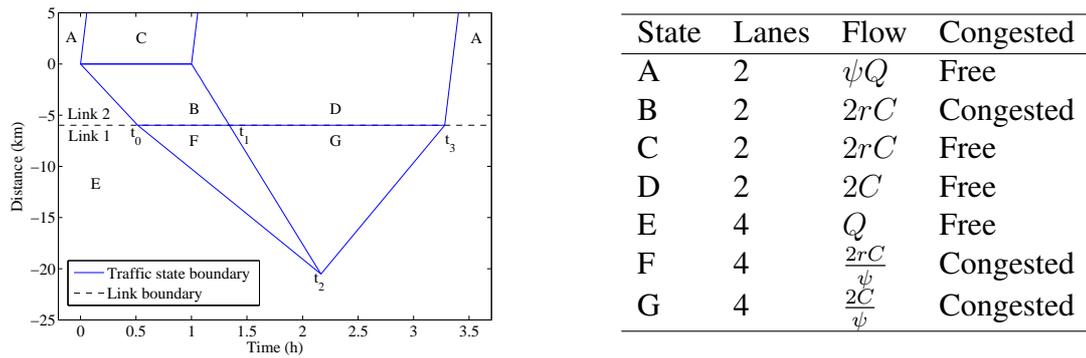


Figure 6.9: Traffic states when the queue is longer than  $x^{\text{inc}}$

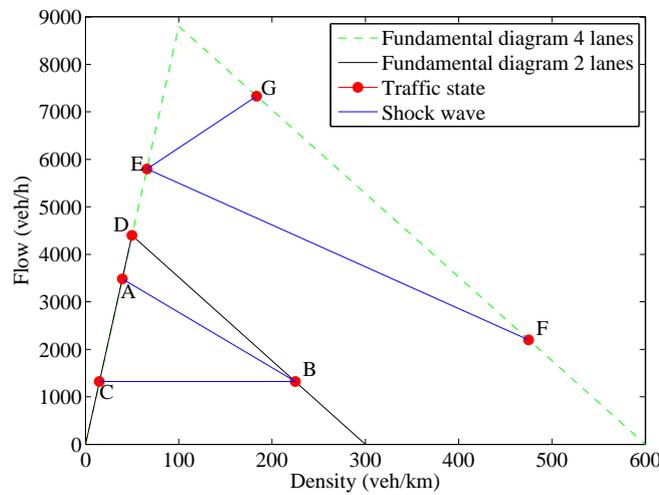


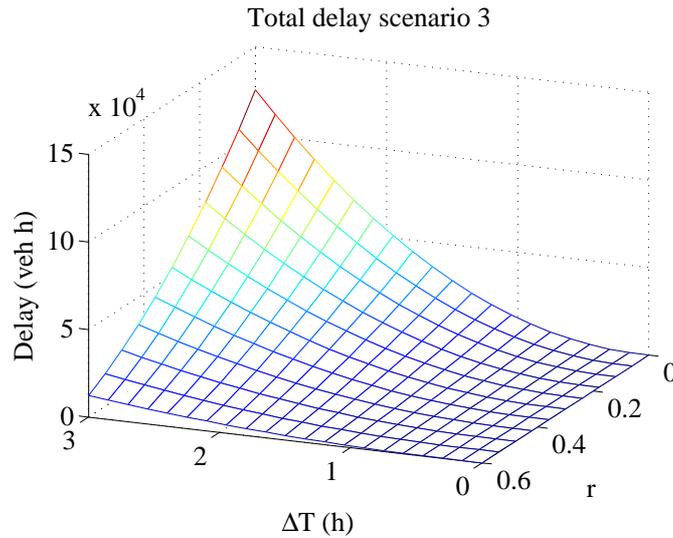
Figure 6.10: The traffic states in case of spillback on the density-flow diagram

cannot accommodate 60% of 8800 vehicles. Instead, the maximum capacity is 2 lanes times 2200  $veh/h/lane$ , hence there is still congestion on link 1. This means that the flow on link 1 is at maximum  $\frac{2 \cdot 2200}{0.6}$ . Depending on the split fraction  $\psi$ , the flow can be restricted by either link 2 or link 3. A general prescription of the flow in area G can be found in equation 6.28. For the layout of our case study, this reduces to:

$$q_G = \min \left\{ \frac{2C_2}{\psi}, \frac{2C_3}{1-\psi}, C_1 \right\} \quad (6.45)$$

The traffic states are separated by boundaries. The speed at which these boundaries propagate can be computed using equation 6.6, the flow values as stated in figure 6.9 and the flow-speed relation in equation 6.5. The traffic states are also indicated as dots in figure 6.10; in this figure, the shock wave speeds are indicated with solid blue lines.

The shock wave speeds determine the size of areas B, F and G in figure 6.9. The delay in these areas can be computed using equation 6.11. Although it is just a simple substitution



**Figure 6.11:** The delay caused by the incident,  $x^{\text{inc}} = 6\text{km}$ .

of variables, the end result is quite long. The resulting delay is expressed as function of the duration and the capacity reduction of the incident. A plot of the delay function is shown in figure 6.11.

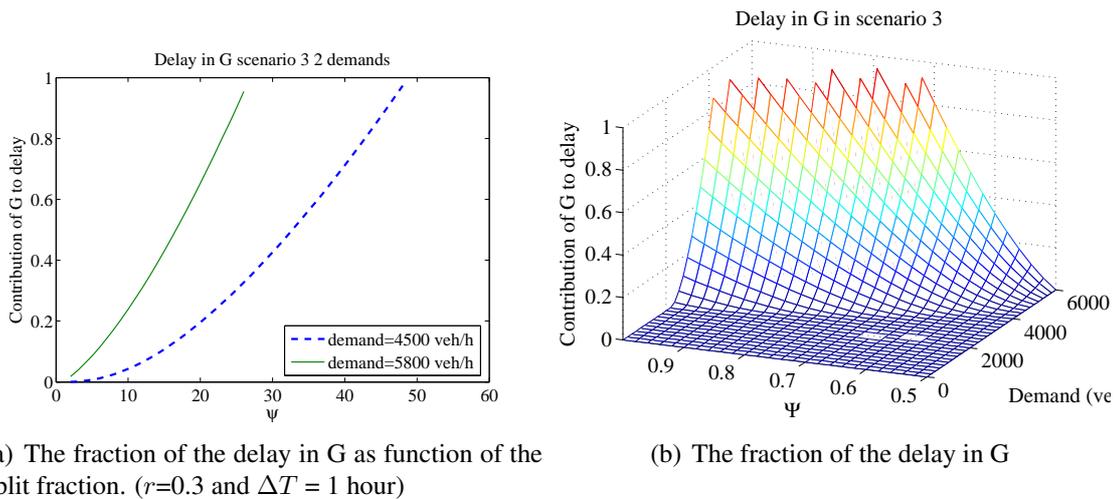
### Influence of junction

In figure 6.12a the influence of  $G$  is plotted as part of the total delay, just like in section 6.5.2. We had to fix the incident capacity and the duration; for the graph we choose  $r = 0.3$  and  $\Delta T = 1$ . We also needed to fix the demand. In the figure we plot two lines: one for a demand of  $4000\text{ veh/h}$  and one for a demand of  $5800\text{ veh/h}$ . The first one, due to the restrictions on  $\psi$ , has a wider range of possible values for  $\psi$  that are possible compared to the demand of  $5800\text{ veh/h}$ . The same plot for the demands of  $4500\text{ veh/h}$  and  $5800\text{ veh/h}$  are similar of shape: at  $\psi = 0.5$  the contribution of  $G$  to the total demand is 0 and it raises (in a convex shape) to 1 (i.e., the full delay is encountered in  $G$ ) for  $\psi = \frac{2C}{Q}$ .

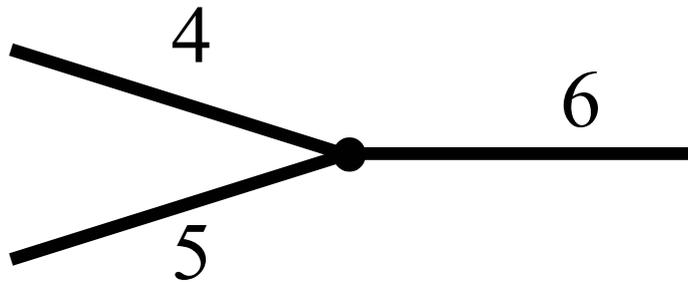
Figure 6.12b shows the same fraction as function of the split fraction and the demand. This fraction raises to 1 in case that the demand approaches the critical demand. The reasoning for increase of this fraction is the same as for scenario 2, explained in section 6.5.2.

## 6.6 Other configurations

This section discusses which elements of the computed configurations are present in other configurations as well. Note that scenarios 2 and 3 are configurations 12A and 12B respectively. From these 2 scenarios, the configurations 11A and 11B can be derived easily, setting the capacity of the lower outflow link 0.



**Figure 6.12: The contribution of the total delay that is encountered in G**



**Figure 6.13: Configuration 21B**

Of the links with multiple input links, configuration 21B is conceptually the most interesting configuration, which will therefore be discussed here. The actual traffic conditions depend on the demand, split fraction and on the capacities of each link, which can not all be covered in this contribution. We will show a typical example indicating the possible variations.

The most important modeling principle with a diverging point is the way traffic gets prioritised when an input node is oversaturated. We use the “fairness principle” as described by Lebacque (1996). This means that when a link downstream of the intersection is oversaturated, this link determines the fraction of the total demand that can be met. From all directions only that fraction of the demands are allowed to flow in the model. The remaining part has to queue.

For illustration, we take a road layout with 2 lanes at each of the three links and a 1 hour blocking. The demand from link 4 is 3300 *veh/h* and the demand from link 5 750 *veh/h* and the fundamental diagram has the same characteristics as in the previous examples (see table 6.1). The resulting fundamental diagram and the traffic states can be found in figure 6.14.

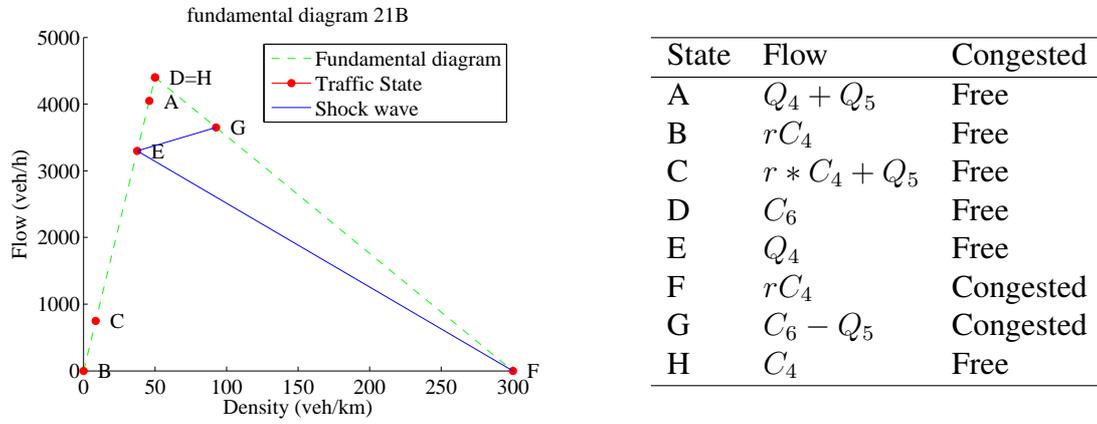


Figure 6.14: The fundamental diagram and the traffic states for configuration 21B

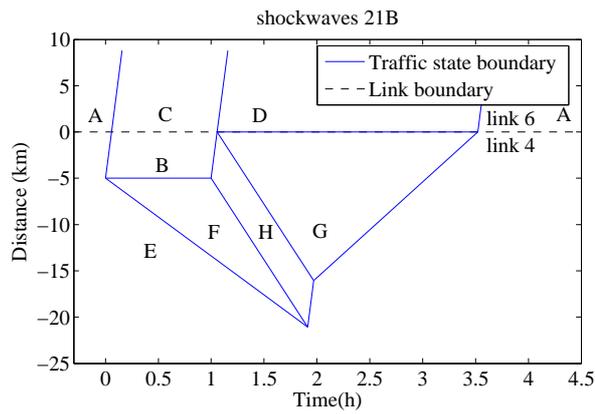


Figure 6.15: Traffic states and shockwaves for configuration 21B

The pattern resembles very much scenario 2. In the beginning a queue builds up upstream of the incident and when the incident is removed and a queue outflow arrives at the junction, the junction cannot handle this increased demand. In this case, a merging point forms the bottleneck.

In the presented case, a flow of 4400 *veh/h* and a flow of 750 *veh/h* arrive at the junction and have to merge, but the capacity of the downstream link is 4400 *veh/h*, which means that only a fraction  $\frac{4400}{4400+750}$  of the flow can flow to link 6. Since the modelling assumption is that only this fraction of the demand of each of the two upstream links can be met. Therefore, congestion is induced on both link 4 and link 5. This, however, immediately increases the demand from link 5 to the outflow capacity of 4400 *veh/h*. Applying the same modelling assumption gives that the fraction that can flow to link 6 is now  $\frac{4400}{4400+4400} = \frac{1}{2}$ , which in turn means that the flow from link 5 to link 6 is  $\frac{1}{2} \times 4400 = 2200$  *veh/h*, which is larger than the initial demand of 750 *veh/h* and therefore the (very short) queue reduces in length until it has disappeared. So in short a very short queue is started on link 5 but as soon as there is a queue, all queued vehicles will merge into the traffic stream. One cycle of queue growing and dissolving will in practice take around a minute. On a large time scale, it can be said that the total demand from link 4 to 5 is met and therefore that the maximum flow from link 4 to link 6 is the capacity of link 6 minus the demand from link 5. This queue will grow until the inflow is reduced from the queue outflow capacity to the initial demand.

Since many variations of the demand and the capacity are possible, there will be no general computations for these cases. However, the basic queuing principles are the same as those of scenario 2. The equations for the queues and the delays (in traffic states F and G) can be used when adapted to the modified demand and outflow.

One of the modifications can be that the demand from link 5 to link 6 is higher than half of the capacity of link 6, also a queue at link 5 will grow. In that case, there will be 2 queuing areas “G”, one at link 4 and one at link 5.

The larger the number of links, the more different possibilities for queuing there are, dependent on relative demands and capacities. However, the basic queuing elements remain the same – it only differs whether elements should be included or not. Therefore, only the scenarios 1-3 are considered numerically in the sequel of this contribution.

## 6.7 Implications

The previous section described the delay caused by an incident. In this section we will rework these formulae to find closed-form expressions for the average delay and the mean delay (section 6.7.1). In section 6.7.2 we show the numerical effects using the statistics of an incident database.

## 6.7.1 Mathematical consequences

### Total delay

Olmstead (1999) derives an equation for the expected delay (equation 6.2) assuming point queues. We will now show that this equation holds whenever the total delay (or any other attribute  $Y$ ) is proportional to the square of the duration (or any other property  $x$ ).

So suppose:

$$Y = cx^2 \quad (6.46)$$

In this equation,  $c$  is a constant. The variation of  $x$ ,  $\text{Var}(x)$ , is:

$$\text{Var}(x) = \langle x^2 \rangle - \langle x \rangle^2 \quad (6.47)$$

In these equations, the angle brackets mean the expectation value. The expectation value of  $Y$  can be written as:

$$\langle Y \rangle = \langle cx^2 \rangle = c \langle x^2 \rangle \quad (6.48)$$

Combining equations 6.48 and 6.47 gives:

$$\langle Y \rangle = c \langle x \rangle^2 + c\text{Var}(x) \quad (6.49)$$

This shows that if  $Y$  is proportional to the square of  $x$  the expectation value of attribute  $Y$  can be derived from the expectation value of  $x$  and the variation of  $x$ .

In scenarios 1 and 2, the delay is proportional to the square of the duration. Therefore, we can formulate an equation for the expected delay as a result of an incident with a stochastic delay, using the above theorem. We can now apply this result on the incident delays. In the case that there are no junctions, the total delay is proportional to the square of the delay (equation 6.24). This means that even if assumption of vertical queues by Olmstead (1999) is relaxed and we introduce realistic queuing, equation 6.2 is valid. Even more surprisingly, also in case of a junction downstream of the incident, the total delay is proportional to the square of the duration (equation 6.32). In case spillback occurs, this condition does not hold any more. We therefore can give an expression for the expected delay in case of no spillback:

$$\langle D \rangle = c \langle \Delta T \rangle^2 + c\text{Var}(\Delta T) = D(\langle \Delta T \rangle) + c\text{Var}\Delta T \quad (6.50)$$

with

$$c = \frac{1}{2} \frac{C_2(r-1)(rC_2-\psi Q)}{C_2-\psi Q} \quad \text{no influence of a junction}$$

$$c = \frac{(r^2\psi * C_1^2 - C_1\psi rQ - rC_1C_2 + C_2Q)}{C_2 - \psi Q} \quad \text{influence of downstream junction} \quad (6.51)$$

Note that these equations are derived for the case that a traffic jam occupies space (so not only for point queues). However, the resulting formulation in case there is no influence of a junction is the same as Olmstead (1999) derives (equation (6.2)).

**Table 6.3: Properties of the incident duration**

Property	value ( <i>min.</i> )	value <sup>2</sup> ( <i>min.</i> <sup>2</sup> )
Mean	77.0	$5.96 * 10^3$
Median	36.8	$1.32 * 10^3$
Standard deviation	106	$1.14 * 10^4$

**Average delay**

If there is no spillback (scenarios 1 and 2), the average delay per driver that encounters congestion is proportional to the duration (equations 6.5.1 for scenario 1 and 6.44 for scenario 2). We can now relate the variation of the delay per (delayed) driver to the variation of the duration. First we start by writing down an equation for the variation of the delay:

$$\text{Var}(A) = \langle A^2 \rangle - \langle A \rangle^2 \quad (6.52)$$

We now substitute  $A$  by  $c\Delta T$ , in which  $c$  is the proportionality constant expressed in equation 6.5.1 or 6.44:

$$\begin{aligned} \text{Var}(A) &= \langle (c\Delta T)^2 \rangle - \langle c\Delta T \rangle^2 \\ &= c^2 \langle (\Delta T)^2 \rangle - \langle \Delta T \rangle^2 = c^2 \text{Var}(\Delta T) \end{aligned} \quad (6.53)$$

This equation gives an expression for the variation of the delay for individual drivers.

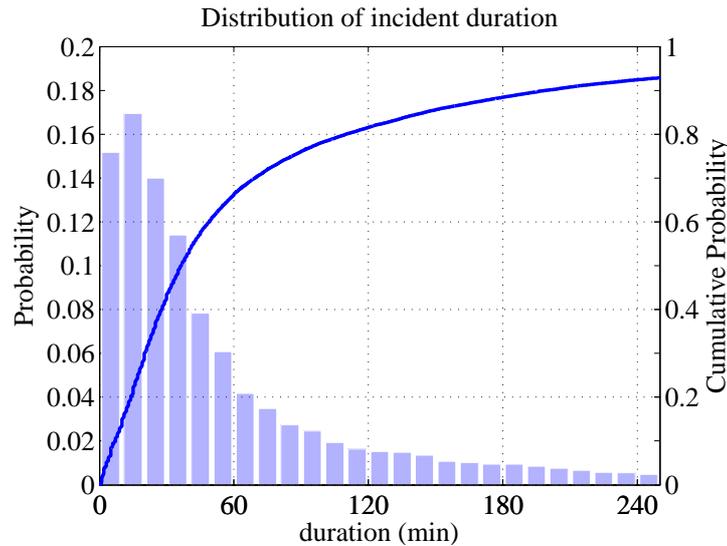
However, none of these simple rules holds for situations where spillback comes into play. Areas F and G (figure 6.9) can be seen as the congestion states for a special case of scenario 2, with a duration of  $(t_1 - t_0)^2$  and a distance to the junction of 0. The delay in area F and G therefore scales with  $(t_1 - t_0)^2$ , which can be computed using the following equations:

$$t_0 = \frac{x^{\text{inc}}}{-\omega_{EF}} \quad (6.54)$$

$$t_1 = \frac{x^{\text{inc}}}{-\omega_{FG}} \quad (6.55)$$

**6.7.2 Real-life incident data**

To show the impact of the equations, we analyse the duration of real-life incidents. We use a database of all incidents in the Netherlands between 1 January 2007 and 1 September 2007 provided by the Dutch Road Authority. The database consists of 55,176 freeway incidents; for 51,050 incidents there is a valid incident duration. Figure 6.16 shows the distribution of the duration these incidents. We see that a majority of the incidents (around 65%) is cleared in less than 60 minutes. There is also a long tail in the distribution: there are some incidents which take a long time to remove.



**Figure 6.16: The duration of incidents**

**Table 6.4: The delays in vehicle hours split up per source**

Scenario	delay mean duration (veh h)	stochasticity (veh h)	total expected delay (veh h)
1	$1.51 * 10^5$	$2.91 * 10^5$	$4.42 * 10^5$
2	$2.64 * 10^5$	$5.10 * 10^5$	$7.74 * 10^5$

Some of the statistics of the incident duration can be found in table 6.3. The total delay of an incident with a duration of the average duration underestimates the expectation value of total delay. According to equation 6.50, the correction that has to be made to this estimate is the variation, or the square of the standard deviation. This is also stated in table 6.3. The relative importance of the variation can be derived from the quotient between the square of the mean value and the square of the standard deviation. The value of this quotient is  $\frac{1.14 * 10^4}{5.96 * 10^3} = 1.93$ . This means that the part caused by the variation is 1.93 times higher than the part of the total delay which would be calculated based on the average delay. This, in turn, means that using the average duration in assessing the total delay would give only  $\frac{1}{1+1.93} = 34\%$  of the total expected delay.

Using the values for demand, capacity and the split fraction as mentioned in section 6.4.1 and assuming a reduction  $r$  of 50% due to the incident we can compute the delays. The following table (6.4) gives the results for applying equation 6.50 using the values in table 6.3. The spread of the duration contributes for almost twice as much to the expected delay as the mean delay in scenarios 1 and 2.

## 6.8 Conclusions

This chapter analyses the traffic states after the occurrence of an incident. Using shockwave theory, the traffic states that result from the incident are calculated. As opposed to a point queue model, the head and the tail of the queue are modeled separately and in this way the spatial extent of the queue is described properly. Using the traffic states, the delay is analytically calculated. This was possible for simple basic cases, and expansion is possible. However, expanding the formulation for a general road layout and general queue patterns is analytically too demanding. It can be applied, though, for a specific road network and a specific incident. For such a specific case, it is then also possible to carry out an experimental validation of the method.

A formula for the total delay was derived. We found that in the scenarios without the tail of the queue blocking the flow on other links (the first two scenarios), the delay is proportional to the square of the duration of the blocking. Because this is not linear in the duration, the expectation value of the delay is not the delay of the incident with the expectation value of the duration. An expression for the expected delay as a function of the variation of the duration of the blocking was formulated. We also formulated the variation of the average delay per involved traveler as function of the variation of the delay.

The delay in case that the queue spills back to an upstream link can be expressed analytically but the result is more complicated and cannot be captured in a simple equation. In that case there is no simple relationship between the expected delay caused by an incident having the expected delay and the expected delay caused by an incident with stochastic delay. We therefore conclude that to analyse the delays in a network, it is needed to analyse the consequences of incidents with various durations.



## Chapter 7

# The influence of spillback modelling when assessing consequences of blockades in a road network

---

Robustness of a network is a main objective for road network managers these days, and has become an important study area for transportation scientists. This chapter discusses one specific aspect in assessing road network robustness: the consequences of the closure of a link. These spillback effects have been examined in a dedicated traffic simulation program in which the representation of spillback can be switched on and off. The impacts are studied in a simulation study of a road network of a regional size in which sequentially links are blocked. Two scenarios for route choice are considered: a fixed route choice based on a daily congestion pattern and a route choice adapted to the actual congestion caused by the closure. The study has also shown the influence of information which makes travellers adapt their routes. Road network robustness and characteristics of vulnerable links are evaluated for both spillback and non-spillback cases. It is found that a valid spillback modelling is a prerequisite for correctly analysing the robustness of the network as a whole, as well as for correctly indicating the locations in the network where a closure causes the largest delays. Furthermore, without simulating spillback, it is not possible to identify correctly the most vulnerable links for the network performance. This chapter is an edited version of: Knoop, V.L., Hoogendoorn, S.P. and Van Zuylen, H.J. (2008) The Influence of Spillback Modelling when Assessing Consequences of Blockings in a Road Network *European Journal of Transportation and Infrastructure Research*, Vol. 8, no. 4, pp. 287-300.

---

**Table 7.1: Reported values of travel time unreliability for car trips – source Van Amelsfort (2009)**

Study	Value of travel time unreliability	Method
Brockfeld et al. (2004)	11-14 €/h (male) <sup>1</sup> 28-30 €/h (female) <sup>1</sup>	RP data using percentile and median approach
Lam and Small (2001)	4.65-12.29 €/h (male) <sup>1</sup> 6.03-27.58 €/h (female) <sup>1</sup>	RP data using percentile and median approach
Senna (1994)	2.50 times VoT	SP data mean-variance approach
Noland et al. (1998)	1.27 times VoT	SP data mean-variance approach
Black and To-wriss (1993)	0.55 times VoT	SP data mean-variance approach
Dutch Road Au-thority (2005b)	0.80 times VoT	Expert meetings
Eliasson (2004)	Travel time variability Morning: 3.86 €/h <sup>1</sup> Afternoon: 1.64 €/h <sup>1</sup> Value of unexpected de-lays Morning: 42.15 €/h <sup>1</sup> Afternoon: 26.70 €/h <sup>1</sup>	SP data mean-variance approach

<sup>1</sup> February 2007 exchange rates were used

## 7.1 Introduction

Reliable road networks are valued by both travellers and network authorities. There are many studies showing this preference. Table 7.1 shows some of them, as well as the value travellers attach to reliability. These studies usually compare the value of travel time (un-)reliability (VoR) with the value travellers attach to the time a trip costs, the value of time (VoT). This can be studied in a questionnaire what people would do in hypothetical situations, giving a stated preference (SP). Alternatively, it can be analysed what choices travellers made in practice, showing the revealed preference (RP).

Chapter 1 defined robustness as the way the network copes with a short-term, strong decrease in capacity or a short-term strong increase of demand. This network property is a corner stone for travel time reliability. The mentioned variations can be caused by normal daily fluctuations in demand and supply as empirically shown by Tu et al. (2005). Another cause for this variation is the closure of a link by an incident or road maintenance. This is not part of the normal daily fluctuations and reduces the capacity.

There are two analyses that are considered in this chapter. The first one is the reduction

of performance of a network caused by an incident. That is, we only consider the impacts on network robustness caused by a fluctuation in the supply side. The different risks on incidents on different links are not considered in this chapter. The second one is the location where a link closure has the biggest impact. We take a look at both these analyses.

Research projects assessing road network robustness use different traffic simulation models. The models differ, among others, in the way the spatial dynamics of traffic flow and congestion are described. Due to the complexity of the network and traffic flow modelling (and thus computation time), sometimes spillback, i.e. congestion propagation to a more upstream link, is not modelled in macroscopic traffic flow models (for instance, Kraan et al. (2008)). This chapter compares the simulation of link network robustness for scenarios with and without spillback modelled in order to assess the need of proper spillback modelling. The method we use works as follows: we sequentially block all links of the network, one at a time, and compare the results with respect to network performance. Furthermore, we analyse the influence of route advice: a distinction is made for cases in which road users adapt their route choice to the new situation with a blocked link and congestion and a situation in which they take their usual routes. Both happen in practice as shown in chapter 5.

It turns out that different modelling assumptions on spillback models affect the results of the assessment of the consequences of link blockings. The results of the assessment can be either a value for network vulnerability or an ordering of most important links. The aim of this chapter is to show the effects of spillback modelling assumption on the outcomes of a study.

It is found that the robustness assessment considerably differs for different approaches. We therefore conclude that it is essential that spillback is correctly modelled in robustness studies.

## 7.2 State of the Art

A considerable part of delays is caused by incidents. Kwon et al. (2006) show that this is around 25% for the USA; a similar number is found for the Netherlands by Dutch Road Authority (2007b). In this chapter, we discuss the consequences of an incident in detail. One could separate the search for important (sometimes called *vulnerable*) links in four approaches, which do have overlaps:

1. Flow characteristics of network links
2. Single assignment (user-equilibrium)
3. Analytical game-theoretical approaches
4. Dynamic simulation of all possible closures

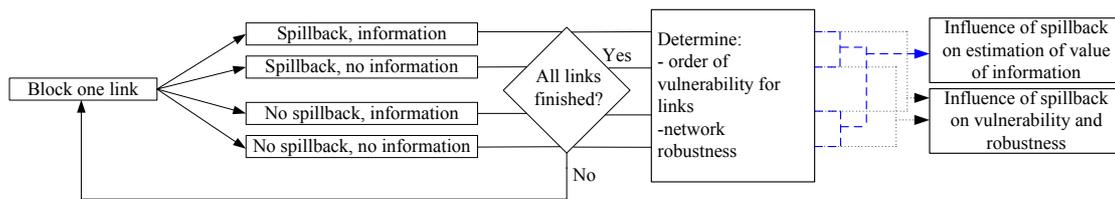
The first approach is the one taken by, amongst others, Tampère et al. (2007). Their paper presents criteria to find vulnerable links. The method accounts for the traffic dynamics.

For example, the time until a backwards growing traffic jam reaches a junction is calculated. A short time to reach a junction and a high flow on the upstream link are two of their criteria. There are more papers which present indicators (for example, volume over capacity ratio) in an equilibrium-assigned network to find the most important links. There are contributions discussing this approach, for instance by Tamminga et al. (2005), Li (2008), or Tampère et al. (2007). Chapter 8 will discuss this approach in more detail.

The second approach is the one proposed by Murray-Tuite and Mahmassani (2004). It starts with an equilibrium assignment, just as Tampère et al. (2007). This approach adds a rerouting for traffic that normally uses blocked links. Sequentially links are blocked and the flow on the blocked links is reassigned. The travel times depend on the flows on the links. Kraan et al. (2008) proposes a scheme which identifies important links in different stages. The first two steps are similar to the method proposed by Murray-Tuite and Mahmassani (2004). In the final step, links that are potentially most important are ordered using a more accurate model. Kraan et al. (2008) use this approach to overcome the computational effort of doing a separate simulation for each blocked link. In addition to the effect of incidents, they also estimate the probability of a closure and calculate the vulnerability as probability multiplied by effect, in which the effect is the total delay.

The third category is introduced by Bell (2000) and Belda et al. (2002). It is an analytical, game theoretical approach. Rather than using the statistic probabilities on an incident, they calculate the maximum disruption of one incident, given that the place is unknown. It is described as an “evil entity” that wants to destroy a network and is given the possibility to destroy a limited number of links. In the second approach, there is just one rerouting step. In this approach, the travellers will account for failing routes when making their route choice. It is iteratively calculated which routes are optimal for the users (users will take that route) and what will be the worst place for a link closure (travellers will count on that situation). Those links that decrease the network performance most if they are blocked are called important. One could make various assumptions for the route choice that is made. Where Bell (2000) and Belda et al. (2002) propose a strict risk-averse route choice, Nagae and Akamatsu (2007) relax this assumption and propose a distribution of risk-averseness. It holds for all of these contributions that the lower the effect of the dropout of links is, the more robust one can consider the network. In this approach, results are based on an analytical approach and a mathematical framework is set; in both articles by Bell (2000) and Belda et al. (2002), a simple network is used as test case.

The fourth approach, described by Knoop et al. (2007a), takes the computational demanding route of computing a dynamic traffic assignment for each possible closure. It identifies important links in a real-size network by running a full dynamic simulation for each possible case, i.e. each place of a closure. It then tries to deduct the properties of the most important links. The study underlines the importance of spillback effects in busy areas. Jenelius et al. (2006) also takes the approach of a full computation. Since they apply the method to an area in East Sweden with less traffic and less congestion (roads with 450 vehicles per day summed over both directions), there is no need for a simulation which accounts for spillback.



**Figure 7.1: Outline of the research approach**

The majority of the articles on important links do not take spillback into account. Knoop et al. (2007b) presented a study to see the importance of spillback for finding the vulnerability of a link. In that study, the route choice was taken fixed. In this chapter, we show that also if the route choice varies over time, spillback effects are important for the robustness and the identification of the important links. This chapter focuses mainly on the importance of spillback on the identification of the most important links in a real-world network.

### 7.3 Research approach

The research approach that is taken in this chapter is outlined in figure 7.1.

We will simulate the traffic flow with all different blockages. Two different traffic flow models will be used. They are identical, except for the modeling of spillback. In one of the models, there is spillback, in the other, there is not. They are described in detail in section 7.3.1. Regarding route choice, we consider two options. In the first option, the route choice is adapted to the recurrent congestion, but travelers will never deviate from their standard routes (even though there is a unexpected queue or a closure). In the second option, they will be informed about the traffic states in the network every 15 minutes and adapt their routes accordingly. A detailed description of the route choice models is given in section 7.3.2.

One could combine the results pair-wise as indicated in figure 7.2b. One pair could be the scenario with spillback and with information and the scenario without spillback and with information. Comparing these will show whether it is needed to model spillback. This is also found in the following pair: the model with spillback without information and the model without spillback without information. Comparing the result with spillback with information and the result with spillback without information shows how much delay can be avoided by informing the travelers. This value can be calculated from a non-spillback model. This gives a new pair (1) a model result (i.e., delay) without spillback and with information and (2) a result without spillback without information. The value of the information predicted by this non-spillback simulator can be compared by the value of the information with spillback modeled.

### 7.3.1 Traffic Flow Modelling and Spillback Modelling

The best way to compare the results of the vulnerability of links in a spillback and in a non-spillback simulator is to use one simulation program that can run simulations both with and without spillback. In this way, there are no differences in systematic errors. Because we are examining effects of the location of queues, we want to have a model with a reasonable queue dynamics. As far as we know, there is no such model on the market in which spillback can be switched off. Therefore, we developed a macroscopic model in which spillback can be switched on and off. The section below briefly states the working of the model.

We use the continuum LWR-model proposed by Lighthill and Whitham (1955) and Richards (1956) that we solve with a Godunov scheme (Godunov, 1959). Lebacque (1996) showed how this is used for traffic flows, yielding a deterministic continuum traffic flow simulation model.

When a queue occurs, the queue may grow further upstream beyond the end of a link. Furthermore it can, depending on traffic conditions, dissolve from the head, while the tail of the queue still moves upstream. The traffic dynamics for a road stretch with a temporal bottleneck – which is typical for an incident – is plotted in Figure 7.2a. The queuing area is shaded in the top figure. In the lower figure, the number of cars in the queue is plotted. This is all present in a LWR model.

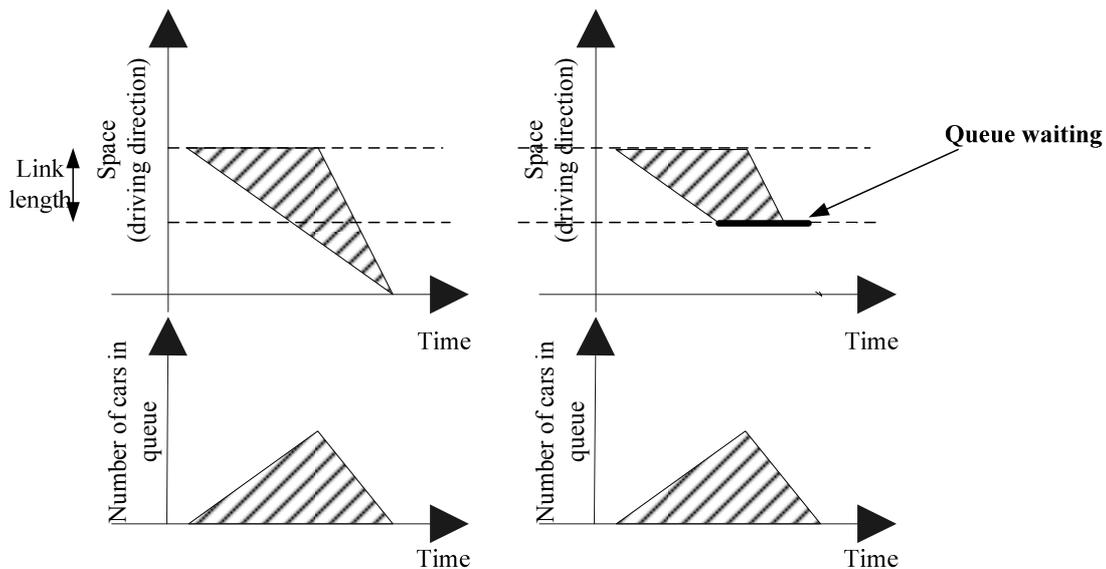
We choose a LWR model exactly for these properties: the queue dynamics are realistic, but it is not unnecessary complicated. For instance, second order effects (e.g., from synchronized flow to wide moving jams) have some minor influences, see Ngoduy (2006). We will neglect these in the remainder of this chapter.

When changing it to a non-spillback model, the flow at the upstream link is (by definition) not influenced by the queues on the (downstream) link. In our representation of a non-spillback model, the queue will grow upstream but does not cross the link border. This is achieved as follows: we choose the same LWR-representation of the traffic flow, but now the inflow in the most upstream cell of a link is not influenced by the density, but only determined by the (static) link capacity. In this way, the traffic dynamics are the same as the in the model with spillback modelled and the only difference is the spillback. Consequently, the density in this cell can reach very (unrealistically) high values. The queue dynamics of this model are plotted in b. In the upper figure, the space-time diagram of the queue is plotted; in the lower figure the number of vehicles in the queue is plotted.

This study compares the situation with queues without spillback (as in figure 7.2b with the situation of full spillback in figure 7.2a).

### 7.3.2 Route Choice Modelling

We consider two route choice possibilities. The first possibility is that the road users choose the routes that are fastest without an incident and do not deviate from it in case an



**Figure 7.2: Congestion dynamics in a space-time plane. From left to right: (a) realistic queue development, including spillback, (b) implementation of a non-spillback model**

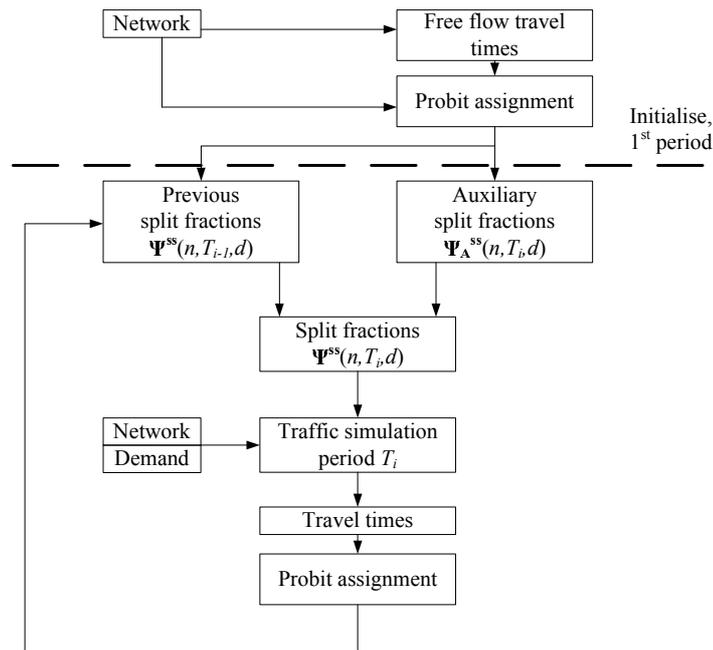
incident occurs. That is, if their route turns out to be congested or blocked, they will have to wait in the queue. The second possibility is that when a road is blocked, travellers will adapt their routes according to the new situation. This implies that travellers are somehow correctly informed about prevailing traffic condition.

### Fixed routes

In this scenario, we assume that travellers use fixed routes to reach their destination. The routes should represent the everyday choice of the travellers. An equilibrium assignment would be suitable, but in our model, it would be too time consuming to simulate. Instead, we choose to assign routes to travellers according to the fastest routes at the moment in normal conditions. The route choice model can also be found in figure 7.3.

Routes are always determined for a 15-minute time interval. At the start of the next time interval, again the fastest routes are determined and vehicles are assigned to these new fastest routes. So we use a stochastic traffic assignment in which the (dynamically derived) travel times of the previous period are used. This is repeated at the end of each time period. Half of the travellers are assigned to the new route, whereas the other half keeps the routes of the previous period. If we would choose to reroute all travellers, this might lead to oscillating effects Fukui et al. (2009). Note that their routes are still based on the everyday travel times without a link closure.

The stochastic routes are found using a probit assignment. This means that for each node, we take 20 random sets of draws of the link travel time. In each of these 20 sets, we determine the fastest route from that node to each destination. These 20 routes to destination  $d$  lead over a few exit links from the node. In a numerical example, suppose that there are three links that flow out of the node and 12 routes start by using link 1 and 6 by link 2 and 5 by link 3. The destination specific split fraction  $\Psi$  is chosen proportional to the number of paths over the exit link from node  $n$ . In the numerical example, that



**Figure 7.3: The route choice module**

means we have a split fraction of 3/4, 1/4. We assume that the link travel times are normally distributed with a standard deviation of 10% of the average travel time. This is the outcome of a basic calibration described in section 7.4.

The routes are stored as split fractions, which differ per node, destination, and time interval. In the numerical example: the numbers 12/20 and 6/20 and 2/20 differ for different nodes, destinations for the traffic and different times based on congestion levels.

After a vehicle has set off for a certain route, at the next node it reaches, it will follow the routes found for that node. This can be another route than the route it was heading for. Even with 20 routes in the probit assignment, more than 20 routes are used from an origin to a destination. This can be best illustrated by the numerical example. Let's consider the the 2 routes using link 3. After link 3, the traffic will arrive at another node, node 2, having 2 out links, link 4 and 5. It is possible that the 2 routes found using link 2 are the same, both for instance using link 4 afterwards. However, also for node 2 a probit assignment was carried out, giving 20 routes for traffic with origin at that node. And suppose 15 of the 20 routes will use link 4, and 5 will use link 5. Then, the traffic will be split according to that split rate: (3/4, 1/4). This means that from the original traffic 2/20 x 15/20 will take the route using link 3 and 4 and 2/20 x 5/20 will use a route using link 3 and 5.

In this scenario, the split factors are fixed and do not change because of the changing traffic operations. The resulting route set is referred to as  $\pi^*(G, ss)$ , in which  $G$  is the network these routes are based on. A network with link  $b$  blocked is denoted as  $G_b$ .  $ss$  denotes the spillback scenario.

The network flows and therefore also the delays  $D$  can be different for the scenario in

which spillback is modelled and the scenario in which no spillback is modelled, so it depends also on the simulation of spillback,  $ss$ .

The performance of the network in this scenario can now mathematically be expressed as:

$$D(\pi^*(G, ss), G_b, ss) \quad (7.1)$$

The most important link  $b^*$  is the link for which the network performance is highest if it is blocked.

$$b^* = \operatorname{argmax}_b (D(\pi^*(G, ss), G_b, ss)) \quad (7.2)$$

In this notation it is assumed that a lower  $D$  equals a better performance, which is the case if delay is chosen as indicator. If another performance indicator  $D$  were chosen for which a higher value would mean a better performance (e.g., total arrivals),  $b^*$  could be found by minimizing  $D$ .

### Route choice with information provision

The basic principles for the route choice are the same as in the previous situation. We still refer to section 7.3. The travellers' choices are still modelled by the same model. The difference is that the network which is put into the route choice module now is the network with the closure on link  $b$ . Consequently, there are also extra queues (caused by this blockings) which are now simulated. The road users adapt their behaviour to information of the new situation with the closure. Therefore, they will also change their routes during the simulated time due to congestion.

Just as in the scenario with fixed routes, routes are chosen based on expected travel time. Routes are updated every time period of 15 minutes based on the congestion, including the congestion caused by the incident. Not all travellers have access to information and some will be unwilling to change their route. Therefore, only a part of the travellers will be assigned to a new route; the other part will choose the old route for the coming period. The path set found in this case, is called  $\pi^*(G_b)$ . It depends on the blocked link  $b$ .

As network flows will differ for scenarios with and without spillback, the network performance can also differ dependent on the simulation of spillback.

For this scenario, the network performance function that needs to be evaluated is

$$D(\pi^*(G_b), G_b, ss) \quad (7.3)$$

This function is calculated for each choice of blocked link  $b$ . The most important link  $b^*$  is:

$$b^* = \operatorname{argmax}_b (D(\pi^*(G_b), G_b, ss)) \quad (7.4)$$

This can be translated in terms of a mathematical game between the road users and an evil entity wanting to harm the network performance most. In this game, this link  $b^*$  from

equation would be the Stackelberg optimum (see Fudenberg and Tirole (1991)) for an evil entity to block if he was given the opportunity to block one link (and given the users' response). In case of the fixed routes, the travellers do not change routes and therefore they are no players; hence, there is no mathematical game any more and the resulting link  $b^*$  from equation is not a Stackelberg optimum.

## 7.4 Case study description

We perform a case study on a regional size network with both motorways and underlying roads. A morning peak period from 6.30 to 9.30 is simulated. 468 links with different link properties (capacity, speed limit) and link connections give insight to which properties are relevant for the network impact of a link being blocked. The network we used is the ring road around the city of Rotterdam (around 600.000 inhabitants). A map of the area is given in figure 7.4. In the peak period the network it is rather busy. The model is qualitatively calibrated for the normal situation: the capacities, demands and the perception error in route choice are chosen to match the daily congestion. Especially near the motorway junctions, some queues develop, but none with a length of more than a few kilometres (less than the distance to the next motorway junction).

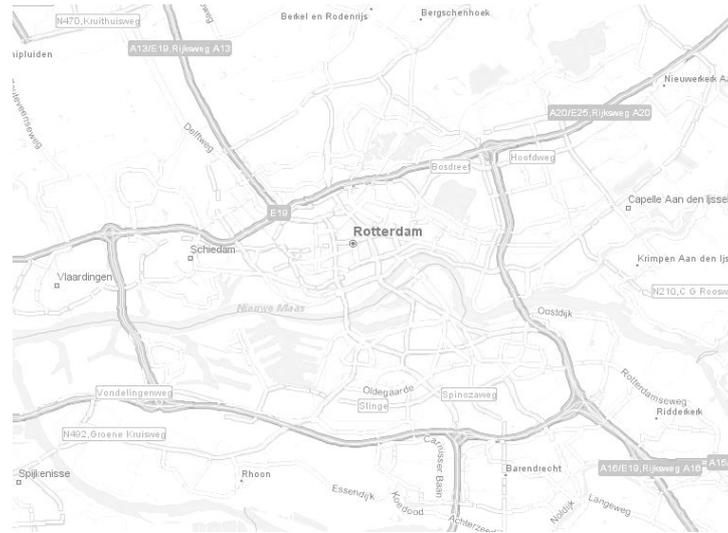
There are two approaches for the route choice modelling and two approaches for the traffic flow modelling. The combination gives four different possibilities:

- no spillback modelled, fixed route choice;
- no spillback modelled, route choice with information provision;
- spillback modelled, fixed route choice;
- spillback modelled, route choice with information provision.

In the case study, we calculate the consequences of the closure of each link in the network in each of these 4 scenarios. There are 4 different scenarios for which each of the links is sequentially blocked. Each scenario gives 468 total performances, one for each of the blocked links.

The route choice with information provision requires the part of the travellers adapting their routes to be set. If the fraction of people that take a different route is too small, the effect of the information provision cannot be seen. On the other hand, a very large share is unrealistic, see for instance Bogers et al. (2005) which shows that users, even when informed, will also stick to their original routes. Furthermore, it needs to be considered that not everyone will get the information. An examination of the most important links would require a careful calibration of the route adaptation of travellers for a specific network. In this chapter, we use a level of 50% of travellers adapting their routes and 50% of the travellers keeping their routes fixed, according to the findings in chapter 5

In this chapter, we choose the *total* or *collective delays* as the main performance indicator. This is in line with the route choice which is also only based on travel time. Based on the demand and the free flow travel times, the free flow arrival pattern can be constructed. Any delay in the arrival pattern contributes to the total delay.



**Figure 7.4:** The case study area, the ring road around Rotterdam, left to right around 25 km

**Table 7.2:** The fit parameters for the relationship between the non-spillback performance and the spillback performance

	Fixed routes scenario	Adaptive routes scenario
$\alpha$	$-1.2 (\pm 0.2) E+6$ veh h	$-2.7 (\pm 0.7) E+06$ veh h
$\beta$	$6.6 \pm 1$	$2.3 \pm 0.3$
$R^2$	0.32	0.33

## 7.5 Case study results

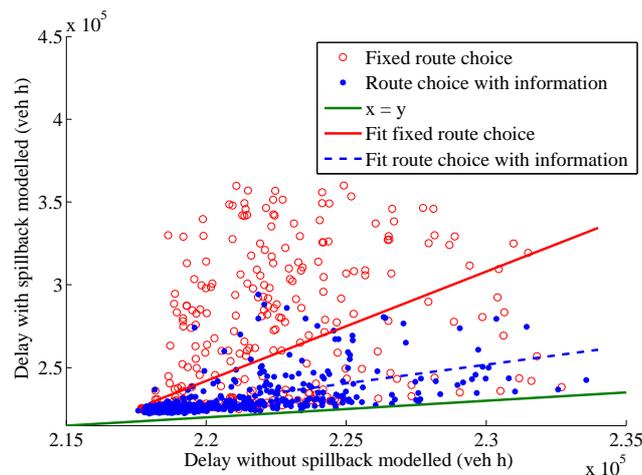
Figure 7.5 shows the performance of the network in the different cases. The x-axis shows the delay without spillback modelled and the y-axis shows the delay with spillback modelled. For each closure, we find a delay with and without spillback modelled. This is represented in the graph. This is done for the case with and without path update. We find the delay to be higher if spillback is included: all points lie above the line  $x = y$ , as we would expect.

One dot indicates one specific blocked link. We tried to fit a linear relationship for both the rerouting and for the fixed route case. Another relationship The correlation shows how poorly a simulation without spillback can foretell the impact of the closure of a link. For this purpose, we fitted the relationship:

$$D_{\text{spillback}} = \alpha + \beta D_{\text{no spillback}} \quad (7.5)$$

and found parameters in table 7.2.

The regression lines are plotted as dotted lines in figure 7.5. The correlation coefficient  $R^2$  indicates how much of the variance in performance reduction in a spillback case can



**Figure 7.5: Comparison of the impacts of link closure in scenarios between spillback simulations and non-spillback simulations**

be explained by the variation in arrivals in the corresponding scenario without spillback. We see that this value is low, around 33%, which can be seen by the large scatter of points in the figure. An important conclusion therefore is that *a static, non-spillback simulation cannot be used to identify the links in a road network having the largest impact when struck by an incident.*

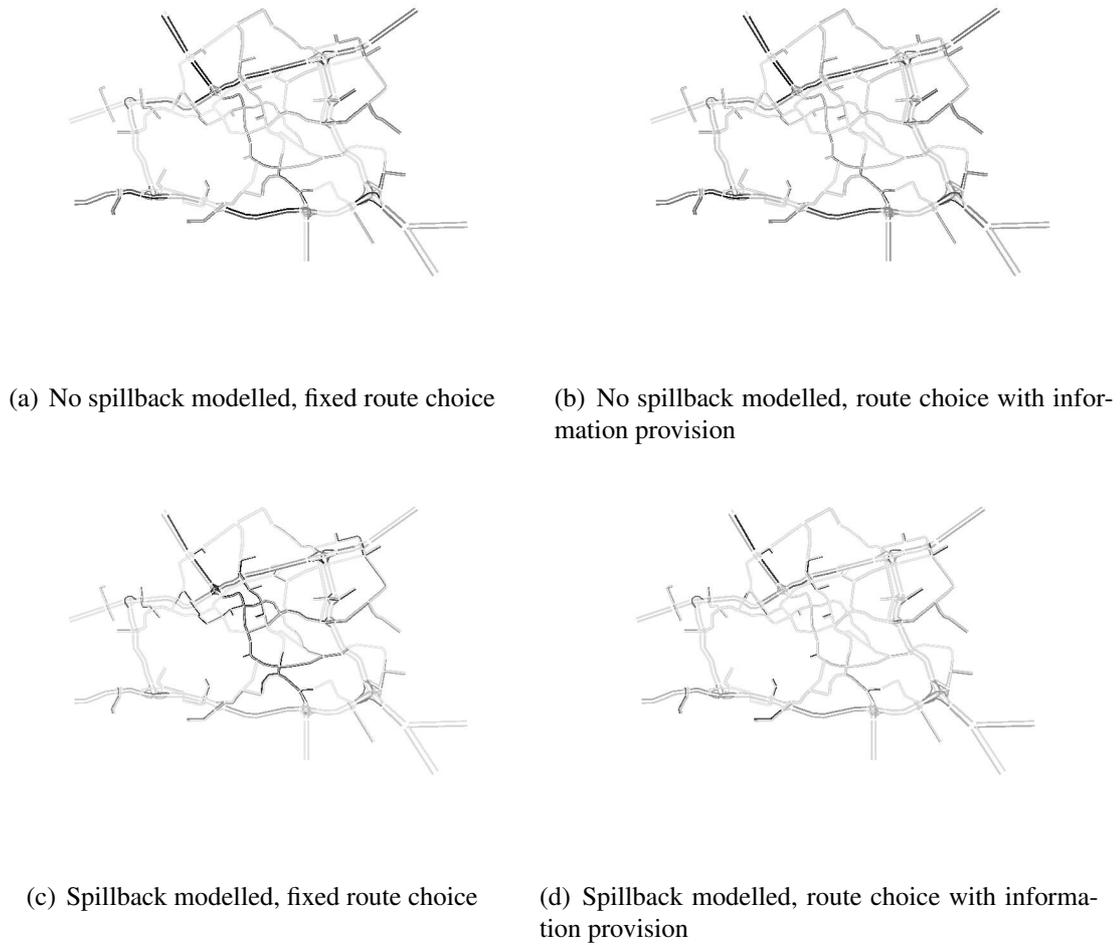
Furthermore, the factor  $\beta$  is larger than 1. It indicates how much larger the consequences are when spillback is modelled compared to the consequences without spillback modelled. It makes sense that this is higher for the scenario with the fixed paths: in this case, the consequences are underestimated by a factor 6.6 if one uses a non-spillback model compared to a spillback model. In the case of adaptive paths, the average underestimation is limited to a factor of 2.3, which is still very substantial.

Figure 7.6 shows the impact of closure for the individual links and where these links are located. If a link is shaded dark, the impact of closure that link is large. Figure 7.5 showed that the magnitude of variance of the performance reductions differ among the different scenarios. Therefore, the gray scales in subplots of figure 7.6 are not the same. Figure 7.6 shows the same area as figure 7.4.

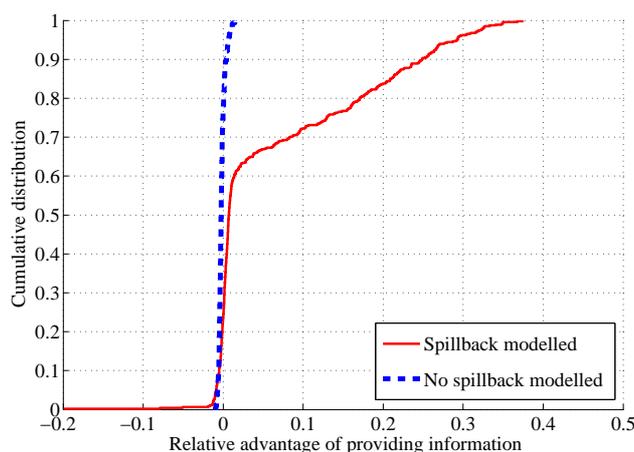
We see that in cases without spillback modelled, the motorways are particularly important for the network performance (the motorways are shaded darker). If one of these links is blocked, the performance reduces most. This makes sense, since the motorways are usually the roads that are most used.

Once spillback is included, the motorways appear to be less critical (compared to the average of all links). When travellers are not informed about the routes, the urban roads are important (in case of realistic spillback modelling). Since it is at maximum two lanes wide, the urban link is completely blocked. Even in low intensity traffic, the queue builds up. As the tail propagates through the network, many links are blocked.

When travellers are informed, they will be rerouted already quite early, since the queue



**Figure 7.6: The impact of closure a link. The shading of a link indicates its impact: delays are large when there is a closure on a dark link; delays are minor if there is a closure on a light link.**



**Figure 7.7: Relative decrease of delay when providing information compared to fixed paths; cumulative distribution for blocking a randomly chosen link**

starts building up immediately. The most important parts in the network in this scenario are not the urban links (as it was without information); the destination links are now important since people cannot exit since there is no alternative for the exit links, as shown by Li (2008).

To show how much performance can be gained by providing information, we calculated the relative advantage  $A$  of updating the paths:

$$A(b, ss) = \frac{D(\pi^*(G, ss), G_b, ss) - D(\pi^*(G_b, ss), G_b, ss)}{D(\pi^*(G, ss), G_b, ss)} \quad (7.6)$$

Each blocked link  $b$  leads to an advantage of updating the route choice. The 468 numbers are ordered and plotted in figure 7.7. So without spillback modelled, this information improves the network performance by at maximum a few percent. When spillback is modelled, the provision of information can improve the network performance much more (in a third of the cases over 10%).

In many cases (i.e., for a large share of the 468 possible locations of an incident), the advantage of route information is in the spillback scenario much larger than in the non-spillback scenario. That could also be derived from figure 7.5, which shows a big performance decrease for the spillback scenario without rerouting. If there is rerouting, the performance reduction is much less. The new advices lead people around the bottleneck and hence reduces the delays, but, more important, also considerably reduces the secondary delays (i.e., the delays induced by spillback).

We also investigated the (relative) performance of the network if a link is blocked. For all 468 possible blocking locations (links), we calculate the relative delay,  $rD$  of the blocking of a link:

$$rD(\pi, b, ss) = \frac{D(\pi^*, G_b, ss)}{D(\pi^*(G, ss), G, ss)} \quad (7.7)$$

This can be calculated for 4 scenarios. Whether spillback is modelled or not influences the numerator and the denominator (ss). The adaptation of paths influences  $p^*$  in the numerator: it either becomes  $\pi^*(G_b, ss)$  when adapted, or  $\pi^*(G, ss)$  when not adapted.

From this analysis we conclude that the closure of a link looks as being of less importance if spillback is not simulated. For both the case with and without rerouting, the distribution of the values for the relative delay (one per link) are approximately the same. In none of the cases, the delay increases more than a few percent. If spillback is taken into account, there are more links causing a large performance drop. If paths are updated, the travel time increases by at most 36%. With fixed routes, the travel times can increase by more than 60% compared to a non-blocking scenario. Robustness of a road network was defined in chapter 1 as the way the network copes with conditions which deviate from the normal conditions. Clearly, if spillback is modelled, the network can handle disturbances worse than if spillback is not modelled. So, robustness is overestimated if it is assessed by a non-spillback simulator and robustness can be increased by giving proper route information.

## 7.6 Conclusions and further research

We simulated a morning traffic flow on a real regional sized, network for which sequentially one of the links was blocked. The traffic simulator had the possibility to simulate both fixed and adaptive choices, and situations with and without spillback: paths could be adapted to the situation or not and, independently, spillback could be switched on and off. This yields four scenarios which have been considered in the simulation study.

An important result is that the links that are considered to be important in terms of their impact on network performance reduction when being blocked differ substantially per scenario. We found that motorways appear to be the most important if spillback is not taken into account. When considering congestion spillback, the impact of a link closure depends on the information given to the drivers, which links are most important. Without dynamic route information, the urban links in the city cause many problems if being blocked; a blocking leads to a total grid lock. If people are informed, the most important links are the links for which there is no route alternative, i.e. the destination links.

The main conclusion of this research is that the links in a network that will cause a major disruption in the network flow operations *cannot be validly identified by a non-spillback simulator*. Hence, the results of previous studies that have been undertaken using models not incorporating spillback have very limited validity. Only a third of the variations of impact of link blocking in realistic spillback simulations can be derived by performing a simplified, non-spillback simulation. Modelling spillback is also important in assessing robustness of a network. Without spillback being modelled, the impact of the closure

of one link is much less and therefore the network is considered more robust in a non-spillback simulation program.

Finally, when a non-spillback simulator is used, the advantage of giving route information is highly underestimated. With spillback being modelled, in around 50% of the link closure locations, route information can increase network performance considerably. This shows that for realistic situations, where spillback will indeed occur, network robustness can be increased substantially by informing road users properly.

# Chapter 8

## Link-level vulnerability indicators for real-world networks

---

Chapter 7 showed it is computationally expensive to find out where vulnerable parts in a network are. In literature a variety of methods were introduced that use relatively simple selection criteria (measured in real-life or calculated in a traffic simulator) to pre-determine the seriousness of the delays caused by the blocking of that link and thereafter perform a more detailed analysis. This chapter reviews the selection criteria proposed in the literature and assesses the quality of these criteria. This chapter is a edited version of: Knoop, V.L., Hoogendoorn, S.P., Snelder, M. and Van Zuylen, H.J. (2010) Reliability of link-based vulnerability indicators, in: *Proceedings of the 89th Annual Meeting of the Transportation Research Board*, 10-14 January 2010, Washington D.C. Accepted for publication.

---

### 8.1 Introduction

Numerous situations can be thought of in which large parts of a road network are blocked due to an event on one single location. For example, an incident in the peak hour in which a truck is involved could cause severe congestion on many roads in the surroundings of the accident location. Other less frequently occurring causes of disruptions are (terrorist) attacks, or disasters or calamities, or the threath thereof which causes the authorities to close the road. All disruptions cause delays, which is undesirable for road users. Road authorities want to know the most vulnerable links of their network because this enables them to protect or to improve those links or parts of their network.

The term “road network robustness” refers to these issues. In the literature different definitions of robustness can be found, but there is not yet a commonly accepted definition for robustness. The cause of disruptions is one of the most important differences. Sometimes only severe and non-recurrent disruptions are considered and sometimes daily

variations are also taken into account. The terms robustness and vulnerability are often used as opposites and this is also done in this chapter. They have a strong relationship, but they are actually each others opposites: vulnerability describes the weakness of a network and robustness describes the strength of a network. In chapter 1 robustness was defined as follows: “Robustness is the ability of the network to maintain its functionality under conditions that deviate from the normal conditions.” In this definition, the normal conditions are conditions in which traffic operations are within the boundaries of the design specifications, i.e. without serious incidents or exceptional demands. In this chapter we focus on (non-recurrent) incidents that block two lanes of a road. This choice was made because incidents on freeways with more than 2 lanes usually do not block the complete freeway. Incidents on roads with 1 or 2 lanes are assumed to block the road completely.

It is difficult to predict the robustness of a road network. Generally, there are two possibilities. Either one simulates all possible link blockings in a road network, which is computationally expensive. Alternatively, one pre-selects potentially vulnerable links based on an equilibrium assignment and certain criteria and performs an additional analysis for the selected links. The second approach raises the following questions. What is the quality of the selection criteria used in the second group? How large should the selection of possible vulnerable links be to be sure that the most vulnerable links are indeed included? And if the selection is good, is a detailed analysis really needed, or could the vulnerability and robustness of a network (or parts of the network) also be determined by applying only the selection criteria (without reducing the capacity for a selected link)? If this is possible, then it would make the modeling of the implications of protective measures much easier. A quick assessment of the vulnerability and the vulnerable parts of a network is also needed for the design of robust road networks with network design models. This problem of network design is very complex and computationally expensive even without the robustness aspect. A very long computation time for the robustness assessment would increase the computation time of the ‘robustness network design problem’ to an unacceptable level. Rather than running one traffic simulation for each road network layout, one has to run many simulation runs in order to assess the consequences of incidents at all locations of the network. Therefore, it would be useful to have indicators showing the most vulnerable parts. The objective of this chapter is to assess the quality and validity of different selection criteria for measuring road network robustness.

This chapter is restricted to assessing link-level criteria that could be calculated from one equilibrium run of the network, or, even better, can be measured from the everyday conditions in a road network. This choice was made because our aim is to limit the computation time and the required data.

The next section of the chapter gives an overview of the state-of-the-art methodologies used to identify the vulnerable links. Then, we present a description of the method that is used for comparing the selection criteria and an overview of the networks on which this comparison is made. The results and conclusions about the quality of the selection criteria are presented in the sections thereafter.

## 8.2 Literature-overview of methodologies to find vulnerable links

The methods for finding vulnerable links can be divided into two groups. The first group contains the “full calculation methods” in which the capacity is reduced for each link separately. In order to find out which links in a network are the most vulnerable, a complete simulation could be made. That is, for each link the capacity could be reduced and an assignment could be made. The effects of the capacity reduction on for instance the total travel time could be regarded as an indicator for the vulnerability of a link. Jenelius (2007) uses the approach of blocking each of the links in a traffic simulation program without traffic jams. In chapter 7 the same approach is used for calculating the consequences for a blocking at each link. However, there it is argued that the network effects including spillback are significant. Hence, a more accurate simulation that represents the dynamics of traffic jams, including spillback, was used. The simulation consequently needs a time dependent OD-matrix. The advantage of the approach, used by Jenelius (2007) and used in chapter 7, a full calculation, is that it gives a complete analysis. However, the computation time of this approach is very high which can be considered as a disadvantage; this brute force method is furthermore lacking a structure for searching weak links.

Several approaches have been introduced in order to overcome this disadvantage. In this second group of approaches a first selection of links that are likely to be vulnerable is made based on certain criteria. For these links a more detailed analysis is made by reducing the capacity and by assessing the vulnerability of these links. Tamminga et al. (2005) was the first method in which this approach was used. Also Tampère et al. (2007) introduced their own selection criteria. These methods are still computationally intensive because simulations for all the selected links are required.

There are also other approaches like the game-theoretical approach presented by Bell (2000). However, this has to the best of our knowledge never been applied in a dynamic simulation environment on a real-size network. These methods show the way people could avoid possible blockings and are more relevant for fully informed travellers.

## 8.3 Calculation of vulnerability indicators

This section describes the approach that is used for determining the quality of different selection criteria that are used in the second group of approaches that were described in the previous section. Section 8.3.1 shows the criteria we used to analyse. In section 8.3.2 is is described which traffic assignment method is chosen.

### 8.3.1 Selection criteria

The different selection criteria are listed below. The selection criteria  $I^1$ - $I^7$  can be found in Tampère et al. (2007), Li (2008) adds  $I^8$  and Tamminga et al. (2005)  $I^9$ . Despite the

**Table 8.1: List of symbols used**

Variable name	Description
Simulation level	
$\Delta t$	Time step
$\mathcal{I}$	The set of all criteria
$R$	The correlation coefficient
Per link $i$	
$I_i^n$	Criterion $n$ for link $i$
$q_i$	Flow, also taken as incident probability
$C_i$	Capacity
$C_i^b$	Remaining capacity at blocking
$v_{f,i}$	Free flow speed
$k_{j,i}$	Jam density
$L_i$	Length
$l_i$	Number of lanes
$\mathcal{R}_k$	The sum of the correlation of criterion $I^k$ with the other criteria

fact that most of these criteria were intended for identifying vulnerable links for different kinds of accidents compared to the kind we used (2 lane blockings), we still included them in order to get a wide range of criteria. As indicated in the introduction, we only assess the criteria that can be evaluated from one equilibrium assignment or that can be measured in a real-life situation without blocking.

Other criteria like the vulnerability index which was introduced by Murray-Tuite and Mahmassani (2004) and the recently introduced criteria of Kurauchi et al. (2007) require more input because they also include the possible route choice if a link is blocked. Compared to the other criteria, these add extra computation time for the rerouting part. Furthermore, they need extra (calibrated) information about users' choices when they face an unexpected blocking. Other criteria mentioned in literature (Tampère et al., 2007; Li, 2008) include the risk of a grid lock (cannot be calculated automatically), the quality on alternative routes (adds computational complexity) and the criterion that all off-ramps ( $I^5$ ) could be vulnerable (this is only one step in the selection process). The reasons for not including these criteria are mentioned between brackets. Finally, some criteria explicitly take the chances of an incident into account. This chapter discusses the possible consequences of an incident given that it happens and therefore also these criteria are excluded.

Below, a short description of each of the used criteria is given. Some of the criteria have been inverted, to get a better comparison. For each of the listed criteria, a higher value means that the predicted impact of the blocking of that link is bigger. The list shows the criteria and indicates the meaning. The symbols are explained in table 8.1.

1.  $I^1 = q / (1 - q/C)$

If the flow ( $q$ ) increases with respect to the capacity ( $C$ ) more travellers have to

queue.  $I^1$  expresses this influence of the flow.

2.  $I^2 = 1/T_b$ .

$T_b$  is the time it take before the tail of a queue reaches the upstream junction. The higher  $T_b$  is, the lower will be the impact of an blockage.  $T_b$  depends on the traffic inflow, the current density of the traffic and the length of the link. Tampère et al. (2007) shows the equation for  $T_b$ :

$$T_b = L_i / q_i (l_i \cdot k_{j_i} - q_i / v_{f_i}) \quad (8.1)$$

3.  $I^3 = I_i^1 \cdot \vartheta(q - 2500)$  Criterion 1 will indicate the links where the queues will be the largest. However, network effects play an important role. Therefore, it is important to also include links with a low capacity. Therefore  $I^3$  is the same as  $I^1$ , but limited to links with a capacity of 2500 pcu/hour. Mathematically, this expression uses the step function  $\vartheta(x)$ , which is 0 for  $x < 0$  and 1 for  $x > 0$ . This criterion should capture the offramps.
4.  $I^4 = I^1 \times q$ .  $I^1$  aims at expressing the effects of an incident. Tampère et al. (2007) argues that vulnerability needs an extra input, being the probability that an incident occurs. In the formulation for  $I^4$  this probability is taken proportional with flow  $q$ .
5.  $I_i^5 = I_i^2 \times q_i \times \sum_{\text{upstream links } j \text{ of } i} I_j^1$ .  $I^5$  is equivalent to  $I^4$ , capturing both effects and incident probability. However,  $I^5$  also takes the possible effect of blocking back into account. It does so by multiplying  $I^4$  by the effect of a blockage on link  $j$ , (estimated as  $I_j^1$ ).
6.  $I_i^6 = I_i^3 \times q_i \times \sum_{\text{upstream links } j \text{ of } i} I_j^1$ .  $I^6$  is the same as  $I^5$ , but restricted to lower-capacity links. This would capture for example risk-prone off ramps just downstream of a motorway junction.
7.  $I^7 = \sum_{\text{upstream links } j \text{ of } i} I_j^1$ .  $I^7$  is a sum of the effects if (estimated by  $I^1$ ) if all upstream links  $j$  of link  $i$  are blocked by spillback. This shows the links that cause large problems in blocking back effects: for example a link just downstream of a motorway junction.
8.  $I^8 = \frac{q}{C}$ . This captures the links that have a large volume compared to their capacity. This ususally is an indication that the link is heavily used, and that if an blockade happens, the queue will grow quickly.
9.  $I^9 = q_i - C_i^b$ . This shows rate at which cars arrive in the queue when an incident occurs on a link and therefore shows the direct consequences; in this chapter, it is assumed that  $C_b$  equals 0.

### 8.3.2 Assignment

Assignments can be divided according to several criteria, like static or dynamic, user equilibrium or no equilibrium, stochastic or deterministic, path based or link based, single user class or multi user class, unimodal or multimodal and en-route route choice possibility or no en-route route choice possibility. For modelling robustness, especially the difference

between static and dynamic assignments and the possibility for en-route assignment are important. It is generally accepted that dynamic assignments are required for correctly modelling robustness. Compared to static assignments, dynamic assignments are better at showing the exact location of congestion and at determining the development over time of congestion. This is important for correctly modelling the effects of variations in demand and capacity (e.g. incidents). The possibility of en-route route choice is important, because in practice a certain percentage of the travellers change their route when they are informed about congestion at a certain location. The importance of en-route route choice for the assessment of the impact of incidents is advocated in the thesis of Li (2008). Tampère et al. (2007) argue that en-route route choice can indeed be of added value, but that it is very difficult to correctly model the en-route route choice of travellers during incidents because of the uncertainty that is inherent to human behaviour, see for instance Bogers et al. (2005). Especially during incidents this uncertainty is important, because it is not known how many people have information about the incident and how they will respond to that information. Besides these two characteristics, Tampère et al. (2007) also claim that a correct modelling of the way in which congestion builds up (at least consistent with first order traffic flow theory) and a correct modelling of intersections is required for vulnerability analysis.

We used the traffic assignment model INDY (Bliemer, 2005, 2007). INDY is a dynamic path based multi-user class assignment model. The model finds an equilibrium route set for three driver types: drivers which use a fixed path, drivers with deterministic route choice and drivers with stochastic route choice. In INDY congestion is modelled in line with the first order traffic flow theory, which means that congestion is based on a capacity reduction. The head of congestion is located at the bottleneck, and the tail of congestion moves upstream. If the capacity at the bottleneck increases, congestion will solve from the head. This congestion modelling is described by Yperman (2007). En-route route choice is not possible in INDY. However, since INDY was used to simulate non-incident situations only, the lack of en-route route choice is not relevant. The package gives a good representation of the network flows without incidents. Therefore, the assignment results can be used for the evaluation of the nine robustness criteria.

Obviously, when facing an incident, it is likely that drivers will deviate from their equilibrium paths. Therefore, for the full calculation a different, dynamic non-equilibrium traffic simulator was used. The macroscopic simulator DSMART introduced by Zuurbier et al. (2006) includes en-route route choice and blocking back. It is a cell transmission model. The assessment of the vulnerability of each link was done by evaluating the impact of blocking single links using this simulator. In this case, blocking means that 2 lanes were blocked (or one if the link only contains one lane). The total travel time (including the delay at the origin) was used as performance indicator. More details can be found in chapter 7. Only for the Rotterdam network the traffic simulation program DSMART is used. For this simulation a calibration has been carried out that was based on link counts on the freeways.

## 8.4 Analysis

Section 8.3 shows how different indicators can be calculated. This section shows how they are compared with each other (section 8.4.1). In section 8.4.2 it is shown how the vulnerability indicators are compared with a assessment of vulnerability by simulation (iteratively block a link and calculate the performance decrease).

### 8.4.1 Redundancy of criteria

Vulnerability indicators for all links are calculated for three different networks (see section 8.5 for the networks). For all three networks, we compared the different criteria. First of all, the mutual cross-correlations were calculated. This indicates how good the correlation between the numbers is. In this chapter,  $R$ -values are presented, which have the advantage over  $R^2$ -values that one can see the sign of the correlation. The value of  $R$  lies between -1 and 1.

This is a linear correlation method of which the underlying assumption is that the numbers might be mutually linearly dependent. Any other relationship that would give the correct order of vulnerability, also non-linear relationships, could make a perfect prediction. Another correlation test, the Spearman Rank Correlation (Spearman, 1904) is a standard test to show how much the ranks are correlated.

The advantage of using this test is that it shows whether the ranking is correct. However, if the values of the criteria are similar, it can be more interesting to know whether there are similarities between the values than to see the differences within that group. In that case the rank correlation might be low, because the ranking within a group is changed, but the indicators might give an reasonable estimate for the vulnerability. In figure 8.4 the scatterplots for the values are given. These show that there is no relationship between the indicators. We choose to fit a linear relationship to show the error, and we expect that fitting any other function would not reveal another relationship.

The correlation coefficients  $R$  between two indicators show the correlation coefficients of one criterion with all other criteria. Now, a sum of these variables,  $\mathcal{R}$ , can be defined, which indicates whether that criterion shows the same trend as others.

$$\mathcal{R}^k = \sum_{l \in \mathcal{I}; l \neq k} R(I^k, I^l) = -1 + \sum_{l \in \mathcal{I}} R(I^k, I^l) \quad (8.2)$$

In this formula,  $k$  and  $l$  are the numbers of the criteria. If criterion  $k$  matches positively and linearly with criterion  $l$ , the value  $R(I^k, I^l)$  equals one. If there is no correlation at all,  $R(I^k, I^l)$  equals zero. Since  $I^k$  matches perfectly with itself, the value  $R(I^k, I^k)$  equals one; this explains the second equal-sign in equation (8.2).

A high value of  $\mathcal{R}^k$  now means that there is a high positive correlation between criterion  $k$  ( $I^k$ ) and the other criteria. That means that its value can represent the average of the

other criteria well, or the other way around, the average of the other criteria already tells something about the value of criterion  $I^k$ .

Apart from the correlations and rank correlations, we examined the top- $n$  links, the  $n$  links predicted as most vulnerable. Each of the criteria orders the links on a vulnerability scale. We compared the orders given by the different criteria. In particular, we analyzed whether the links that are indicated as most vulnerable are the same. For that purpose, we calculated the relative overlap between the top- $n$  most vulnerable links.

Note that for the small network and the Delft network the results of statistics can be influenced by the small number of links that is available. The statistical analyses are most interesting for the larger Rotterdam network.

## 8.4.2 Predictive value of criteria for simulation result

For one network, the combined selection power compared to a simulation result was examined. Since the criteria are intended to complement each other, the minimum number of links that is to be selected by each criterion in order to get the complete top- $n$  of the full analysis (found in the same way as in chapter 7) was determined. It is assumed that this method gives a correct representation of the vulnerability of each of the link.

In an example it is now shown how this method works. If for instance link number 10 is the most vulnerable link according to the full analysis, then the position of link 10 is determined in the link ordering of the different criteria. Thereafter the minimum is determined. It could be that  $I^3$  is the criterion that gives link 10 the highest rank: position 3. From this, it would be concluded that at least 3 links are to be selected by each criteria. Since it is likely that there is an overlap in the selected links by each criterion, the number of uniquely selected links is also presented.

Also the correlation coefficients and rank correlation coefficients for each indicator and the full calculation are determined for the simulation result.

## 8.4.3 Multi-linear fit of criteria

For one network, the delay caused by the blocking of the link (D) is known as well as all indicators. We propose a linear model to predict the delay-values for each of the links

$$\widetilde{D}_i = \sum_{k \in \mathcal{K}} \beta^k I_i^k \quad (8.3)$$

In this equation,  $\mathcal{K}$  is a set of criteria. The most simple models would take a set of 1 parameter, e.g.  $\mathcal{K} = \{1\}$ . The maximum complexity of the model is if all 9 criteria are included. Now, vector  $\beta$  is optimised in order to minimise the error,  $\varepsilon$

$$\beta = \underset{b}{\operatorname{argmin}} \varepsilon = \underset{b}{\operatorname{argmin}} \sum_{i \in \mathcal{K}} \left( TDL_i - \widetilde{D}_i \right)^2 \quad (8.4)$$

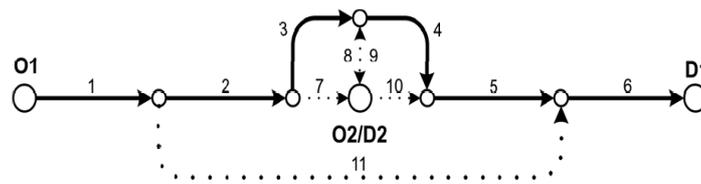


Figure 8.1: Simple test network Source Li (2008)

The aim of the fit is that the predicted value for the delay ( $\widetilde{D}_i$ ) is similar to the simulated total delay ( $D_i$ ). For a set of validation links, we compute the residual error,

$$\varepsilon = \widetilde{D}_i - D \quad (8.5)$$

Ideally, this would be zero. It is most interesting to analyse the variations in the errors, rather than a constant offset. Therefore we assess the quality of the prediction model by the standard deviation of  $\varepsilon$ .

The vector  $\beta$  is estimated based on a calibration set of links which is a sub-set of all the links in the network. One third of the links will be kept out of the calibration. These links are used for the validation. In fact, it is calculated what is the predicted delay for a blocking on each of these links. These predicted delays are compared with the delays calculated in the dynamic traffic simulation program.

## 8.5 Networks

For the comparison of the selection criteria, we used three differently sized networks. We used a simple test network to show clearly the characteristics of the different indicators. The second test network is a bit more detailed and shows the effects of on and off ramps. The simulation of traffic in a real-world, medium-sized network shows how the effects work out in practice (third network). The characteristics of the three networks are presented in table 8.2.

The first network studied is a test network with 11 directional links (figure 8.1). It can be seen as a freeway that passes a city. There are connections to the city (links 7, 8, 9 and 10) and there is a local road that passes the city (link 11). All local connections have a maximum speed of 50 km/h, whereas the freeway has a maximum speed of 120 km/h. As congestion sets in, more drivers take the local road around the city.

The second network is a test network that is based on the network of Delft in the Netherlands (figure 8.2). The freeways around the city are included as well as the largest two roads through the city. All local roads, which usually have a very low traffic volume and for which there are, due to the network structure, many alternatives, are excluded. The on and off ramps are modelled in detail. Since the capacity and location of on and off ramps is likely to be of relevance for the robustness of a road network, this is an important addition compared to the first test network.

**Table 8.2: Network characteristics**

	Simple	Delft network	Rotterdam network
Links	11	150	454
Nodes	5	90	239
Centroids	3	12	44
Paths	9	478	2071

The network around the city of Rotterdam (about 600,000 inhabitants, see figure 8.3) is the third network considered here. The freeways around the city are modelled as well as the most important corridors through the city. The network is used for local traffic and for transit traffic. The period from 6.30 to 9.30 in the morning was simulated.

## 8.6 Results

In the section, we present the interesting results for all three networks.

### 8.6.1 Simple Network

All indicators are formulated chosen in such a way that bigger values indicate a higher vulnerability for the network. It is therefore remarkable that some of the correlation indices are negative, meaning that a best fit is a negative relationship.

$\mathcal{R}$  is even negative for  $I^3$  and  $I^6$ . For  $I^3$ , it can be explained by the exclusion of the freeway links. When the freeway links are vulnerable according to the other criteria and (by exclusion) they are not any more according to  $I^3$ , the correlation coefficient becomes negative.  $I^6$  uses  $I^3$  as input, so it was expected that it would follow the trend of  $I^3$ . As that counteracts the average, so will  $I^6$ . The cross correlation of  $C3$  and  $C6$  is relatively high (0.81). It is also the only combination with the same top-1, top-2, top-3 and top-5 of vulnerable links.

The correlation of the  $I^1$  and the  $I^9$  is the highest of all with an R of 0.99. It is, apart from  $I^3$  and  $I^6$ , the only combination that produces the same top-5 (though not in the same order). Other related combinations are:  $I^2-I^4$ ,  $I^1-I^5$ ,  $I^1-I^9$ ,  $I^2-I^8$ , and  $I^4-I^9$ .

### 8.6.2 Delft network

The strong correlations are the same in the Delft network. The cross correlation values are in the same order of magnitude, but the accordance of the top- $n$  values is lower. Due to the higher number of links, there is less chance of accidentally including the same links in the top- $n$  ( $n$  is chosen as a percentage of the total number of links).

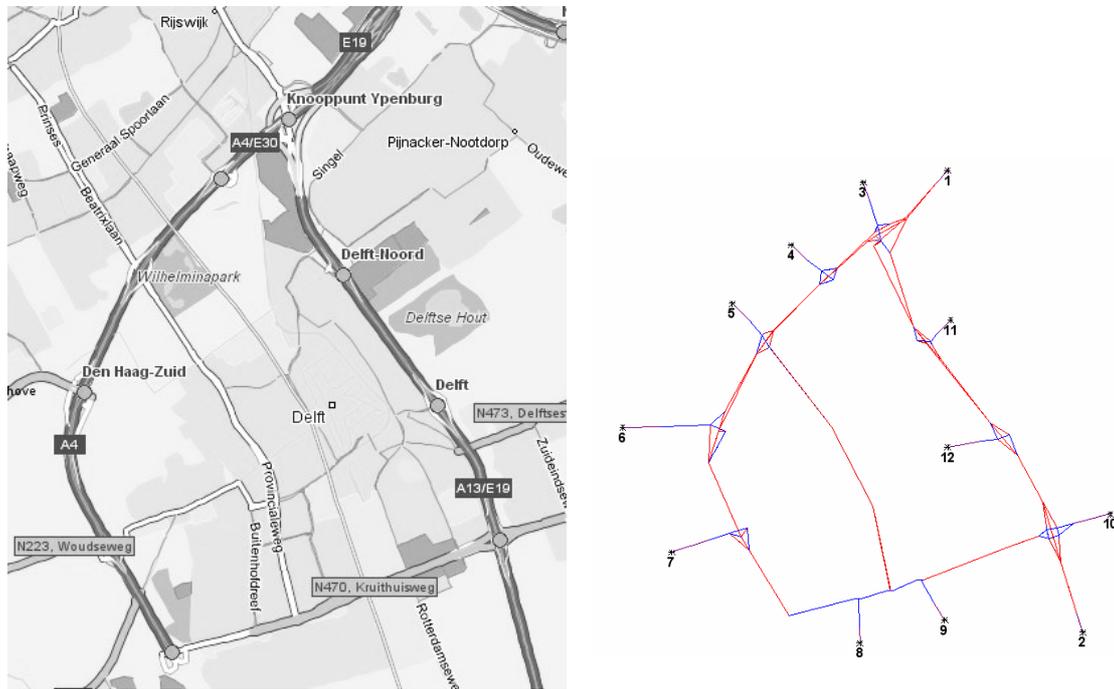


Figure 8.2: Network of Delft, the Netherlands

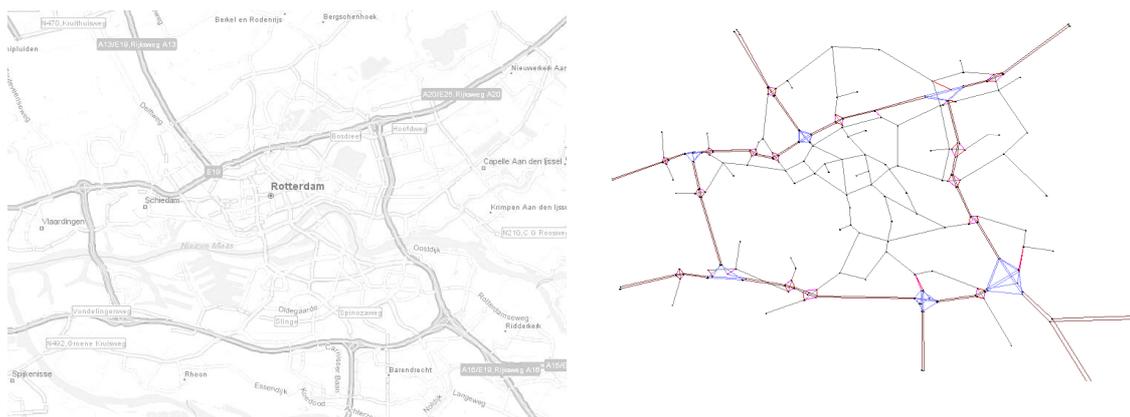


Figure 8.3: Network of Rotterdam, the Netherlands

**Table 8.3: The correlation coefficients**

Correlation	Linear	Rank
Full calculation and $I^1$	0.15	0.10
Full calculation and $I^2$	-0.01	-0.01
Full calculation and $I^3$	0.078	-0.05
Full calculation and $I^4$	0.13	0.15
Full calculation and $I^5$	0.15	0.11
Full calculation and $I^6$	0.09	-0.06
Full calculation and $I^7$	0.052	0.02
Full calculation and $I^8$	0.15	0.12
Full calculation and $I^9$	0.10	0.08

Here, we find strong correlations in the following combinations:  $I^1$ - $I^9$  and  $I^3$ - $I^6$ . The value for  $\mathcal{R}$  varies from 1.7 ( $I^2$ ) to 4.4 ( $I^1$ ).

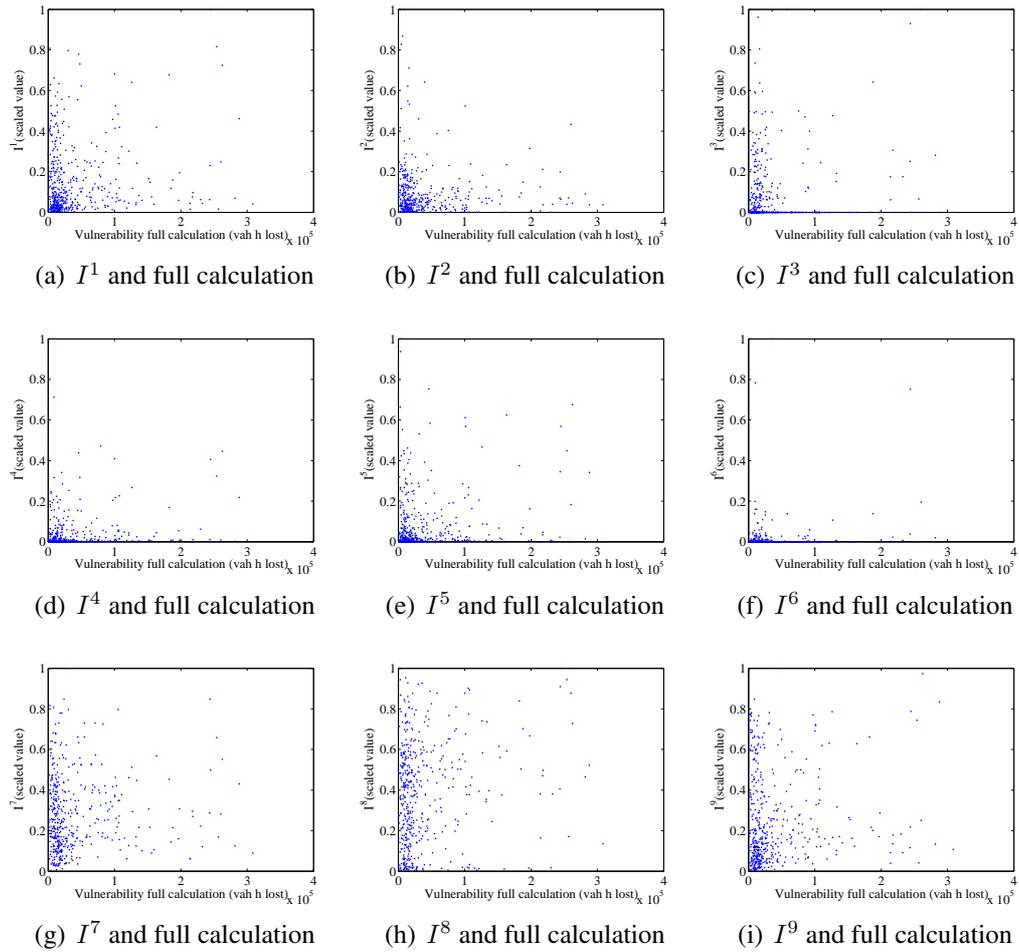
### 8.6.3 Rotterdam network

#### Analysis of indicators

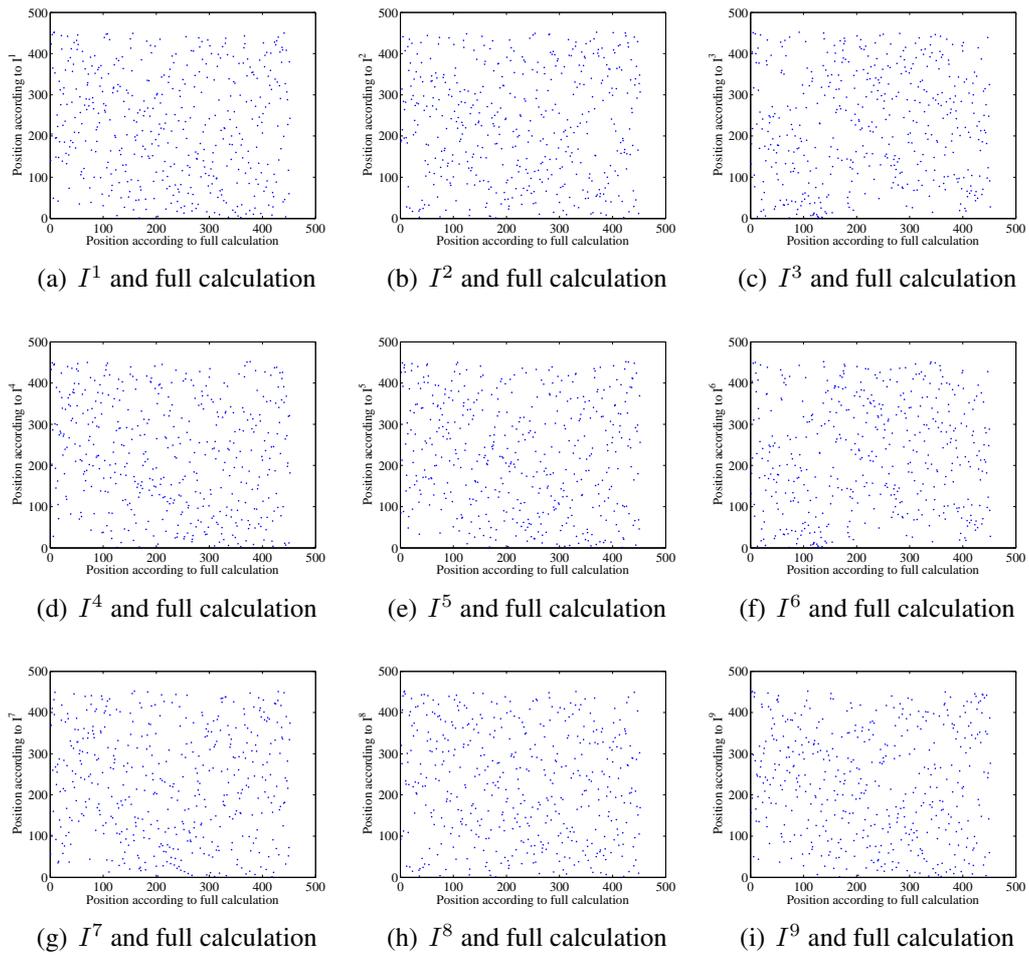
Since the statistics on this 454-link network have the least random error, for this network all results for the comparison with the full calculation are presented. Figure 8.4 shows the correlation between each of the indicators with the full computation. Figure 8.5 shows the ranks of the links (the lower the number the more vulnerable it is) for each of the indicators and for the full computation. Clearly, the correlation in values and rank is low: in fact, the points are all over the area. The correlation results as proposed in section 8.4.1 are presented in table 8.3. None of the indicators can properly predict the consequences of a blocking. The highest  $R$  is 0.15.

In this real-world network, the same combinations of indicators are related as in the other networks. There is one relationship that correlates more than in the other networks,  $I^1$ - $I^5$ . The cross correlation value  $R$  is 0.85.

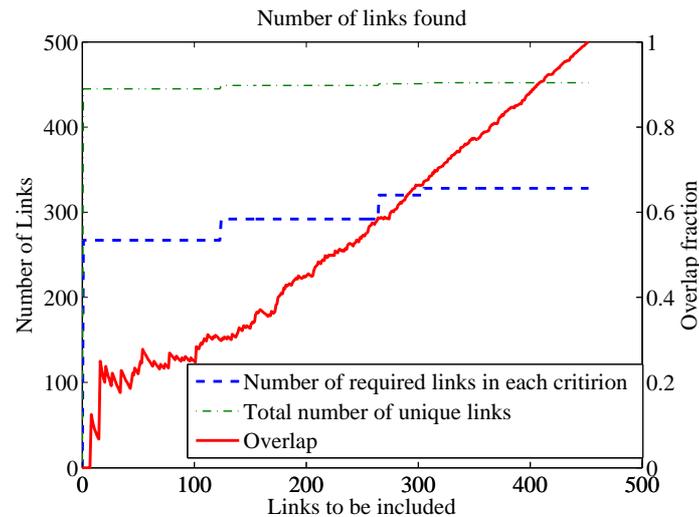
The combined selection power of the links is shown in figure 8.6. The figure shows three lines which are explained in this paragraph. First, the figure shows the number of unique links ( $y$ ) that are to be selected by each criterion in order to get the complete top- $n$  of the actual impact analysis. From the line ‘Number of links required to select per criteria’ it can be concluded that more than 250 links (55% of all links) need to be selected in order to include the most vulnerable links. Secondly, the figure also shows the number of unique links that result from selecting  $y$  links by each criterion. In the example given above this corresponds to the union of the top-250 links from all criteria. Already for a very small search area (top-1 vulnerable link), a very large subset of links needs to be considered (almost all). This implies that, at least for this case, pre-selecting links has hardly any



**Figure 8.4:** The scatterplot of the results of the full calculation compared with the calculated criteria



**Figure 8.5: Scatterplot of the ranks of the full calculation and the criteria**



**Figure 8.6:** The number of links that should be evaluated after pre-selecting using all criteria

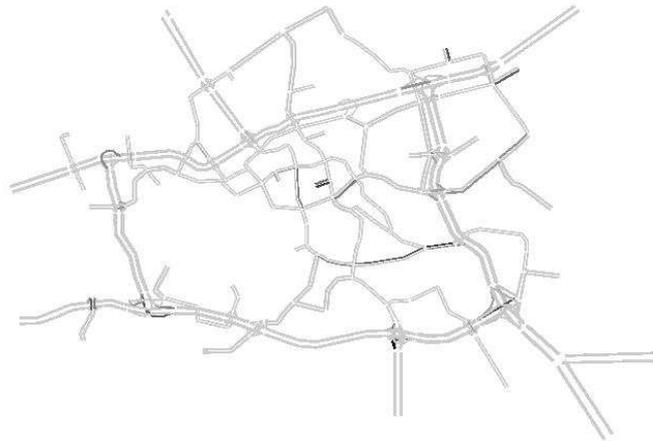
added-value. Finally, the figure shows the overlap between the top- $n$  links selected by the criteria and the top- $n$  links based on the actual impacts. If the high level of the required links to include in the further analyses is caused by a few links that are not selected by the criteria, or in other words, if most of the vulnerable links are captured by the criteria, this would appear in the overlap. We selected the most vulnerable  $n$  links according to the full calculation and we analyzed which percentage of these links also appears in the top- $n$  of any criteria. This is the overlap percentage. The line will go to 100% for all links: all links belong to the set of vulnerable links if there is no threshold. The ‘overlap line’ shows that 10% of the top-10 of most vulnerable links are included in the selection of the criteria based top-10. For the top-150 this is 33%. This implies that it is not just 1 link that is missing in the criteria selection.

Figure 8.7 shows which links are the most difficult to find. A dark shading implies that the maximum of the criteria scores for that link is higher than its ranking (easier to find). The other links are shaded with an intensity according to the difficulty of finding these links. From this figure it can be concluded that especially the freeway junctions, the links downstream of junctions and the main urban arterial are not well covered by the criteria. This could imply that spillback effects are not properly included in the criteria.

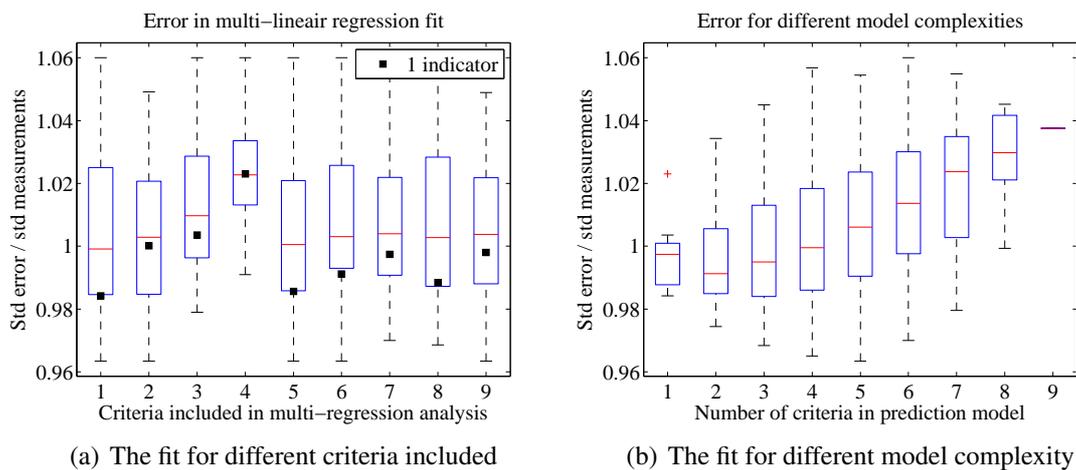
### Multi-linear fit

For the Rotterdam network, the full computation results are known. We fit a multi-linear model on the indicators to approximate the full computation results, as explained in section 8.4.3.

Note that there are  $2^9 = 512$  possibilities to fit a multi-linear model if each of the 9 criteria could be included or not. Exactly half of them includes  $I^1$  and the other half does not. For each of the criteria there is a set of models if it is included. Figure 8.8a shows the distribution of these models. It shows that generally the performance is very poor. This is derived from the fact that the standard deviation of the residual error (equation 8.5)



**Figure 8.7: Links that are difficult to pre-select with the selection criteria**



**Figure 8.8: The goodness of fit of the multi-linear models**

of the validation set is approximately the same as the standard deviation of the delays resulting from the simulation. So in fact, the model could not find an explanation for the deviations. In some cases, the standard deviation of the model becomes even worse, meaning the model overfitted the results.

The boxes in figure 8.8a show the distribution of the model fits with different complexity. A square shows the performance of the linear fit including only one indicator (and none of the others). This is generally a slightly better fit than the fit which included more parameters.

This is also seen in figure 8.8b which shows the fitness for models with different complexity, i.e. the number of criteria that are included in the fit. The standard deviation of the residual error for models with one parameter lies slightly under the standard deviation

of the the measurements. That shows that including 1 indicator is slightly better than not having any information at all. However, when more criteria are included in the model, the error increases, which means the results are overfitted. Obviously, if one would plot the fit results for the calibration set rather than the validation set, the error will go down.

## 8.7 Conclusions and recommendations

This chapter compares different criteria that have been proposed in literature to indicate the most vulnerable links in a network. We found that the different criteria indicate different links as most vulnerable. They should therefore be seen as complementary. Excluding freeways gives a completely different list of vulnerable links. This implies that the freeways are usually (i.e., by the other indicators) indicated as vulnerable. The Incident Impact,  $q/(1 - q/C)$ , gives the best correlation with the other criteria. However, when comparing it to the fully calculated results, it is not better than the others.

In fact, none of the indicators on their own gives a good representation of the full consequences of the blocking of a link. It is also insufficient to take the top-level numbers and analyze them in depth, as there is no indication that the indicated top-level vulnerable links are indeed the most vulnerable. Apart from that, they differ among the criteria. Furthermore, a combination of the criteria also did not result in a good prediction of the list of most vulnerable links. The combined selection power of the criteria in the network appeared to be minimal. Especially, the freeway junctions, the links after the junctions and the main urban arterial are not well covered by the criteria. This could imply that spillback effects are not properly included in the criteria.

From these results it can be concluded that the quality of these criteria is not good enough to properly identify the most vulnerable links in a network. Also a linear model combining the criteria cannot predict the right vulnerability of links.

Future research should analyse the quality of the selection criteria for identifying vulnerable links for other disruptions than 2 lane blockings. Furthermore, the conclusion that the existing criteria are insufficient, should lead to new research to find out whether new criteria can be introduced that enable us to identify vulnerable links without doing a full calculation. For instance, the indicators based on routes instead of links, for example mentioned by Murray-Tuite and Mahmassani (2004), can provide an interesting approach to this problem which is to be integrated in a future study.



# Chapter 9

## Risk-averse traffic assignment using a dynamic simulator

---

Chapter 5 discusses the route choice of travellers after an incident has happened. Chapter 7 showed that the consequences of an incident can be severe. Therefore, some risk-averse travellers might choose their route based on the risk of possible incidents and possible queues. This chapter combines a dynamic traffic simulation program and the framework of a risk-averse traffic assignment. In a dynamic simulator with accurate representation of spillback, a disruption at one place can cause traffic jams at others or at other times. Any suitable traffic simulator can be fitted in the framework presented here. The risk-averseness of drivers is represented by a two-parameter route choice model. A layered solution algorithm prevents the computation time becoming too large.

In the case study presented, the consequences of an accident on the motorway are large and risk-averse drivers are assumed to assign a large weight to the scenarios which give large delays. In a way, this weight therefore is also an indicator of vulnerability as discussed in chapter 8. The case-study shows that risk-adverse travellers will leave the motorway while risk-neutral users still assume the motorway to be faster. With increased demand, the risk-neutral users also change to the slower but quieter and more certain route.

---

### 9.1 Introduction

Often, one needs to be certain that one arrives before a certain time, regardless of the road conditions. That is one reason why road authorities and governments nowadays are also setting goals for reliability, for example the Netherlands in the Nota Mobiliteit (Dutch Road Authority, 2005a).

If people want reliable routes, they tend to avoid the routes that can have long delays, even if such delays are seldom. This chapter discusses a traffic assignment which takes

risk avoidance into account. However, the probability distribution of travel times on links and the probability of incidents on links is not known to the travellers beforehand. They therefore have to make assumptions based on their risk-attitude.

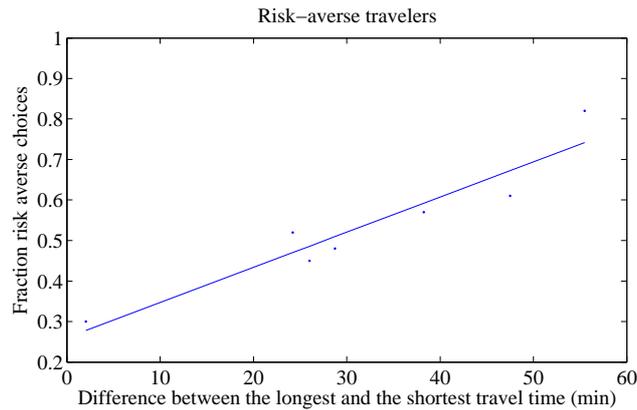
Risk-averse traffic assignment has been studied by, for instance, Lo and Tung (2003). They propose a way to calculate a probabilistic user equilibrium assignment. This assumes that the average travel time of travellers, including the occasional delay due to an incident, is equal over all routes which are in use for a certain OD pair. Recently, Marchal and de Palma (2008) showed that an equilibrium can also be found for a dynamic traffic simulation. However, in both approaches, the anticipated probability of an incident is taken fixed.

A different approach to risk-averse route choice behaviour is introduced by Bell (2000). He assumes that risk-averse drivers anticipate worst cases and minimize their exposure to these; note that the worst case is influenced by the traffic loads. Until now, this game theoretical approach to risk-averse assignment has only been combined with static models. However, it is known that the travel times and the location of congestion in these models are not correctly calculated statically and that these models are therefore not able to calculate the duration of congestion and the location of the queues, as shown in chapter 7. This matters especially for cases in which there is unexpected blocking back, causing severe traffic jams. More importantly, in static models the metering effect of bottlenecks, which cause queues to grow or decline, is not represented. This time-dependency and spillback (or blocking back) can be captured in dynamic traffic simulation models. This chapter combines a game theoretical approach to risk-averse traffic assignment with dynamical traffic simulation.

The risk-averse route choice model proposed here can be used in traffic assignment models. In the future, this risk-averseness might be implemented in journey planners. One can conceive of an on-line journey planner with a slide bar for risk-averseness. In general, risk-minimizing behaviour is relevant for transport of special goods or persons, such as hazardous materials or VIPs. For other drivers, risk-averseness can depend on the purpose of the trip. When trying to catch an airplane, one is likely to be more risk-averse than when making a family trip, for example. Nowadays, just-in-time delivery requires freight transport to be reliably in time. When a truck misses the time window for delivery, it often needs to return a day later to make the delivery. Freight transport is therefore likely to be interested in risk-averse route choice. Bogers and Van Zuylen (2004) give evidence that truck drivers become more risk avoiding the larger the uncertainty of a route, as shown in figure 9.1.

The result of a risk-averse traffic assignment can also be used in the evaluation of the vulnerable parts of a network. It is essential to know which scenarios cause large delays when trying to improve the network reliability. When one specific scenario gets a large weight in the decision-making process of risk-averse travellers, that scenario contains a risk of large travel times.

The chapter is set-up as follows. First, section 9.2 reviews the literature on network robustness and risk-averse traffic assignment. Section 9.3 describes how a dynamic traffic



**Figure 9.1:** The fraction of risk sensitive truck drivers in category 3 plotted against the difference in range - from Bogers and Van Zuylen (2004)

simulator can be fitted into a framework of risk-averse traffic assignment. The model is applied to a case study which is presented in section 9.4. Section 9.5 shows the results of the case study and the final section presents the general conclusions of the modelling framework.

## 9.2 Literature review

The issue of reliability has become a subject of research over the last decade. Liu et al. (2004) show that users highly value reliability. The studies mentioned in table 7.1, presented in section 7.1 show that people avoid routes with possible high delays. Bogers and Van Zuylen (2004) present a study which shows that experiences of extreme travel times, the travel times typically caused by incidents, influence the route choice.

Extreme travel times can be caused not only by accidents on the link itself, but also by blocking caused by congestion spillback from another link as shown in chapter 7 using simulation. On the other hand, if a trip has alternative route options, the impact of an incident on one of the routes will be smaller as shown by Snelder et al. (2008a).

Chapter 7 uses two types of routes. First, there are routes independent of the blocking based on a network without a risk of blocking. Then, there are routes with en-route assignment – dependent on the location of the blocking, people will take another route. Jenelius et al. (2006) present measures for the importance of links for network performance. They define a measure of importance and exposure for a link, based on the flows and the delays if the link is blocked. This also requires intensive calculation. In contrast to chapter 7 of this thesis, his measures focus on the connectivity of the road network. In particular, this method can be used to improve network planning.

Liu et al. (2002) looked at the route choice process as determined by a stochastic perception of the travel time and its variability by users. Uncertainty of travel time is in their

model a part of the route disutility. They assume that the travel time and travel time uncertainty are terms in a utility function and solve the stochastic network (SN) combined with stochastic perception dynamic user optimum (SDUO). The disutility in their model depends on the travel time, observation error, uncertainty of the travel time and the risk attitude of the traveller. Liu et al. took a normally distributed travel time function without specifying the mechanism that causes the travel time uncertainty. Tu (2008) derived travel time reliability functions for routes and used that function as a component of the utility function, next to travel time and scheduled delay.

Bell (2000) was the first to approach link-vulnerability via a game-theoretical approach. He assumes the road users do not know the chances of blockage of a link. They will count on the worst case scenario. The users change their route for this scenario. Due to the changed routes, the possibility for the occurrence of queues change, and the network performance in each scenario changes. As a consequence, another scenario is now perhaps the worst case scenario. Users now anticipate other scenarios more. Bell allows for a mixed strategy for the worst case, i.e. there is a distribution of blocking probabilities over different links, adding up to 1. The opposite would be to require a pure strategy: the blocking probability is either 1 or 0. This is for instance done by Brown et al. (2006) for a network of oil supply and an electrical transmission system.

In the concept by Bell (2000) there are just two costs for a link possible: when it is unblocked the costs are low and when it is blocked the costs are high. Belda et al. (2002) extend the concept. In their paper, a blocked link has a lower capacity than an unblocked link. The traffic demand then determines the travel costs, using a link-based relationship between flow and travel time. However, the traffic representation is still static and the consequences are link-based.

Nagae and Akamatsu (2007) continue on this line of research. They point out that it might be too extreme to expect the worst case situation to happen. They relax the assumption of people being completely risk-averse. They add two extra terms to spread the breakdown chances over the different scenarios. The term they add is comparable with an entropy term: it moves the solution towards a more evenly spread solution. This changes the perspective on route choice behaviour slightly, but it also makes the mathematical framework much easier to solve (see also section 9.3.1). Bell et al. (2008) use the same but, compared to the strictly risk-averse simulation, they just relax the perception of link failure probabilities (and not the route choice). The mathematical advantage then still holds.

In chapter 7 it is pointed out that, especially in cases with a link blocking, the network effects (including spillback) are important. Some studies try to identify the most vulnerable links in a network without a thorough analysis with a valid traffic flow model. Kraan et al. (2008), for instance, provide indicators which estimate the vulnerability based on characteristics of a link. However, it is shown in chapter 8 that these link-level robustness indicators are insufficient to indicate which links are really vulnerable, mainly due to network effects.

None of the studies using the game-theoretical approach to link-level vulnerability have incorporated spillback or any dynamics of the traffic pattern. The main contribution of this chapter is that it combines the risk-averse assignment with a dynamic traffic simulation.

It finds the vulnerable parts in the network using a risk-averse traffic assignment. All articles that handle risk-averseness in a game-theoretical way so far assume that all people have the same route choice behaviour. In this chapter, different user classes are considered each with their own route preferences, like introduced in Knoop et al. (2008a).

## 9.3 Methodology

This methodology can be used to identify vulnerable links. This is done via a risk-averse route assignment. In this section, we will first present our model and then relate the outcome of the model to link vulnerability, the magnitude of the problems occurring when an incident happens at a link.

### 9.3.1 Mathematical modelling

#### Mathematical formulation

The method presented here is an extension of the one presented by Bell (2000). He applied a static method to find the scenarios (for instance a link blocking) which were better avoided. Furthermore, in Bell's work, each scenario changed the travel times on just one link. In our formulation, there is a time dimension in the traffic operations and the influence of a blocking can extend to other links and other time slices then in which the blocking takes place. For the sake of simplicity, we formulate the model for one destination. A multi-destination network is easily fit into this model. All route-related variables should then be considered for each destination separately.

Table 9.1 lists the symbols used in the formulation of the game. The key variables in the problem formulation are  $\mathbf{h}_{U,t}$  and  $\mathbf{f}_U$ .  $\mathbf{h}_{U,t}$  is the route choice that the travellers make (and changes for different user classes  $U$  and time intervals  $t$ ). It can be seen as a vector of split fractions. The  $z^{th}$  element indicates how many travellers choose route  $z$ . There are different scenarios possible for the network state (indicated by  $j$ ). One scenario could for instance be a blocking at a specific part in the network.  $\mathbf{f}_U$  indicates the anticipation of user class  $U$  on these different scenarios. It therefore is a vector whose  $j^{th}$  element indicates how likely users of class  $U$  estimate the occurrence of scenario  $j$  is.  $\mathbf{f}_U$  is taken time-independent for this study.

Users want to minimize the cost of travelling, which is expressed here in total travel time. If necessary, also extra cost (route preference, tolls) can be included. The travel time is known for each link  $i$  under each scenario  $j$ . We introduce a variable  $a_{iz}$ , which is the path parameter which is one if  $i$  is on path  $z$  and 0 otherwise. We can now construct the link travel times,  $\text{tt}_{jzt}^{\text{path}}$ .

$$\text{tt}_{jzt}^{\text{path}} = \sum_i a_{iz} \cdot \text{tt}_{ijt}^{\text{link}} \quad (9.1)$$

We now take a weighted average over the routes  $z$  (using the vector  $\mathbf{h}$ ) and over the different time intervals  $t$  (using the demand vector  $\mathbf{Q}$ ). This gives us the cost under scenario  $j$ ,

**Table 9.1: The symbols used**

Symbol	Meaning
$U$	User class
$t$	Time, discretized in time slices numbered by the integer $t$
$\Delta t$	The duration of one time slice
$z$	Path number
$\mathbf{f}_U$	Vector of link failure probabilities for user class $U$
$\mathbf{h}_{U,t}$	Route choice vector for user class $U$ : fraction of travellers in class $U$ for each route $z$
$i$	Link number
$j$	Scenario number
$\mathcal{B}$	Set of scenarios with a blocking (i.e., all scenarios except a free network)
$a_{iz}$	Link incidence matrix (1 if link $i$ is on path $z$ , 0 otherwise)
$tt_{ijt}^{\text{link}}$	Travel time in time slice $t$ on link $i$ under scenario $j$
$tt_{jzt}^{\text{path}}$	Travel time in time slice $t$ on path $z$ under scenario $j$
$\tau_{j,U}$	Travel time of a user in class $U$ under scenario $j$
$\langle tt_{it} \rangle_U$	The anticipated travel time in time slice $t$ on link $i$ for users in class $U$
$\langle T \rangle_U$	The anticipated travel time for a user in user class $U$
$\vartheta$	Constant indicating the risk-averseness of a user class
$n$	Iteration number of the algorithm
$m$	The reciprocal frequency [iteration numbers] for calculation the Total Cost of all scenarios
$w$	Number of scenarios assessed when iteration number is not a multiple of $m$
$D$	Operator for the distance between 2 vectors
$th$	Threshold value for converging algorithm
$p_U$	Probability assigned to the blocking scenarios for the risk-neutral users ( $0 < p_U < 1$ )

$\tau_j$ :

$$\tau_{j,U} = \sum_t \sum_z Q_t h_{z,U} \mathbf{tt}_{jzt}^{\text{path}} = \sum_t \sum_z \sum_i Q_t h_{z,U} a_{iz} \mathbf{tt}_{ijt}^{\text{link}} \quad (9.2)$$

The anticipated cost for travelling depends on the anticipated likelihood for each scenario,  $f_{j,U}$ . The total anticipated cost is the weighted average of the costs for each scenario:

$$\langle T \rangle_U = \sum_j f_{j,U} \tau_{j,U} = \sum_j \sum_t \sum_z \sum_i f_{j,U} Q_t h_{z,U} a_{iz} \mathbf{tt}_{ijt}^{\text{link}} \quad (9.3)$$

The key of a risk-averse approach, as introduced by Bell (2000), is that a risk-averse user would give the largest probability in their anticipation on the worst case scenarios. In other words: he anticipates on the scenarios  $j$  in such a way that the impact for him is largest, so  $\langle T \rangle_U$  is the largest. Because we consider just one OD-pair and, this equals maximising the cost of travellers of that class for the whole the network. The cost equals the travel time. Therefore,  $\mathbf{f}_{\text{risk-averse}}$  (the anticipation on the scenarios for the risk-averse users) is determined by *maximising* the travel time for the risk-averse travellers.

Mathematically, this is expressed in equations 9.4 to 9.7. The minimization and maximisation means we have to perform the following simultaneous optimisation, based on equation 9.3:

$$\max_{\mathbf{f}} \left( \min_{\mathbf{h}} (\langle T \rangle_U) \right) = \max_{\mathbf{f}} \left( \min_{\mathbf{h}} \left( \sum_j f_{j,U} \tau_{j,U} \right) \right) \quad (9.4)$$

Nagae and Akamatsu (2007) pose that a worst case is perhaps too extreme and add an entropy term to give a higher probability to the non-extreme cases. Following their principle, the maximisation process now changes from equation 9.4 to 9.7:

$$\mathbf{f}_{\text{risk-averse}} = \operatorname{argmax}_{\mathbf{f}} \left( \min_{\mathbf{h}} \left( \sum_j f_{j,U} \tau_{j,U} - 1/\vartheta \sum_j f_j \ln(f_j) \right) \right) \quad (9.5)$$

where  $\vartheta$  indicates the influence of the entropy term. The extra term  $-1/\vartheta \sum_j f_{U,j} \ln(f_{U,j})$  favours  $\mathbf{f}_U$ s which have a more spread distribution of the risks over the scenarios. A low  $\vartheta$  means that more weight is given to the entropy term. In these cases, it is less likely that there is one scenario that is assigned a very high probability. Nagae and Akamatsu (2007) also prove that equation 9.5 has a closed solution for  $f_j$ :

$$f_{j,\text{risk-averse}} = \exp(\theta \tau_{j,\text{risk-averse}}) / \sum_j \exp(\theta \tau_{j,\text{risk-averse}}) \quad (9.6)$$

Note the way the parameter  $\vartheta$  works out in the solution: it indicates how smoothly the incident chance is distributed over the scenarios.  $\vartheta = 0$  gives an equal probability to each scenario and  $\vartheta = \infty$  gives only weight to the scenario with the highest disruption.

This formula is generalised, as presented by the authors earlier (Knoop et al., 2008a). In that paper, the risk-attitude was modelled in a static framework, whereas here the combination of the risk-averse behaviour and the dynamic traffic situation is important. Since the risk-averse behaviour is important in this combination, the different user classes are explained in this chapter as well.

It is argued that a not risk-averse user does not always consider an incident to happen. In contrary, optimistic users will assign a probability 0 to all scenarios in which a blocking occurs. We therefore introduce a two-parameter model.

$$\begin{aligned} f_{U,j} &= p \exp(\theta \tau_{U,j}) / \sum_j \exp(\theta \tau_{U,j}) & \forall j \in \mathcal{B} \\ f_{U,j} &= 1 - p_U & \forall j \notin \mathcal{B} \end{aligned} \quad (9.7)$$

This is a generalised form of equation 9.6. The two parameters  $p$  and  $\vartheta$  are user class specific. Parameter  $\vartheta$  still indicates how extreme the users' anticipation on the scenarios is.

A user group with  $\vartheta = 0$  considers each scenario with a blocking similarly likely. The newly introduced parameter  $p$  indicates the total weight that is given to any incident scenario. If  $p = 0$ , users assumed there will be no disruption in the network. Setting  $p = 1$  gives back the formulation of Nagae and Akamatsu (2007). Setting  $p = 1$  and  $\vartheta = \infty$  gives back the formulation of Bell (2000). By adapting the parameters, different user classes can be introduced. These different user classes share the same network and thus interact with each other.

The risk-averse users expect the worst case to happen: the scenarios where the disruption is the biggest are taken more likely. If one would formulate the model for multiple OD-pairs, the formulation has to be destination specific. Instead of the total cost, the risk-averse route choice would then depend on the cost the travellers on a specific OD pair encounter. It sums the expected link travel times of all links that are in route  $z$ .

Equation 9.7 gives the solution for the anticipated link failure  $f$ . Now the part on the route choice  $h$  is discussed. The expected travel time on link  $i$  in period  $t$  depends on the anticipated likelihood of the scenarios,  $\mathbf{f}_U$ .

$$\langle \mathbf{tt}_{it} \rangle_U = \sum_j f_{j,U} \mathbf{tt}_{ijt}^{\text{link}} \quad (9.8)$$

A user will choose path  $z$  if it minimizes his expected travel time. The Wardrop equilibrium, Wardrop (1952), is described by

$$\begin{aligned} \sum_i a_{iz} \langle \mathbf{tt}_{it} \rangle_U &= \mathbf{tt}_{\min} \forall z : h_{U,t,z} > 0 \\ \sum_i a_{iz} \langle \mathbf{tt}_{it} \rangle_U &\geq \mathbf{tt}_{\min} \forall z : h_{U,t,z} = 0 \end{aligned} \quad (9.9)$$

Note that we choose for a user optimum for the network loading and not a system optimum. In general a user optimum traffic assignment is different from a distribution of

the traffic over the network such that the total travel time is minimized, a system optimum. Earlier studies on all risk-averse formulations, including Bell (2000) and Nagae and Akamatsu (2007), describe the search for a system optimum, but the algorithm used minimizes individual travel time and in this way it leads to a user optimum. A user optimal assignment is a better way to describe the traffic operations, since travellers do not have an incentive to reach the system optimum.

Also in incident situations, a Wardrop equilibrium on costs is often used in traffic models, as for instance by Marchal and de Palma (2008). However, one might even argue that travellers take a real-option approach as proposed by Frejinger and Bierlaire (2007). In this approach travellers will account for the events that might occur, but also they will account for the possibilities that they have to change their intended path en-route to avoid these problems. Section 9.3.2 points out which near-equilibrium route choice model is chosen in our implementation.

### Interpretation of the variables

The used variables are listed in table 9.1. This section briefly explains how the key variables can be interpreted.

$\mathbf{h}_{U,t}$  is a vector; the number of elements equals the total number of routes. Each of the elements of  $\mathbf{h}_U$  shows which fraction of the travellers of user class  $U$  takes route  $z$ . An element of  $\mathbf{h}_{U,t}$  is therefore indicated by  $h_{U,t,z}$ . The sum of  $\mathbf{h}_U$  is 1.

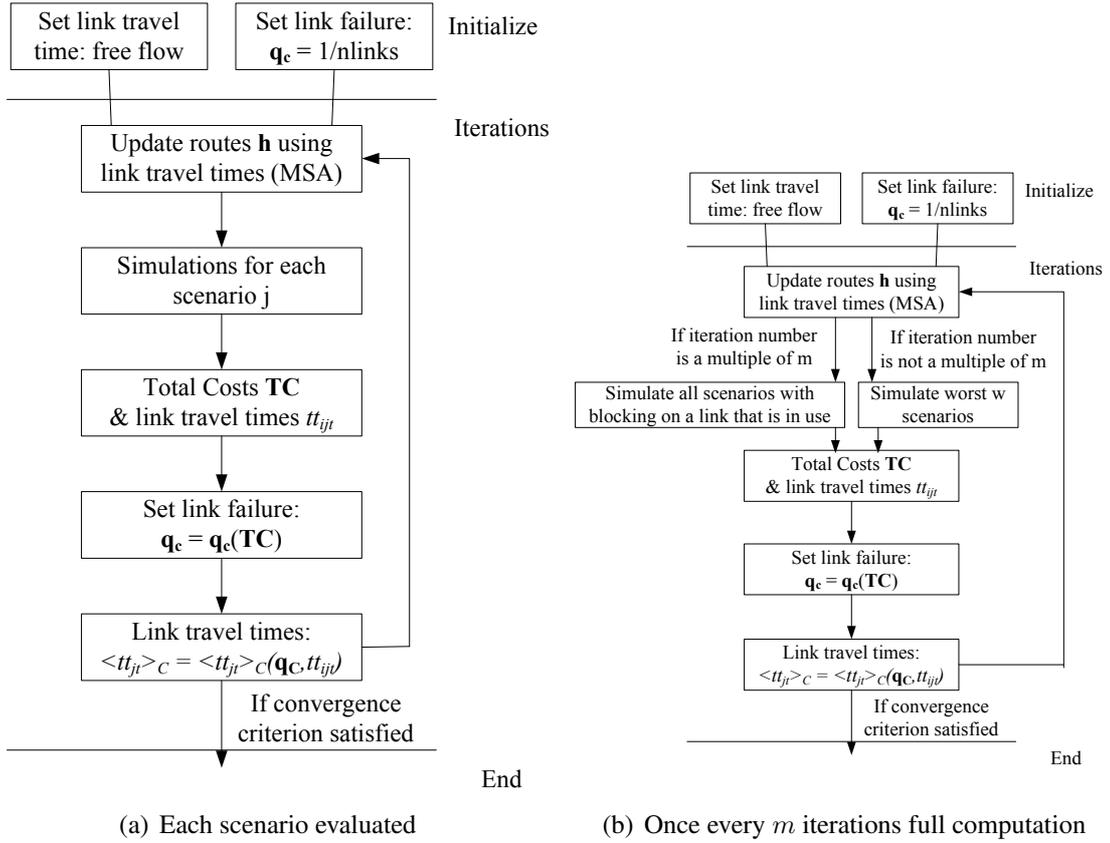
$p_U$  is the total number of incidents that a user class expects to happen. A very pessimistic user would not only expect that there are disturbances on the worst place, but he might even expect several disturbances. The extent to which he counts on disturbances is indicated with  $p_U$ . This is not necessarily an integer number: one can, for instance, think that there is 50% probability on an accident.

$\mathbf{f}_U$  is a vector with  $j$  elements, one for each scenario. It is the anticipation on a (disturbance) scenario according to user class  $U$ . Therefore, the elements  $f_{j,U}$  are between 0 and 1 and the sum of  $\mathbf{f}_U$  equals  $p_U$ .

$\vartheta$  is the risk-averseness of the user class. User classes with a large  $\vartheta$  are risk-averse to a large extent. If the expected travel time rises in a scenario, the scenario gets a larger weight, according to equation 9.7. Each increase in travel time of  $\vartheta$  adds an extra weight factor to that scenario of a factor exponential  $e$ . To find the anticipated likelihood of the scenario, this is normalised in the end.

### 9.3.2 Solution algorithm

To solve the equations, we use an iterative method based on the method described by Bell et al. (2008). However, the fundamental difference is that, the scenarios influence the traffic supply not only in the time period in which there is a blockage or the link on which there is a blocking; its influence extends to other links and later time periods. The consequence is that for each scenario a full simulation has to be carried out to evaluate the total cost for the scenario.



**Figure 9.2: Schemes of different solution algorithms**

To calculate the route choice  $\mathbf{h}_{U,t}$ , the anticipated link failure probability vector  $\mathbf{f}_U$  is needed (equations 9.8 and 9.9) and to calculate  $\mathbf{f}_U$ , one needs the route choice  $\mathbf{h}_{U,t}$  (equation 9.7 and 9.3). This can be calculated iteratively. We implemented this process as follows. First the traffic is assigned in a free flow network. We use a time-sliced static traffic assignment with travel times determined from the dynamic traffic simulator; see for details section 9.4.3. Then, all scenarios are calculated using this fixed route choice. For each scenario, the total cost is calculated and the link travel times (for each period) are stored. Following equation 9.7, the link travel times are now used to update the anticipation on each scenario to happen. The anticipation on the scenarios and the link travel times under scenarios lead to an expected travel time on links. This is now used to construct the routes again. A graphical representation of the solver algorithm can be found in figure 9.2a. One iteration of the game is the combined move of two players: the computation of the link failure probability vector  $\mathbf{f}_U$  and the reassignment of traffic, vector  $\mathbf{h}_{U,t}$ .

We use a Method of Successive Averages, introduced by Sheffi (1985). This means that the fraction of the traffic that is reassigned to a route in iteration  $n$  is  $1/n$ ; the route choice of the remaining traffic is unchanged. Consequently,  $\mathbf{h}_{U,t}$  converges.  $\mathbf{f}_U$  depends only on  $\mathbf{h}_{U,t}$ , so a converging  $\mathbf{h}_{U,t}$  means  $\mathbf{f}_U$  will converge too. For one user class, it can be proven that the optimal  $\mathbf{f}_{\text{risk-averse}}$ , equation 9.9, is unique and is found using this method, see Bell et al. (2008). In our formulation, the route choice of one user class interacts with the route

choice of another class. This interaction takes place via the assignment of traffic and  $\mathbf{f}$ . For multiple user classes, it cannot be proven that this solution method yields a unique solution for all cases. However, in an application in a practical case in section 9.4 shows that the algorithm, in practice, will converge to a solution.

### 9.3.3 Reducing computation time

In the scheme presented above, the traffic simulation needs to run many times: the number of scenarios times the number of iterations needed for convergence. When trying to find the vulnerable links, the number of scenarios equals the number of links. However, the number of links that possibly are most vulnerable is usually limited to just a few.

The algorithm is changed to reduce the number of runs. Initially, all scenarios are simulated. Then, the  $w$  most harmful scenarios are selected. Now, several iterations are run in which only these  $w$  scenarios are assessed. Of the anticipated scenarios, only the  $f_{U,j}$ 's change for which  $j$  is part of  $w$ . All anticipated probabilities  $f_{U,j}$  for scenarios  $j$  that do not belong to  $w$  remain constant. In this way, we do not calculate the effect of all scenarios in each iteration step. This saves the time of calculating the network performance for many scenarios in which the traffic flow is hardly changed compared to the situation without a blockage. Only once in  $m$  runs, the cost of all scenarios is calculated. Additional saving is made by skipping scenarios which only entail a blocking on an unused link. The whole algorithm is represented in figure 9.2b.

Two issues are relevant when applying this modified method: (1) the convergence criterion to be used and (2) whether every solution still can be reached even when not all scenarios are evaluated.

First the convergence criterion which stops the iteration is discussed. The method we use is presented here, and will be illustrated with results from the case study in section 9.5. The most varying vector is the vector  $\mathbf{f}$ . We define the change of the vector as a distance in Euclidian space:

$$D(\mathbf{f}_1, \mathbf{f}_2) = \sqrt{\sum_j (f_{1j} - f_{2j})^2} \quad (9.10)$$

We consider the algorithm to be converged if the distance between two vectors  $\mathbf{f}$  computed in subsequent iterations does not differ more than a threshold:

$$D(\mathbf{f}_{\text{iteration } n}, \mathbf{f}_{\text{iteration } n+1}) < \text{th}_1 \quad (9.11)$$

Only once every  $m$  iterations, all network scenarios are computed. Convergence is likely to be reached when the difference between the two vectors  $\mathbf{f}$  computed after such a computation of the network cost for all scenarios do not differ more than a threshold value:

$$D(\mathbf{f}_{\text{textiteration } mn}, \mathbf{f}_{\text{textiteration } m(n+1)}) < \text{th}_m \quad (9.12)$$

This threshold  $th_m$  is not necessarily the same as threshold  $th_1$ ; ultimately, the convergence as required in equation 9.11 should be checked. The risk perception  $f_j$  is likely to be oscillating between two values with decreasing amplitude. At each iteration the value will jump from the higher to the lower value or vice versa. It should be avoided that the distance between two of the highest points of  $f_j$  is used for the convergence criterion. It therefore is advisable to choose an odd number for  $m$ .

Since equation 9.12 only gives a indication of fulfilling the convergence criterion based on two non-subsequent iterations, a full computation of all scenarios is needed to check the distance between two subsequent iterations,  $m(n+1)$  and  $n(m+1)+1$  to check equation 9.11. So the check that needs to be carried out now is as follows:

$$D(\mathbf{f}_{\text{iteration}m(n+1)}, \mathbf{f}_{\text{iteration}m(n+1)+1}) < th_1 \quad (9.13)$$

One could choose the threshold  $th_m$  different from threshold  $th_1$ . In the end, the only requirement is that the distance of the vector between two subsequent iterations is not too large, so  $th_1$  indicates the desired convergence scale. In  $m$  iterations, the vector differs generally more than in one iteration. When one wants the best results in the smallest number of iterations, it is probably better to choose the threshold  $th_m$  higher. In that case, one can check the convergence, equation 9.13, after fewer iterations. However, after equation 9.12 is satisfied, one still needs to check equation 9.13. There is a risk that one evaluates all scenarios (required for equation 9.13) in vain. The disadvantage of choosing  $th_m$  higher than  $th_1$  is that a few times a simulation of all scenarios is done for two subsequent iterations without having convergence of the two subsequent terms because  $th_m$  was too large compared to  $th_1$ . In terms of computation power, it is therefore questionable whether different thresholds increase computation speed, which is why we opted for equal threshold values.

The next question to be discussed is the reachability of all solutions. In the newly proposed scheme, the scenarios that are initially unlikely to cause many problems are also included on a regular basis, namely once every  $m$  times. The reason is that some effects (e.g., spillback to upstream links) can cause larger problems than initially predicted. Some routes are just assigned every  $m^{\text{th}}$  iteration. The weight of each of these assignments is  $1/mn$ , in which  $n$  is an integer number used to indicate the iteration number. It now needs to be proven that the total weight of these assignments equals at least one. So it needs to be proven that:

$$\sum_{n=mn_0}^{\infty} \frac{1}{mn} > 1 \quad (9.14)$$

This is shown by equations 9.15 to 9.17.

The sum of series  $1/n$  is infinity:

$$\sum_{n=0}^{\infty} \frac{1}{n} = \infty \quad (9.15)$$

That means that from any starting point in the series,  $n_0$ , the limit of the sum from  $n_0$  to

infinity of the series is infinity:

$$\sum_{n=n_0}^{\infty} \frac{1}{n} = \infty \quad (9.16)$$

And so, the weight for a route  $z$  even if just assigned every  $m$  steps can reach 100%:

$$\sum_{n=mn_0}^{\infty} \frac{1}{mn} = \infty > 1 \quad (9.17)$$

This completes the proof and shows that even if a scenario is not considered in each iteration step, it is still possible that the total weight of anticipation is assigned to that scenario.

### 9.3.4 Interpretation for vulnerability

The result of the methodology presented above, is a distribution of scenario probabilities for each user group. The most interesting is a risk-averse user group, a user group defined by parameter  $p = 1$ . The vector  $\mathbf{f}_{\text{risk-averse}}$  indicates which would be the worst places to have a blocking.

Chapter 7 computes the impact of a blockade on each link without considering the consequences of the impact on route choice. This route choice needs to be modified to incorporate the findings of Bogers and Van Zuylen (2004). They point out that people take the risk into account when choosing their route. We introduce different user classes which are all to some extent, risk-averse. The extent to which the user classes are risk-averse is expressed in two parameters,  $p$  and  $\vartheta$  – see equation 9.7. For illustration of the methodology and the influence of the composition, we use several compositions in this study. However, for a practical application as for instance a journey planner, the parameters  $\vartheta$  and  $p$  need to be calibrated for a specific network and a specific user group.

## 9.4 Case study

This section presents an application of the method presented in section 9.3. The case study shows the trade-off between travel time and reliability in a grid network. The first section introduces the scenarios. Section 9.4.2 presents the network and discusses its characteristics. Then section 9.4.3 describes the traffic simulation model used here and how exactly the method presented in this chapter is modelled in this simulator.

### 9.4.1 Scenarios

The scenarios in our case study are related to accidents. For every link in the network, a scenario was created with an accident on the link. The location of the accident is always

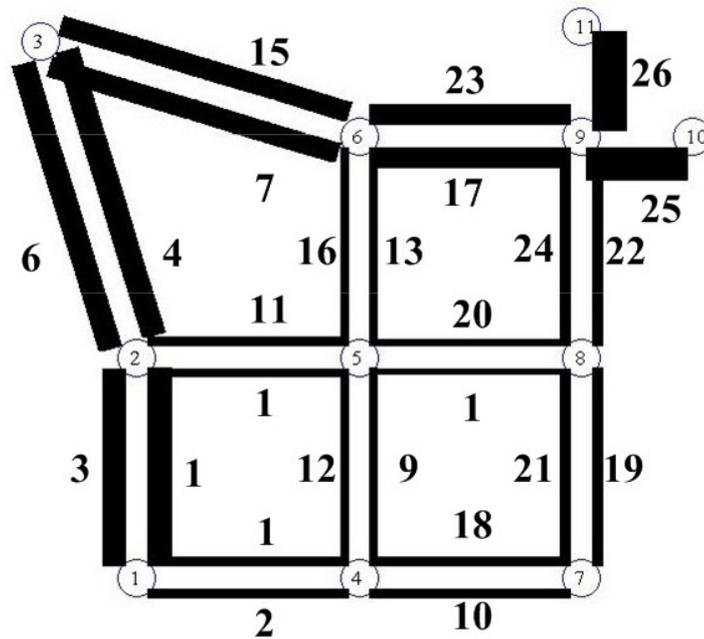


Figure 9.3: The network used in the case study

at the middle of the link. The time when the accident happens is always the same. Within the same framework, it would also be possible to define multiple scenarios for each time an accident could happen. For reasons of computation time, we choose just to change the location of the accident. Otherwise, having  $i$  links and  $t$  time steps, it would result in  $i$  times  $t$  scenarios.

The duration of the simulation is three hours. The first hour traffic flows normally. After one hour, an accident happens somewhere in the network. The capacity is reduced to 10% of the original capacity for one hour. Then, the capacity increases again to 70% of the original capacity. One can think that there is an accident and after the wrecks are removed, but there are still irregularities (e.g., workers at the road side, skid marks) which influence the capacity.

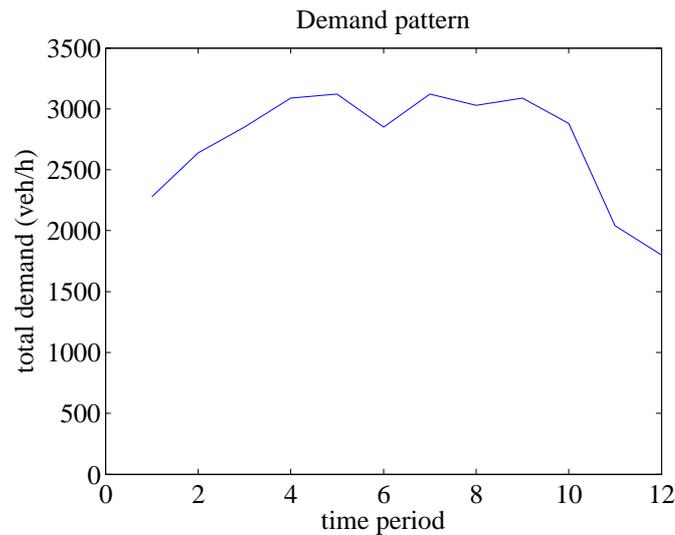
#### 9.4.2 Network and OD-matrix

The network we use is a hypothetical test network based on a grid network. Figure 9.3 shows the network; the width of the lines is the capacity of the links. The traffic flows from the origin in node 1 to the destination in node 9. Table 9.2 shows the properties of all links. The basic idea is to have a network with three types of routes. One can interpret the network as follows. The centre of the network is a city centre. There are just local roads (low capacity, low speed limit) through this centre. The roads at the bottom and right hand side are 80 km/h highways.

The motorway around the city (left and top) is a bit longer, but quicker and has a higher capacity. This is probably the link that is also avoided by the risk-averse people. Since

**Table 9.2: Properties of the links**

Link nr	From node	To node	Speed limit (km/h)	Capacity (veh/h)	Nr. of lanes	Queuing density (veh/km)
1	1	2	100	4200	2	250
2	1	4	80	1750	1	125
3	2	1	100	4200	2	250
4	2	3	100	4200	2	250
5	2	5	70	1750	1	125
6	3	2	100	4200	2	250
7	3	6	100	3800	2	250
8	4	1	80	1750	1	125
9	4	5	50	1250	1	125
10	4	7	80	1750	1	125
11	5	2	70	1750	1	125
12	5	4	50	1250	1	125
13	5	6	50	1250	1	125
14	5	8	50	1250	1	125
15	6	3	100	3800	2	250
16	6	5	50	1250	1	125
17	6	9	100	3800	2	250
18	7	4	80	1750	1	125
19	7	8	80	1750	1	125
20	8	5	50	1250	1	125
21	8	7	80	1750	1	125
22	8	9	80	1750	1	125
23	9	6	100	3800	2	250
24	9	8	80	1750	1	125
25	9	10	100	6000	1	150
26	9	11	100	6000	1	150



**Figure 9.4: The traffic demand**

there are many people taking this high-capacity road, the delay caused by a blocking will be high. The motorway links 1 and 4 are chosen to have a slightly higher capacity than links 7 and 17. In this way an oversaturated flow will lead to congestion within the network. Alternatively, these cars would queue outside the network and the motorway itself would be uncongested. This would lead to low travel times and an advantage for the motorway. An alternative is a rural road (bottom and right). This network layout will show the trade-off that the different user classes make between the fast routes reliable routes.

The demand of the network is shown in figure 9.4. The total value for this network is chosen to match about the capacity of the motorway, the profile is adopted from a morning peak in the Netherlands. The network is loaded without congestion occurring at the beginning.

Three user groups are introduced: a risk-averse user group,  $\vartheta = 50(h)^{-1}$  and  $p = 1$ , a group with neutral risk behaviour,  $\vartheta = 0$  and  $p = 1$ , and an optimistic group which does not consider any risk at all,  $p = 0$ .

The total demand is spread over the three different user classes. The different distributions that we used are shown in table 9.3. Classifying all travellers as risk-averse would be the same as in Bell (2000). This is a hypothetical situation to show the possibilities of the model. To what extent travellers take risks into account in practice is not analysed in this chapter.

Also route choices made for user classes that are not present in the traffic composition are relevant. It could be useful for the following case. Suppose one had to plan a special transport through a network with non risk-averse users. That transport, being one vehicle, has a negligible influence on traffic jams. In this way, the framework can also be used to plan an important transport which has to be in time.

**Table 9.3: The distribution of the travellers over the users classes for each scenario**

Number	Risk-averse	Neutral risk	Optimistic
1	0.3333	0.3333	0.3333
2	1	0	0
3	0	1	0
4	0	0	1
5	0	0.5	0.5
6	0.5	0	0.5
7	0.5	0.5	0
8	0.5	0.25	0.25
9	0.25	0.5	0.25
10	0.25	0.25	0.5

To analyse the sensitivity of the results to the demand, two more simulations with a different demand are carried out. We increase the demand of traffic to 150% or 200% of the original traffic demand.

### 9.4.3 Simulation model

In this paragraph, we focus first on the traffic flow operations. In the second part the implementation of the route choice model is explained.

The mathematical framework we developed can fit any simulation model. Chapter 7 states that in situations with delays due to incidents, it is important to have a simulation program which captures spillback well. Therefore, a static simulation model with a travel time function is considered insufficient. A microscopic simulator models the behaviour of all individual cars. This can represent the traffic flows well when all parameters are properly calibrated and validated. Since all vehicles have to be simulated individually, it is time-consuming to run a microscopic simulator.

A macroscopic simulator does not simulate individual cars, but traffic flows, average speeds on links and densities on links. This reduces computation time. Furthermore, it is always required to run a microscopic model several times due to the stochastic effects. This is not needed for macroscopic model without stochasticity. There are several properties of the simulation program important for the validity of this study: (1) queues occur upstream of the bottleneck, (2) queues spill back to upstream links, (3) queues can solve from the head and from the tail. A combination of (2) and (3) also makes that shockwaves can travel through the network.

A first-order traffic simulator fulfils all these requirements. According to Daganzo (1995), the best way to represent traffic models is a first-order model. This is an ongoing debate (Aw et al., 1999; Helbing, 2009). If one prefers another model, this is possible in the framework of the route assignment methodology presented here. In this chapter, we choose

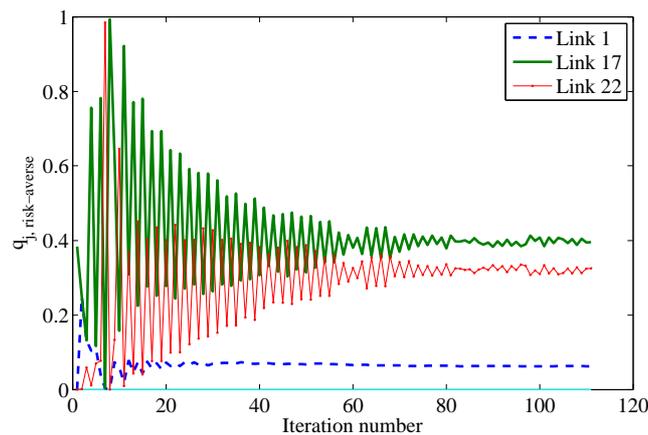
for a (first-order) LWR model, presented by Lighthill and Whitham (1955) and Richards (1956), which is solved by a Godunov scheme (Godunov, 1959). The use of this solution scheme for traffic flow is described by Daganzo (1994) and Daganzo (1995) and by Lebacque (1996). This model can handle different types of flow-density diagrams. The most simple is a triangular flow-density relationship is used, see for instance Daganzo (1997). With this relationship, the three requirements with respect to queuing can be fulfilled. This is all incorporated in the model DSMART, developed at the Delft University of Technology (Zuurbier et al., 2006).

The essence of the model proposed here, is that blockages on one link can influence the traffic flow operations outside the link which is blocked and outside the time frame in which there is a physical blocking. It is essential that the traffic simulator used here can cope with these dynamics. There are additions to the simple traffic flow model used here which make the traffic flow model more realistic. For instance, a capacity drop can be added: the outflow capacity at the front of a queue is lower than the free flow capacity, see for instance Hall and Agyemang-Duah (1991). When needed in real-life cases, however, it can be fitted within this framework. In fact, every simulation model can be fitted in the framework presented in the previous section.

There are several possibilities to model the route choice. When facing unexpected congestion, users are likely to change their route during a trip as shown in chapter 5. Therefore, we had to choose for a model with time-dependent route choice (which incorporates time-dependent delays). Each time travellers arrive at a node, they will reconsider their route according to the last traffic conditions that are available at that moment. Assigning new routes is in fact equivalent to redistributing traffic via a (destination specific) split fraction: also when a new route is assigned to the vehicles, it will change at the next node. We choose to store the routes as split fractions.

All travellers have a different error on the perception of travel time. This needs to be modelled to prevent all people taking the same path; this will make the final route choice stochastic. The route choice is made with a probit assignment based on the expected link travel times in the last completed period. Note that this expectation value is based on different scenarios (equation 9.8). These times are disturbed by a normally distributed error of 10% of the travel time. From each node the fastest path to the destination is determined. This is repeated a large number of times. For each link that leads away from the node, it is counted how many routes leave the node via that link. The fractions of the flow to all links together form the split fraction for that specific node for that specific period; this is stored. It is possible that they will drive back along the same route for a short distance and then choose a path on which they avoid the congestion.

Within one time period, this is a traffic assignment without feedback of the traffic intensities or travel times. The traffic conditions, including speed, are based on these intensities as well as the different scenarios. These speeds in different scenarios give the next expected travel time, as posed in equation 9.8. In the next iteration of the game-theoretical framework, these intensities are fed back and routes will be adjusted based on these new speeds. The same en-route traffic assignment is for instance used in Dynasmart (of Maryland, 2003) and Integration (Van Aerde, 2005).



**Figure 9.5:** The convergence of assumed link blocking probabilities in composition 7; the other links have a probability of approximately 0.

## 9.5 Results

In the network of the case study, the fastest route is the route passing over the motorway (links  $1 \Rightarrow 4 \Rightarrow 7 \Rightarrow 17$ ). Almost all traffic in the optimistic user class will travel over the motorway network, as expected. In the specific case of our network and demand, the risk-neutral user group has a route choice similar to the optimistic user group. The route choice of the risk-averse user class is therefore most interesting.

Figure 9.5 shows how the probabilities for the scenarios for the risk-averse user classes converge. Note that there are just a few links that are vulnerable. This supports the proposed idea of not calculating the vulnerability for each link in every iteration. It also shows the oscillation of the probabilities for one scenario between two values.

All travellers in all user classes start off at a motorway route, even the risk-averse users. This is because at the first period there is no risk of a blockade yet. Figure 9.6 shows the percentage of risk-averse users taking the motorway over time. This decreases for the time periods in which they expect a link will be blocked. We divided the traffic compositions in groups with the same number of risk-averse people. Note that within these groups the patterns of motorway usage are similar. Figure 9.7a shows the route set in time period 5 for a complete risk-averse traffic composition: the width of the links shows the traffic volume (of risk-averse travellers) over that link.

Figure 9.6 shows that as the percentage of risk-averse users increases, the part of this user class that takes the motorway decreases. Table 9.4 presents the share of the *total* demand travelling over link 4. If all users are risk-averse, 46% travels by motorway. The total number of travellers using the unreliable path increases as the composition of the user classes gets less risk-averse. It also depends on the composition of the other drivers. Optimistic drivers will use the motorway more than the risk-neutral drivers. Therefore, the risk-averse users will use the motorway less in the scenario where other users are optimistic compared to the scenario where other users are risk-averse.

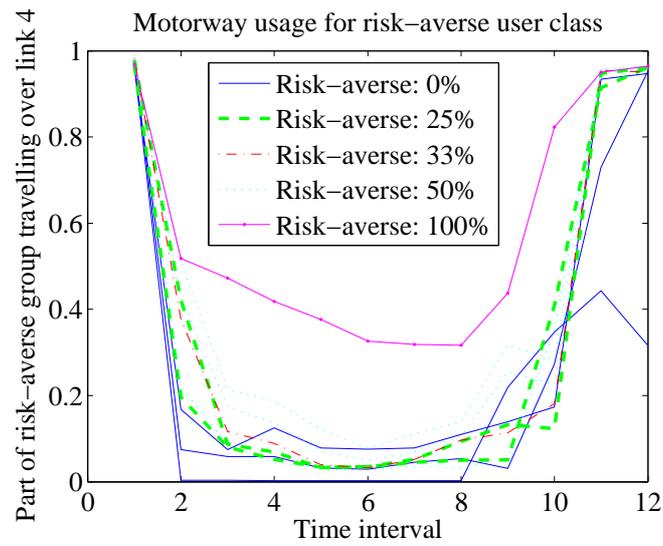
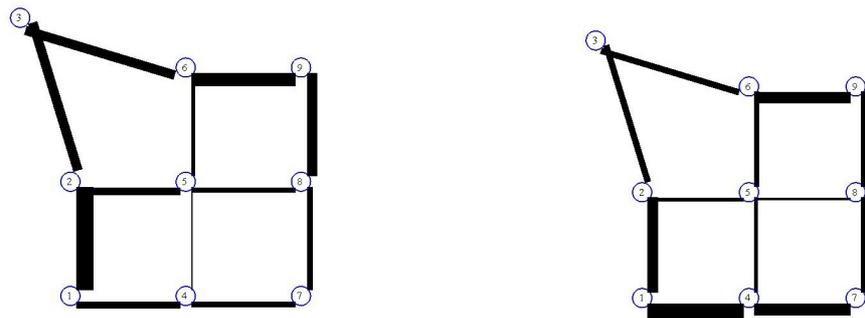
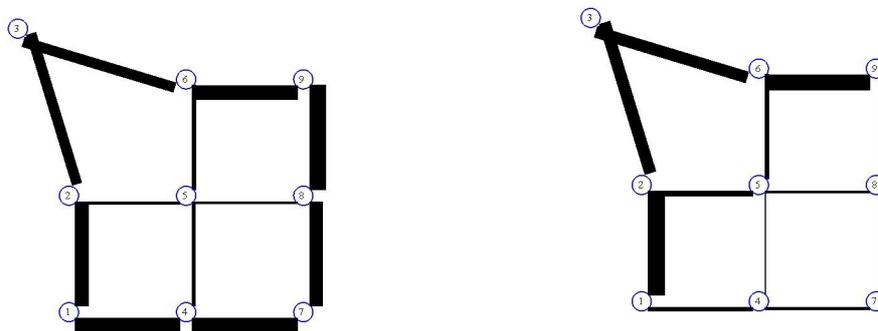


Figure 9.6: The fraction of risk-averse travellers taking the motorway (link 4) over time in the different scenarios of table 9.3



(a) Risk-averse travellers in composition 2

(b) Risk-averse travellers in composition 3 with 50% increase in demand



(c) Risk-averse travellers in composition 2 with 100% increase in demand

(d) Risk-averse travellers in composition 3 with 100% increase in demand

Figure 9.7: Relative traffic volumes over each of the links in period 5

**Table 9.4: The fraction of all travellers traversing link 4**

Number	Use of link 4
1	62%
2	49%
3	54%
4	98%
5	59%
6	51%
7	51%
8	51%
9	62%
10	63%

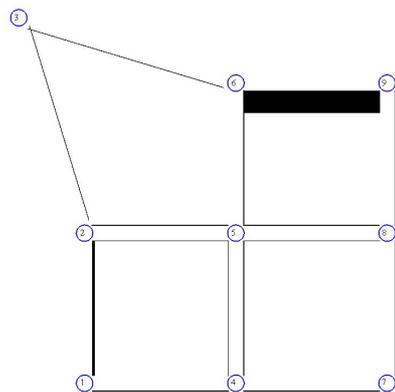
**Figure 9.8: The anticipated blocking probability  $f$  for the risk-averse user class in composition 2 – the width of the line indicates the anticipated blocking probability**

Figure 9.8 shows the probability in the risk-averse view for a link to be blocked by the width of the links. The links that are closer to the destination have a higher probability to be blocked. This is because the moment of blocking is taken fixed. This fixed time has as side-effect that a blockade of a link closer to the destination cause a larger total delay than the links close to the origin. Imagine one link which can be blocked at any time. The earlier it is blocked, the worse the consequences are (since the capacity is not fully restored after the accident). If now the location of the blockade is variable between two places and the time is fixed, it would be worse if the location of the blocking was further downstream. All cars between the two possible link locations are hindered if the downstream location is chosen and they are unhindered if the upstream location is chosen.

The route choice of the risk-averse drivers deviates even more in case of an increased demand of 50% (see figure 9.7b) or 100% (figure 9.7c). With a higher demand, the traffic jam caused by a blockade will cause more delays. Therefore, the risk-averse traffic is more likely to deviate from the initially fast route, the motorway. With an increased demand, we found that also the risk-neutral user class starts to take alternative routes. This is shown in figure 9.7d.

## 9.6 Conclusions and possible application

This chapter introduces a risk-averse traffic assignment model. It combines risk-averse traffic assignment with dynamical traffic simulation models. In the presented framework, the effect of a scenario is not necessarily restricted to one link. Also, a disruption on one link can disturb the flows in other time periods. These characteristics come at a cost: a dynamic simulation takes more simulation time than a static simulation. The chapter proposes a method to overcome this problem. Rather than calculating all scenarios in each iteration of the game-theoretical framework, in each iteration only a fraction of all the scenarios are evaluated.

The users' route choice is generalised in a two-parameter equation, accounting for risk-averseness. In the simulation program, different user classes, characterized by different parameters, can use the network simultaneously. The different user classes interact by their route choice.

The method was applied to a case study, in which there was a network with a busy motorway which is fast in free flow and alternative routes which were slower. Blocking a large motorway would lead to large delays. The case study showed that the more risk-averse drivers are, the more they are likely to avoid the busy motorway and opt for the alternative routes. They exchange the free flow speed of the motorway for the reliability of the slower routes.

This chapter showed the basis for combining a game-theoretical framework with a non-static traffic flow model. The framework provides opportunities to use all different types of traffic flow simulation (macroscopic, mesoscopic and microscopic). It is required that the traffic simulation program stores the travel times at the links for all simulated times.

Furthermore, the simulation program needs to adapt the route choice provided by this framework and it is necessary that it has an option to be called iteratively from another program.

The goal of the chapter was to show that the risk-averse traffic assignment was possible. In this chapter, three groups of travellers are proposed. Future work includes calibrating the risk-averseness of travellers in practice, for instance by questionnaires or a detailed analysis of traffic data. This should also give the distribution of travellers over these user classes. Within the framework proposed here, one could also assign different driving behaviour to each of these groups.

The simulation shows the anticipated blocking probabilities for each link (different for each class). Risk-averse drivers anticipate most the scenarios with the largest risk, in order to avoid the possible consequences. However, it is questionable whether individual travellers are able to assess the risks of a certain route. They will therefore anticipate most scenarios that will harm the network performance most. An automated program, for which the foundations are laid down in this chapter, could take over this task and calculate the possible worst outcomes. The result of this calculation is the anticipated probability of a scenario. A high anticipated probability for risk-averse people also means that that link harms the network performance considerably. The information which link can degrade the network performance most is useful for road authorities. Based on that knowledge, they can improve their plans for incident management.

The method could also be applied in a route planner. The route planner has to be loaded with a network and the (calibrated) demand and route preference of other travellers. Users of the route planner can select their own route preferences. The route planner will then present a personalised route advice. Like any dynamic route planner, it would need the total demand as input. This risk-averse journey planner needs (an estimate) of the risk-attitude of all users as well.

Future research further includes a dynamic traffic assignment model which is not time-sliced but also incorporates future conditions. Only in that way, the routes could be transferred to a user equilibrium where users will take recurrent congestion in future time periods into account. In such a model, also the time of the blocking can be made flexible.



# Chapter 10

## Synthesis and conclusions

### 10.1 Aim and set-up of the chapter

This final chapter concludes the thesis. To this end, it will first present the findings of the research, based on the research questions introduced in chapter 1 and discussed throughout this thesis in the different chapters. In section 10.3 these findings are generalized, where possible, and the answers to different sub-questions are combined to get a broader conclusions. Section 10.4 shows the implications of the findings for future scientific research and practical work: what lessons have been learnt in this thesis? Finally, section 10.5 shows the research questions and new research directions that have come up in this thesis and which, based on the work presented in this thesis, are interesting to carry out.

### 10.2 Findings

This section shows the findings which have been presented in this thesis. To structure the findings, we use the division we made in chapter 1 in table 1.1 which is repeated here in table 10.1. The findings here are presented according to that scheme, which means that they are not ordered according to the chapters in which they have been answered. This section thereby answers the research questions formulated in section 1.3. How these answers can be connected to answer the main research question is described in section 10.3.

#### 10.2.1 Traffic properties at incident location

##### Microscopic effects

First the effects at the location of the incident are discussed, starting with the microscopic effects, I in table 10.1. This part of the research, related to the change in driving behaviour at the incident location (research question 1 is described in chapter 3). There we found

**Table 10.1: The different levels of analysis and spatial scopes**

Level	microscopic	macroscopic
Spatial scale		
Incident location	I	II
Network	III	IV

that the reaction time of people decreases at an incident site and that people reduce their speeds when passing an incident site. Consequently, the time headways at the head of the queue increase. For the two incident sites we studied at a microscopic level, the headways increased with a factor 2 compared to normal conditions. It is remarkable that the incidents affect driving behaviour in both directions equally. For traffic in both directions of the motorway, a similar increase in reaction time is observed, as well as a similar increase in time headway

### Macroscopic effects

At the macroscopic level, the capacities at the incident location were studied to answer research question 2: what is the capacity of incident sites at motorways. In chapter 4 it is described that the capacity substantially decreases. We determined the outflow rate out of a queue for different locations. The queue discharge rate (in veh/h/lane) is around 30% lower at the head of the queue at an incident site compared to normal conditions. The reduction of the maximum flow (queue discharge rate) at the incident site is even more if it is compared to the free flow capacity. For traffic in the opposite direction, at the other side of the crash barrier, similar reductions were found.

## 10.2.2 Network effects

### Microscopic effects

On the microscopic level (III in table 10.1) the route choice behaviour of travellers is studied (chapter 5). For the studied incident cases, a difference was found between incidents causing a small delay and incidents causing a large delay. During an incident with small delays, the amount of travellers that change their route was very low. However, for the larger incidents causing longer queues, a considerable fraction of travellers (30% to 50%) decided to take a detour. This answers research question 3. However, a unique cause for this change in route choice could not be identified.

Question 7 concerns the routes taken by risk-averse travellers. Chapter 9 describes how the risk-averse attitude of travellers influences their route choice. For the first time, a dynamic traffic simulator is combined with a risk-averse traffic assignment. It is found that risk-averse people tend to avoid busier roads, particularly if these roads are taken anyway by non risk-averse travellers. In this way, the collective, macroscopic, behaviour interacts with the individual microscopic behaviour of the travellers. Therefore this finding can be categorised between III and IV in the scheme in table 10.1.

### **Macroscopic effects**

Question 4 is about the macroscopic traffic modelling for queues propagating through junctions. This is discussed in chapter 6. Shock wave theory is used to compute the delay of an incident. This method is useful because it shows the variables of which the total delay depends most. It shows for instance that the duration of an incident is particularly important. Therefore, a few long-lasting incidents cause a large part of the delay.

Then, there are findings on the network level for macroscopic properties, which is IV in table 10.1. Research question 5 asked at which locations did incidents cause the largest delays. Chapter 7 considered this question and found that main motorways and links leading traffic away from the motorways were two locations where incidents cause large delays. The second type of link is vulnerable because queues that are generated there grow upstream to the motorways and then delay the through traffic on the motorway. The same chapter also shows that the impact of this spillback is important in assessing the vulnerability of the road network. In fact, an assessment of vulnerability made in a simulation program without spillback cannot give the vulnerability levels for a realistic situation with spillback. Question 6 concerns indicators for vulnerability. The literature suggests that one could use link-based indicators to find the vulnerability of links, instead of simulating all possible incidents. These indicators were tested in chapter 8, but it was found that none of the existing indicators gives a good measure for the relative vulnerability of all the links in the network. Also, a combination of indicators cannot rank the links in order of vulnerability.

## **10.3 Synthesis and conclusions**

The previous section summarized the findings as directly found from the analysis. This section synthesises these findings and presents the conclusions: what can be learnt from these findings. It also connects the findings from different chapters to each other.

### **10.3.1 Research methodology**

In chapter 6 shock wave theory was applied and solved analytically to compute the delay caused by an incident. The chapter shows that it is possible to use this convenient technique and that it also provides a quick and analytical solution. The solution of the equations describing the traffic flow processes is exact and is better than the numerical approximations thereof which result from simulation models like the LWR-model used in chapter 7; both are of course subject to other constraints such as the fundamental diagram and traffic demand. If the analytical model is solved, it can give a very quick answer about the traffic conditions. Furthermore, it can also show the influence of different variables on the delay.

A model was developed in which spillback could be switched on and off. Based on simulations with different blockages, it can be concluded that static models do not provide

a good alternative since they do not include spillback (chapter 7) and also link-based indicators are not good at predicting the vulnerability of a link (chapter 8). Therefore, it is better to use a dynamic model for these purposes. The analysis of finding the vulnerable links for a 468-link network takes about a week of computation time with a dynamical simulation model in case route choice is included. A computation time of one week is still feasible for an analysis, but for larger networks the computation time, even with multiple computers running, becomes too long. Future improvements of the model will enable researchers to run the model in feasible time. Alternatively, the network size could be reduced. It should be large enough to include the rerouting possibilities are, but if the network of interest is larger, it can be split up into different sub-networks.

Based on the framework shown in chapter 9 it can be concluded that a framework of risk-averse traffic assignment can be combined with dynamic traffic assignment. Moreover, it is also shown that the risk-averseness of travellers can be varied and groups of travellers with all different levels of risk-averseness can be combined in one traffic simulator.

### 10.3.2 Incident location

Chapters 3 and 4 show that the traffic behaviour at the incident site differs from normal behaviour. In some of the studied cases there would not have been a queue at all if traffic behaviour had not changed. Most remarkable in the microscopic behaviour is the reduction of speed, which might induce the traffic jam, and the increase of reaction time, which makes that the traffic jam does not dissolve. The capacity reduction resulting from this changed behaviour on these two locations is significant, around 50% compared to free flow capacity. This is in agreement with the 30% - 50% capacity reduction found for a larger set of incidents. Therefore we conclude that behavioural changes (speed reduction and reaction time increase) found at the two incident sites used in the microscopic analysis also take place at any other incident site.

### 10.3.3 Network effects

Based on the results in chapter 5 we conclude that people change their route if they have to pass an incident which causes large delays. This means that for identifying the vulnerable links, route choice have to be included.

We also conclude that, apart from route choice, duration of an incident has a considerable impact in the accumulation of delay (chapter 6. If queue does not cross a junction, the delay scales with the square of the delay. But if the tail of the queue crosses a junction, the delays grow even more. Therefore we conclude spillback is an important phenomenon in determining the delays from an incident.

The studied network in chapter 7 is not expected to influence the type of links that come out as most vulnerable. Therefore, we conclude that in general, it is expected that main motorways and, due to spillback effects, the links leading traffic away from the main motorway are the most vulnerable. Although these can be indicated beforehand by someone

with a good knowledge of the network, it cannot be automated; furthermore, this identification method might not capture all the links, which is why a full simulation is still advised.

Chapter 9 shows that the risk-averseness of travellers has a large influence on the route choice, and therefore on the link loads. The framework shown in chapter 9 did not include a calibration of the risk-averseness of the travellers in real life on a real-world network. However, it shows this phenomenon can be important to predict link loads, congestion and travel times. Furthermore, from the results it is shown that travellers will generally avoid the important motorways more if there is risk-avoidance. Therefore, it can be concluded that these main motorways are less vulnerable if traffic takes risk-averse routes.

## 10.4 Implications for policy makers and road authorities

This section continues on the conclusions presented in section 10.3 and presents the implications of the research. This section therefore presents recommendations on what should be done or should be avoided in the future. There are two areas where practitioners could use the conclusions from this thesis: in the operations and in the planning stage. Both are discussed in this section.

### 10.4.1 Operations

The cause of the delay at incidents is, apart from closed lanes, the changed driving behaviour of drivers. This is probably caused by the distraction of drivers: the drivers' attention is not focussed on the driving process anymore, but they are distracted by the incident. This means that once the road authority succeeds in taking away the distraction, the queues can be largely reduced. The Dutch road authority experiments at the moment with screens which block the view on the incident site in order to take away the distraction. This should increase the capacities.

The robustness of a road network depends on the route choice. If people adapt their routes when there is a temporal blockade, the extra delay is much less than if people do not change their routes, as shown in chapter 7. We also found that people change their route if the alternative route is quicker and it is advised (chapter 5). That means that routing information, for instance on Dynamic Route Information Panels, can change the route choice and thereby increase the robustness of the network. So, if policy makers wish to improve the network reliability, routing advice systems can, to a certain extent, provide an alternative to building new roads and lanes.

The analytical analysis in chapter 6 showed that the duration of the incident is particularly important. Particularly when spillback effects start occurring there is a risk of gridlock effects. This could be reduced by minimising the time spent dealing with the incident or giving advice on alternative routes or a combination thereof.

## 10.4.2 Planning

Snelder et al. (2008b) set several criteria for robustness which should be considered during the design. These are:

- Redundancy – spare capacity can be added. This is not only roads to supply for flow, but also buffer lanes to facilitate storage in congestion. This is excellent solution to prevent spillback queues, as discussed in chapter 7.
- Compartmentalising – different subnetworks are defined which are independent. Designers have to make sure that congestion in one subnetwork does not influence another. These subnetworks can be different geographical area, but also networks with different functions (motorways, underlying road network). This can for instance be done by adding spare capacity (to avoid spillback as discussed in chapter 7). Especially adding capacity to keep the traffic on the motorways flowing would be useful. Alternatively, road for traffic flowing in each of the directions could be regarded as separate subnetworks. For instance, there could be a road design with screens in the median strip already prepared such that when there is an incident, the rubbernecking as described in chapter 3 would occur less. Alternatively, they could design the median strip so wide with bushes that viewing the other side is not possible at any time. Another way to create compartments is, for instance, separating different flows on one motorway by a hard barrier. An example is given below.
- Resilience – the possibility of the system to get traffic operation back to normal after a disturbance. In chapters 7 and 8 it is discussed how vulnerable links could be found. These are for instance locations where emergency vehicles are best placed, in order to remove a wreck as soon as possible.
- Flexibility – the possibility to use the network differently in case the circumstances change. Chapter 5 showed that people take different routes if they know that there is an incident and so decrease total delay. So, informing people could increase the network robustness. For the road planning, this matters since during the planning phase, the authorities could decide on placing Dynamic Routing Information Panels at strategic locations in the network. Furthermore, the robustness of the network could be increased if the drivers are taught how to use the information. Advertisements could show drivers that it is beneficial for them to follow the advised routes on the Dynamic Information Panels or to switch on the radio at times when there is traffic information.
- Balance – the balance between the capacities of different alternative links in the network. If the road planner decides to create one large motorway and no alternative routes, the network depends heavily on that motorway, because in case there is an incident, there are no rerouting options. This means that this motorway is a critical link, as discussed in chapter 7. Moreover, if the *strategy* of the network planning is to create the network this way, all motorways are critical links. This is not a good planning strategy to create a robust network

A way of compartmenting is a hard barrier between two lanes. One could for instance think of an urban ring road where local traffic is physically separated from through traffic, for instance by a barrier. This way, congestion spilling back from the city does not

influence the through traffic. The downside of this construction is that once an incident happens on either side of the barrier, there is no possibility to guide traffic that has already passed the diversion point to the other, uncongested, lanes. This disadvantage can be overcome if the flows are not separated by a barrier, but by road markings. However, this is only possible if the traffic discipline is very good, and drivers stay on the correct side of the line. This alternative option also has a downside for safety if traffic on one side is heavily congested and the speeds are low and on the other side is flowing freely with high speeds.

Apart from this, there are several other advices for traffic planning. It is shown to be impossible to find the vulnerable links in a network using a static model or link-based indicators. This means that the only way of finding vulnerable links is to simulate traffic with a realistic traffic model, including spillback and en-route route choice, which is computationally expensive.

If delays are unavoidable, the road authority should give proper route guidance. However, from the perspective of total network performance, these routes do not need to be faster routes for the travellers themselves, but they should lead the drivers away to places where the queued cars do not harm the traffic flow to other directions. These places could be dedicated “buffer zones” or “buffer lanes”.

Finally, there are conclusions about dynamic traffic simulation and traffic assignment. Chapter 9 shows that the risk-averseness of travellers can considerably change the link loads. Suppose that travellers in practice are to a certain extent risk-averse, this could be an important addition to dynamic traffic simulators. Currently, traffic simulators without this routing principle are used. However, since the traffic situation is sensitive for the risk-averseness of travellers, the traffic assignment could be improved a lot if this feature were included in the programs. However, calibrating and validating this property of traffic for a specific road network still requires some effort. When this is completed, the road planners can have a much better

## 10.5 Future research

This PhD study answered some questions, but also other, related questions were raised. These are summarized here.

### 10.5.1 Methodological advancements

On the microscopic scale, at the site of the incident the behavioral changes were remarkable. However, the psychological basis for these behavioral changes cannot be observed from remote-sensing traffic data. To measure the effects, it is necessary that *drivers'* reactions can be quantified, rather than the resulting movements of the vehicles. To know the drivers reactions, they have to be monitored while driving past an incident site. This can for instance be done in a driving simulator experiment where the physical reaction of

drivers is measured when they pass a virtual accident location. It is interesting to know how drivers react to such an event.

The distraction reduces capacity and also increases the probability of a secondary incident, which are both undesired. It would be interesting to find out whether there are some measures that can prevent these negative effects happening. One could for instance think of screens to prevent people looking at the incident site, as is applied now, but also a police officer encouraging people to speed up might help in increasing the capacity. Future research should show whether these measures are effective; it might also come up with other measures.

In this thesis, the change of route choice is derived from the flow values on a day with an incident and a day without an incident. However, this method does not give certainty about the route choice since a changed split fraction could also be caused by a different OD-matrix. Modern techniques enable researchers to find the actual route choice of individual vehicles. One could think for instance of number plate recognition to track a route of a vehicle or even the tracking of GSM or GPS signals which give the route. This will be valuable information concerning the routes of drivers take.

However, the method of tracking devices could be biased. It is likely that travellers with GPS are better informed than the average driver about the traffic conditions by queuing messages on the navigation device. Furthermore, they are quite certainly better informed about the alternative routes as they are guided to their destination, even if they take a route which is unfamiliar to them. For instance, most navigation devices have a graphical representation of the road network such that the drivers can see the alternatives. This bias does not hold for the number plate tracking.

When route choice is analysed in more detail it would be interesting to see the influence of the queue messages and alternative routes that are given. In that study one could distinguish between different messages that are given: “incident”, “complete road blocking”, or even an advised alternative route. One could hypothesise that the worse the message is, the larger the fraction of people changing their route. One could also analyse whether the effect is influenced by the medium by which the messages are given, for instance on Dynamic Route Information Panels, broadcasted over the radio or sent to GPS devices. It can be expected that the more accurate and up-to-date the people expect this message to be, the higher the influence of the message is on their actual route choice.

The modern media like smart phones or GPS devices gain a higher and higher penetration rate. If these media turn out to be particularly suited to providing route advice which is followed, it is possible that the fraction of drivers changing their routes will increase over time. Whether this is true, is a topic for further study. The possible consequences of rapidly changing route choice behaviour also need to be investigated.

Instead of changing their route because of the incident, people might change their trip planning. If people are informed about possible delays, they might postpone a trip, change the order of different stops on a trip with different destinations, change a mode of transportation or even cancel the complete trip. How this process works is a topic for future

study. Also, the amount of travellers actually deciding to change their schedule has to be studied in the future.

### 10.5.2 Modelling advancements

It has been shown that for analysing vulnerability it is necessary to have a fast simulation tool which captures dynamic effects and spillback. These simulation tools are usually not fast for large networks. However, the methodology with shockwave theory, applied in chapter 6, proved very useful there. It could cope with network elements although combining them can be analytically hard work. It might be worthwhile looking into the possibilities for a simulation that tracks the boundaries of the shock waves, a method similar to the one proposed by Wong and Wong (2002), to increase the simulation speed compared to another macroscopic model, like the LWR-model used in chapter 7.

Alternatively, it would be interesting to have a method which can predict which links are vulnerable or exclude links from a list of possible most vulnerable links. The current indicators did not prove to be very useful in that respect, as shown in chapter 8, and it is, based on these results, not expected that other indicators would do much better. However, future research might come up with better alternatives. These indicators, however, should explicitly account for spillback effects and the possibility that travellers take an alternative route once the traffic conditions are bad.

Finally, the thesis showed that the concept of risk-averse traffic behaviour can, in principle, be combined with a dynamic traffic simulator. However, it is unclear how risk-averseness can best be quantified. This should be studied before it can be implemented in a traffic simulator which can be calibrated on the current road layout and a drivers' population or even be validated for unknown road layouts. After this problem has been solved, it will be a challenge to integrate these findings into a journey planner. The current state of the art is that people, at best, get an advice based on the most likely traffic conditions predicted for the duration of the trip. This new simulator, however, would also incorporate the risks on each route. Based on their individual, purpose-dependent risk-averseness, travellers then get route advice. Much research is needed, however, before this application can be used.



# Bibliography

- Al-Kaisy, A. and Hall, F. (2001). Examination of the effect of driver population at freeway reconstruction zones. *Transportation Research Record: Journal of the Transportation Research Board*, No. 1776., pages 35–42.
- Angel, A., Hickman, M., and Mirchandani, P. (2003). Methods of Analyzing Traffic Imagery Collected From Aerial Platforms. *IEEE Transactions on Intelligent Transportation Systems*, 4(2):99 – 107.
- Aw, A., Rascle, M., Buisson, W. T. C., and Lesort, O. (1999). Resurrection of “second order” models of traffic flow? *SIAM Journal of Applied Mathematics*, 60:916–938.
- Bates, J., Polak, J., Jones, P., and Cook, A. (2001). The valuation of reliability for personal travel. *Transportation Research Part E: Logistics and Transportation Review*, 37(2-3):191–229.
- Belda, E., Cambres, C., and Toms, V. R. (2002). Traffic management in case of emergency in the serti area of the mediterranean corridor, valencian community. In *Proceedings of Intelligent Transport Systems*. Chicago.
- Bell, M. G. H. (2000). A game theory approach to measure the performance reliability of transport networks. *Transportation Research Part B*, 34(6):533–545.
- Bell, M. G. H., Kanturska, U., Smcker, J.-D., and Fonzona, A. (2008). Attacker-Defender Models and Road Network Vulnerability. *Philosophical Transactions of the Royal Society A*, 366:18931906.
- Black, I. G. and Towriss, J. G. (1993). Demand effects if travel time reliability. Technical report, Centre for Logistics and Transportation, Cranfield Institute of Technology, Great Britain.
- Bliemer, M. J. C. (2005). INDY 2.0 Model Specifications. Technical report, Delft University of Technology.
- Bliemer, M. J. C. (2007). Dynamic Queuing and Spillback in an Analytical Multiclass Dynamic Network Loading Model. *Transportation Research Record: Journal of the Transportation Research Board*, No. 2029.

- Blumentritt, C. W., Ross, D. W., Glazer, J., Pinnell, C., and McCasland, W. R. (1981). Guidelines for selection of ramp control systems. Technical Report 232, Transportation Research Board.
- Bogers, E. A. I. (2009). *Traffic Information and Learning in Day-to-Day Route Choice*. TRAIL Thesis Series T2009/5, Delft University of Technology, TRAIL Research School, Delft, the Netherlands.
- Bogers, E. A. I. and Van Zuylen, H. J. (2004). The importance of reliability in route choices in freight transport for various actors on various levels. In *Proceedings of the European Transport Conference*. Strasbourg.
- Bogers, E. A. I., Viti, F., and Hoogendoorn, S. P. (2004). Joint modeling of ATIS, habit and learning impacts on route choice by laboratory simulator experiments. In *Proceedings of the 83th annual meeting of the Transportation Research Board*. Washington.
- Bogers, E. A. I., Viti, F., and Hoogendoorn, S. P. (2005). Joint Modeling of Advanced Travel Information Service, Habit, and Learning Impacts on Route Choice by Laboratory Simulator Experiments. *Transportation Research Record: Journal of the Transportation Research Board*, No. 1926, pages 189–197.
- Börjesson, M. (2007). Departure time modelling: Applicability and uncertainty. In *Proceedings of European Transport Conference*. Leiden, the Netherlands.
- Boyles, S., Fajardo, D., and Waller, S. T. (2007). A Nave Bayesian Classifier for Incident Duration Prediction. In *Proceedings of the 86th Annual Meeting of the Transportation Research Board*. Washington D.C.
- Brackstone, M., Sultan, B., and McDonald, M. (2002). Motorway driver behaviour: studies on car following. *Transportation Research Part F: Traffic Psychology and Behaviour*, 5(1):31–46.
- Brilon, W., Geistefeldt, J., and Regler, M. (2005). Reliability of Freeway Traffic Flow: A stochastic Concept of Capacity. In Mahmassani, H., editor, *Proceedings of 16th International Symposium on Transportation and Traffic Theory*, pages 125 – 144. College Park, Maryland.
- Brockfeld, E., Kühne, R. D., and Wagner, P. (2004). Calibration and Validation of Microscopic Traffic Flow Models. *Transportation Research Record: Journal of the Transportation Research Board No.1876*, pages 62 – 70.
- Brown, G., Carlyle, M., Salmerón, J., and Wood, K. (2006). Defending Critical Infrastructure. *Interfaces*, 36(6):530 – 544.
- Cassidy, M. J. and Bertini, R. L. (1999). Some traffic features at freeway bottlenecks. *Transportation Research Part B: Methodological*, 33(1):25–42.

- Chakravarti, I. M., Laha, R. G., and Roy, J. (1967). *Handbook of methods of applied statistics: techniques of computation, descriptive methods, and statistical inference.* John Wiley and Sons, New York.
- Chandler, R. E., Herman, R., and Montroll, E. W. (1958). Traffic dynamics: studies in car following. *Operations Research*, 6:165–184.
- Chen, R. B. and Mahmassani, H. S. (2004). Travel Time Perception and Learning Mechanism in Traffic Networks. In *Proceedings of the 83th Annual Meeting of the Transportation Research Board*. Washington D.C.
- Chiabaut, N., Buisson, C., and Leclercq, L. (2008). Fundamental Diagram Estimation through Passing Rate Measurements in Congestion. In *Proceedings of 87th Annual Meeting of the Transportation Research Board*. Washington D.C.
- Cho, Y. and Rice, J. (2006). Estimating Velocity Fields on a Freeway from Low-Resolution Videos. *IEEE Transactions on intelligent transportation systems*, 7(4):463–469.
- Chorus, C. G., Molin, E. J. E., and Van Wee, B. (2006). Use and Effects of Advanced Traveller Information Services (ATIS): A Review of the Literature. *Transport Reviews*, 26(2):127149.
- Choudhury, C., Akiva, M. E. B., Toledo, T., Rao, A., and Lee, G. (2006). Cooperative Lane changing and Forced Merging Model: NGSIM Final Report. Technical report, Federal Highway Administration.
- Chung, K., Rudjanakanoknad, J., and Cassidy, M. J. (2007). Relation between traffic density and capacity drop at three freeway bottlenecks. *Transportation Research Part B: Methodological*, 41(1):82–95.
- Daganzo, C. F. (1994). The Cell Transmission Model: a Dynamic Representation of Highway Traffic Consistent With the Hydrodynamic Theory. *Transportation research part B*, 28B(4):269–287.
- Daganzo, C. F. (1995). The Cell Transmission Model 2. Network Traffic. *Transportation Research Part B-Methodological*, 29(2):79–93.
- Daganzo, C. F. (1997). *Fundamentals of Transportation and Traffic Operations*. Pergamon.
- De Jong, G., Tseng, Y., Kouwenhoven, M., Verhoef, E., and Bates, J. (2007). The value of travel time and travel time reliability. Technical report, Significance - Prepared for The Netherlands Ministry of Transport, Public Works and Water Management.
- De Palma, A. and Picard, N. (2005). Congestion on Risky Routes with Risk Averse Drivers. In *Proceedings of the 84th Annual Meeting of the Transportation Research Board*. Washington D.C.

- Dicke-Ogenia, M., Rademaker, R., and Brookhuis, K. (2008). Safe and innovative in-car travel information. In *Proceedings of 10th TRAIL congress*. Rotterdam, the Netherlands.
- Dijker, T., Bovy, P. H. L., and Vermijs, R. G. M. M. (1997). Car-following under non-congested and congested conditions. Technical Report VK 2206.301, Delft University of Technology.
- Dixon, K. K., Hummer, J. E., and Lorscheider, A. R. (1996). Capacity for north carolina freeway work zones. *Transportation Research Record: Journal of the Transportation Research Board*, No. 1529.
- Dutch Road Authority (2005a). Nota Mobiliteit. Technical Report 2004-2005, 29644 nr. 13, Ministerie van Verkeer en Waterstaat & VROM.
- Dutch Road Authority (2005b). Verkeerskundige effecten varianten “anders betalen voor mobiliteit” (the effects on traffic of different implementations of road pricing). Technical report.
- Dutch Road Authority (2007a). Bereikbaarheidsmonitor hoofdwegenet 2006. Technical report.
- Dutch Road Authority (2007b). Inventarisatie beleidseffecten incidentmanagement. Technical report.
- Dutch Road Authority (2007c). Nieuwe Ontwerprichtlijn Autosnelwegen. Technical Report ISBN 90 3693 636 5, Rijkswaterstaat Adviesdienst Verkeer en Vervoer.
- Eliasson, J. (2004). Car drivers’ valuations of travel time variability, unexpected delays and queue driving. In *Proceedings of the European Transport Conference*. Strasbourg, France.
- Frejinger, E. and Bierlaire, M. (2007). Stochastic Path Generation Algorithm for Stochastic Networks. In *Proceedings of TRISTAN VI*. Phuket, Thailand.
- Fu, L. and Rilett, L. R. (1997). Real-Time Estimation of Incident Delay in Dynamic and Stochastic Networks. *Transportation Research Record: Journal of the Transportation Research Board*, No. 1603.
- Fudenberg, D. and Tirole, J. (1991). *Game Theory*. MIT Press.
- Fukui, M., Nishinari, K., Yokoya, Y., and Ishibashi, Y. (2009). Effect of real-time information upon traffic flows on crossing roads. *Physica A: Statistical Mechanics and its Applications*, 388(7):1207–1212.
- Garib, A., Radwan, A. E., and Al-Deek, H. (1997). Estimating magnitude and duration of incident delays. *ASCE Journal of Transportation Engineering*, 123:459–466.

- Geroliminis, N. and Daganzo, C. F. (2008). Existence of urban-scale macroscopic fundamental diagrams: Some experimental findings. *Transportation Research Part B: Methodological*, 42(9):759–770.
- Godunov, S. K. (1959). A difference scheme for numerical computation of discontinuous solutions of equations of fluid dynamics. *Math. Sb.*, 47:271 – 290.
- Goolsby, M. E. (1971). Influence of Incidents on Freeway Quality of Service. In *Proceedings of The 50th Annual Meeting of the Transportation Research Board*. Washington D.C.
- Greenshields, B. D. (1934). A Study of Traffic Capacity. *Proceedings Highway Research Board*, 14:448–477.
- Hall, F. L. and Agyemang-Duah, K. (1991). Freeway Capacity Drop and the Definition of Capacity. *Transportation Research Record: Journal of the Transportation Research Board No.1320*, pages 91–98.
- Hamdar, S. H. and Mahmassani, H. S. (2008). From Existing Accident-Free Car-Following Models to Colliding Vehicles: Exploration and Assessment. In *Proceedings of the 87th Annual Meeting of the Transportation Research Board*. Washington D.C.
- Heaslip, K., Louisell, C., and Collura, J. (2008). Driver population adjustment factors for the highway capacity manual work zone capacity equation. In *Proceedings of the 87th Annual Meeting of the Transportation Research Board*. Washington, D.C.
- Heikoop, H., Hoogendoorn, S. P., and Martens, G. J. (2007). Onderzoek Verkeersafwikkeling en capaciteitswaarden discontinuïteiten Autosnelwegen. Technical Report 2007-049, Rijkswaterstaat, Adviesdienst Verkeer en Vervoer.
- Helbing, D. (1997). *Verkehrsdynamik*. Springer, Heidelberg.
- Helbing, D. (2009). On the controversy around daganzo’s requiem for and aw-rasclé’s resurrection of second-order traffic flow models. *The European Physical Journal B - Condensed Matter and Complex Systems*, 69.
- Hoogendoorn, S., Van Zuylen, H. J., Schreuder, M., Gorte, B. G. H., and Vosselman, M. (2003). Microscopic traffic data collection by remote sensing. *Transportation Research Record: Journal of the Transportation Research Board No.1885*, pages 121 – 128.
- Hoogendoorn, S. P. (2005). Unified approach to estimating free speed distributions. *Transportation Research Part B: Methodological*, 39(8):709–727.
- Hoogendoorn, S. P. and Van Lint, J. W. C. (2007). Estimation of Car-Following Models using Prior Information. In *Proceedings of the IEEE conference on Intelligent Transport Systems*.

- Jenelius, E. (2007). Incorporating Dynamics and Information in a Consequence Model for Road Network Vulnerability Analysis. In Van Zuylen, H. J., editor, *Proceedings of Third International Symposium on Transport Network Reliability*. The Hague, the Netherlands.
- Jenelius, E., Petersen, T., and Mattsson, L.-G. (2006). Importance and exposure in road network vulnerability analysis. *Transportation Research Part A: Policy and Practice*, 40(7):537–560.
- Jou, R.-C., Lam, S.-H., Liu, Y.-H., and Chen, K.-H. (2005). Route switching behavior on freeways with the provision of different types of real-time traffic information. *Transportation Research Part A: Policy and Practice*, 39(5):445–461.
- Kaplan, E. L. and Meier, P. (1958). Nonparametric estimation from incomplete observations. *Journal of the American Statistical Association*, 53(282):457–481.
- Kesting, A. and Treiber, M. (2008). Calibrating Car-Following Models using Trajectory Data: A Methodological Study. In *Proceedings of the 87th Annual Meeting of the Transportation Research Board*. Washington D.C.
- Khattak, A., Schofer, J., and Wang, M. (1995). A simple time-sequential procedure for predicting freeway incident duration. *IVHS Journal*, 2:113–138.
- Khattak, A. J., Yim, Y., and Prokopy, L. S. (2003). Willingness to pay for travel information. *Transportation Research Part C: Emerging Technologies*, 11(2):137–159.
- Kim, T., Lovell, D. J., and Paracha, J. (2001). A new methodology to estimate capacity for freeway work zones. In *Proceedings of the 80th Annual Meeting of the Transportation Research Board*. Washington D.C.
- Knibbe, W. J. J., Alkim, T. P., Otten, J. F. W., and Aidoo, M. Y. (2006). Automated Estimation of Incident Duration on Dutch Highways. In *Proceedings of IEEE ITS Conference*. Toronto, Canada.
- Knoop, V. L., Bell, M. G. H., and Van Zuylen, H. J. (2008a). Route-Advice Based On Individual Risk-Attitude. In *Proceedings of IEEE Conference on Infrastructure Systems - Building Networks for a Better Future*. Rotterdam, the Netherlands.
- Knoop, V. L., Hoogendoorn, S. P., and Van Zuylen, H. J. (2007a). Approach to Critical Link Analysis of Robustness for Dynamical Road Networks. In Schadschneider, A., Pöschel, T., Kühne, R., Schreckenberg, M., and Wolf, D., editors, *Traffic & Granular Flow*, pages 393–402. Springer Verlag, Berlin.
- Knoop, V. L., Hoogendoorn, S. P., and Van Zuylen, H. J. (2007b). Quantification of the impact of spillback modeling in assessing network reliability. In *Proceedings of 86th annual meeting of the Transportation Research Board*. Washington D.C.
- Knoop, V. L., Hoogendoorn, S. P., and Van Zuylen, H. J. (2008b). Capacity Reduction at Incidents: Empirical Data Collected from a Helicopter. *Transportation Research Record: Journal of the Transportation Research Board*, No. 2071, pages 19–25.

- Koo, R. and Yim, Y. (1998). Commuter Response to Traffic Information on an Incident. Technical Report UCB-ITS-PWP-98-26, University of California, Berkely.
- Koshi, M., Iwasaki, M., and Ohkura, I. (1981). Some findings and an overview on vehicular flow characteristics. In *Proceedings of the 8th International Symposium on Transportation and Traffic Theory*, pages 403–426. Univ. of Toronto Press, Toronto.
- Kraaijeveld, R. (2008). *Onderzoek naar de identificatie van kwetsbare wegvakken*. Master's thesis, Delft University of Technology, Delft, the Netherlands.
- Kraan, M., Schreuder, M., and Tamminga, G. (2008). Vulnerability of a National Road Network. In *Proceedings of the 87th Annual Meeting of the Transportation Research Board*. Washington D.C.
- Kraan, M., Van der Zijpp, N., Tutert, B., Vonk, T., and Van Megen, D. (1999). Evaluating Networkwide Effects of Variable Message Signs in the Netherlands. *Transportation Research Record: Journal of the Transportation Research Board*, No. 1689, pages 60–67.
- Kurauchi, F., Sumalee, A., Tamura, H., and Uno, N. (2007). Bilevel programming problem for analysing capacity vulnerability in a transportation network under limited damage. In *Proceedings of third International Symposium on Transport Network Reliability*. The Hague, the Netherlands.
- Kwon, J., Mauch, M., and Varaiya, P. (2006). The components of congestion: delay from incidents, special events, lane closures, weather, potential ramp metering gain, and excess demand. In *Proceedings of the 85th annual meeting of the Transportation Research Board*, volume CD-ROM. Washington.
- Kyte, M., Dixon, M., Urbanik, T., Nayak, V., and Abdel-Rahim, A. (2008). Whats New in the Queue: Discovering More about Queue Discharge Characteristics and Their Effect on Signal Timing Using the New NGSIM Data Set. In *Proceedings of the 87th Annual Meeting of the Transportation Research Board*.
- Lam, T. C. and Small, K. A. (2001). The value of time and reliability: Measurement from a value pricing experiment. *Transportation Research Part E: Logistics and Transportation Review*, 37(2-3):231–251.
- Lebacque, J. P. (1996). The Godunov Scheme and What it Means for First Order Traffic Flow Models. In Lesort, J., editor, *Proceedings of the 13th International Symposium on Transportation and Traffic Theory*.
- Li, H. (2009). *Reliability-based Dynamic Network Design with Stochastic Networks*. TRAIL Thesis Series T2009/3, Delft University of Technology, TRAIL Research School, Delft, the Netherlands.
- Li, J., Lan, C.-J., and Gu, X. (2006). Estimation of Incident Delay and Its Uncertainty on Freeway Networks. *Transportation Research Record: Journal of the Transportation Research Board*, No. 1959.

- Li, M. (2008). *Combining DTA Approaches for Studying Road Network Robustness*. TRAIL Thesis Series T2008/14, Delft University of Technology, TRAIL Research School, Delft, the Netherlands.
- Lighthill, M. J. and Whitham, G. B. (1955). On Kinematic Waves. II. A Theory of Traffic Flow on Long Crowded Roads,. *Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences*, 229(1178):317 – 345.
- Liu, H. X., Ban, X., Ran, B., and Mirchandani, P. (2002). An Analytical Dynamic Traffic Assignment Model with Probabilistic Travel Times and Perceptions. In *Proceedings of the 81st Annual Meeting of the Transportation Research Board*. Washington D.C.
- Liu, H. X., Recker, W., and Chen, A. (2004). Uncovering the contribution of travel time reliability to dynamic route choice using real-time loop data. *Transportation Research Part A: Policy and Practice*, 38(6):435–453.
- Lo, H. K. and Tung, Y.-K. (2003). Network with degradable links: capacity analysis and design. *Transportation Research Part B: Methodological*, 37(4):345–363.
- Marchal, F. and de Palma, A. (2008). Measurement of uncertainty costs with dynamic traffic simulations. In *Proceedings of the 87th Annual Meeting of the Transportation Research Board*. Washington D.C.
- Muller, T. H. J., Miska, M. P., and Van Zuylen, H. J. (2005). Monitoring Traffic under Congestion: Base for Dynamic Assignment in Online Prediction Models. In *Proceedings of the 84th Annual Meeting of the Transportation Research Board*. Washington D.C.
- Murray-Tuite, P. M. and Mahmassani, H. S. (2004). A Methodology for the determination of Vulnerable Links in a Transportation Network. In *Proceedings of the 83rd annual meeting of the Transportation Research Board*.
- Nagae, T. and Akamatsu, T. (2007). An Efficient Algorithm for a Maxmin Routing Model for Hazardous Materials. In *Proceedings of the Third International Symposium on Transport Network Reliability*. The Hague, the Netherlands.
- Next Generation Infrastructures (2009). [www.nginfra.nl](http://www.nginfra.nl).  
<http://www.nginfra.nl/index.php?pageID=23>.
- Ngoduy, D. (2006). *Macroscopic Discontinuity Modeling for Multiclass Multilane Traffic Flow Operations*. TRAIL Thesis Series T2006/3, Delft University of Technology.
- NGSIM (2008). Home of the next generation simulation community.  
[www.ngsim.fhwa.dot.gov/](http://www.ngsim.fhwa.dot.gov/).
- Noland, R. B., Small, K. A., Koskenoja, P. M., and Chu, X. (1998). Simulating travel reliability. *Regional Science and Urban Economics*, 28(5):535–564.
- of Maryland, U. (2003). DYNAMSMART-X Users Guide. Technical report.

- Olmstead, T. (1999). Pitfall to Avoid when Estimating Incident-Induced Delay by Using Deterministic Queuing Models. *Transportation Research Record: Journal of the Transportation Research Board*, No. 1683.
- Ossen, S., Hoogendoorn, S., and Gorte, B. G. H. (2006). Interdriver Differences in Car-Following: A Vehicle Trajectory-Based Study. *Transportation Research Record: Journal of the Transportation Research Board* No.1965, pages 121–129.
- Ossen, S. J. L. (2008). *Longitudinal Driving Behavior: Theory and Empirics*. Trail thesis series 2008/8, Delft University of Technology, TRAIL Research School, Delft, the Netherlands.
- Qin, L. and Smith, B. L. (2001). Characterization of accident Capacity Reduction. Technical Report STL-2001-02 University of Virginia, University of Virginia.
- Richards, P. I. (1956). Shock waves on the highway. *Operations Research* 4, 4:4251.
- Schaap, N., Van Arem, B., and Van der Horst, R. (2008). Drivers' Behavioural Reactions to Unexpected Events. In *Proceedings of 10th TRAIL Congress - TRAIL in Perspective*. Delft, the Netherlands.
- Schrijver, J., Immers, B., Snelder, M., and De Jong, R. (2006). Effecten van de landelijke invoering van incidentmanagementmaatregelen op de voertuigverliestijd in het netwerk (effects of nation-wide introduction of incident management measures for vehicle delay in the network). Technical Report 06.34.15/N079/034.65116/JS/LK—, TNO, Delft.
- Senna, L. A. D. S. (1994). The influence of travel time variability on the value of time. *Transportation*, 21(2):203–228. 10.1007/BF01098793.
- Sheffi, Y. (1985). *Urban transportation networks: Equilibrium analysis with mathematical programming methods*. Prentice-Hall, Englewood Cliffs, New Jersey.
- Sinha, P., Hadi, P. E. M., and Wang, E. I. A. (2007). Modeling Reductions in Freeway Capacity due to Incidents in Microscopic Simulation Models. In *Proceedings of 86th Annual Meeting of the Transportation Research Board*. Washington D.C.
- Smith, B. L., Qin, L., and Vankatanarayana, R. (2003). Characterization of Freeway Capacity Reduction Resulting from Traffic Accidents. *Journal of Transportation Engineering*, 129(4):362 – 368.
- Snelder, M., Immers, B., and Van Zuylen, H. J. (2008a). The Importance of Alternative Routes for the Robustness of a Road Network. In *Proceedings of IEEE International Conference on Infrastructure Systems - Building Networks for a Better Future*. Rotterdam, the Netherlands.
- Snelder, M., Schrijver, J., Landman, R., Mak, J., and Minderhoud, M. (2008b). De kwetsbaarheid van randstedelijke vervoernetwerken uit verkeerskundig perspectief. Technical Report 2008-D-R0882/A, TNO, The Netherlands Organisation of Applied Scientific Research.

- Sodnik, J., Dicke, C., Tomazic, S., and Billingham, M. (2008). A user study of auditory versus visual interfaces for use while driving. *International Journal of Human-Computer Studies*, 66(5):318–332.
- Spearman, C. (1904). The proof and measurement of association between two things. *The American Journal of Psychology*, 15(1):72–101.
- Tamminga, G. F., Maton, J. C., Poorteman, R., and Zee, J. (2005). De Robuustheidsscanner - Robuustheid van netwerken: een modelmatige verkenning. Technical Report I&M-99366053-GT/mk, Grontmij Nederland, i.o.v. Adviesdienst Verkeer & Vervoer.
- Tampère, C. M. J. (2004). *Human-kinetic multiclass traffic flow theory and modelling*. TRAIL Thesis Series T2008/11, Delft University of Technology, TRAIL Research School, Delft, the Netherlands.
- Tampère, M. J., Stada, J., and Immers, L. H. (2007). Methodology for Identifying Vulnerable Section in a National Road network. In *Proceedings of 86th Annual Meeting of the Transportation Research Board*. Washington D.C.
- Thiemann, C., Kesting, A., and Treiber, M. (2008). Estimating Acceleration and Lane-Changing Dynamics Based on NGSIM Trajectory Data. In *Proceedings of the 87th Annual Meeting of the Transportation Research Board*. Washington D.C.
- Toledo, T., Koutsopoulos, H. N., and Ahmed, K. I. (2007). Estimation of Vehicle Trajectories with Locally Weighted Regression. In *Proceedings of the 86th Annual Meeting of the Transportation Research Board*. Washington D.C.
- Transportation Research Board (2000). Highway Capacity Manual. Technical report.
- Treiterer, J. and Myers, J. A. (1974). The hysteresis phenomenon in traffic flow. In *Proceedings of the 6th International Symposium on Transportation and Traffic Theory*, pages 13–38. Elsevier.
- Tu, H. (2008). *Monitoring Travel Time Reliability on Freeways*. Trail thesis series 2008/7, Delft University of Technology, TRAIL Research School, Delft, the Netherlands.
- Tu, H., Van Lint, J. W. C., and Van Zuylen, H. J. (2005). Real-time modeling travel time reliability on freeway. In Jaskiewicz, A., Kaczmarek, M., Zak, J., and Kubiak, M., editors, *Proceedings of 10th Jubilee Meeting of the EURO Working Group of Transportation*, pages 540–546. Poznan.
- Tu, H., Van Lint, J. W. C., and Van Zuylen, H. J. (2008). Travel time reliability modelling on freeways. In *Proceedings of the 87th Annual Meeting of the Transportation Research Board*. Washington D.C.
- Van Aerde, M. (2005). INTEGRATION Release 2.30 for Windows: User's Guide-Volume II. Technical report.

- Van Amelsfort, D. H. (2009). *Behavioural Responses and Network Effects of Time-varying Road pricing*. TRAIL Thesis Series T2009/4, Delft University of Technology, TRAIL Research School, Delft, the Netherlands.
- Van Toorenburg, J. A. C. and Nijenhuis, T. A. (2007). Capaciteitsbeperking door incidenten. Technical report, Transpute, Gouda, the Netherlands.
- Vickrey, W. S. (1969). Congestion theory and transport investment. *The American Economic Review*, 59(2):251–260.
- Wang, W., Chen, H., and Bell, M. C. (2005). Vehicle Breakdown Duration Modeling. *Journal of Transportation and Statistics*, 8(1):75–84.
- Wardrop, J. (1952). Some theoretical aspects of road traffic research. *Proceedings of the Institute of Civil Engineers, Part II*, I(2):325–378.
- Wong, S. C. and Wong, G. C. K. (2002). An analytical shock-fitting algorithm for LWR kinematic wave model embedded with linear speed-density relationship. *Transportation Research Part B: Methodological*, 36(8):683–706.
- Wu, N. (2002). A new approach for modeling of Fundamental Diagrams. *Transportation Research Part A: Policy and Practice*, 36(10):867–884.
- Yperman, I. (2007). *The Link Transmission Model for Dynamic Network Loading*. Ph.D. thesis, Katholieke Universiteit Leuven.
- Zhang, L., Kovvali, V., Clark, N., Sallman, D., and Alexiadis, V. (2007a). NGSIM-VIDEO User's Manual. Technical Report FHWA-HOP-07-009, US Department of Transportation, Federal Highway Administration.
- Zhang, N., Jiyang, B., Sun, L., Hoogendoorn, S. P., and Hoogendoorn-Lanser, S. (2007b). Estimation of Incident Duration with a Combined Classification Tree and Regression Method. In *Proceedings of Sino Dutch cooperation on Intelligent Transport Systems and Road Safety*. Beijing, China.
- Zuurbier, F. S., Van Zuylen, H. J., Hoogendoorn, S. P., and Chen, Y. (2006). A generic approach to generating optimal controlled prescriptive route guidance in realistic traffic networks. In *Proceedings of the 85th annual meeting of the Transportation Research Board*, volume CD-ROM. Washington.



# Appendix A

## Processing traffic data collected by remote sensing

---

Video data is used more and more often to study traffic operations. However, extracting vehicle trajectories from video by using current methods is a difficult process, typically resulting in many errors. Additionally, this requires much labor to manually correct the trajectories. This chapter proposes a new method to process video data from traffic operations. We do not detect a vehicle in each picture of the video separately, but instead transform the video data such that the trajectories of the vehicles (their position over time) become visible in a single image. In this single image, the trajectories can be found by detecting lines. The difference with other methods is trajectories are detected rather than vehicles. Trajectory (line) detection is more robust than vehicle (rectangle) detection, seen from the fact that applying this method around 95% of the trajectories are detected correctly and, more importantly, the length of the segments of each of the trajectories is much longer, compared to results from other methods reported in literature. Also, the detection is a very quick process since only a single image is required to be analysed. For a dataset of 5 minutes in length, transforming costs several minutes, and automatically detecting and tracking costs 40 to 50 minutes per lane. Further manually correcting is then necessary which costs about 10 minutes per lane. In contrast, using a different method the total processing time for analysing the traffic operations costs approximately one week for all lanes together. This appendix is an edited version of Knoop, V.L., Hoogendoorn, S.P. and Van Zuylen, H.J.J. Processing Traffic Data collected by Remote Sensing. Accepted for publication in *Transportation Research Records*.

---

### A.1 Introduction

Traffic congestion forms an increasing problem in modern society. Traffic engineers try to improve the situation, but knowledge of traffic characteristics in general and at the

bottleneck location are crucial. In early days, most traffic-related studies were carried out using macroscopic data. The first attempt to find the relation between density and speed was made 75 years ago by Greenshields (1934). Chandler et al. (1958) argued the necessity of data on individual driving behavior in 1958, using wired vehicles on a test-track for their experiments. Treiterer and Myers (1974) chose a different approach and were among the first to capture traffic operations on video in 1974. In that study, the data was processed manually.

With modern techniques it is possible to gain better knowledge of traffic on this detailed level. Observations can be made of driving behavior of individual drivers. A detailed understanding of the individual driving behavior of road users can help improve traffic flow and thus reduce delays. The data of individual drivers, i.e. microscopic data, leads to insights into, for instance, headway and traffic dynamics from which, for instance, the capacity or the stability of a traffic flow can be derived. There exist different ways of collecting microscopic data. In this contribution, we focus on video data. In this chapter, the word “video” indicates a consecutive series of pictures illustrating traffic operations with a fixed time-interval, regardless of the picture size and frequency. The word “picture” here always indicates a single time-frame of the video.

The time-dependent coordinates of all vehicles give a complete overview of all traffic operations. From the video, these coordinates over time are known, which gives insights in the dynamics of driving. This chapter discusses the process of extracting this information from given video images. The newly proposed method gives more robust results than existing methods. This is the consequence of the fact that in other methods independent vehicles are detected, leading to difficulties in linking the vehicles found in consecutive pictures in order to get vehicle trajectories. When the program fails to recognise a vehicle in a subsequent picture, that particular vehicle needs to be indicated by the user. Obviously, these manual corrections are time-consuming. However, in the method proposed here, the duration of the parts of found trajectories are much longer, which is an advantage. Therefore, in using the proposed method, the time needed to manually complete the set of trajectories is much less. Another advantage of the newly proposed method is that it provides insight into the traffic operations without actually tracking the vehicles. That is, relevant properties of the traffic are put directly into a single image. In this way, the remote sensing data are made more easily accessible.

In the next section, an overview is given of data collecting and processing methods for individual vehicles reported in literature. The main section of the chapter (section A.3) then describes how the video can be transformed. It shows which operations can be applied to combine all relevant information in a single image. Section A.4 discusses how trajectories can be detected automatically from this (transformed) image. The focus in this process is on obtaining the trajectory data. The final section then summarises the presented research and gives concluding remarks.

## A.2 Literature Review

This section provides an overview of how microscopic traffic information can be gathered. Several methods and the advantages and disadvantages of methods are shown first. Which data collection methods are most appropriate and which data can be extracted is briefly discussed. This section shows that video data is a good option for several applications. Then, the section A.2.2 discusses the methods literature provides to collect and process video data as well as the difficulties in processing video data into trajectory data.

### A.2.1 Microscopic Traffic Data

Various studies require traffic data on the level of individual vehicles, which can be collected using a variety of methods. An exhaustive overview thereof as well as their applications can be found in Appendix C of the thesis of Ossen (2008). Here, we only briefly discuss the most frequently used methods.

One of the ways of collecting microscopic data is recording the information obtained from (double) loop detectors in the road at vehicle level which provides passing times (and therefore headways) and speeds at a specific location (e.g., Hoogendoorn (2005)). However, since the cars are not followed over time, this does not give information about the dynamics of traffic.

A second way of data collecting is by GPS which can be used to record information on the speed and location of an instrumented vehicle over time. This provides the dynamics for that particular instrumented car (e.g., Brockfeld et al. (2004)), but does not give insights into the interaction of vehicles. Only when there is a platoon of equipped vehicles, there is insight into the relative positions and relative speeds of the vehicles. However, this is only realised in an experiment at a test-track or in an artificially created situation on the road.

A third way of data collecting is using an instrumented vehicle. For instance, vehicles equipped with a radar can provide a more representative view on the vehicle dynamics (e.g., Brackstone et al. (2002)). The radar can measure the distance between two vehicles (distance to leading vehicle, or to the following vehicle, or both). Note that the radar can be used to monitor the driving behavior of the driver of the equipped vehicle, which is called the active mode, or to monitor the behavior of the following vehicle, which is called the passive mode. In the active mode, only a few drivers can be studied, whereas in the passive mode, every driver that follows the instrumented vehicle can be studied. The disadvantage of the passive mode is that the test vehicle influences the traffic. Additionally, the distances measured by the radar can be combined with a GPS-signal to see the influences of the location on the car-following behavior. The disadvantage of using an instrumented vehicle is that the observations for different cars are made in different conditions because the instrumented vehicle itself is moving. That means that either all observations are made at different locations, or, if the instrumented vehicle returns to the

same location, observations are made at different time instances and under different traffic conditions.

A fourth way of data collecting is capturing video from a high point of view. This provides insights into the driving behavior of drivers both longitudinally (e.g., car-following), and laterally (e.g., lane selection). Consequently, one can analyse the headway choice, like using an instrumented vehicle, but also, for instance, the influence of the distance to the leader of the leader on the acceleration. So, one advantage is that the information has both a spatial and temporal dimension. Another advantage is that many drivers can be observed under similar driving conditions.

## A.2.2 Collecting and Processing Video Data

It is only a few years ago that projects started using remote sensing data on a larger scale, for instance Tracing Congestion Dynamics at the Delft University of Technology or The NGSIM (Next Generation Simulation) project which provides data of individual vehicles. These are used in many projects in the US, for a wide variety of applications, from analysing signalised intersections (Kyte et al., 2008) to lane-change behavior (Choudhury et al., 2006) and emergency behavior (Hamdar and Mahmassani, 2008). Also in Europe, these US data are used in several projects, for instance to show properties of the fundamental diagram (Chiabaut et al., 2008) or to calibrate car-following models (Kesting and Treiber, 2008).

Before these video data can be used, these need to be processed. Hoogendoorn et al. (2003) propose a method of first finding vehicles by comparing pixel values in a picture with the median value of that pixel (i.e., that position) over time, which can be interpreted as the brightness of the roadway background. Then, clusters of pixels deviating from the values identified for the background are marked as vehicles. This process is repeated for each picture. In the final step, the vehicles are tracked by comparing the location of vehicles in two subsequent images. A detection of over 90% is claimed.

However, this method has shown to have some flaws when we applied it to a data set with different shades of tarmac and different light intensities. In practice, it turned out that the trajectory of a vehicle was interrupted over time. This makes it necessary to combine partial short trajectories. A normal filter procedure would discard too short trajectories, which could be interpreted as noise. Because the trajectories of vehicles were interrupted and short, such a filter procedure could not be applied now. Additionally, other objects than vehicles were detected (e.g., road lines), which we could not simply ignore. This gave some serious trouble regarding the combination of detected parts since it was unclear whether the parts were (moving) vehicles or other (stationary) objects. In short, the main problem was the tracking.

A similar methodology is stated by Angel et al. (2003). In each picture of video taken from a helicopter, the intensity is compared with a predefined pixel value. A concentration of deviating pixels is marked as a vehicle. In contrast to Hoogendoorn et al. (2003) where only the information on the pictures is used to find reference points, Angel et al. (2003)

use added information collected during the flight: the position (GPS-coordinates) and the angle of the helicopter is saved together with the video data. The main problem indicated here is the “matches” of vehicles, i.e. to connect the same detected vehicle over different frames. They also report a matching percentage of approximately 90%.

In the NGSIM-project (NGSIM, 2008) a subproject deals with extracting the trajectory information from the video, NGSIM-video (Zhang et al., 2007a). There, vehicles are found in each image, but also origins and destinations are to be provided. In the tracking process, this information on flow direction is used. The tool has a possibility to track both forward and backward in time. No reliability value is given for the automated procedure.

However, given the similar methodology to both Hoogendoorn et al. (2003) and Angel et al. (2003), it is likely that similar problems occur. Additionally, the manual states that there is a need to post-process the tracks manually, for which a tool is provided. This is time-consuming work: ‘Experienced NGSIM-VIDEO operators can effectively process video at a rate of 40 - 100 seconds of tracking per work day for a 500-meter motorway section.’ (Zhang et al., 2007a).

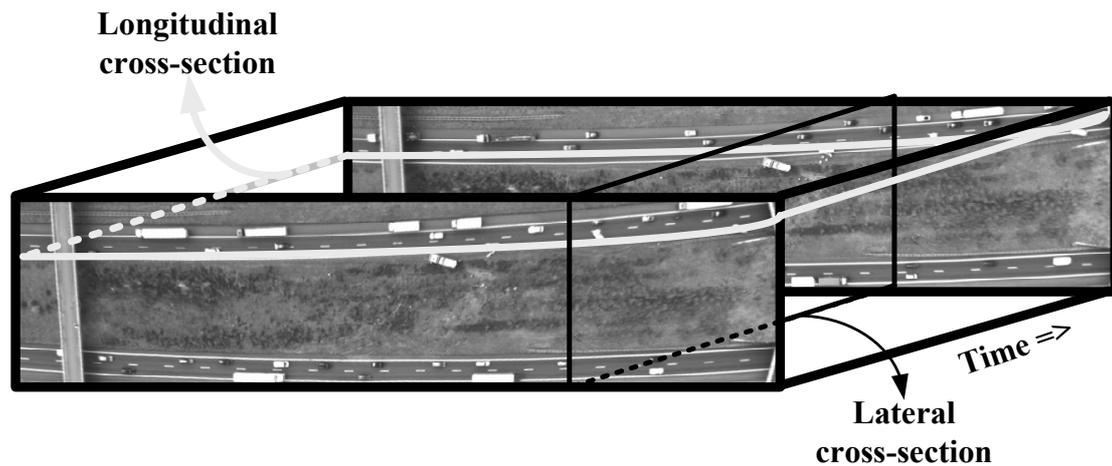
Cho and Rice (2006) present a method to estimate the velocity field from video data, based on the image intensity in each lane. To find a vehicle, they take the maximum value for the image intensity across a single lane. This method detects bright vehicles, but darker vehicles cannot be identified since the maximum intensity value on a cross section is the intensity of the pavement. However, this is not a problem for determining the velocity field, which can be derived from the bright vehicles only. For the same reason, unmatched parts of a trajectory do not form a problem, since it is sufficient to use the parts that are available. However, a more detailed analysis requires the full trajectories of all vehicles.

Therefore, there is a need for a method which gives the trajectories of vehicles from remote-sensing data reliably for longer time intervals without much manual post-processing.

### **A.3 Transformations**

This chapter proposes to reduce the problem with one dimension and to capture all relevant information in one image. This section explains which transformations can be made in order to construct these images.

If the video is captured from a moving helicopter, the image need to be stabilised and rectified to get an picture with a perspective like viewing with angle is not perpendicular to the road surface. If it is taken from a stable platform, one could start with the transformation (as proposed later in this section) and perform the rectification later on the transformed image. This will save computing time. Due to space restrictions, we will not discuss this process in this chapter. For this chapter, the starting point is a series of stabilised and rectified pictures. Usually, vehicles are identified in each picture, and this process is repeated for each picture. Note that we will use the word picture for a time slice of the video (the snap-shot of the traffic conditions in one moment in time), whereas the word image is used for graphical representations which are constructed otherwise.



**Figure A.1:** A three-dimensional box of pictures with the lateral (black) and longitudinal (gray) cross-section

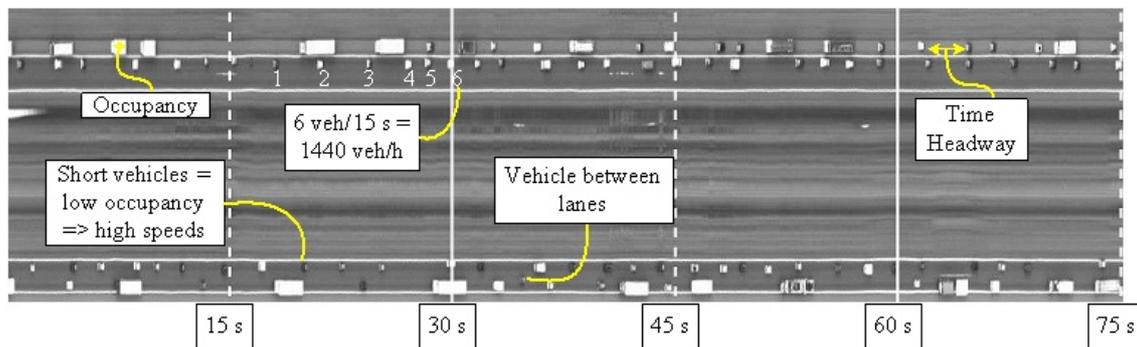
The time series of pictures can be combined to a three-dimensional box, with two space dimensions and a time dimension (figure A.1). It is proposed to take a cross-section (i.e., a slice or cross-cut of this box) in one of the space dimensions. The first subsection shows the transformation in the lateral direction, perpendicular to the driving direction, indicated by a black frame in figure A.1. For one point along the road, the image shows the occupancies over time and the flows. The second subsection shows the transformations in the longitudinal direction, along the driving direction, indicated by a gray frame in figure A.1. This transformation gives the occupancy over space and time for one lane, directly indicating the vehicle trajectories.

The video used here as example and to quote computation times originate from a experiment described in chapter 3. The pictures with a size of 1392x1040 pixels are recorded at a frequency of 15.1 Hz.

### A.3.1 Lateral Cross-Section

Figure A.2 shows a lateral cross-section of a series of pictures, or, expressed differently, it is a sectional plane perpendicular to the driving direction, like black red plane in figure A.1. Note that the size of the image is a distance, 50 meters from bottom to top, times a time, 75 seconds from left to right. The solid lines indicate a time step of half a minute, the dotted lines are indicate a time step of 15 seconds.

This image can be interpreted as the occupancy of a virtual loop-detector over time. It essentially shows when, at the horizontal axis, a vehicle passes a point along the road. Data from detection loops can lead to a similar image. However, this image also has a space component. Therefore, the image shows where on the lateral y-position a vehicle passes the point of the cross-section. This image can be constructed in a few minutes' time.



**Figure A.2: A lateral cross section, size 1.25 min x 50 m**

The image shown here is easily confused with a normal picture of traffic. This is because the structure of the road seems well visible. However, the length of the vehicles at the horizontal axis varies strongly since in this image the length of a vehicle is the time it would occupy the virtual detector and therefore faster driving cars are drawn shorter, as is visible in figure A.2.

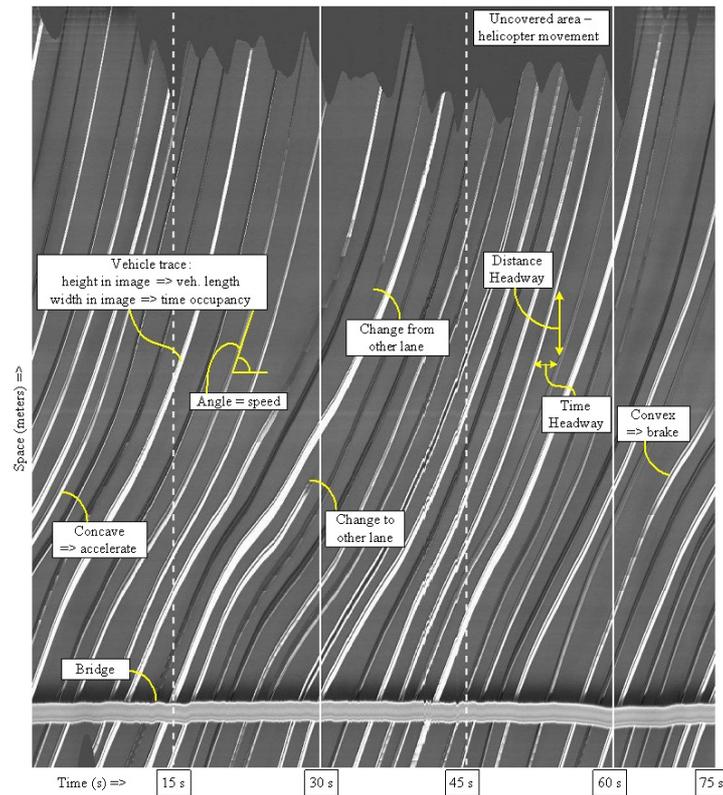
In this image, cars in both directions have the nose at the left and the back at the right. Each vehicle first reaches the cross section location with the nose of the vehicle. Since left in the image is an early point in time, the nose of the vehicle is drawn at the left hand side. This is most easily recognised by articulated lorries where the truck is at the left and the lorry at the right.

The transformation proposed in this subsection is ideal for getting time headways and flow values out of remote sensing data. The time headway can be read out as the distance between two passing cars as indicated in the figure. To obtain the flow, it is required to count the number of vehicles that pass by in a time interval, either manually or automatically, which is preferred for a longer dataset.

This type of images is very informative for the lateral position on the road. It can directly be seen whether cars within a platoon drive at the same lateral position or whether that depends on their position in the platoon or the headway or the type of car they are following. It furthermore can be derived whether the headway depends on the type of car (passenger car, truck) drivers are following.

There is one pitfall. Vehicles are not detected if they are upstream of the virtual detector at one picture and they are downstream of the detector in the next. However, when recording pictures at a time interval of (in our case) 15 images per second or more, that is very unlikely. In the worst case, there is a small (4 meter = 13 ft) vehicle that just does not reach the cross section point when the first image is taken. When the next picture is taken, it should have passed the point completely. So it should have covered 4 meters in  $1/15^{\text{th}}$  of a second, and thus should drive 60 m/s, which equals over 200 km/h or over 120 mph. For a frame rate of 10 images/second, the critical speeds is 145 km/h (90 mph). For longer cars these speeds are proportionally higher.

From this image, one cannot derive the dynamics of the car-following behavior. This can be derived from a longitudinal cross section discussed in the next subsection.



**Figure A.3: A longitudinal cross-section**

### A.3.2 Longitudinal Cross-Section

Figure A.3 shows a lateral cross-section image along the axis of the lane of a road, like indicated with a gray frame in figure A.1. On the horizontal axis, the time is plotted which progresses from left to right. Again, the vertical lines indicate 15 seconds (dotted lines) and 30 seconds. On the vertical axis, one finds the distance along the road with the traffic flowing from bottom to top. The cross-section is not necessarily taken along a straight line. A curved road requires that one takes a section along a curved line, as is shown in figure A.1.

Cho and Rice (2006) already propose a method filtering one lane out of a set of pictures. However, they take one *lane* out which is several tens of pixels wide instead of one pixel row as we proposed. Then the maximum of the pixels in that lane across the road is taken, along a lateral cross-section of the lane. This gives a clear image of the bright cars, but the dark cars disappear since the maximum intensity value of the pavement is higher than those of a dark car. We propose to directly narrow the data to a one-pixel wide cross-section. This is, to the best of the authors' knowledge, the first time such an approach is taken in the field of traffic flow.

Figure A.3 is constructed from the video data in a few minutes. It shows the intensity values of the (grayscale) images. The background of the image is the gray shade of the pavement. Vehicles move from the bottom (downstream) part of the image to the upper

part. A bright line indicates a bright vehicle moving, a dark line the passing of a dark vehicle. This is a graphical representation of the trajectory that is indicated by the word “trace” in this chapter. The length of the vehicle determines the width of the line in this image, the occupancy time at one location determines the height of the line.

At the bottom part, a bridge blocks the view on the road which is (obviously) present during the whole observation. It therefore forms a line from left to right. Because the bridge is higher than the road surface, the part of the road that is made invisible by the bridge changes slightly when the helicopter moves and therefore the line is not straight.

The angle that the trace of a vehicle makes with the horizontal line is its speed. In this way, the image shows the speeds of all cars at all times. Also the changes of speeds are visible since concave traces equal a reduction of speed and therefore braking, whereas convex traces equal an increase of speed and therefore accelerating.

Lane changes are represented by an appearance or disappearance of the trace of a vehicle in figure A.3. Examples thereof indicated in the figure. In this example, the following vehicle accelerates and closes the gap to the predecessor. Another vehicle subsequently merges into the gap that is now created behind the accelerating vehicle. The lane-changing vehicle enters this stream and thus appears in the image.

Using this longitudinal cross-section image, the dynamics of the traffic operations can be made visible without advanced detection and tracking algorithms. The image is constructed in typically several minutes on one 1.5 GHz core of a dual core pc. The constructing time is only determined by the speed of reading the images because here are no advanced image processing algorithms required. Note that in this way in a few minutes the dynamics of vehicle movements are made visible without fitting a model.

Contrary to the lateral transformation, this longitudinal transformation provides no information of the lateral position of the vehicles, other than the lane in which it is driving. In most cases, however, the only lateral information that is needed is the lane in which a vehicle is driving. Apart from that, each lane is shown in a different image. Nevertheless, one can track vehicles that change lane by combining the tracks of the vehicles in different lanes.

## A.4 Vehicle Detection

This section describes how vehicles can be detected and tracked in the cross-sectional images. As example we take the longitudinal cross-section image, because the trajectory information is the information that is more often required in studies using remote sensing data. The method of finding the cars (described in subsection threshold values) is applicable for the lateral cross-section image as well. However, the smoothing described in the subsequent section is only needed for the trajectories found in the longitudinal cross-section images. In practice it turned out that if we use a longitudinal cross-section image we can track a single vehicle for a longer time than if we use a procedure which detects

and tracks vehicles in each picture separately. Also the manual corrections are much easier than in the existing methods. Both the automatic part and the manual corrections are much quicker than the methods proposed so far. Namely, an image of a dataset of 5 minutes in length can be processed (detecting and tracking) in 40-50 minutes. The manual corrections cost around 10 minutes.

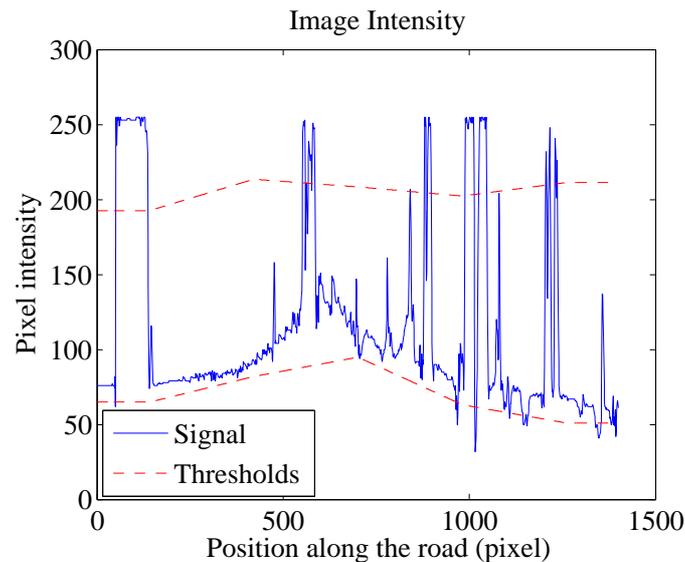
The images themselves provide a lot of information on traffic flows which can easily be extracted manually. The sequel of this section describes how this can be done automatically. However, fitting a polynomial function through the lines of a longitudinal cross section image gives the trajectories very accurately. In fact, the extent to which the polynomial curve follows the trace can be influenced by the number of points one uses to fit the trace. Generally, the number of points will be higher for vehicles with more fluctuation in speed. Therefore, choosing for either an automatic detection or a manual identification is a trade-off between programming time and optimization of the algorithm at one hand and working time (clicking points in the image) at the other hand.

#### **A.4.1 Threshold Values**

The images need to be changed from a gray value to a Boolean value (true or false) indicating vehicle presence. Compared to existing approaches, e.g. by Hoogendoorn et al. (2003) or Angel et al. (2003), we make two changes in the process of detecting vehicles. First, we do not take the values of one pixel over the whole time as basis for the threshold values because this gives some problems with the change of light intensity due to clouds passing by. It works much better if a shorter time (10 seconds) is considered. However, this also has a drawback, since, for example, a truck could be at the same spot for more than half of the time. In that case, using a median for the background, the intensity value of the truck would actually be considered to be the background and therefore the pavement would be seen as vehicle. To avoid this problem, we clustered the pixels in blocks of around 50 meters (around 150 pixels), which gave many more pixels to extract the distribution of intensities. This gives “blocks” of around (10 seconds =) 150 pixels times (50 meters =) 150 pixels to find the background intensity. The second difference with existing methods (e.g., Hoogendoorn et al. (2003)) is that we did not determine the background and found a deviation of it. In contrast, with upper and lower percentile values are used which turned out to give more reliable results.

For each of the “blocks” mentioned in the previous paragraph a lower and upper value are computed. This can be seen in figure A.4 where the intensity along one line is plotted. Note, the background intensity profile varies along the road which for instance can be caused by the different reflection of the road at various angles. The threshold values for each of the blocks are interpolated linearly to obtain the threshold values for each pixel individually. This pattern is also plotted in A.4.

Note that the percentage of pixels marked as car is fixed by choosing the percentile values. Moreover, it also determines whether the dark or the light deviations will be marked as vehicle. Therefore, the choice for the best percentile values depends, among others, on



**Figure A.4: The intensity of the image over space**

the traffic conditions and the brightness of the pavement relative to the brightness of the vehicles. For example, in case there is much traffic, the lower percentile value should be chosen low and the high percentile value should be chosen high to assure that a large part of the pixels is marked as vehicles. And in case the pavement is light, the majority of the vehicles will be darker than the pavement and therefore both the lower and the higher percentile values can be increased. In practice it is best to try different settings of the percentile values. In our case, it was best to choose 15% for the lower percentile value and 90% for the higher percentile value.

Special attention should be given to shadows since they might interfere with the vehicle detection. For instance, if the shadows lie over the adjacent lane, it is advisable to choose a pixel line of the video more to the side of the lane where the shadows does not lie. This was possible for our test data set. If it is unavoidable to take a pixel line with shadows, the percentile values have to be changed. In particular, the lower percentile value has to be increased to avoid that the shadows are detected as cars. Similarly, shadows could be detected as part of a car if they lie along the road which could increase length of the area being detected as different from the background and therefore increase the vehicle length and thus the inter-vehicle spacing. To solve this, the lower threshold value should be adjusted.

Another problem arises when there are vehicles with a colour which gives a pixel intensity similar to or equal to the pavement. Vehicles with the exact same reflection as the pavement cannot be detected. However, in our test case there were no such vehicles, as could be derived from the shadows. Let us finally remark that distinguishing vehicles from the pavement is easier if color video is used.

In case colour video is used, it is expected that detecting vehicles will be much easier and the results are expected to be better (more accurate, longer trajectories). Similar

methodologies can be applied, but the thresholds can be set to the hue and/or saturation to distinguish vehicles from the background. In case hue and saturation are used, no problems will arise with shadows, because a shadows on the road surface differs only in intensity from the lit road, and not in hue or saturation.

Vehicle presence now has to be filtered. First, the image with traces (in the space-time plane of the longitudinal cross-section) is filtered by applying an exponential moving average filter with an exponent of 2 pixels (2/15 of a second) in the time direction and 1 pixel (30 cm) in the space dimension to filter out the small areas that deviate from the neighboring pixels. The result is a smoother image. Then the image is filtered based on the properties of cars. Any vehicle has to have a length of 3 (10 feet) to 20 meters (66 feet). If the vehicle is too small, the signal is ignored because this is noise. Additionally, a gap between two vehicles has a minimum size of 3 meters (10 feet). If the gap is too small, the two parts are combined since this is either noise or for instance an articulated lorry.

Now, the vehicles found in each time step should be matched to create trajectories. In two subsequent lines, the expected displacement of the vehicle can be calculated based on speed; in case there is no speed known, for instance in the first time step, a fixed estimate can be used. If the match of size and location is better than a threshold value, the detected vehicles are matched to make a trajectory.

This will give the trajectory in some distinct parts. In some time instants namely vehicles are not detected properly which means there are many “partial” trajectories which have to be coupled in the next step. Of trajectories which end at another location than at the edge of the image have, a majority has to be coupled to another partial trajectory; only if the vehicle changes lane it disappears from the image. The coupling of this partial trajectory works as follows. Each partial trajectory is extrapolated, based on the position and speed slightly before it ends. In this process, it is better not to use the very last detected position since the measurement gets noisy at the last measurement before it is not detected anymore. This extrapolated trajectory is compared with every starting partial trajectory. If one of them matches in space and speed, both partial trajectories are coupled. The maximum allowed error in this matching process can be chosen. If we choose this maximum high, more partial trajectories are coupled, but this increases the probability of coupling partial trajectories which do not belong to the same vehicle, whereas a lower threshold leaves some partial trajectories uncoupled, but there are fewer errors in the coupling. We choose for a relatively strict option.

This choice means that several partial trajectories remain uncoupled: on average, each vehicle trajectory now consists of 2 partial trajectories. This is now corrected manually. To this end, the partial trajectories and a unique number identifying the partial trajectory are plotted (figure A.5). In an Excel sheet we now indicate which partial trajectories should be combined. This takes around 10 minutes for a video of 5 minutes in length. From this point, the further processing is automatic.

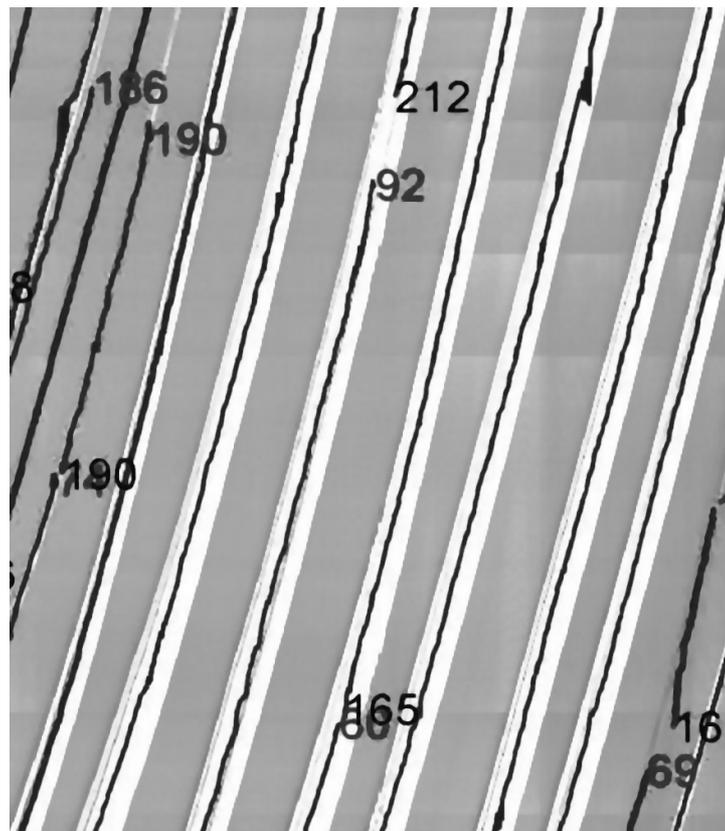
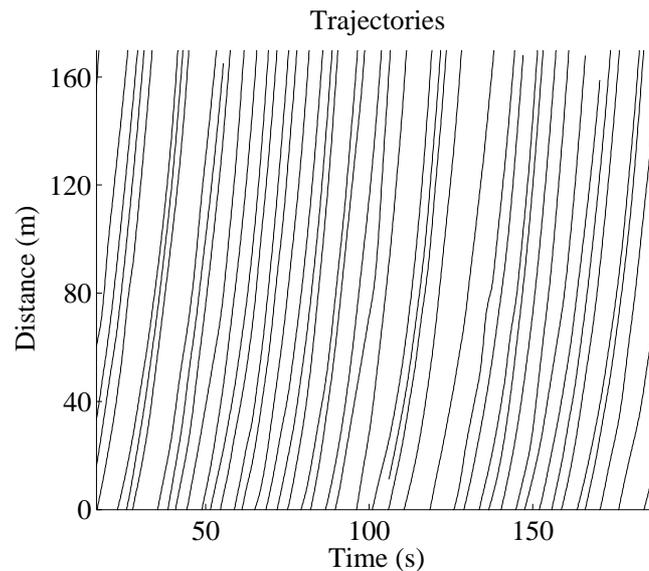


Figure A.5: The trajectories found on top of the image of the longitudinal cross-section



**Figure A.6: Trajectories for several vehicles**

### A.4.2 Smoothing

The procedure described in the last subsections gives us for each vehicle the trajectory. Nevertheless, these trajectories are not smooth yet, as shows. This is mainly caused by the spatial and temporal resolution of the images, but also by detecting and tracking errors. To represent the reality, they need to be smoothed, without smoothing out the dynamics of traffic. To this end, the method of Toledo et al. (2007) is applied. This method averages the position at a time step with the position at neighboring time steps, with a larger weight for time steps which are closer by. The smoothing filter takes time frames into account with times that differ at maximum 1.5 seconds from the current time frame. An example of the end result, the trajectories, can be found in figure A.6.

## A.5 Conclusions

Video data from traffic operations contain much valuable information. However, extracting, for instance trajectory data using existing methods is a difficult and often erroneous process. The approach proposed here increases the length of the parts of the trajectories, thereby overcoming the shortcomings of existing methods. To this end, instead of analysing each picture in a video sequence separately, the data of the video is transformed in such a way that the vehicle trajectories (vehicle positions over time) become visible within one single image.

For instance, the cross-section of the three-dimensional box along the axis of the road provides insight in a quick and easy manner into the dynamic car-following behavior. Within a few minutes, the traces of the cars, i.e. the vehicle trajectories, are plotted in a

space-time diagram. Then, from the constructed image, many dynamic properties of the traffic are apparent, such as shockwaves, reaction times, and driver over-reaction. Using image detection techniques, the vehicle positions can be quantified. This is a very reliable method of constructing trajectories.

In other methods, the vehicle trajectories are often divided into many different segments. In the presented method, the length of the segments can be longer, as well as the segments are more easily coupled. Close to 95% of the vehicle presence (in the space-time frame) is detected, while the remaining 5% can be manually reconstructed very quickly. The complete process consisting of transforming, detecting, tracking, and manually correcting of one single lane can be performed in approximately 1 hour for a 5-minute dataset using a PC equipped with 1.5 GHz core processor.

Another cross-section of the three-dimensional box, perpendicular to the driving direction, shows other traffic properties. Headways and occupancies can namely be derived directly from this single image. For these purposes, it is not necessary to detect all vehicles in all pictures, but only the vehicles presently passing the location of the virtual detector.

In our analyses, black and white images were used. For the case of using color images, the contrast between vehicles and pavement increases, making it easier to detect these vehicles. The tracking process also benefits from color since vehicles detected in consecutive pictures can then be identified and linked by color. Thus, in the future, the use of color images is expected to further increase the length of the automatically discerned vehicle trajectories.



# Appendix B

## Statistics

This appendix explains briefly the statistic measures that are used in this thesis. A complete description can be found in Chakravarti et al. (1967); this appendix only gives a brief explanation of the interpretation of the concepts.

- $R^2$  indicates how two stochastic variables relate. The value of  $R^2$  indicates which part of the variation in one variable ( $y$ ) can be explained by a variation in the other,  $x$ .  $R^2$  is the square of the Pearson square, indicated with  $\mathbf{R}$ , which is calculated as follows:

$$R = \frac{\sum (x - \bar{x})(y - \bar{y})}{s_x s_y}. \quad (\text{B.1})$$

In this equation,  $\bar{x}$  and  $\bar{y}$  indicate the mean values,  $s_x$  and  $s_y$  indicate the standard deviations, and  $n$  is the size of the sample.

- A  **$P$  value** indicates the probability of an event more extreme than the current draw of a statistic. Since it is a probability, the  $P$  value is always between 0 and 1. Depending on the desired significance level, the  $P$  value is chosen. Typically, values of 0.01 or 0.05, indicating that 99% or 95% of the cases would be less extreme are chosen as threshold to decide whether an event is significant.
- A two-sided **t-test** is used to test whether there are significant differences in mean value between two series of numbers, indicated by  $x$  and  $y$ , consisting of  $n$  and  $m$  elements. The T-value is computed using the following equation:

$$T = \frac{\bar{x} - \bar{y}}{s \sqrt{\frac{1}{n} + \frac{1}{m}}}. \quad (\text{B.2})$$

In this equation,  $s$  is the pooled sample standard deviation, which, in turn is a weighted average of the standard deviations in  $x$  and  $y$ , computed by the following equation:

$$s^2 = \frac{(n-1)s_x^2 + (m-1)s_y^2}{n+m-2}. \quad (\text{B.3})$$

In this equation,  $s_x$  and  $s_y$  are the standard variations of  $x$  and  $y$ . The outcome of the t-test shows how much pooled standard deviations the means of the two stochastic

errors are apart. This can be translated into with a  $P$  value.

- A **z-test** is the continuous variant of the t-test. Rather than using a discrete equation, both samples are assumed to have a normal distribution. Again, the outcome will show how much standard deviations the means are apart.

# Summary

## Road Incidents and Network Dynamics Effects on driving behaviour and traffic congestion

Incidents are responsible for large number of traffic jams and the consequent delays are a large part of the total delay. Moreover, these delays are unexpected and therefore the associated discomfort is higher. It is therefore worthwhile studying the effects of incidents on driving behaviour and the delay.

The main topic of this thesis is traffic flow after an incident. This is discussed at two levels: both at the microscopic level and at the macroscopic level. The change in individual driving behaviour is analysed using trajectory data for two incident locations. It turned out that the reaction time of drivers increases. There is a group of drivers which have a slightly increased reaction time (approximately 2 seconds) and a group of drivers with a very large reaction time (approximately 5 seconds). Such a bimodal distribution is also found for the headways at the bottleneck location. For the two incident situations that are recorded on video, the capacity per available lane turned out to be 30-50% lower than under normal conditions. This finding was consistent with loop detector data collected at 90 other incident sites. Most remarkable is that the reduction of flow in the opposite direction on the motorway, on the other side of the guard rail, is similar in magnitude. If the road authority succeeded in reducing the distraction, this could considerably reduce the total congestion. Currently, the Dutch road authority has a test to set up screens around the incident site which prevents the other road users from looking at the incident site. By this measure they try to reduce the distraction in order to increase the capacity.

Also the change of route choice was studied for five incident cases. Using traffic counts the split fractions towards good, sometimes even advised, alternative routes were computed. These were compared with the split fractions at the same points for comparable days without an incident. There were two incidents with a high remaining capacity and small delays. For these incidents, the fraction of people that took another route was negligible. There were also three incidents with a low remaining capacity and large delays; in one case, the road was completely blocked. In these cases, typically 30-50% of the drivers changed route.

In a traffic simulator the total delay of an incident is calculated. That shows that if more travellers take another route when an incident occurs, the delay for all travellers together

will decrease much. Therefore, the road authority can improve the “network robustness” (the way the network copes with a short-term, strong decrease in capacity or a short-term strong increase of demand) by advising travellers of alternative routes, for instance by dynamic route information panels.

Using the simulator and the real fraction of travellers changing their route, it was investigated at which location in a traffic network a capacity reduction results in the largest reduction in network performance. These locations are called “vulnerable” network parts. This analysis is carried out with a computer simulation which dynamically predicts the traffic flows. It turned out that the most vulnerable links are (1) main motorways and (2) links leading traffic away from the main motorways. If one of these links is blocked, the traffic jam grows until the tail reaches a main motorway, which will give large delays as well. A special simulator was developed to quantify these spillback effects which showed that the effects of blocking back or spillback play an important role. In fact, simulators without proper spillback modelling are not suitable for carrying out a valid vulnerability analysis. In practice, this conclusion means that it can be beneficial to separate traffic streams, for instance separate through traffic and the local traffic on a ring road by a barrier. This way, the situation can be avoided where a queue originating in the city centre hinders the through traffic. Alternatively, dedicated buffer zones could be considered to store the queuing vehicles.

Particularly when there is a risk of spillback, but also in non-spillback situations, the duration of an incident plays an important role in the total delay. A quick incident recovery can therefore reduce the total delay considerably.

The simulator indicates which links are vulnerable, but it is computationally expensive to analyse all the links. It would be useful if there are indicators which can be computed from one run of the simulator or, even better, they can be computed from the traffic flow values from measurements in real-life. However, a comparison of indicators available in the literature showed that the predictive value of these indicators is very poor. Also a linear model of all these indicators would not give any sensible results.

A completely different approach to finding vulnerable links is to analyse which links drivers can easily avoid (for instance, if they expect a blockage) and which ones are hard to avoid. This approach consists of a game-theoretical approach applying a risk-averse traffic assignment. Since dynamic queuing effects, like spillback, are shown to be important, a recipe was developed to combine risk-averse traffic assignment with a dynamic traffic simulator with proper spillback. It shows that the risk-averseness of drivers plays a major role in the distribution of traffic over different routes. The next step to be taken now is to test this recipe in practice and calibrate the risk-averseness for travellers in a real-life network.

Victor L. Knoop, 2 December 2009

# Samenvatting

## Incidenten op de weg en netwerkdynamiek Effecten op rijgedrag en congestie

Een groot deel van de files en vertragingen op de weg wordt veroorzaakt door incidenten. Deze vertragingen zijn vaak vervelend omdat deze onverwacht zijn. Daarom is het nuttig de effecten van incidenten op weggedrag en vertraging te bestuderen.

Het onderwerp van deze dissertatie is de verkeersstroom na incidenten. Het effect wordt op twee niveaus behandeld: zowel op microscopisch niveau als op macroscopisch niveau. De verandering van individueel rijgedrag wordt geanalyseerd aan de hand van trajectorie-data bij twee ongevalslocaties. Daaruit blijkt dat de reactietijd van weggebruikers hoger is dan in normale omstandigheden. Een deel van de bestuurders heeft een reactietijd die slechts iets langer is dan normaal (ongeveer 2 seconde), maar een andere groep heeft een reactietijd van ongeveer 5 seconde – veel langer dan normaal. De volgtijden bij de bottleneck zijn verdeeld in twee groepen, een groep met een grote volgtijd en een groep met een kleine volgtijd. Op twee ongevalslocaties waarvan video-opnamen zijn gemaakt, was de wegcapaciteit per beschikbare strook 30 tot 50 procent lager dan in normale omstandigheden zonder ongeluk. Deze bevinding is in lijn met de analyses uitgevoerd met tellussen in de weg voor 90 andere incidenten. Het meest opvallend is dat er een even grote reductie is in de maximale stroom verkeer in de tegengestelde richting, aan de andere kant van vangrail. Het is verstandig dat de wegbeheerder de afleiding voor deze mensen probeert weg te nemen, zoals Rijkswaterstaat dat nu doet met verplaatsbare “kijkschermen” die opgezet worden als er een ongeluk is.

Ook is de routekeus bestudeerd voor vijf ongelukssituaties. Met behulp van tellussen in de weg is bepaald welk deel van de weggebruikers afslaat naar een goede alternatieve route, zoals die enkele gevallen geadviseerd werd. Het aantal weggebruikers dat een bepaalde kant op gaat als deel van het geheel heel de splitfractie. De splitfractie tijdens een ongeluk bij het begin van een alternatieve route is vergeleken met de splitfractie op vergelijkbare dagen zonder ongeval. Bij twee ongelukken waar de capaciteit op de plek van het ongeluk nog hoog was, was het aantal weggebruikers dat een alternatieve route zocht verwaarloosbaar. Bij de andere drie ongelukken was de wegcapaciteit veel lager en waren er grote vertragingen; in een geval was de weg zelfs geheel afgesloten. Bij deze ongelukken veranderde 30 tot 50 procent van de bestuurders hun route.

In een verkeerssimulatieprogramma is de vertraging van een incident berekend. Deze berekening liet zien dat de robuustheid van een netwerk groter is als meer mensen een alternatieve route nemen tijdens een ongeluk. Daarom kunnen wegbeheerders de robuustheid vergroten door weggebruikers alternatieve routes te adviseren, bijvoorbeeld door alternatieve routes dynamische route-informatiepanelen te tonen.

Met behulp van het verkeerssimulatieprogramma en het gegeven welk percentage bestuurders omrijdt bij een ongeluk, was het mogelijk te onderzoeken op welke plekken in een netwerk een capaciteitsbeperking de grootste totale vertraging oplevert. Deze locaties heten “kwetsbare” delen. De analyse is uitgevoerd met een computersimulatieprogramma dat de verkeersstromen dynamisch doorrekent. Het bleek dat de meest kwetsbare weggedelen autosnelwegen waren en wegen waarop verkeer weggrijdt van de autosnelweg. Als er een blokkade is op een van deze laatste weggedelen, groeit de file van achteren aan en bereikt de staart van de file de autosnelweg die dan geblokkeerd raakt. Dit heeft vervolgens veel vertraging tot gevolg. Een speciaal verkeerssimulatieprogramma is ontwikkeld om deze fileterugslageffecten te kunnen kwantificeren. Daaruit bleek dat fileterugslag een belangrijke rol speelde in de totale vertraging. Voor analyses betekent dit dat programma's zonder een goede modellering van fileterugslag onbruikbaar zijn om kwetsbaarheid van een wegennetwerk te bepalen. Voor de praktijk betekent dit dat het nuttig kan zijn verkeersstromen te scheiden, bijvoorbeeld doorgaand verkeer op een rondweg fysiek scheiden van lokaal verkeer. Op deze manier heeft een file die in de stad ontstaat en tot de snelweg groeit geen invloed op het doorgaande verkeer op de autosnelweg. Een alternatieve oplossing is het creëren van bufferzones. Dit is extra ruimte op de weg waar auto's die vanwege een stroomafwaarts gelegen bottleneck niet kunnen doorstromen kunnen wachten zonder daarbij andere stromen auto's te hinderen.

In alle gevallen, maar vooral als er een risico op fileterugslag bestaat, is het belangrijk hoe lang een ongeluk de weg blokkeert. Een snelle afhandeling van een ongeluk kan daarom de totale vertraging sterk verminderen.

De verkeerssimulator toont ook welke weggedelen kwetsbaar zijn. Het duurt alleen lang voordat blokkades op alle mogelijke links geanalyseerd zijn. Daarom zou het handig zijn als er een betrouwbare kwetsbaarheids-indicator was die bepaald kan worden met één run van het simulatieprogramma of, beter nog, die bepaald kan worden met meetgegevens van de weg. Een vergelijking van beschikbare indicatoren gaf echter aan dat de voorspellende waarde van de huidige indicatoren erg slecht is. Ook een lineaire combinatie van (een deel van) deze indicatoren geeft geen bruikbare resultaten.

Een andere benadering om kwetsbare weggedelen te vinden is te bepalen welke weggedelen bestuurders makkelijk kunnen vermijden (bijvoorbeeld als ze een blokkade verwachten) en welke delen moeilijk te vermijden zijn. Dit is een speltheoretische benadering waarbij het verkeer volgens risicomijdende principes wordt toegedeeld. Omdat is aangetoond dat dynamische effecten van een rij wachtende auto's, zoals fileterugslag, belangrijk zijn, is een methode ontwikkeld waarbij verkeer risicomijdend kan worden toegedeeld in een verkeerssimulatieprogramma met een correcte representatie van fileterugslag. Een toepassing van deze methode toont aan dat de risico-attitude van automobilisten een belangrijke rol speelt in de verdeling van het verkeer over de verschillende routes. De volgende stap

is deze methode in de praktijk te testen en de risico-attitude van reizigers te kalibreren in een echt netwerk.

Victor L. Knoop, 2 december 2009



# Curriculum vitae



Victor L. Knoop was born on 10 July 1981 in Arnhem (the Netherlands). In 1999 he graduated from his high school, “Stedelijk Gymnasium Arnhem”. He then started two major studies at the science faculty of the Leiden University, physics and mathematics. After he finished successfully his foundation year for both studies, he continued with a master in physics. As part of his study he spent 10 month at the École Polytechnique Fédérale de Lausanne in Switzerland. There he carried out a research project on fluid dynamics. The last part of his master studies consisted of an internship. He worked on a project of modelling the noise of traffic using microscopic traffic simulation at TNO, the Netherlands Organisation for Applied Scientific Research.

During his studies, Victor participated in the academic community of the Leiden University. He chaired for instance the association for students in Physics, Mathematics, Astronomy and Computer Science, “De Leidsche Flesch” from February 2001 to April 2002. From September 2002 to September 2003 he was member of the executive board of the faculty of science of the Leiden University.

In May 2005 Victor started his PhD research at the Transport and Planning department at the Delft University of Technology. Next to his PhD research, he worked on several projects. As part of his affiliation at the TU Delft, Victor joined Arane traffic consultancy for 3 months in 2006. There, he worked on a evaluation of dynamic traffic signs. He also worked on several traffic projects for the Dutch Road Authority and the Port of Rotterdam.

Victor taught in the master course on Traffic Flow Theory and Simulation in 2009. During his PhD he also supervised several master students during their master thesis.

Victor carried out a part of his PhD research at Imperial College, London. From October 2007 to January 2008 he developed a method for risk-averse traffic assignment. For this stay Victor was awarded a research grant from the IDEA league, a network of leading European universities of science and technology.

Victor actively participated in the research school TRAIL. He chaired for instance the PhD council from March 2007 until now. Victor was also active in the society outside the academia. He joined the social liberal political youth party “Jonge Democraten” in 2004. There, he was executive board member of the international committee from June 2007 to October 2008. He chaired this committee from June to October 2007.

Victor is affiliated at the Delft University of Technology where he continues working on traffic flow in the Transport and Planning department. He was recently awarded a personal (Rubicon) grant from the Netherlands Organisation for Scientific Research (NWO) for a research proposal entitled: “How Traffic Jams Start”. He will carry out this research project at the École Nationale des Travaux Publics de l’État in Lyon.

# Author's publications

## Journal Articles

Knoop, V.L., Hoogendoorn, S.P. and Van Zuylen, H.J. (in press) Processing Traffic Data collected by Remote Sensing. *Transportation Research Records*

Knoop, V.L., Hoogendoorn, S.P. and Adams, K. (in press) Incidents at Motorways: Capacity Reductions and the Effects of Incident Management. *European Journal of Transportation and Infrastructure Research*

Hoogendoorn, S.P., Landman, R., Knoop, V.L., Pel, A.J., Huibregtse, O. and Hoogendoorn, R. (2009) Verkeersmanagement bij exceptionele omstandigheden (Traffic Management at Exceptional Conditions). *NM Magazine*, Vol. 2, pp. 24-28.

Hoogendoorn, S.P., Knoop, V.L. and Van Zuylen, H.J. (2008) Robust Control of Traffic Networks under Uncertain Conditions. *Journal of Advanced Transportation*, Vol. 42, no. 3, pp. 357-377.

Hoogendoorn, S.P., Van Lint, J.W.C. and Knoop, V.L. (2008) Macroscopic Modeling Framework Unifying Kinematic Wave Modeling and Three-Phase Traffic Theory. *Transportation Research Record: Journal of the Transportation Research Board*, No. 2088, pp. 102-108.

Knoop, V.L., Hoogendoorn, S.P. and Van Zuylen, H.J. (2008) Capacity Reduction at Incidents: Empirical Data Collected from a Helicopter. *Transportation Research Record: Journal of the Transportation Research Board*, No. 2071, pp. 19-25.

Knoop, V.L., Hoogendoorn, S.P. and Van Zuylen, H.J. (2008) The Influence of Spillback Modelling when Assessing Consequences of Blockings in a Road Network *European Journal of Transportation and Infrastructure Research*, Vol. 8, no. 4, pp. 287-300.

## Peer Reviewed Conference Proceedings

Knoop, V.L., Hoogendoorn, S.P. and Van Zuylen, H.J. (in press) The Influence of Variable Incident Duration on Delay. *Proceedings of Traffic and Granular Flow 09*. Springer.

Knoop, V.L., Hoogendoorn, S.P., Snelder, M. and Van Zuylen, H.J. (2010) Reliability of link-based vulnerability indicators, in: *Proceedings of the 89th Annual Meeting of the Transportation Research Board*, 10-14 January 2010, Washington D.C. Accepted for publication.

Knoop, V.L., Hoogendoorn, S.P. and Van Zuylen, H.J. (2010) Stochastic Incident Duration - Impacts on Delay, in: *Proceedings of the 89th Annual Meeting of the Transportation*

*Research Board, Washington D.C.*, 10-14 January 2010, Washington D.C. Accepted for publication.

Knoop, V.L., Hoogendoorn, S.P. and Van Zuylen, H.J. (2009) Route Choice Under Exceptional Traffic Conditions, in: *International Conference on Evacuation Management*, 23-25 September 2009, The Hague, the Netherlands.

Hoogendoorn, S. P., Van Lint, J. W. C. and Knoop, V. L. (2009) Dynamic first-order modeling of phase-transition probabilities. In Appert-Rolland, C., Chevoir, F., Gondret, P., Lassarre, S., Lebacque, J.-P. and Schreckenberg, M. (Eds.) *Proceedings of Traffic and Granular Flow 07*. p. 85-92, Springer, New York.

Knoop, V.L., Hoogendoorn, S.P. and Van Zuylen, H.J. (2009) Empirical Differences between Time Mean Speed and Space Mean Speed. In Appert-Rolland, C., Chevoir, F., Gondret, P., Lassarre, S., Lebacque, J.-P. and Schreckenberg, M. (Eds.) *Proceedings of Traffic and Granular Flow 07*. p. 351-356, Springer, New York.

Knoop, V.L., Van Zuylen, H.J. and Hoogendoorn, S.P. (2009) Microscopic Traffic Behaviour near Accidents. IN Lam, W.H.K., Wong, S.C. and Lo, H.K. (Eds.) *Proceedings of the 18th International Symposium of Transportation and Traffic Theory*. 16-18 June, Hong Kong, China. Springer, New York.

Knoop, V.L., Hoogendoorn, S.P. and Van Zuylen, H.J. (2009) Traffic behavior near incidents, in: *Proceedings of Sino-Dutch seminar on transport development*, 30 June 2009, Beijing, China.

Knoop, V.L., Van Zuylen, H.J. and Hoogendoorn, S.P. (2009) Microscopic traffic behaviour near Incidents, in: *Proceedings of Second Sino-Dutch Joint Workshop in Transportation and Traffic Study*, 26 June 2009, Shanghai, China.

Adams, K., V.L.Knoop, Hoogendoorn, S.P., Stoop, J.A.A.M. and Van Loon, A. (2009) Safety Impacts of Incident Management at Incident Sites, in: *Proceedings of the 88th Annual Meeting of the Transportation Research Board*, 11-15 January 2009, Washington D.C.

Knoop, V.L., Hoogendoorn, S.P. and Van Zuylen, H.J. (2009) Processing Traffic Data collected by Remote Sensing, in: *Proceedings of the 88th Annual Meeting of the Transportation Research Board*, 11-15 January 2009, Washington D.C.

Knoop, V.L., Bell, M. G. H. and Van Zuylen, H.J. (2008) Route-Advice Based On Individual Risk-Attitude in: *Proceedings of IEEE Conference on Infrastructure Systems - Building Networks for a Better Future*, 10-12 November 2008, Rotterdam, the Netherlands.

Hoogendoorn, S.P., Van Lint, J.W.C. and Knoop, V.L. (2008) Modeling Traffic Phase Transitions, in: *Proceedings of CHAOS2008 - Chaotic Modeling and Simulation International Conference*, 3-6 June 2008, Chania Crete Greece.

Knoop, V.L., Hoogendoorn, S.P. and Van Zuylen, H.J. (2008) Capacity Reduction at Incidents: Empirical Data Collected from a Helicopter, in: *Proceedings of 87th Annual Meeting of the Transportation Research Board*, 13-17 January 2008, Washington D.C.

Knoop, V.L., Hoogendoorn, S.P. and Van Zuylen, H.J. (2007) Approach to Critical Link Analysis of Robustness for Dynamical Road Networks. IN Schadschneider, A., Pöschel, T., Kühne, R., Schreckenberg, M. and Wolf, D. E. (Eds.) *Proceedings of Traffic and Granular Flow 05*. Berlin, Springer Verlag.

Hoogendoorn, S.P., Knoop, V.L. and Van Zuylen, H.J. (2007) Model-based Stochastic Control of Traffic Networks, in: *Proceedings of the Third International Symposium on Transportation Network Reliability*, 19-20 July 2007, Den Haag, the Netherlands.

Knoop, V.L., Snelder, M. and Van Zuylen, H.J. (2007) Comparison of Link-Level Robustness Indicators, in: *Proceedings of the Third International Symposium on Transportation Network Reliability (INSTR)*, 19-20 July 2007, Den Haag, the Netherlands.

Knoop, V.L. and Hoogendoorn, S.P. (2007) Need for realistic spillback modeling in assessing network reliability, in: *Proceedings of Tristan VI*, 10-15 June 2007, Phuket, Thailand.

Knoop, V.L., Hoogendoorn, S.P. and Van Zuylen, H.J. (2007) Quantification of the impact of spillback modeling in assessing network reliability, in: *Proceedings of 86th annual meeting of the Transportation Research Board*, 21-25 January 2007 Washington D.C.

Knoop, V.L., Hoogendoorn, S.P. and Van Zuylen, H.J. (2006) Network robustness assessment: the need for spillback modelling, in: *Proceedings of 9th TRAIL Congress, TRAIL in motion*, 21 November 2006, Rotterdam, the Netherlands.

Knoop, V.L., Hoogendoorn, S.P. and Van Zuylen, H.J. (2006) Network robustness assessment: the need for spillback modelling. IN ZUYLEN, H.J. V. (Ed.) *Selected Papers of the 9th TRAIL Congress, TRAIL in motion*. Rotterdam, TRAIL Research School.

Knoop, V.L., Hoogendoorn, S.P. and Van Zuylen, H.J. (2006) Importance of spillback modeling in assessing ITS, in: *Proceedings of IEEE 9th International Conference on Intelligent Transport Systems*, 17-20 September 2006, Toronto, Canada.

Zurbier, F.S., Van Lint, J.W.C. and Knoop, V.L. (2006) Traffic network state estimation using extended Kalman filtering and DSMART, in: *Proceedings of Eleventh IFAC Symposium on Control in Transportation Systems*, 29-31 August 2006, Delft, the Netherlands.

### **Reports**

Knoop, V.L., F. Viti, H. Tu, E. Jagtman, Van Zuylen, H.J. and Smit, H. (2007) Accessibility of the harbour. Delft, Delft University of Technology.



# TRAIL Thesis Series

The following list contains the most recent dissertations in the TRAIL series. See for a complete overview the TRAIL website: [www.rsTRAIL.nl](http://www.rsTRAIL.nl).

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

Knoop, V.L., *Road Incidents and Network Dynamics: Effects on driving behaviour and traffic congestion*, T2009/13, December 2009, TRAIL Thesis Series, the Netherlands

Baskar, L.D., *Traffic Control and Management with Intelligent Vehicle Highway Systems*, T2009/12, November 2009, TRAIL Thesis Series, the Netherlands

Konings, J.W., *Intermodal Barge Transport: Network Design, Nodes and Competitiveness*, T2009/11, November 2009, TRAIL Thesis Series, the Netherlands

Kusumaningtyas, I., *Mind Your Step: Exploring aspects in the application of long accelerating moving walkways*, T2009/10, October 2009, TRAIL Thesis Series, the Netherlands

Gong, Y., *Stochastic Modelling and Analysis of Warehouse Operations*, T2009/9, September 2009, TRAIL Thesis Series, the Netherlands

Eddia, S., *Transport Policy Implementation and Outcomes: the Case of Yaounde in the 1990s*, T2009/8, September 2009, TRAIL Thesis Series, the Netherlands

Platz, T.E., *The Efficient Integration of Inland Shipping into Continental Intermodal Transport Chains. Measures and decisive factors*, T2009/7, August 2009, TRAIL Thesis Series, the Netherlands

Tahmasseby, S., *Reliability in Urban Public Transport Network Assessment and Design*, T2009/6, June 2009, TRAIL Thesis Series, the Netherlands

Bogers, E.A.I., *Traffic Information and Learning in Day-to-day Route Choice*, T2009/5, June 2009, TRAIL Thesis Series, the Netherlands

Amelsfort, D.H. van, *Behavioural Responses and Network Effects of Time-varying Road Pricing*, T2009/4, May 2009, TRAIL Thesis Series, the Netherlands

Li, H., *Reliability-based Dynamic Network Design with Stochastic Networks*, T2009/3, May 2009, TRAIL Thesis Series, the Netherlands

Stankova, K., *On Stackelberg and Inverse Stackelberg Games & their Applications in the Optimal Toll Design Problem, the Energy Markets Liberalization Problem, and in the Theory of Incentives*, T2009/2, February 2009, TRAIL Thesis Series, the Netherlands

Li, T., *Informedness and Customer-Centric Revenue*, T2009/1, January 2009, TRAIL Thesis Series, the Netherlands

Agusdinata, D.B., *Exploratory Modeling and Analysis. A promising method to deal with deep uncertainty*, T2008/17, December 2008, TRAIL Thesis Series, the Netherlands

Kreutzberger, E., *The Innovation of Intermodal Rail Freight Bundling Networks in Europe. Concepts, Developments, Performances*, T2008/16, December 2008, TRAIL Thesis Series, the Netherlands

Taale, H., *Integrated Anticipatory Control of Road Networks. A game theoretical approach*, T2008/15, December 2008, TRAIL Thesis Series, the Netherlands

Li, M., *Robustness Analysis for Road Networks. A framework with combined DTA models*, T2008/14, December 2008, TRAIL Thesis Series, the Netherlands

Yu, M., *Enhancing Warehouse Performance by Efficient Order Picking*, T2008/13, October 2008, TRAIL Thesis Series, the Netherlands

Liu, H., *Travel Time Prediction for Urban Networks*, T2008/12, October 2008, TRAIL Thesis Series, the Netherlands

Kaa, E.J. van de, *Extended Prospect Theory. Findings on Choice Behaviour from Economics and the Behavioural Sciences and their Relevance for Travel Behaviour*, T2008/11, October 2008, TRAIL Thesis Series, the Netherlands

Nijland, H., *Theory and Practice of the Assessment and Valuation of Noise from Roads and Railroads in Europe*, T2008/10, September 2008, TRAIL Thesis Series, the Netherlands

Annema, J.A., *The Practice of Forward-Looking Transport Policy Assessment Studies*, T2008/9, September 2008, TRAIL Thesis Series, the Netherlands

Ossen, S.J.L., *Theory and Empirics of Longitudinal Driving Behavior*, T2008/8, September 2008, TRAIL Thesis Series, the Netherlands

Tu, H., *Monitoring Travel Time Reliability on Freeways*, T2008/7, April 2008, TRAIL Thesis Series, the Netherlands

D'Ariano, A., *Improving Real-Time Train Dispatching: Models, Algorithms and Applications*, T2008/6, April 2008, TRAIL Thesis Series, the Netherlands

Quak, H.J., *Sustainability of Urban Freight Transport. Retail Distribution and Local Regulations in Cities*, T2008/5, March 2008, TRAIL Thesis Series, the Netherlands