

Traffic Flow at Sags: Theory, Modeling and Control

Bernat Goñi-Ros

Delft University of Technology

This thesis is one of the outcomes of a research project funded by Toyota Motor Europe.



Cover illustration: Pedro Simón

Traffic Flow at Sags: Theory, Modeling and Control

Proefschrift

ter verkrijging van de graad van doctor

aan de Technische Universiteit Delft,

op gezag van de Rector Magnificus prof. ir. K.C.A.M. Luyben,

voorzitter van het College voor Promoties,

in het openbaar te verdedigen op

maandag 21 maart 2016 om 10:00 uur

door

Bernat GOÑI-ROS

M.Sc. Transport, Infrastructure and Logistics

Technische Universiteit Delft

geboren te Barcelona, Catalonië, Spanje

This dissertation has been approved by the
promoters: Prof. dr. ir. B. van Arem and Prof. dr. ir. S.P. Hoogendoorn
copromotor: Dr. V.L. Knoop

Composition of the doctoral committee:

Rector Magnificus	Chairman
Prof. dr. ir. B. van Arem	Delft University of Technology
Prof. dr. ir. S.P. Hoogendoorn	Delft University of Technology
Dr. V.L. Knoop	Delft University of Technology

Independent members:

Prof. dr. ir. J. Hellendoorn	Delft University of Technology
Prof. Dr.-Ing. M. Papageorgiou	Technical University of Crete, Greece
Prof. D. Eng. H. Kita	Kobe University, Japan
Dr. C. Buisson	University of Lyon, France
Prof. dr. ir. J.W.C. van Lint	Delft University of Technology, reserve member

TRAIL Thesis Series no. T2016/2, the Netherlands TRAIL Research School

TRAIL
P.O. Box 5017
2600 GA Delft
The Netherlands
Phone: +31 (0) 15 278 6046
E-mail: info@rsTRAIL.nl

ISBN 978-90-5584-198-1

Copyright © 2016 by Bernat Goñi-Ros.

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

Printed in the Netherlands

Time passes unhindered. All we can do is use the present well.
- Tenzin Gyatso, Dalai Lama

Acknowledgments

This dissertation is the main outcome of a research project that I conducted during the last four and a half years at TU Delft. The dissertation would not have been possible without my willingness to explore, learn and develop my own ideas; however, I am aware that I have not done this alone. Many people helped and supported me during this process. I would like to take the opportunity to thank them all most sincerely.

First of all, I would like to thank my supervisors – Bart, Serge and Victor – for giving me the opportunity to pursue a PhD and for guiding me through the process. It was an intense and rewarding journey. Special thanks to Victor for his (contagious) enthusiasm and for keeping the door of his office open to discuss any issue at any time.

I would also like to thank Takahashi san, Kitahama san and Sakata san (Toyota Motor Corporation) for always looking at my work with a critical eye and for making me think constantly about the practical implications of my research. Working together has been an honor.

Many thanks to all the doctoral students of the Department of Transport & Planning for the interesting discussions we had and for the moments we shared. One of the best things about doing a PhD is that it gives one the opportunity to interact with and learn from bright people every day.

I would also like to thank the independent members of my doctoral committee and the anonymous reviewers of my papers for their valuable comments and suggestions. Revising a manuscript based on the comments of the reviewers is not an easy task, but it usually improves the quality of the manuscript substantially.

During the fourth year of my PhD, I had some a medical problem. I would like to thank the IRITEB team (Barcelona), particularly Dr. Minoves and Mage, for helping me to recover. I always felt I was in good hands. I also would like to acknowledge and thank Paula Meesters (TU Delft) for providing counseling when I needed it.

Many thanks to Michael Cassidy, Anthony Patire and Robert Bertini for helping to organize my visit to UC Berkeley in the first year of my PhD. The visit was very fruitful from a scientific point of view and was a great overall experience.

Also, many thanks to the technical and administrative staff of the Department of Transport & Planning and TRAIL Research School for taking care of many practical issues, which allowed me to focus on my research project.

I would also like to thank Iván Muñiz (Autonomous University of Barcelona), Oriol Homs, Núria Benach (University of Barcelona), Lourdes Benería (Cornell University) and Rob van Nes (TU Delft) for encouraging me to pursue a PhD at various moments during the past years. There is a time for everything.

I would like to thank Antonio and Simeon – my two paranymphs – for their friendship and for their support during the thesis writing process, and Olga for translating the summary of my thesis into Dutch. Many thanks also to all my friends in Delft for the good times we had during the past years, particularly to the members of the pop-rock band The Trains.

Special thanks go to my classmates in the Department of City and Regional Planning at Cornell University (New York). Their willingness to live fully and to make a difference in the world remains a source of inspiration to me.

Finally, I would like to thank my parents for all the time and energy they devoted to my education and wellbeing, my sister for being the person she is, and Nadjla for her deep love and patience.

Contents

List of figures	ix
List of tables	xi
1 Introduction	1
1.1 Freeway traffic dynamics, modeling and control	1
1.1.1 Freeway traffic dynamics	1
1.1.2 Freeway traffic modeling	3
1.1.3 Freeway traffic control	5
1.2 Traffic dynamics, modeling and control at sags	6
1.2.1 Traffic dynamics at sags	7
1.2.2 Traffic modeling at sags	7
1.2.3 Traffic control at sags	8
1.3 Challenges with regard to the development of dynamic traffic manage- ment measures for sags	9
1.4 Research objective and research questions	10
1.5 Research approach	10
1.6 Research scope	11
1.7 Main contributions	12
1.7.1 Scientific contributions	12
1.7.2 Contributions to practice	13
1.8 Outline of the dissertation	14

2 Empirical analysis of the causes of stop-and-go waves at sags	17
Abstract	17
2.1 Introduction	18
2.2 Background	19
2.3 Data characteristics	21
2.3.1 Study site	21
2.3.2 Vehicle trajectories	21
2.4 Data analysis methods	22
2.4.1 Calculation of vehicle speeds	22
2.4.2 Identification of the causes of formation and growth of stop-and-go waves	23
2.4.3 Test of statistical significance	26
2.5 Results	27
2.5.1 Identification of stop-and-go waves	27
2.5.2 Causes of stop-and-go wave formation	28
2.5.3 Causes of stop-and-go wave growth	28
2.6 Discussion	28
2.7 Conclusions	31
3 Modeling traffic at sags	33
Abstract	33
3.1 Introduction	33
3.2 Causes of congestion at sags	34
3.3 Model requirements and suitability of existing models	35
3.4 Microscopic traffic model	37
3.4.1 Car-following model	37
3.4.2 Lane change model	39
3.4.3 Model inputs	39
3.5 Model verification methodology	40
3.5.1 Simulation setup	40
3.5.2 Sensitivity analysis	43

3.5.3	Indicators	43
3.6	Results	45
3.6.1	Simulated traffic flow patterns	45
3.6.2	Sensitivity analysis	46
3.7	Conclusions	52
4	Optimization of traffic flow at freeway sags by controlling the acceleration of some vehicles equipped with in-car systems	55
	Abstract	55
4.1	Introduction	56
4.2	Optimal control framework	58
4.2.1	Traffic system elements	58
4.2.2	Assumptions	58
4.2.3	Control problem formulation	59
4.3	Experimental setup	64
4.3.1	Objectives	64
4.3.2	Scenarios	65
4.3.3	Solution method	66
4.3.4	Performance Indicator	67
4.4	Results	67
4.4.1	Overview of the optimization results in all scenarios	67
4.4.2	Primary strategy: deceleration - acceleration - deceleration - acceleration (DADA) maneuver in the sag area	70
4.4.3	Supporting strategy: deceleration - acceleration maneuvers upstream of the sag area	76
4.5	Conclusions	78
4.6	Appendix	80
5	Mainstream traffic flow control at sags	81
	Abstract	81
5.1	Introduction	82
5.2	Background	83

5.2.1	Sags as freeway bottlenecks	83
5.2.2	Control measures to mitigate congestion at sags	84
5.3	Control strategy	85
5.3.1	Control concept: mainstream traffic flow control	85
5.3.2	Control law: proportional feedback	86
5.4	Method of performance evaluation	87
5.4.1	Model of longitudinal driving behavior	88
5.4.2	Simulation settings	89
5.4.3	Performance indicator: TD	92
5.5	Results	92
5.5.1	Reference scenario	92
5.5.2	No-control scenario	92
5.5.3	Control scenario	93
5.6	Sensitivity analysis	94
5.7	Conclusions	95
6	Conclusions	97
6.1	Main findings	97
6.1.1	Causes of stop-and-go waves at sags	97
6.1.2	Microscopic modeling of traffic flow at sags	97
6.1.3	Optimal acceleration behavior of vehicles equipped with in-car systems at sags	98
6.1.4	Mainstream traffic flow control using variable speed limits at sags	99
6.2	Conclusions	99
6.3	Implications for practice	102
6.4	Recommendations for future research	102
	References	105
	Summary	112
	Samenvatting (Summary in Dutch)	119

About the author	125
Bio	125
Publications	126
TRAIL Thesis Series	129

List of Figures

1.1	Research approach	11
1.2	Overview of the dissertation structure	15
2.1	Factors triggering the formation of stop-and-go waves on the uphill section of sags (according to the scientific literature)	20
2.2	Layout of the study site	22
2.3	Steps to identify the causes of formation and growth of stop-and-go waves	23
2.4	Examples of stop-and-go wave propagation, formation and growth	24
2.5	Identification of the causes of stop-and-go wave formation and growth	26
2.6	Speed deviation (from 50 km/h) of vehicles passing each camera location on the center and median lanes	29
3.1	Vertical profile of the freeway stretch	41
3.2	Speed contour plots of the shoulder lane, the center lane and the median lane in the base scenario (first simulation replication)	46
3.3	Speeds and flows over time at location $x=8.2$ km (i.e., near the end of the sag vertical curve) in the base scenario (first simulation replication)	47
3.4	Sensitivity analysis: time-to-breakdown in each scenario	48
3.5	Sensitivity analysis: average total exit flow after breakdown in each scenario	49
3.6	Sensitivity analysis: average vehicle delay after breakdown in each scenario	50
4.1	Example of vehicle stream	58
4.2	Overview of the inputs needed to calculate the state dynamics	61
4.3	Vertical profile of the freeway stretch	66

4.4	Average vehicle delay in every scenario and differences in comparison with the no-control scenario	69
4.5	Deceleration-acceleration-deceleration-acceleration (DADA) maneuver of the controlled vehicle in the scenario with $M = \{75\}$: acceleration and speed profiles	71
4.6	Deceleration-acceleration-deceleration-acceleration (DADA) maneuver of the controlled vehicle in the scenario with $M = \{75\}$: position and net distance headway over time	72
4.7	Effect of the DADA maneuver of the controlled vehicle on the behavior of the first 90 following vehicles in the control scenario with $M = \{75\}$	73
4.8	Speed contour plots of the no-control scenario and the control scenario with $M = \{75\}$	74
4.9	Effect of the DADA maneuver of the controlled vehicle on the total travel time in the control scenario with $M = \{75\}$	75
4.10	Speed and time headway of every vehicle on the first part of the sag in the no-control scenario, in the control scenario with $M = \{150\}$ (optimal solution) and in the virtual scenario without supporting strategy corresponding to that control scenario	77
4.11	Speed and time headway of every vehicle at the end of the sag in the no-control scenario, in the control scenario with $M = \{150\}$ (optimal solution) and in the virtual scenario without supporting strategy corresponding to that control scenario	78
5.1	Flows within network in two scenarios: without controlled section and with controlled section	86
5.2	Vertical alignment of network (from $x = 25.0$ km to $x = 30.0$ km)	90
5.3	Simulation results	93

List of Tables

2.1	Causes of stop-and-go wave formation and growth	28
3.1	Parameters of the car-following model and the lane change model (base scenario)	42
3.2	Traffic composition	42
3.3	Sensitivity analysis: model inputs in each scenario	43
4.1	Values of the parameters that describe the traffic system	66
5.1	Parameter values	91
5.2	Performance of controller, including sensitivity analysis	95

Chapter 1

Introduction

1.1 Freeway traffic dynamics, modeling and control

A *freeway* is a road designed for high-speed vehicular traffic that has no at-level crossings with other roads and in which the two traffic directions are segregated. Entrances to and exits from freeways are provided by ramps, which allow for speed adaptation. Freeways are generally the backbone of regional transport networks. Their main purpose is to provide short and reliable travel times between key locations of a region while ensuring safe operations (Findley et al., 2015). However, whether or not freeways can properly perform this function depends on the traffic conditions along them.

1.1.1 Freeway traffic dynamics

Traffic can be considered a self-driven many-particle system in which each particle corresponds to a driver-vehicle unit (Helbing, 2001); therefore, traffic dynamics are determined by the individual behavior of the driver-vehicle units. This behavior can be disaggregated into a longitudinal component and a lateral component. The longitudinal component corresponds to the way driver-vehicle units accelerate while they are on a specific lane. The lateral component corresponds to the way they: *a*) choose lanes; and *b*) change lanes whenever it is required or desirable.

With regard to longitudinal driving behavior, empirical observations show that there is a general relation between the speed of a vehicle and the distance to its predecessor. If this distance is long enough, vehicles move at the driver's desired speed. However, below a certain distance threshold, vehicle speeds are lower – the lower the distance, the lower the speed (Tilch & Helbing, 2000). The relation between spacing and speed differs between drivers and vehicle types (Helbing, 2001). Furthermore, for individual driver-vehicle units, this relation depends on the weather (Hoogendoorn et al., 2011), the freeway geometry (Koshi, 2003) and the traffic state (Tilch & Helbing, 2000).

With regard to lateral driving behavior, empirical observations indicate that drivers choose lanes based on route requirements, lane preferences and the prevailing traffic conditions. Leaving aside route considerations, drivers select a lane based on various strategies ranging from maintaining the desired speed to adapting to the traffic speed on a particular lane (Keyvan-Ekbatani et al., 2015). If drivers decide to move to another lane, they do so only when a sufficiently long gap is available on the target lane. If such a gap is not available, drivers may indicate their desire to change lanes and cooperate with other drivers in order to create one (Schakel et al., 2012).

The general traits of driving behavior mentioned above have crucial implications for traffic dynamics. Most importantly, they determine the capacity of freeways and the characteristics of congested traffic. Specifically, as a result of the general characteristics of longitudinal driving behavior, there exists a fundamental relation between traffic flow, density and speed on a lane. Empirical observations show that, at low densities, vehicles move at the free speed; therefore, traffic flow increases with density. Instead, above a certain density (critical density), vehicles move at a lower speed than the free speed; as a result, traffic flow decreases as density increases. The maximum flow (capacity) is reached at the critical density.

There are several models that describe the fundamental relation between traffic flow, density and speed. The first model was proposed by Greenshields et al. (1935), who assumed a linear relation between speed and density, and a concave parabolic relation between flow and density. The model proposed by Koshi et al. (1981) assumes that the relation between traffic flow and density looks like a reversed λ , with a concave branch for densities below the critical density, a convex branch for higher densities, and a discontinuity between the two branches. The model proposed by Newell (1993), which is widely used in traffic flow modeling, assumes that the flow-density relation is continuous and looks like an inversed V, with two linear branches.

Capacity is not constant along freeways. Neither is the demand. Freeway networks often have *bottlenecks*, i.e., sections where the demand exceeds the capacity. Typical examples of potential bottlenecks are lane drops, sections with on-ramps (where the demand equals the mainline flow plus the on-ramp flow), tunnels and weaving sections (where the capacity is limited because vehicles keep longer headways than expected given their speed). Transitory events such as accidents and adverse weather conditions can also cause a local decrease in freeway capacity (TRB, 2010).

In general, when the total demand exceeds the capacity of a freeway section, traffic becomes congested upstream of the bottleneck (Daganzo, 1997). It is important to note that the capacity of freeways in congested traffic conditions is significantly lower than in free-flow conditions (Hall & Agyemang-Duah, 1991; Tilch & Helbing, 2000). This phenomenon is known as *capacity drop*. The causes of the capacity drop are not fully understood yet, but they seem to be related to the bounded acceleration capability of lane-changing vehicles (Leclercq et al., 2011), differences in longitudinal driving behavior between drivers, and intra-driver differences in longitudinal driving behavior depending on the traffic conditions (Yuan et al., 2016).

Empirical observations indicate that there are four main types of congested traffic patterns (Helbing et al., 2009): *a*) oscillating congested traffic; *b*) homogeneous congested traffic; *c*) pinned localized clusters; and *d*) moving localized clusters. Oscillating and homogeneous congested traffic are spatially extended patterns. The first is characterized by oscillating traffic speeds, which are due to the presence of *stop-and-go waves* (i.e., spatially-confined regions of low traffic speed that propagate upstream at a constant velocity of 15-25 km/h). The second is characterized by stable (low) traffic speeds. Localized clusters are congested traffic patterns in which traffic is congested only on a short freeway subsection whose length is stable over time (i.e., they are spatially confined congested patterns). The main difference between pinned and moving localized clusters is that the first stay at a fixed location over a long period, whereas the second propagate upstream with a characteristic speed of 15-25 km/h (hence, a moving localized cluster can be seen as an “isolated” stop-and-go wave).

There is empirical evidence that stop-and go waves can be triggered by: *a*) instabilities in longitudinal vehicle interactions; and *b*) lane-changing maneuvers. The formation of stop-and-go waves as a result of instabilities in longitudinal vehicle interactions has been observed primarily in single-lane facilities (Smilowitz et al., 1999; Sugiyama et al., 2005). The causes of these instabilities seem to be related to drivers’ maneuvering errors, anticipation to downstream traffic conditions, and personal characteristics such as aggressiveness (Yeo & Skabardonis, 2009; Laval & Leclercq, 2010). Instead, in multi-lane freeways, empirical observations show that lane-changing maneuvers are the primary factor triggering the formation of stop-and-go waves (Ahn & Cassidy, 2007; Zheng et al., 2011).

Finally, it is important to note that in multi-lane freeways the proportion of flow across lanes varies significantly with respect to the total flow. Typically, the proportion of flow on the median lane increases as the total flow increases, whereas the proportion on the other lanes, particularly the shoulder lane, decreases (Wu, 2006; Knoop et al., 2010). The main reason for this lane flow distribution pattern appears to be the existence of fundamental differences between drivers with regard to their desired speed and level of aggressiveness, which influences their lateral driving behavior (Daganzo, 2002).

1.1.2 Freeway traffic modeling

Two main types of traffic models can be distinguished based on their level of detail: *a*) microscopic; and *b*) macroscopic. Microscopic models describe traffic as a system of interacting particles, whereas macroscopic models describe it as a compressible fluid. The main characteristics of these two types of models are described below. Other modeling approaches, such as sub-microscopic models (Ludmann, 1998), cellular automata (Nagel & Schreckenberg, 1992) and mesoscopic models (Jayakrishnan et al., 1994), are not discussed here.

Microscopic traffic models

Microscopic traffic models describe the movements of individual driver-vehicle units, which are considered to be the result of: *a*) characteristics of the drivers and the vehicles; *b*) characteristics of the interactions between driver-vehicle units; *c*) characteristics of the freeway; *d*) environmental conditions (weather, light); and *e*) traffic control measures. Microscopic traffic models generally have two main components: a longitudinal driving behavior model and a lateral driving behavior model.

Different modeling approaches have been proposed in the literature to describe longitudinal driving behavior (Chandler et al., 1958; Gipps, 1981; Bando et al., 1995; Treiber et al., 2000). These models are generally defined by ordinary differential equations that describe the dynamics of a vehicle based on the known dynamics of the preceding vehicle (Ossen, 2008). In general, the equation of motion of a vehicle is characterized by an acceleration function whose input stimuli are the vehicle own speed (v_i), the net distance to the preceding vehicle (s_i), and the speed of the preceding vehicle (v_{i-1}):

$$\ddot{x}_i(t) = \dot{v}_i(t) = f(s_i(t), v_i(t), v_{i-1}(t)) \quad (1.1)$$

In order to take the drivers' multi-anticipative behavior into account, some longitudinal driving behavior models also include as stimuli the dynamics of vehicles located in front of the immediate predecessor (Bexelius, 1968; Treiber et al., 2006).

There exist several modeling approaches to describe lateral driving behavior (Gipps, 1986; Kesting et al., 2007; Toledo et al., 2007; Schakel et al., 2012). These models typically consider three tasks: *a*) lane-change desirability evaluation; *b*) target lane selection; and *c*) lane-change execution (Toledo et al., 2007). The first two tasks are generally modeled by the use of deterministic priority rules or utility maximization models, whereas lane-change execution is typically modeled by the use of gap-acceptance models. Most lateral driving behavior models classify lane changes as either mandatory or discretionary, which influences the conditions in which lane changes are considered desirable and are executed. Mandatory lane changes are motivated by the necessity to move to another lane in order to follow the desired route. Discretionary lane changes are motivated by a perceived improvement of the driving conditions in the target lane.

Macroscopic traffic models

Macroscopic traffic models formulate the relationships between traffic flow characteristics such as density (k), flow (q) and mean speed (u) of a traffic stream. All macroscopic traffic models share two fundamental relations. The first one is the vehicle conservation law (see Eq. 1.2, where x denotes distance and t denotes time) and the second one is the flow-density-speed relation (see Eq. 1.3):

$$\frac{\partial k}{\partial t} + \frac{\partial q}{\partial x} = 0 \quad (1.2)$$

$$q = k \cdot u \quad (1.3)$$

There are two main types of macroscopic traffic models: first-order and higher-order models. *First-order models* (Lighthill & Whitham, 1955; Richards, 1956) assume that the flow-density-speed relation observed in stationary traffic conditions applies also to dynamic traffic conditions. This type of models can be solved in discrete time and space using the Cell Transmission Model (Daganzo, 1994; Lebacque, 1996). Recently, it has been shown that this type of models can be solved in three different two-dimensional coordinate systems arising in the space of vehicle number, time and distance (Laval & Leclercq, 2013). *Higher-order models* assume that the flow-density-speed relation observed in stationary traffic conditions does not hold in dynamic traffic conditions. As a result, they can reproduce phenomena that can generally not be reproduced by first-order models, such as smooth shocks, the capacity drop and stop-and-go waves. An example of this type of models is the Payne model (Payne, 1971), which can be solved using the numerical scheme proposed by Kotsialos et al. (2002).

1.1.3 Freeway traffic control

Traffic congestion has important impacts on society, including delays, travel time unreliability, increased fuel consumption, increased air pollution and lower levels of traffic safety (Van Wee et al., 2013). For this reason, public authorities dedicate substantial resources to mitigate congestion. A traditional approach is to expand the freeway infrastructure in order to increase its capacity. However, due to limitations of space and economic resources, this solution is often not feasible or not desirable.

In recent decades, a more flexible and less costly approach known as *dynamic traffic management* (DTM) has grown in importance. This approach aims to optimize the utilization of the available freeway infrastructure by applying various types of control measures that influence travel and driving behavior (Rijkswaterstaat, 2003). The goal of these measures is to minimize the total time that vehicles spend in the network, which is equivalent to maximizing the network exit flow (Papageorgiou et al., 2003). In general, the effectiveness of DTM measures depends on the availability of adequate real-time traffic data and the use of efficient control methodologies.

Important types of DTM measures for freeway networks are ramp metering, variable speed limits and route guidance (Papageorgiou et al., 2003; Hegyi et al., 2009). These three types of measures are discussed in the remainder of this section. Other types of measures – not discussed here – include reversible lanes (Wolshon & Lambert, 2006) and lane advice systems (Xing et al., 2014; Schakel & van Arem, 2014).

Ramp metering measures regulate the flow entering a freeway via traffic signals located at the end of on-ramps. Their main objective is to prevent traffic on the mainline from becoming congested, since that leads to decreased freeway capacity. The rate at which vehicles are released onto the mainline can be determined off-line based on historical

demand data (Wattleworth, 1967) or on-line based on the prevailing traffic conditions on the mainline (Masher et al., 1975; Papageorgiou et al., 1991). Also, ramp metering installations can be operated in isolation or in coordination (Papamichail et al., 2010). Several empirical studies show that ramp metering measures can substantially mitigate freeway traffic congestion (Papageorgiou et al., 1997; Levinson & Zhang, 2006).

Dynamic speed limits are used to regulate the traffic speed on a freeway section in order to: *a*) homogenize vehicle speeds; or *b*) limit the flow. The first approach aims to reduce speed differences between vehicles so that traffic flow becomes more stable. In this case, the speed limits are typically above the critical speed; therefore, traffic flow is not reduced, but the average speed and the density slightly decrease and increase, respectively. This approach can delay the formation of traffic congestion (Smulders, 1990), but there is no empirical evidence that it can significantly increase the traffic flow (Van den Hoogen & Smulders, 1994; Weikl et al., 2013). The second approach aims to limit the inflow into a certain freeway section (e.g., a bottleneck) in order to dissolve or prevent congestion there. Since traffic remains uncongested, flow can be higher. This approach has been applied to dissolve moving localized clusters (Hegyi et al., 2008) and to prevent congestion at on-ramp bottlenecks (Carlson et al., 2010, 2013). Field tests show promising results (Hegyi & Hoogendoorn, 2010) but the effectiveness of this approach needs to be evaluated based on more empirical data.

Route guidance measures aim to influence the route choice of drivers when multiple routes are available. The main goal is to spread traffic efficiently over the network and minimize the total travel time (Papageorgiou & Messmer, 1991). Typically, route guidance systems display traffic information – such as congestion length or travel time to the next common point – for each alternative route on variable message signs. There is empirical evidence that route guidance measures can significantly reduce congestion in freeway networks (Kraan et al., 1999). One of the main challenges is to accurately predict the traffic conditions that drivers will encounter along each route in order to provide them with reliable information (Papageorgiou et al., 2003; Hegyi et al., 2009).

1.2 Traffic dynamics, modeling and control at sags

Sag vertical curves (or *sags*) are roadway sections along which the gradient increases gradually in the direction of traffic. Sags provide a transition between two sloped roadway sections, allowing vehicles to negotiate the change in elevation rate in a gradual manner. The design of sag vertical curves considers factors such as design speed of the roadway, slope, acceptable rate of gradient change and drainage (TRB, 2010). On freeways, sags have important effects on traffic dynamics. They are one of the most common types of freeway bottleneck in hilly regions. For example, in Japanese inter-city freeways, 60% of traffic jams occur at sags (Xing et al., 2014). Despite this fact, analysis of the characteristics of traffic flow at sags and development of DTM measures for this type of bottlenecks are still relatively underdeveloped research fields.

1.2.1 Traffic dynamics at sags

Empirical observations indicate that, at sags, most drivers reduce their desired speed (Furuichi et al., 2003; Brilon & Bressler, 2004) and keep longer distance headways than expected given their speed (Koshi, 2003). These changes in longitudinal driving behavior are generally unintentional: their main reason appears to be related to inadequate acceleration behavior (Yoshizawa et al., 2012). However, in the case of trucks the reason may sometimes be related to insufficient acceleration capability (Laval, 2009). The characteristics of the longitudinal driving behavior of drivers at sags have important implications for traffic dynamics.

Previous research shows that the capacity of sags is 10-25% lower than that of flat freeway sections having the same number of lanes (Okamura et al., 2000; Xing et al., 2010). As a result, traffic often becomes congested at sags in high demand conditions (Koshi et al., 1992; Patire & Cassidy, 2011). It should be noted that, although the presence of heavy vehicles has a negative impact on the capacity of sags (Brilon & Bressler, 2004), congestion also occurs when traffic demand consists mainly of passenger cars (Furuichi et al., 2003). Based on the classification proposed by Helbing et al. (2009), the congested traffic patterns most often observed at sags are moving localized clusters and oscillating congested traffic (Patire & Cassidy, 2011; Zheng et al., 2011).

Typically, the process of formation of persistent congestion at sags consists of two phases. In the first phase, congestion forms on the median lane (Koshi et al., 1992; Patire & Cassidy, 2011). The main reason why congestion emerges first on the median lane is related to the characteristics of lane flow distribution: with high demand and uncongested traffic, flows tend to be higher (and closer to capacity) on the median lane than on the other lanes (Hatakenaka et al., 2006; Xing et al., 2010). In the second phase, congestion spreads from the median lane to the other lanes (Koshi et al., 1992; Patire & Cassidy, 2011). This process can be described as follows. When traffic becomes congested on the median lane, some vehicles migrate from that lane to the less crowded lanes in order to avoid queuing (Hatakenaka et al., 2006; Patire & Cassidy, 2011). When the flow on those lanes exceeds their capacity, traffic also breaks down there. At that point, traffic is congested on all lanes, which causes a significant decrease in total outflow – due to the capacity drop – and the formation of a queue (Koshi et al., 1992; Patire & Cassidy, 2011). In general, the head of the queue stays on the first 500 to 1000 m downstream of the bottom of the sag (Brilon & Bressler, 2004).

1.2.2 Traffic modeling at sags

In the past decades, some researchers have developed mathematical models aimed at reproducing the characteristics of traffic flow at sags, basically microscopic models. These models include car-following models that take into account in some way the influence of the freeway vertical profile on vehicle acceleration. These car-following models can be grouped into two categories based on whether they assume that drivers

explicitly compensate for the limiting effect that an increase in freeway gradient has on vehicle acceleration (Yokota et al., 1998; Oguchi & Konuma, 2009) or they do not (Koshi et al., 1992; Komada et al., 2009). The first assumption is more consistent with empirical observations than the second one; for this reason, the models based on that assumption are able to reproduce longitudinal driving behavior and traffic dynamics at sags more accurately. To the author's knowledge, there is no lane change model that takes the freeway vertical profile into account. To the author's knowledge, there is also no evidence that decisions to change lanes are directly influenced by the vertical profile of the freeway.

1.2.3 Traffic control at sags

During the last two decades, several DTM measures aimed at mitigating congestion at sags have been developed. The (potential) effectiveness of those measures has been investigated through field tests or simulation experiments. Generally, the goals of these measures are: *a*) to increase the free-flow capacity of sags; *b*) to prevent the formation of congestion at sags in nearly-saturated traffic conditions; and/or *c*) to increase the queue discharge capacity of sags.

An example of a measure from the first category is equipping vehicles with adaptive cruise control systems, which perform the acceleration task more efficiently than human drivers (Ozaki, 2003). Another example is distributing the traffic flow more evenly across lanes to use the bottleneck capacity more efficiently (Hatakenaka et al., 2006; Xing et al., 2010). The second category comprises measures such as preventing the formation of long vehicle platoons (Hatakenaka et al., 2006) and discouraging drivers from performing lane changes to the busiest lanes (Hatakenaka et al., 2006; Patire & Cassidy, 2011). The third category comprises measures such as giving information to drivers about the location of the head of the queue and encouraging them to recover speed after leaving congestion (Xing et al., 2007; Sato et al., 2009).

Most existing measures utilize variable message signs (VMS) located alongside the freeway to communicate with drivers and influence their driving behavior. However, nowadays, there is growing interest in developing DTM measures based on the use of *in-vehicle* systems (Hatakenaka et al., 2006). This type of technology allows traffic management centers to communicate with vehicles equipped with adequate receivers in order to send them information or take (partial) control of certain driving tasks, which opens up new possibilities for dynamic traffic management.

1.3 Challenges with regard to the development of dynamic traffic management measures for sags

Although significant achievements have taken place in the past years, several challenges need to be addressed if we want to be able to manage traffic at sags more effectively. These challenges include: *a)* gaining a better understanding of the causes of congestion at sags; *b)* modeling traffic at sags in a more realistic way; and *c)* identifying the most effective strategies to manage traffic at sags.

The first challenge is to properly explain why traffic becomes congested at sags from a theoretical viewpoint. It is well known that, in high demand conditions, traffic generally becomes congested in the form of stop-and-go waves. However, there is no consensus in the literature on what the main cause of these waves is. As explained in Section 1.2.1, there is evidence that the car-following behavior of most drivers changes at sags (Koshi, 2003). Some researchers suggest that stop-and-go waves form mainly as a result of the cumulative effect of these local changes in car-following behavior (Koshi et al., 1992; Koshi, 2003). Instead, other researchers suggest that lane changes – some of which may be indirectly caused by the above-mentioned changes in car-following behavior – may be the primary cause of stop-and-go waves at sags (Brlon & Bressler, 2004; Zheng et al., 2011). Elucidating the causes of traffic congestion at sags is crucial for the development of effective DTM measures.

The second challenge is to improve our ability to model traffic on freeways with sags. As explained in Section 1.2.2, there are various car-following models that take account of the influence of the freeway vertical profile (Koshi et al., 1992; Yokota et al., 1998; Komada et al., 2009; Oguchi & Konuma, 2009). However, it is currently unclear whether microscopic traffic models composed of the existing car-following and lane change models can accurately reproduce all the traffic phenomena that characterize traffic flow dynamics at sags, particularly as regards the traits of acceleration behavior and the exact location of the bottleneck. Therefore, it is necessary to identify the most relevant traffic phenomena that occur at sags and determine whether existing models can reproduce them accurately. If that is not the case, a new modeling approach should be developed based on theoretical assumptions supported by evidence. The ability to accurately model traffic at sags is essential for the development and evaluation of DTM measures for this type of bottlenecks. Also, we can gain insight into the characteristics of traffic flow dynamics at sags by analyzing the properties of a valid traffic model.

The third challenge is to identify the most effective strategies to manage traffic at freeway sags. As mentioned in Section 1.2.3, various specific DTM measures for sags have been proposed in the past years. Nevertheless, it is not evident what (combination of) strategies could generate the greatest reduction in congestion using different types of actuators and in different traffic conditions. These strategies could be identified using mathematical optimization methods. Note, however, that this type of method requires

a valid traffic flow model to produce relevant results. Identifying the most effective strategies to manage traffic at sags is particularly relevant for the development of DTM measures based on the use of in-car systems. These type of measures are mostly in early phases of development. Insight on how equipped vehicles should behave at sags in order to minimize congestion would be very valuable to guide the development of these type of measures towards promising directions. Furthermore, there are some types of DTM measures that have not been applied to sags yet, but have been shown to be very effective in mitigating congestion at other type of bottlenecks. An example is mainstream traffic flow control using variable speed limits (Carlson et al., 2011, 2013). It is necessary to evaluate the potential effectiveness of these measures in reducing congestion at sags.

1.4 Research objective and research questions

The main objective of this thesis is to develop new concepts for dynamic traffic management at freeway sags based on a thorough understanding of the causes of congestion at this type of bottlenecks. In order to achieve this objective, the thesis provides answers to the following research questions:

1. What is the primary cause of stop-and-go waves at freeway sags?
2. How should traffic be modeled on a microscopic level in order to accurately reproduce its characteristics on freeways with sags?
3. How should vehicles equipped with in-car systems behave at freeway sags in order to minimize the severity of congestion?
4. To what extent can mainstream traffic flow control (MTFC) measures based on the use of variable speed limits mitigate congestion at freeway sags?

1.5 Research approach

A five-stage approach is used to answer the research questions (see Figure 1.1). To answer the first research question, a detailed analysis of empirical trajectory data from a sag in a Japanese freeway is carried out. On the basis of this analysis, it is established whether the primary cause of congestion at the study site is related to car-following phenomena or disruptive lane changes. This advances the theoretical understanding of traffic flow dynamics at sags.

Four steps are taken to answer the second research question. First, the requirements of a valid traffic model are specified. Second, a literature review is conducted in order to determine the suitability of existing models for this research (based on the model

requirements). Third, a new microscopic traffic model is developed on the basis of the results of the literature review as well as the findings on the causes of congestion at sags. Finally, the face-validity of the model is tested by means of a simulation study.

Three steps are taken to answer the third research question. First, an optimal control problem is formulated. The problem consists in finding the acceleration profiles of a set of vehicles that minimize the total travel time of the traffic stream. Traffic dynamics are modeled using the previously developed traffic modeling approach. Second, the problem is solved for various scenarios. Finally, the main strategies that define the optimal acceleration behavior of vehicles equipped with in-car systems at sags are identified by analyzing the results.

To answer the fourth question, two steps are taken. First, a DTM measure based on the concept of mainstream traffic flow control (MTFC) is designed. The measure regulates the traffic density at the end of the sag (bottleneck) through a variable speed limit section located upstream of the vertical curve. Second, the effectiveness of the proposed measure in mitigating congestion is evaluated by means of a simulation study using the previously developed traffic model.

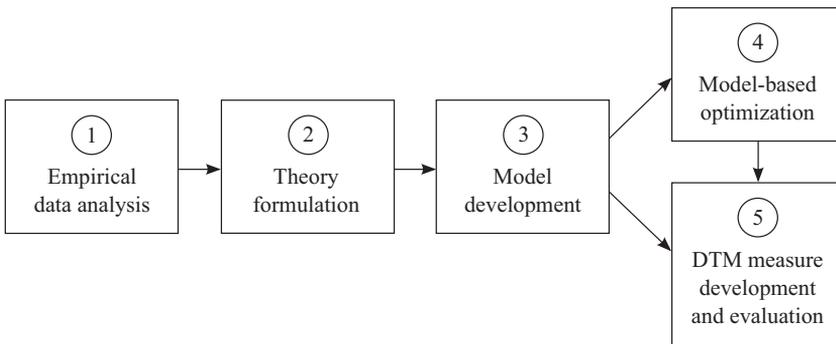


Figure 1.1: Research approach.

1.6 Research scope

The thesis focuses on freeway traffic (thus, it considers traffic flows in one direction not hindered by any at-level crossings). The reason is that sags are a common type of bottleneck in freeway networks (in hilly regions). Instead, in urban road networks, traffic flow rarely breaks down as a consequence of the presence of sags, because other types of bottlenecks (e.g., intersections) are normally dominant.

This thesis considers networks consisting of a freeway section with a sag vertical curve with a realistic profile. It is assumed that no other type of spatial inhomogeneity that

can constitute a bottleneck (e.g., ramps, weaving sections, lane drops, horizontal curves and tunnels) is present. This allows the study of the individual effects of sags on traffic flow and the development of specific traffic management strategies for sags. However, it is worth noting that the traffic model and control strategies developed in this thesis could be extended in order to make them operational in freeway sections with a more complex layout.

The thesis focuses on developing traffic management strategies aimed at reducing total delay. Other relevant traffic control objectives such as increasing safety or reducing environmental impacts are not explicitly considered. Moreover, the thesis focuses on the effects that traffic management strategies can have on traffic flow and the reasons why they are able to reduce total delay. Topics like detailed design, implementation, enforcement and maintenance are beyond the scope of this thesis. Total traffic demand is taken as given; therefore, route guidance strategies are not investigated.

1.7 Main contributions

This section summarizes the main contributions of this thesis. A distinction is made between contributions that are of a scientific nature (either theoretical or methodological, see Section 1.7.1) and contributions that are of a practical nature (Section 1.7.2).

1.7.1 Scientific contributions

The main scientific contributions of this thesis are as follows:

1. *Synthesis of current knowledge on the characteristics of driving behavior and traffic flow at freeway sags* (Chapter 2). The synthesis focuses on the phenomena that cause congestion at sags. It identifies various knowledge gaps in traffic flow theory regarding the effects of sags on traffic dynamics. One gap is the reason why stop-and-go waves form at sags.
2. *Empirical identification of the primary cause of stop-and-go waves at sag vertical curves* (Chapter 2). The thesis shows that stop-and-go waves form mainly as a result of car-following phenomena instead of lane changes. This finding has important implications for traffic flow modeling (it indicates that a valid car-following model is crucial for modeling traffic at sags) and control.
3. *Review of car-following models that take account of the influence of sag vertical curves* (Chapter 3). The review shows that most existing models are based on assumptions about driving behavior that are inconsistent with empirical observations. This hinders their validity. The only modeling approach that is based on accurate assumptions is the one presented by Oguchi and Konuma (2009).

4. *New car-following model that reproduces longitudinal driving behavior on freeways with sags* (Chapter 3). Our model generalizes the Intelligent Driver Model (Treiber et al., 2000) to take account of the influence of vertical curves. The model is based on a similar principle to that of Oguchi and Konuma's model but has a more generic formulation.
5. *New microscopic traffic model that reproduces traffic flow dynamics on freeways with sags* (Chapter 3). This model consists of the newly developed car-following model and a regular lane change model. The thesis shows that the proposed traffic model is face-valid. Scientific applications of this model include development and evaluation of DTM measures for sags.
6. *Generic optimization-based method to determine the optimal acceleration behavior of vehicles equipped with in-car systems at bottlenecks* (Chapter 4). The method is generic in the sense that traffic flow can be optimized at various types of bottlenecks (including sags) and multiple optimization objectives can be specified. The proposed method can be used to develop new DTM concepts.
7. *Identification of the main strategies that define the optimal behavior of vehicles equipped with in-car systems at sags* (Chapter 4). The thesis shows the importance of motivating drivers to accelerate fast through sags and limiting the inflow into the vertical curve in order to reduce congestion. Also, it suggests ways to do that by regulating the acceleration of vehicles equipped with in-car systems.
8. *Evaluation of the effectiveness of mainstream traffic flow control using variable speed limits as an approach to mitigate congestion at sags* (Chapter 5). The thesis proposes a new traffic management approach for sags that consists in regulating the inflow into the sag through variable speed limits. The goal is to prevent congestion at the bottleneck while maximizing outflow. The potential effectiveness of this approach is demonstrated through traffic simulation.

1.7.2 Contributions to practice

The main contributions to practice of this thesis are as follows:

1. *New microscopic traffic model that reproduces traffic flow dynamics on freeways with sags* (Chapter 3). This model allows for improved representation of traffic flow in microscopic simulation tools. Thus, it can support policy makers, road authorities and traffic engineering companies in making better evaluations of potential traffic control measures or infrastructure adjustments.
2. *Guidelines for developing DTM measures for sags based on the use of in-car systems* (Chapter 4). The thesis determines how equipped vehicles should move at sags in order to minimize congestion. These findings provide valuable guidance

to automotive and traffic engineering companies with regard to the development of traffic control measures using in-vehicle systems as actuators.

3. *New DTM measure for sags based on the concept of mainstream traffic flow control (MTFC)* (Chapter 5). The thesis presents a specific control measure that regulates the traffic density at the end of the sag (bottleneck) by changing the speed limit upstream of the curve on the basis of a feedback control law. The thesis demonstrates that this measure can substantially reduce congestion at sags. Therefore, it provides traffic engineering companies and road authorities with a new and promising alternative to manage traffic at this type of bottlenecks.

1.8 Outline of the dissertation

The remainder of this thesis is structured as shown in Figure 1.2. Chapter 2 investigates the causes of stop-and-go waves at sags by analyzing trajectory data from a Japanese freeway. This chapter describes the characteristics of the data analysis method and discusses the results of the analysis. It is found that stop-and-go waves form primarily as a result of local car-following phenomena rather than lane changes.

Chapter 3 presents a microscopic traffic model that is able to reproduce traffic dynamics on freeways with sags. The findings presented in the previous chapter indicate that a valid car-following model is crucial to model traffic at sags. This chapter analyzes the validity of existing car-following models and, based on the results of this review, formulates a new model. This car-following model constitutes the main novelty of the proposed microscopic traffic model. The face-validity of the traffic model is verified by comparing simulated traffic flow patterns with patterns observed in empirical data.

Chapter 4 identifies the strategies that define the optimal acceleration behavior of vehicles equipped with in-car systems at sags. The chapter formulates an optimal control problem in which the acceleration of some vehicles is regulated with the objective of minimizing total delay. Traffic dynamics are modeled by using the modeling approach developed in the previous chapter. The control problem is solved for various scenarios. The results show that motivating drivers to accelerate fast through sags and limiting the inflow into the vertical curve are the most effective traffic management strategies. Furthermore, the results indicate ways to apply those strategies by making equipped vehicles perform specific acceleration-related maneuvers.

Chapter 5 presents a new DTM measure for sags based on the concept of mainstream traffic flow control (MTFC). The previous chapter concludes that limiting the inflow into sags is a key strategy for managing traffic at this type of bottleneck. The measure presented in this chapter uses variable speed limits to limit the inflow into the sag and regulate the traffic density at the bottleneck, thereby preventing congestion. A proof of concept is carried out by means of a traffic simulation study using an adapted version of the model presented in Chapter 3.

Finally, Chapter 6 presents the conclusions of the thesis. This chapter summarizes the main research findings and discusses their implications. In addition, it proposes directions for future research.

As a final note, it is important to remark that this thesis is written in the format of collection of articles. Chapters 2, 3 and 5 correspond to articles that have already been published in scientific journals, and Chapter 4 corresponds to an article that has been submitted for publication. Therefore, the chapters contain repeated information, particularly in the abstracts and introductory sections. The reader may want to skim those sections.

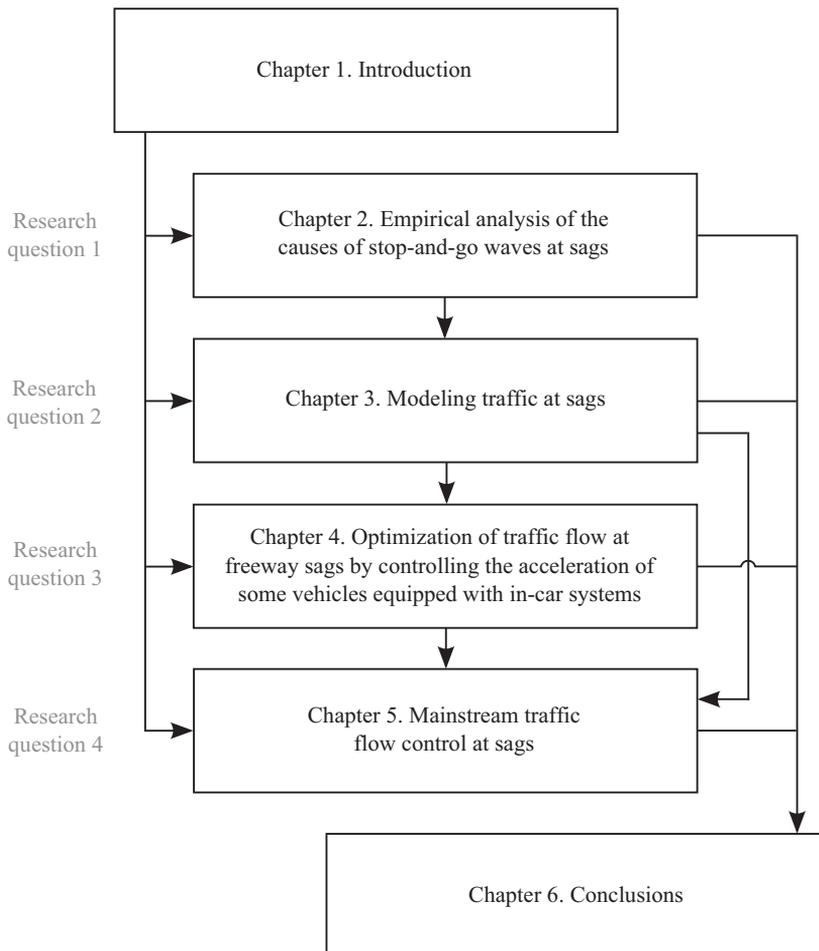


Figure 1.2: Overview of the dissertation structure (including indication of relationships between chapters, research questions addressed in every chapter, and stages of the research approach corresponding to every chapter).

Chapter 2

Empirical analysis of the causes of stop-and-go waves at sags

Goñi-Ros, B., Knoop, V.L., van Arem, B., Hoogendoorn, S.P. (2014). Empirical analysis of the causes of stop-and-go waves at sags. *IET Intelligent Transport Systems*, 8(5), 499–506. © Institution of Engineering and Technology.

Abstract

Stop-and-go waves are spatially-confined regions of low traffic speed that propagate upstream at a constant velocity. The occurrence of stop-and-go waves on freeways has negative impacts on both travel time and traffic safety. Sags are freeway sections along which gradient changes significantly from downwards to upwards. Stop-and-go waves often emerge on the uphill section of sags, both in uncongested and congested traffic conditions. According to previous studies, the formation of stop-and-go waves at sags can be caused by local changes in car-following behaviour as well as disruptive lane changes. However, it is not clear which of those two causes is more frequent. This paper aims to identify the primary factor triggering stop-and-go waves at sags. To this end, the authors analyse vehicle trajectories collected by means of video cameras on a three-lane sag of the Tomei Expressway (Japan), identifying the causes of formation and growth of stop-and-go waves on the study site. The results show that the primary factor triggering stop-and-go waves is related to car-following behaviour. This finding shows the relevance of developing systems to assist drivers in performing the acceleration task at sags.

2.1 Introduction

Stop-and-go waves are spatially-confined regions of low traffic speed that propagate upstream at a constant velocity of 15-25 km/h (Schönhof & Helbing, 2007; Schreiter et al., 2010). The occurrence of stop-and-go waves on freeways has negative impacts on both travel time and traffic safety (Hegyi et al., 2005). Sags are freeway sections along which gradient changes significantly from downwards to upwards in the direction of traffic (Furuichi et al., 2003). Traffic flow capacity is significantly lower at sags than at flat sections (Koshi et al., 1992; Okamura et al., 2000). Because of that, sags become bottlenecks in freeway networks, causing the formation of congestion in conditions of high traffic demand. The mechanism of formation of congestion at sags has two phases (Koshi et al., 1992; Hatakenaka et al., 2006; Patire & Cassidy, 2011). In the first phase, congestion forms on the fast lane(s) of the uphill section in the form of stop-and-go waves. In the second phase, congestion spreads from the fast lane(s) to the slow lane(s). Stop-and-go waves also emerge within congested traffic (Zheng et al., 2011).

The primary factor triggering the formation of stop-and-go waves at sags has not been clearly identified yet. Several studies show that drivers unintentionally change their car-following behaviour on the uphill section of sags. More specifically, drivers tend to reduce speed (Koshi et al., 1992; Furuichi et al., 2003) and keep longer distances to the leading vehicle than expected (Koshi, 2003; Yoshizawa et al., 2012). Some authors suggest that those local changes in car-following behaviour are the dominant factor triggering stop-and-go waves on the uphill section of sags (Koshi et al., 1992; Koshi, 2003). In contrast, other authors suggest that disruptive lane changes may be the primary triggering factor (Brilon & Bressler, 2004).

The objective of this paper is to determine whether the primary factor triggering stop-and-go waves at sags is related to car-following behaviour or to lane changes, both in congested and uncongested traffic conditions. This is important to understand the mechanism of stop-and-go wave formation at sags, and to develop effective measures to improve traffic flow efficiency and traffic safety at that type of bottlenecks.

To that end, we analyse a set of vehicle trajectories collected by means of video cameras on a three-lane sag of the Tomei Expressway (Japan) during the start of the morning rush hour, before and after the formation of persistent congestion. We analyse the evolution of speed over time on each lane at several locations as well as individual vehicle trajectories, identifying and classifying the causes of stop-and-go wave formation and growth.

This paper will show that the primary factor triggering stop-and-go waves at sags is related to car-following behaviour; lane changes seem to be a less significant triggering factor. This finding shows the relevance of developing systems to assist drivers in performing the acceleration task at sags (Goñi-Ros et al., 2012).

The rest of this paper is structured as follows: Section 2.2 describes the main causes of stop-and-go waves at sags according to the scientific literature. Section 2.3 presents the characteristics of the study site and the trajectory data. Section 2.4 describes the data analysis methods used to identify the main factor triggering stop-and-go waves at the study site. Section 2.5 reports the results of the analysis. Section 2.6 discusses the implications of the results, taking into account the limitations of the data. Section 2.7 presents the conclusions of this study.

2.2 Background

In general, the occurrence of traffic congestion on freeways is caused by a combination of three elements (Helbing et al., 2009): a) high traffic volume; b) a spatial inhomogeneity on the freeway that generates a capacity bottleneck; and/or c) a temporary disturbance of the traffic flow. Several empirical studies show that traffic flow capacity can be significantly lower at sags than at flat sections having the same number of lanes (up to 30% lower) (Koshi et al., 1992; Okamura et al., 2000; Schönhof & Helbing, 2007). Because of their lower capacity, sags become bottlenecks in freeway networks, causing the formation of congestion in conditions of high traffic demand. Generally, in freeways with keep-left or keep-right rules, the mechanism of congestion formation at sags consists of two phases: i) formation of congestion on the inner (fast) lane(s) of the uphill section; and ii) spreading of congestion to the outer (slow) lane(s) (Koshi et al., 1992; Hatakenaka et al., 2006; Patire & Cassidy, 2011).

In the first phase, congestion forms on the fast lane(s) of the uphill section. The main reason why congestion emerges first on the fast lane(s) instead of the slow lane(s) is related to the characteristics of lane flow distribution. In freeways with keep-left or keep-right rules, with high demand and uncongested traffic, flows tend to be much higher on the fast lane(s) than on the slow lane(s), hence flows are closer to capacity on the fast lane(s) (Koshi et al., 1992; Wu, 2006; Xing et al., 2012). In those conditions, small perturbations can destabilise traffic flow and trigger the formation of congestion in the form of stop-and-go waves (Koshi et al., 1992; Patire & Cassidy, 2011).

In the second phase, congestion spreads from the fast lane(s) to the slow lane(s). Some authors suggest that the spreading of congestion is caused by the fact that, when stop-and-go waves emerge on the fast lane(s), some drivers migrate to the less crowded slow lane(s) in order to avoid stopping (Koshi et al., 1992; Hatakenaka et al., 2006; Patire & Cassidy, 2011). The spreading of congestion to all lanes causes a significant decrease in total outflow rates (due to the capacity drop phenomenon (Helbing, 2001)) and the formation of a queue upstream of the bottleneck (Koshi et al., 1992; Patire & Cassidy, 2011). The head of the queue stays on the first 500-1000 m downstream of the bottom of the sag (Koshi et al., 1992; Brilon & Bressler, 2004). Within the queue, small traffic flow disturbances can trigger the formation of stop-and-go waves (Zheng et al., 2011).

The primary factor triggering stop-and-go waves at sags (both in uncongested and congested traffic conditions) has not been clearly identified yet (see Fig. 1). Several studies show that two significant changes in car-following behaviour occur when vehicles reach the uphill section. First, drivers tend to reduce speed (Koshi et al., 1992; Furuichi et al., 2003; Koshi, 2003; Brilon & Bressler, 2004). Second, some drivers keep longer headways than expected given their speed (Koshi, 2003; Yoshizawa et al., 2012). These local changes in car-following behaviour seem to be unintentional (Yoshizawa et al., 2012) and caused by a combination of two factors: increase in the resistance force due to the increase in freeway slope, and insufficient acceleration operation by drivers (Koshi et al., 1992; Yoshizawa et al., 2012). Some authors suggest that the above-mentioned changes in car-following behaviour can be the direct cause of the disturbances that trigger the formation of stop-and-go waves on the uphill section of sags (Koshi et al., 1992; Koshi, 2003).

However, other authors suggest that at sags stop-and-go waves can also be triggered by disruptive lane changes (Brilon & Bressler, 2004; Zheng et al., 2011). Some of those disruptive lane changes may indirectly be caused by the above-mentioned changes in car-following behaviour, since vehicle decelerations on the uphill section may locally modify the differences in traffic speed between lanes, incentivising lane changes (Brilon & Bressler, 2004). However, no empirical evidence of a causal relationship between changes in car-following behaviour and lane changes at sags has been presented in the scientific literature.

To summarise, it is not clear whether the primary factor triggering stop-and-go waves at sags is related to car-following behaviour or to lane changes. Identifying the primary factor is important to understand the mechanism of stop-and-go wave formation at sags and to develop effective measures to improve traffic flow efficiency and traffic safety at that type of bottlenecks.

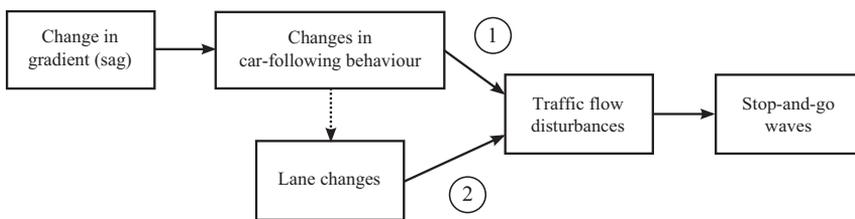


Figure 2.1: Factors triggering the formation of stop-and-go waves on the uphill section of sags (according to the scientific literature): 1) changes in car-following behaviour; and 2) disruptive lane changes.

2.3 Data characteristics

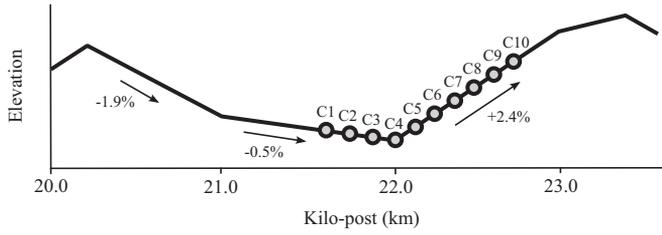
This section describes the empirical data used to analyse the causes of stop-and-go waves at sags. The data consist of a set of vehicle trajectories on a freeway sag section in Japan during the morning peak hour. Section 2.3.1 describes the study site, and Section 2.3.2 describes the characteristics of the vehicle trajectories.

2.3.1 Study site

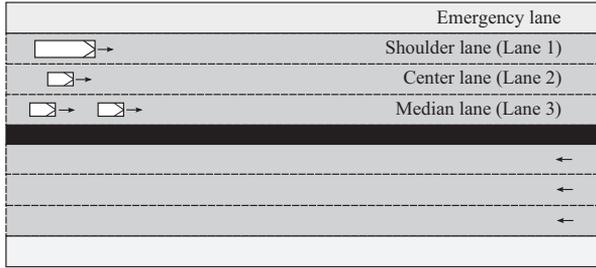
The study site is a westbound stretch of the Tomei Expressway (near Tokyo, Japan) located between kilo-posts (KP) 20 and 23.5 km. It contains a downhill section followed by an uphill section. Figure 2.2(a) shows the vertical alignment profile of the study site. The downhill approach is 1.8 km long, and it consists of a steeper section (-1.9% gradient) followed by a gentler section (-0.5% gradient). The bottom of the sag is located at KP 22.03 km. The uphill section is around 1.5 km long. Its gradient is +2.4% on the first 1000 m, but it decreases on the last 500 m. The study site has three lanes used by regular traffic (median, centre and shoulder) plus an emergency lane (see Figure 2.2(b)). The median lane is the fastest and right-most lane (note that in Japan drivers have the obligation to drive on the left unless they are overtaking). There are no ramps nor lane drops in or near the site. The expressway curves gently to the right at the study site, which may restrict the drivers line-of-sight, but only to a limited extent.

2.3.2 Vehicle trajectories

The study site is equipped with 10 video-cameras located in sequence between KP 21.67 and 22.75 km, capturing the last 360 m of the downhill approach, the bottom of the sag and the first 720 m of the uphill section (see Figure 2.2(a)). The distance between consecutive cameras is around 120 m and the exact locations are known. Using a software tool developed by Patire (2010), individual vehicles were identified in the video recordings of each camera, obtaining one passing time and lane per vehicle per camera location. Vehicle trajectories were constructed by combining the passing time and lane of each vehicle (rear bumper) at each camera location. This was done for the period 6:40 h-7:05 h on Friday, 23 December 2005, resulting in 2284 vehicle trajectories during the start of the morning peak hour, before and after the formation of persistent congestion on the study site. Note that the space and time resolution of the trajectories are limited due to the characteristics of the data collection and processing methods. Cameras are located around 120 m apart, and the passing time and lane of each vehicle are recorded only once per camera location. Therefore, space resolution is 120 m and time resolution varies between 4 and 12 s depending on vehicle speed.



(a) Vertical alignment profile and camera locations.



(b) Lane layout.

Figure 2.2: Layout of the study site.

2.4 Data analysis methods

To determine whether the primary factor triggering stop-and-go waves at the study site is related to car-following behaviour or to lane changes, we first analysed the evolution of vehicle speed over time at all camera locations, identifying the locations where stop-and-go waves form or grow in amplitude. Next, we analysed individual vehicle trajectories in order to determine the reasons why the vehicles that cause the formation or growth of stop-and-go waves decelerate, disrupting traffic. Insight on the causes of stop-and-go wave growth may help us to understand the causes of stop-and-go wave formation.

2.4.1 Calculation of vehicle speeds

The speed of a given vehicle between two consecutive locations was calculated as the distance between those two locations divided by the time that the vehicle takes to travel between them. Therefore, the speed of vehicle n between the locations of cameras $i - 1$ and i is:

$$v_{n,i} = \frac{x_i - x_{i-1}}{t_n(x_i) - t_n(x_{i-1})} \quad (2.1)$$

where x_{i-1} and x_i are the locations of cameras $i - 1$ and i , and $t_n(x_i)$ and $t_n(x_{i-1})$ are the passing times of the rear bumper of vehicle n at locations x_i and x_{i-1} .

2.4.2 Identification of the causes of formation and growth of stop-and-go waves

A speed disturbance is a temporary decrease in the speed of vehicles passing a particular location on a particular lane. In this study, we defined a speed disturbance as a decrease in speed of 7 km/h or more within a short period of time, in line with Ahn and Cassidy (2007). A threshold of 7 km/h guarantees that small variations in vehicle speed are filtered out, whereas significant disturbances are taken into account. A sensitivity analysis of that threshold was beyond the scope of this study. The occurrence of a speed disturbance is generally a sign of the occurrence of a traffic flow perturbation. In uncongested traffic, if a disturbance does not destabilise traffic flow, it generally propagates downstream. If a disturbance destabilises traffic flow or traffic flow is already congested, the disturbance typically propagates upstream at a constant speed of 15-25 km/h, creating a stop-and-go wave (Schreiter et al., 2010). Sometimes, the amplitude of stop-and-go waves increases or decreases as they propagate (Ahn & Cassidy, 2007). To determine the triggering factors for the formation and growth of stop-and-go waves on the study site, we followed a multi-step method based on Ahn and Cassidy (2007) (see Figure 2.3).

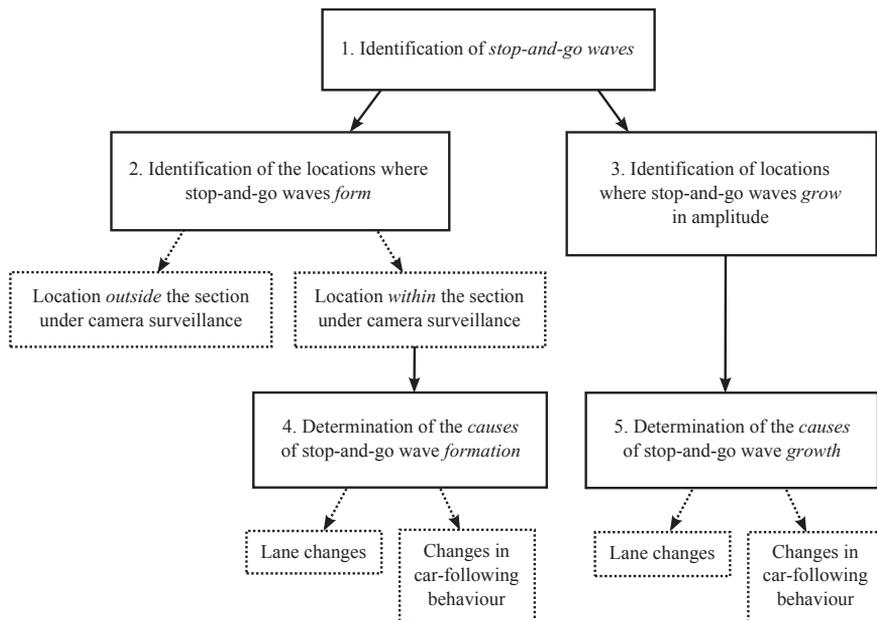
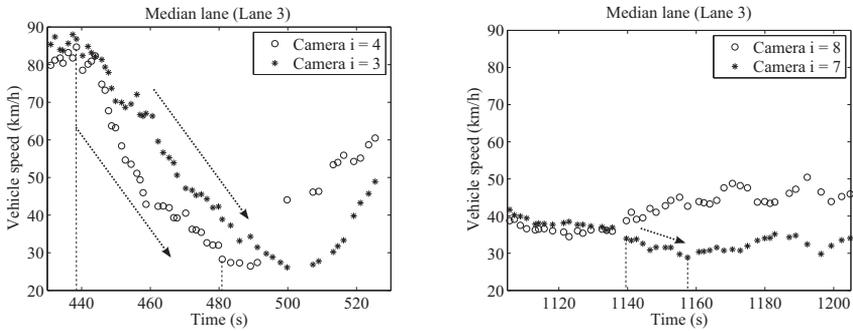


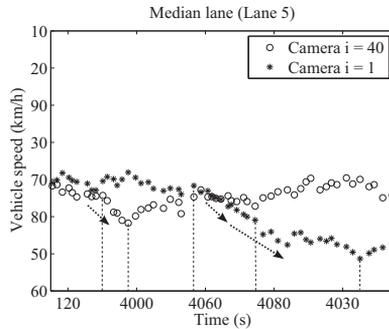
Figure 2.3: Steps to identify the causes of formation and growth of stop-and-go waves.

Identification of stop-and-go waves

To identify the presence of stop-and-go waves, we compared the evolution of vehicle speeds over time in all pairs of consecutive camera locations. This analysis was done separately for each lane. The presence of a stop-and-go wave results in a temporary decrease in the speed of vehicles passing a particular camera location over time. At the study site, the distance between consecutive video cameras is around 120 m. Therefore, if a stop-and-go wave propagates upstream at a speed of 15-25 km/h, a similar vehicle speed pattern is observed on the same lane at the next camera location in the upstream direction after 15-30 s. Figure 2.4(a) shows an example of stop-and-go wave. At the location of Camera 4, vehicle speed decreases from 85 to 28 km/h between $t = 438$ s and $t = 481$ s. A similar speed pattern is observed at the location of Camera 3 (upstream) with a time lag of around 15 s.



(a) Stop-and-go wave propagating upstream. Camera 3 is located upstream of Camera 4. (b) Formation of a stop-and-go wave. Camera 7 is located upstream of Camera 8.



(c) Stop-and-go wave propagating upstream and growing in amplitude as it propagates. Camera 9 is located upstream of Camera 10.

Figure 2.4: Examples of stop-and-go wave (a) propagation, (b) formation and (c) growth.

Identification of the locations where stop-and-go waves form or grow in amplitude

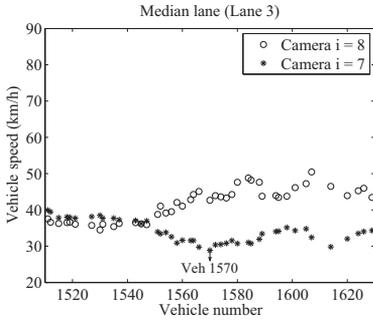
Once the presence of a stop-and-go wave was identified, we determined the location where the wave originated (if that location is within the area under camera surveillance). At that location, vehicle speed decreases without the presence of a similar speed pattern 15-30 s earlier on the same lane at the previous downstream camera location. We defined that a stop-and-go wave forms at a particular location if the speed at that location reaches values more than 7 km/h lower than at the preceding downstream camera location during the previous 30 s (Ahn & Cassidy, 2007). Figure 2.4(b) shows an example of formation of a stop-and-go wave, which later will propagate upstream (the latter cannot be observed in the figure). Vehicle speed stays between 35 and 40 km/h at the locations of Camera 7 and Camera 8 between $t = 1110$ s and $t = 1140$ s. At $t = 1140$ s, speed starts to decrease at the location of Camera 7, reaching 28 km/h at $t = 1158$ s. However, speed does not decrease at the location of Camera 8 (downstream) during the previous 30 s. Actually, first, speed stays constant and later it increases. Therefore, a stop-and-go wave forms between the locations of Camera 7 and Camera 8 without the influence of any downstream trigger.

We also determined the locations within the camera surveillance area where stop-and-go waves grow in amplitude as they propagate upstream. The growth of a stop-and-go wave at a particular location results in a significantly greater decrease in speed than at the previous camera location. We defined that a stop-and-go wave grows at a particular location if it causes the speed at that location to decrease to values more than 7 km/h lower than on the same lane at the previous downstream camera location (Ahn & Cassidy, 2007). Figure 2.4(c) shows an example of stop-and-go wave growth. At the location of Camera 10, vehicle speed drops from 47 to 38 km/h between $t = 990$ s and $t = 998$ s. At the location of Camera 9 (upstream), speed follows a similar pattern with a time lag of around 30 s: speed drops from 49 to 39 km/h between $t = 1017$ s and $t = 1034$ s. However, after $t = 1034$ s, speed does not increase (as it does at the location of Camera 10 after $t = 998$ s), but it keeps decreasing, reaching 28 km/h at $t = 1065$ s. This indicates that the stop-and-go wave grows in amplitude between the locations of Camera 9 and Camera 10.

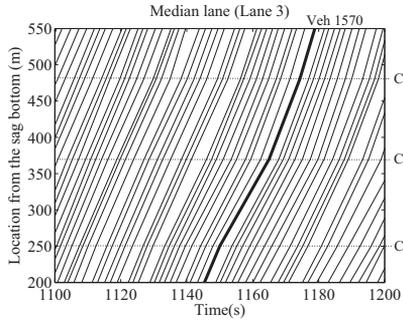
Determination of the causes of stop-and-go wave formation and growth

Once we identified the locations where stop-and-go waves form or grow, we determined the causes why they do so. To this end, we followed a two-step approach. First, we identified the vehicles that decelerate and cause the formation or growth of each stop-and-go wave. This was done by manually comparing the speeds of each individual vehicle in each pair of consecutive camera locations on the same lane (see Figures 2.5(a) and 2.5(c)). Second, we determined the cause why those vehicles decelerate. This was done by manually analysing individual vehicle trajectories to check whether any vehicles move to the subject lane in front of the vehicles that decelerate

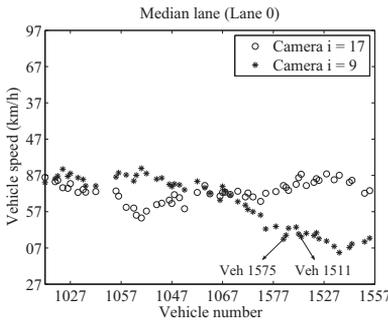
and cause the formation or growth of each stop-and-go wave. If that is the case, we concluded that the cause is one or more lane-changing manoeuvres (see Figure 2.5(d)). If that is not the case, we concluded that the cause is related to car-following behaviour (see Figure 2.5(b)). Note that we assumed that both types of causes are mutually exclusive.



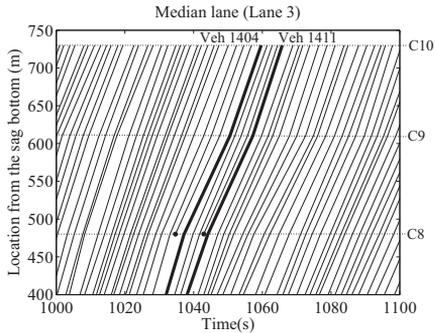
(a) Example of formation of a stop-and-go wave. Speed per vehicle number at the locations of cameras 7 and 8 on the median lane. Camera 7 is located upstream of Camera 8.



(b) Example of formation of a stop-and-go wave. Vehicle trajectories between the locations of cameras 6, 7 and 8 on the median lane.



(c) Example of stop-and-go wave growth. Speed per vehicle number at the locations of cameras 9 and 10 on the median lane. Camera 9 is located upstream of Camera 10.



(d) Example of stop-and-go wave growth. Vehicle trajectories between the locations of cameras 8, 9 and 10 on the median lane. Small circles represent lane changes performed before a particular camera location.

Figure 2.5: Identification of the causes of stop-and-go wave formation and growth.

2.4.3 Test of statistical significance

The primary factor triggering stop-and-go waves at the study site was identified in two steps. First, we counted the number of instances in which the causes of stop-and-go wave formation or growth are related to: i) car-following behaviour; and ii) disruptive lane changes. Thus, we identified the most frequent cause. Second, we

tested whether the frequency of stop-and-go waves caused by car-following behaviour phenomena was significantly different from the frequency of stop-and-go waves caused by lane changes. This is comparable to testing whether a coin is fair: from a series of outcomes, it is tested whether the relative frequency of one of the sides is so high that it can be considered unlikely. A two-tailed binomial test is suitable to do so. Based on a series of outcome observations, that type of test determines whether: a) two outcome categories are equally likely to occur (null hypothesis); or b) two outcome categories are not equally likely to occur (alternative hypothesis) (Moore et al., 2011). Applied to our case, the test shows whether observed differences in frequency between the two causes of stop-and-go waves result purely from chance or from the fact that one cause is actually more frequent than the other. The p -value indicates how unlikely the outcome should be to reject the null hypothesis. If the p -value is higher than the significance level (for which we chose 5%), the null hypothesis cannot be rejected. Instead, if the p -value is lower than the significance level, the null hypothesis is rejected.

2.5 Results

In this section, we present the results of the analysis of the causes of stop-and-go wave formation and growth at the study site.

2.5.1 Identification of stop-and-go waves

In total, 13 stop-and-go waves have been identified in the data, all of them on the centre and median lanes. No stop-and-go wave has been identified on the shoulder lane. All the identified stop-and-go waves have their origin either on the uphill section or farther downstream (see Fig. 2.6). Five stop-and-go waves have been identified on the centre lane, of which four have their origin within the section under camera surveillance and the other one originates farther downstream (see Figure 2.6(a) and Table 2.1). Eight stop-and-go waves have been identified on the median lane, of which five have their origin within the area under surveillance and the rest originate farther downstream (see Figure 2.6(b) and Table 2.1). Note that some of the stop-and-go waves seem to occur simultaneously on the centre and median lanes (compare Figure 2.6(a) with Figure 2.6(b)). Seven of the 13 stop-and-go waves that have been identified on the centre and median lanes form in uncongested traffic (waves 1, 2, 6, 7, 8, 9 and 10 in Figures 2.6(a) and 2.6(b)). The remaining six waves form in congested traffic (waves 3, 4, 5, 11, 12 and 13 in Figures 2.6(a) and 2.6(b)). The traffic speed indicates whether traffic flow is congested or uncongested (see also Patire (2010)).

2.5.2 Causes of stop-and-go wave formation

The cause of formation of eight of the nine stop-and-go waves that have their origin within the section under camera surveillance (both in the centre and median lanes) is related to car-following behaviour (see Figures 2.5(a), 2.5(b), 2.6(a), 2.6(b) and Table 2.1). One stop-and-go wave on the centre lane is triggered by lane-changing vehicles coming from the shoulder lane (see Figure 2.6(a) and Table 2.1). The two-tailed binomial test shows that the causes of stop-and-go wave formation are more frequently related to car-following behaviour than to lane changes at the 5% significance level (p -value = $0.04 < 0.05$).

Table 2.1: Causes of stop-and-go wave formation and growth (CF stands for car-following behaviour; LC stands for lane changes).

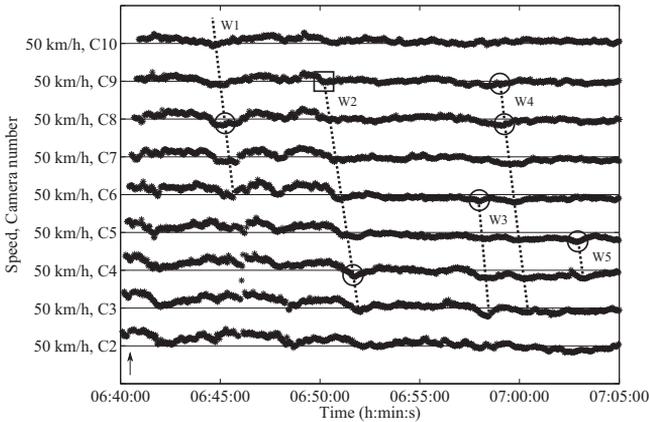
	Center lane			Median lane		
	Origin identified	Origin not identified	Total	Origin identified	Origin not identified	Total
Stop-and-go waves	4	1	5	5	3	8
Cause of formation (CF/LC)	3/1			5/0		
Stop-and-go wave amplifications			3			6
Cause of amplification (CF/LC)			3/0			5/1

2.5.3 Causes of stop-and-go wave growth

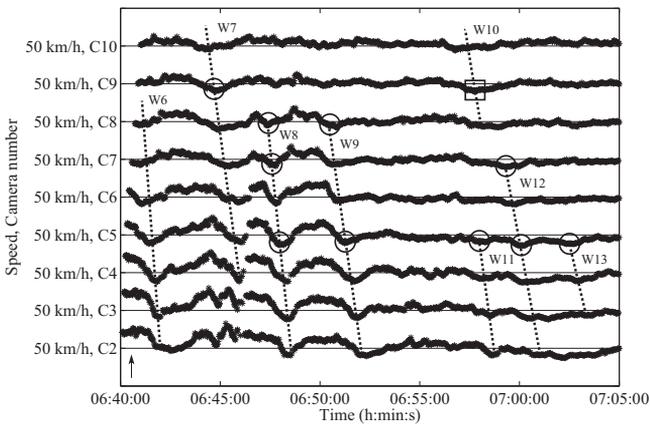
The 13 stop-and-go waves observed in the data set have been identified to grow in amplitude in nine cases. In eight of those nine cases, the cause of growth is related to car-following behaviour (see Figures 2.6(a), 2.6(b) and Table 2.1). One stop-and-go wave grows on the median lane as a result of lane-changing manoeuvres performed by vehicles coming from the centre lane (see Figures 2.5(c), 2.5(d) and 2.6(b) and Table 2.1). The two-tailed binomial test shows that the causes of stop-and-go wave growth are more frequently related to car-following behaviour than to lane changes at the 5% significance level (p -value = $0.04 < 0.05$).

2.6 Discussion

The results presented in Section 2.5 must be interpreted with caution, due to the reduced sample size and the limited scope of the data. First, the sample size is small (only 13 stop-and-go waves were identified in the data), which may reduce the significance of the results. The results would be more conclusive if additional trajectory data with more stop-and-go waves were available. Second, we analysed microscopic flow data of one particular sag. The causes of stop-and-go waves can be site-specific, therefore our findings can be generalised to other sags only to a certain extent. An



(a) Center lane.



(b) Median lane.

Figure 2.6: Speed deviation (from 50 km/h) of vehicles passing each camera location on the center and median lanes. Dashed lines indicate stop-and-go waves, which are numbered (W1, W2, etc.). Circles and squares indicate locations where stop-and-go waves form or grow in amplitude. Circles indicate that the cause for stop-and-go wave formation/growth is related to car-following behaviour. Squares indicate that the cause for stop-and-go wave formation/growth is related to disruptive lane changes. Camera 4 is located at the bottom of the sag.

analysis of data from additional sites is necessary to draw a final conclusion regarding the transferability of our findings to other sags. If possible, the additional sites should be from other countries and include sags with different vertical alignment profiles and different number of lanes than our study site.

In spite of those data limitations, the findings presented in this study are important to understand traffic flow dynamics at sags. Our results indicate that the primary factor triggering stop-and-go waves at sags is related to car-following behaviour, whereas lane changes are a less significant triggering factor. We found evidence that lane-changing vehicles can disrupt traffic and trigger stop-and-go waves at sags, as suggested by other authors (Brilon & Bressler, 2004). However, at our study site, the formation and growth of stop-and-go waves seem to be caused by lane changes only in a few cases (11%). In most cases (89%), the causes of stop-and-go wave formation and growth are related to car-following behaviour. This difference in frequency is statistically significant. Our results seem to be in line with the findings of Zheng et al. (2011), who found that in most cases (66%) the cause of stop-and-go waves forming on the uphill section of a US freeway was related to car-following behaviour. The remaining waves (33%) were triggered by lane changes. That site also contains a merge and a diverge, which may explain why the percentage of stop-and-go waves triggered by lane changes is higher than in our study site.

Some studies found evidence that car-following behaviour changes significantly on the lower part of the uphill section at sags. Drivers tend to decelerate and keep longer headways than expected given their speed (Koshi et al., 1992; Furuichi et al., 2003; Koshi, 2003; Yoshizawa et al., 2012). Goñi-Ros et al. (2013a) found evidence of those changes in car-following behaviour in the same trajectory data set used in this paper. Some authors suggest that those local changes in car-following behaviour are the primary triggering factor of stop-and-go waves at sags (Koshi et al., 1992; Koshi, 2003). It is important to note that due to the characteristics of our data and methodology, we could only determine whether the causes of stop-and-go wave formation and growth are related to lane changes or to car-following behaviour. The hypothesis that the primary factor triggering stop-and-go waves at sags are the changes in car-following behaviour caused by the change in freeway gradient seems plausible. However, further research is necessary to identify the characteristics of the car-following behaviour phenomena that trigger stop-and-go waves at sags. For instance, some of the stop-and-go waves identified in this paper seem to occur simultaneously on the centre and median lanes; the cause for that may be that drivers adapt their car-following behaviour to the traffic conditions on the adjacent lanes, triggering stop-and-go waves on their lane.

Alternatively, the cause for the simultaneous occurrence of stop-and-go waves on different lanes may be related to lateral driving behaviour. As explained in Section 2.2, when a stop-and-go wave emerges on a given lane, some drivers move to the other lanes in order to avoid stopping. Those lane changes may spread congestion to the target lanes (Koshi et al., 1992; Hatakenaka et al., 2006; Patire & Cassidy, 2011). Even if those lane changes do not directly disrupt traffic on the target lanes, they do cause an increase in flow. With higher flow, perturbations related to car-following behaviour are more likely to destabilise traffic flow and trigger stop-and-go waves.

2.7 Conclusions

Stop-and-go waves frequently emerge on the uphill section of sags, both in uncongested and congested traffic conditions. Previous studies suggest that the causes of stop-and-go waves at sags are related to: a) car-following behaviour; and/or b) disruptive lane changes. However, it is not clear which of those two triggering factors is dominant. We analysed vehicle trajectories collected by means of video cameras on a three-lane sag of the Tomei Expressway (Japan), identifying the causes of stop-and-go wave formation and growth on the study site. The results show that in most cases the factor triggering stop-and-go waves is related to car-following behaviour; disruptive lane changes are a less significant triggering factor. The car-following behaviour phenomena that trigger stop-and-go waves at sags may be related to the changes in car-following behaviour caused by the change in freeway slope, as suggested in the scientific literature. However, further research is necessary to validate that hypothesis. Our findings have important implications for the development of measures aimed at mitigating oscillatory traffic. Particularly, our findings show the relevance of developing systems to assist drivers in performing the acceleration task at sags (Goñi-Ros et al., 2012).

Chapter 3

Modeling traffic at sags

Goñi-Ros, B., Knoop, V.L., Shiomi, Y., Takahashi, T., van Arem, B., Hoogendoorn, S.P. (2016). Modeling Traffic at Sags. *International Journal of Intelligent Transportation Systems Research*, 14(1), 64–74. © Springer Science+Business Media.

Abstract

Freeway capacity decreases at sags due to local changes in car-following behavior. Consequently, sags are often bottlenecks in freeway networks. This article presents a microscopic traffic model that reproduces traffic flow dynamics at sags. The traffic model includes a new car-following model that takes into account the influence of freeway gradient on vehicle acceleration. The face-validity of the traffic model is tested by means of a simulation study. The study site is a sag of a Japanese freeway. The simulation results are compared to empirical traffic data presented in previous studies. We show that the model is capable of reproducing the key traffic phenomena that cause the formation of congestion at sags, including the lower capacity compared to normal sections, the location of the bottleneck around the end of the vertical curve, and the capacity drop induced by congestion. Furthermore, a sensitivity analysis indicates that the traffic model is robust enough to reproduce those phenomena even if some inputs are modified to some extent. The sensitivity analysis also shows what parameters need to be calibrated more accurately for real world applications of the model.

3.1 Introduction

The *gradient* of a freeway is the inclination of the freeway surface to the horizontal. Sags or sag vertical curves are freeway sections along which the gradient increases

gradually in the direction of traffic. The capacity of sags is generally lower than that of normal freeway sections (Koshi et al., 1992; Okamura et al., 2000). Consequently, sags often cause congestion in high traffic demand conditions (Koshi et al., 1992). The main reason why the freeway capacity decreases at sags is related to local changes in car-following behavior (Furuichi et al., 2003; Koshi, 2003; Brilon & Bressler, 2004; Yoshizawa et al., 2012; Goñi-Ros et al., 2014a).

In order to evaluate the effectiveness of possible traffic management measures aimed at mitigating congestion at sags, it is necessary to be able to simulate traffic at that type of bottlenecks in a realistic way. This paper presents a microscopic traffic model that reproduces traffic flow dynamics at sags. The model consists of two sub-models: a car-following model and a lane change model. The main novelty is the car-following modeling approach, which is based on assumptions about the way in which sags affect longitudinal driving behavior that are more realistic than those of existing models, and has a very generic formulation.

The face-validity of the traffic model is tested by means of a simulation study. The study site is the Yamato sag (Tomei Expressway, Japan). The traffic flow patterns obtained from simulation are compared to the patterns observed in empirical traffic data according to previous studies. Furthermore, we carry out a sensitivity analysis aimed at evaluating the robustness of the model to changes in key model inputs, and at determining the way in which modifying each of those model inputs influences the characteristics of the simulated traffic.

The rest of this article is structured as follows: Section 3.2 describes the main causes of congestion at sags according to the scientific literature; Section 3.3 specifies the requirements of the traffic model and analyzes the suitability of existing models; Section 3.4 presents the microscopic traffic model; Section 3.5 presents the methodology used to test the face-validity of the traffic model; Section 3.6 reports the results of the analysis; and Section 3.7 presents the conclusions of this study.

3.2 Causes of congestion at sags

Various empirical studies show that the capacity of sags is considerably lower than the capacity of normal sections (up to 30% lower) (Koshi et al., 1992; Okamura et al., 2000). Because of that, sags are often bottlenecks in freeway networks, hence they cause the formation of congestion in conditions of high traffic demand. In general, the bottleneck is located 500 to 1,000 m downstream of the bottom of the sag (Brilon & Bressler, 2004). The factors reducing the capacity of sags seem to be related primarily to two changes in car-following behavior that occur when vehicles go through the vertical curve: i) drivers tend to reduce speed (Koshi et al., 1992; Furuichi et al., 2003; Brilon & Bressler, 2004); and ii) drivers tend to keep longer headways than expected given their speed (Koshi, 2003; Yoshizawa et al., 2012; Goñi-Ros, et al., 2014a). These changes in car-following behavior seem to be unintentional (Yoshizawa et al., 2012).

They are caused by a decrease in vehicle acceleration resulting from the combination of two factors: i) increase in resistance force; and ii) insufficient acceleration operation by drivers (Koshi et al., 1992; Yoshizawa et al., 2012). Drivers fail to accelerate sufficiently even though they generally perceive the increase in slope (Furuichi et al., 2003; Yoshizawa et al., 2012). The reason why drivers do not accelerate sufficiently seems to be related to their throttle operation behavior: drivers generally push down the throttle pedal at the beginning of the vertical curve but it takes time for them to adjust the throttle position so as to fully compensate for the increase in resistance force (Yoshizawa et al., 2012). Drivers are generally able to re-accelerate and recover their desired speed once they leave the vertical curve (Brilon & Bressler, 2004; Yoshizawa et al., 2012).

Typically, in freeways with keep-left or keep-right rules, the process of congestion formation at sags consists of two phases. In the first phase, congestion forms on the median lane (Koshi et al., 1992; Hatakenaka et al., 2006; Patire & Cassidy, 2011). The main reason why congestion emerges first on the median lane is related to the characteristics of lane flow distribution: with high demand and uncongested traffic, flows tend to be higher (and closer to capacity) on the median lane than on the other lanes (Hatakenaka et al., 2006; Xing et al., 2010). In the second phase, congestion spreads from the median lane to the other lanes (Koshi et al., 1992; Hatakenaka et al., 2006; Patire & Cassidy, 2011). That process can be described as follows. When traffic becomes congested on the median lane, some vehicles migrate from that lane to the less crowded lanes in order to avoid queuing (Hatakenaka et al., 2006; Patire & Cassidy, 2011). When the flow on those lanes exceeds their capacity, traffic also breaks down there. At that point, traffic is congested on all lanes, which causes a significant decrease in total outflow (due to the capacity drop phenomenon (Hall & Agyemang-Duah, 1991) and the formation of a queue upstream of the bottleneck (Koshi et al., 1992; Patire & Cassidy, 2011).

3.3 Model requirements and suitability of existing models

As discussed in Section 3.2, the main traffic phenomena involved in the formation of congestion at sags are: a) decrease in vehicle acceleration and local changes in car-following behavior; b) low capacity of sags compared to normal sections (the bottleneck being the end of the vertical curve); c) congestion-induced capacity drop; d) queue dynamics; e) uneven lane flow distribution in high demand conditions; and f) migration of vehicles from congested to uncongested lanes. Any model that aims to reproduce traffic at sags in a realistic way should be capable of reproducing all the traffic phenomena mentioned above.

In order to reproduce phenomena e and f, the traffic model needs to include a lane change model. Note that there is no empirical evidence suggesting that lane-changing

behavior is different at sags than in normal sections. For that reason, we conclude that it is sufficient to use a regular lane-change model, in which the set of variables that determine the decision to change lanes does not include the freeway gradient (it includes variables such as the difference in speed between lanes).

In order to reproduce phenomena a, b, c and d, the traffic model needs to include a car-following model. As discussed in Section 3.2, drivers significantly change their car-following behavior when they go through a sag, which is the main cause of the difference in capacity between sags and normal sections. Therefore, the car-following model needs to include the freeway gradient as separate variable in addition to the variables that determine regular car-following behavior (e.g., speed and distance to the preceding vehicle). Furthermore, the car-following model needs to adequately reproduce the influence of the freeway gradient on vehicle acceleration. Particularly, the model needs to reproduce: i) the limiting effect that an increase in gradient has on vehicle acceleration; ii) the inability of drivers to fully compensate for that limiting effect immediately; iii) the resulting changes in car-following behavior at sags (i.e., speed reduction and increase in headways); and iv) the creation of a capacity bottleneck at the end of sag vertical curves. In addition, the car-following model also needs to reproduce the ability of drivers to gradually regain their normal car-following behavior once they leave the vertical curve.

Several car-following models have been developed in the last decades with the objective of reproducing longitudinal driving behavior at sags. Those models can be grouped into two broad categories based on whether they assume that drivers do or do not explicitly compensate for the limiting effect that an increase in freeway gradient has on vehicle acceleration. Examples of models that assume no explicit compensation are those proposed by Koshi et al. (1992) and Komada et al. (2009). Both models assume that a constant positive slope has a constant negative influence on vehicle acceleration. However, that assumption is not consistent with empirical observations, which show that drivers generally regain their normal car-following behavior as they climb an uphill section (at least if the uphill section is not too steep) (Brilon & Bressler, 2004; Yoshizawa et al., 2012).

Examples of models that assume that drivers explicitly compensate for the limiting effect that an increase in gradient has on vehicle acceleration are those proposed by Yokota et al. (1998) and Oguchi and Konuma (2009). Yokota et al. (1998) present a car-following model that assumes that drivers are able to fully compensate for changes in gradient with a certain time delay. Hence a constant slope has a decreasing influence on vehicle acceleration. This is more in line with empirical observations. However, an important disadvantage of that model is that it does not accurately reproduce the location of the bottleneck at sags. The model generates a bottleneck around the bottom of the sag (Yokota et al., 1998), whereas empirical observations show that the bottleneck is generally located around the end of the vertical curve (Brilon & Bressler, 2004; Patire & Cassidy, 2011).

Oguchi and Konuma (2009) present a model that assumes that drivers compensate for

the increase in resistance force caused by an increase in freeway gradient in such a way that the sensitivity of the vehicle acceleration to that increase in resistance force decreases over time as the vehicle moves along the vertical curve and the subsequent uphill section. The authors show that their model reproduces the car-following behavior of drivers (as observed in a driving simulator environment) better than models assuming that drivers do not explicitly compensate for the limiting effect that an increase in gradient has on vehicle acceleration (Oguchi & Konuma, 2009). Oguchi and Konuma do not present simulated traffic data, but one may expect that with appropriate parameter values their model generates a bottleneck at the end of sag vertical curves, which is the location where the limitation on vehicle acceleration caused by the gradient term is stronger. However, one drawback of the model proposed by Oguchi and Konuma (2009) is that it is not sufficiently generic. More specifically, the values of some parameters (such as T_a and T_W) necessarily depend on the vertical profile of the sag.

We conclude that in order to model traffic at sags, it is sufficient to use a regular lane change model, but it is necessary to develop a new car-following model. That model should reproduce the characteristics of car-following behavior at sags and their effect on freeway capacity in a realistic way. To that end, the main principle of the modeling approach proposed by Oguchi and Konuma (2009) seems adequate. However, that principle should be re-formulated in such a way that the model parameters are independent of the vertical profile of the freeway, in order to make the model generic for all sags.

3.4 Microscopic traffic model

The proposed microscopic traffic model consists of two sub-models: i) a car-following model; and ii) a lane change model. The two sub-models are presented in Section 3.4.1 and Section 3.4.2, respectively. The inputs required by the traffic model are described in Section 3.4.3.

3.4.1 Car-following model

We developed a new car-following model that takes into account the influence of vertical curves on vehicle acceleration. The model determines the vehicle acceleration by means of a two-term additive function:

$$\dot{v}(t) = f_r(t) + f_g(t) \quad (3.1)$$

The first term in Eq. 3.1 (f_r) describes regular car-following behavior. Its formulation is based on the Intelligent Driver Model (IDM) (Treiber et al., 2000). This term accounts for the influence of speed (v), relative speed (Δv) and spacing (s) on vehicle

acceleration:

$$f_r(t) = a \cdot \left[1 - \left(\frac{v(t)}{v_{\text{des}}} \right)^4 - \left(\frac{s_{\text{des}}(v(t), \Delta v(t))}{s(t)} \right)^2 \right] \quad (3.2)$$

where the dynamic desired spacing (s_{des}) is:

$$s_{\text{des}}(v(t), \Delta v(t)) = s_s + v(t) \cdot T(v(t)) + \frac{v(t) \cdot \Delta v(t)}{2\sqrt{a \cdot b}} \quad (3.3)$$

and the safe time headway (T) depends on the traffic state:

$$T(v(t)) = \begin{cases} T_F & \text{if } v(t) \geq v_{\text{crit}} \\ \gamma \cdot T_F & \text{if } v(t) < v_{\text{crit}} \end{cases} \quad (3.4)$$

The parameters in Eqs. 3.2-3.4 are: desired speed (v_{des}); maximum acceleration (a); maximum comfortable deceleration (b); net spacing at standstill (s_s); safe time headway in uncongested traffic conditions (T_F); congestion factor on safe time headway ($\gamma > 1$); and critical speed (v_{crit}). The critical speed is used as threshold to differentiate between congested and uncongested traffic.

The second term in Eq. 3.1 (f_g) accounts for the influence of vertical curves on vehicle acceleration. At a given time t , that influence is the difference between the gradient at the location where the vehicle is at that time ($G(x(t))$) and the gradient compensated by the driver until that time ($G_c(t)$), multiplied by a sensitivity parameter (θ):

$$f_g(t) = -\theta \cdot [G(x(t)) - G_c(t)] \quad (3.5)$$

The compensated gradient (G_c) is a variable that accounts for the fact that drivers have a limited ability to compensate for the negative effect that an increase in gradient has on vehicle acceleration. The model assumes that drivers compensate for any increase in freeway gradient linearly over time with a maximum gradient compensation rate defined by parameter c . The value of parameter c does not depend on the vertical profile of the sag. This assumption is based on findings by Yoshizawa et al. (2012), who found that, when drivers go through a sag, they push down the throttle pedal at a similar rate regardless of the vertical profile of the sag. Note that the model assumes that drivers are able to fully compensate for any decrease in gradient (e.g., at crest vertical curves) immediately. Therefore, the compensated gradient at a given time t is defined as follows:

$$G_c(t) = \begin{cases} G(x(t)) & \text{if } G(x(t)) \leq G(t_c) + c \cdot (t - t_c) \\ G(t_c) + c \cdot (t - t_c) & \text{if } G(x(t)) > G(t_c) + c \cdot (t - t_c) \end{cases} \quad (3.6)$$

where t_c is the time when the driver could no longer fully compensate for the gradient change:

$$t_c = \max(t \mid G_c(t) = G(x(t))) \quad (3.7)$$

The properties of the model are as follows. If the gradient profile of a sag is such that the rate at which the freeway gradient increases is lower than the driver's maximum

gradient compensation rate (c), then $G - G_c = 0$ (hence $f_g = 0$) at any time t . Consequently, at very gentle sags, the increase in freeway gradient has no effect on vehicle acceleration, so acceleration is determined only by the regular car-following behavior term (f_r). However, at sags where the rate at which the gradient increases is higher than the driver's maximum gradient compensation rate (c), then $G - G_c > 0$ for a certain period of time. During that period, the compensated gradient (G_c) increases linearly over time but f_g is negative, which limits vehicle acceleration. It is important to remark two things about the latter type of sags. First, the sharper the sag, the greater the value of $G - G_c$ along the sag, hence the stronger the vehicle acceleration limitation. Second, of all locations along the freeway, vehicle acceleration limitation is maximum at the location where the gradient increase rate becomes lower than the driver's maximum gradient compensation rate (i.e., the end of the sag vertical curve), because $G - G_c$ is maximum at that location.

Note that the formulation of the second term of Eq. 3.1 (f_g) is based on a similar principle to that of the gradient term of the model presented by Oguchi and Konuma (2009). However, in our model, the values of the parameters of the gradient term (i.e., c and θ) do not necessarily depend on the vertical profile of the freeway. This makes our model more generic than the model presented by Oguchi and Konuma.

3.4.2 Lane change model

To model lane-changing behavior we used the Lane Change Model with Relaxation and Synchronization (LMRS) (Schakel et al., 2012). Note that we modified the LMRS with regard to the calculation of the speed gain desire incentive: in our model, the speeds of all preceding vehicles that are not farther ahead than x_0 meters have the same weight in calculating the anticipated speed of downstream traffic. This modification allows drivers to anticipate more accurately the speed of downstream traffic when they drive into a queue and they have to decide whether to change lanes.

3.4.3 Model inputs

The traffic model requires the user to provide the following inputs: i) simulation period; ii) characteristics of the freeway; iii) traffic demand profile; and iv) traffic composition. The freeway characteristics include length, number of lanes, vertical profile (i.e., degree of gradient over distance) and speed limits. The traffic demand profile specifies the flow entering the simulated freeway stretch over time (per lane). The traffic composition describes the characteristics of the vehicle-driver units that enter the freeway. The vehicle characteristics taken into account by the model are vehicle type (passenger car, truck, etc.) and vehicle length (l). The driver characteristics are the parameters of the car-following model and the lane change model.

3.5 Model verification methodology

The face-validity of the traffic model described in Section 3.4 was tested by means of a simulation study. The objective was to determine the model's ability to realistically reproduce traffic flow dynamics at sags. First, we defined a freeway stretch containing a sag and we simulated traffic using our traffic model. We analyzed the traffic flow patterns obtained from simulation and we compared them with the patterns observed in empirical data. Second, we performed a sensitivity analysis. The objectives of that analysis were to determine: i) the robustness of the model to changes in key model inputs; and ii) the way in which modifying each of those model inputs influences the simulation results. The general simulation setup is presented in Section 3.5.1. The scenarios of the sensitivity analysis are defined in Section 3.5.2. Finally, Section 3.5.3 describes the indicators used to analyze the simulation results.

3.5.1 Simulation setup

Freeway characteristics The simulated freeway stretch is 10 km long and has three lanes (median, center and shoulder). Vehicles enter the freeway at location $x = 0$; the exit point is at location $x = 10$ km. The freeway stretch has three sections: i) constant-gradient downhill section; ii) sag vertical curve; and iii) constant-gradient uphill section. The first and third sections have a constant gradient equal to -0.5 and 2.5%, respectively. The vertical profile of the sag is equivalent to that of the Yamato sag (Tomei Expressway, Japan): the gradient increases linearly from -0.5 to +2.5%, and the length of the vertical curve is 600 m (see Fig. 3.1). The constant-gradient downhill section is long enough to ensure that if congestion forms at the sag, the queue does not spill back to the freeway entry point. The speed limit is 100 km/h for cars and 85 km/h for trucks on the whole freeway stretch. There are no on-ramps, off-ramps, lane drops or horizontal curves.

Simulation period and traffic demand profile The simulation period is 100 min. The total traffic demand (on all lanes) increases linearly from 3,000 to 5,200 veh/h between $t = 0$ and $t = 75$ min. From $t = 75$ to $t = 100$ min, the total inflow stays at 5,200 veh/h. Given the total demand, the demand per lane was defined based on a lane flow distribution model presented by Wu (2006). We used the same parameter values provided in Wu (2006) for three-lane freeways (see Table 6 of that reference), except for a and e on the center lane (we used $a = 0.39$ and $e = 0.30$). We slightly modified the values of those two parameters to make the lane flow distribution model more accurate for Japanese freeways, based on empirical data presented by Xing et al. (2010). Note that after entering the freeway, vehicles are free to change lanes in accordance with the lane change model.

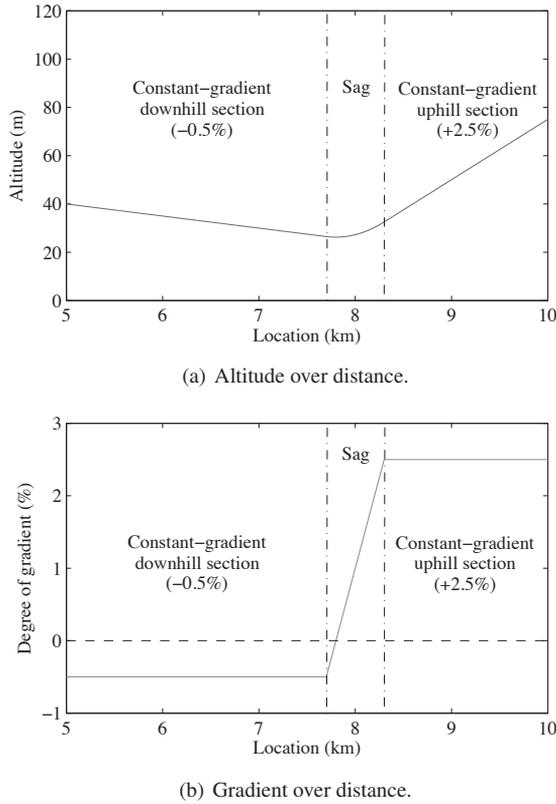


Figure 3.1: Vertical profile of the freeway stretch.

Traffic composition We defined two types of vehicles (cars and trucks), which have different vehicle length (4 and 15 m, respectively). Furthermore, we defined three types of car drivers (each corresponding to one lane) and one type of truck driver. Defining one type of car driver per lane was necessary to take into account the differences in desired speed and target time headway between lanes observed in empirical traffic data (Patire & Cassidy, 2011). The parameters of the car-following model and the lane change model are different for each driver type (see Table 1). For the interpretation of the lane change model parameters, we refer to Schakel et al. (2012). Table 3.2 shows the percentage of vehicles entering the freeway on each lane that belong to each vehicle and driver type. Note that we defined some parameters of the car-following model and the lane change model as stochastic parameters (v_{des} , a , b , T_F , c , T_{min} for car drivers, and v_{des} for truck drivers). The value of those parameters differs between drivers belonging to the same driver type. For car drivers, those parameters depend on the stochastic factor δ , which differs between drivers (it is normally distributed with mean δ^* and standard deviation σ_δ): $v_{\text{des}} = \delta \cdot v_{\text{des},0}$; $a = \delta \cdot a_0$; $b = \delta \cdot b_0$; $T_F = T_{F,0}/\delta$; $c = \delta \cdot c_0$; and $T_{\text{min}} = T_{\text{min},0}/\delta$. Note that the factor δ is defined per driver, hence all stochastic parameters corresponding to a particular driver are correlated. For

trucks, only the desired speed parameter is stochastic. That parameter is assumed to be normally distributed with mean $v_{des,t}^*$ and standard deviation $\sigma_{vdes,t}$.

Data collection system We placed virtual loop detectors every 100 m all along the freeway stretch, on all lanes. Those detectors measure flow and average speed. The data aggregation period is 30 s.

Table 3.1: Parameters of the car-following model and the lane change model (base scenario)

Vehicle type	Car			Truck
	Car driver 1	Car driver 2	Car driver 1	Truck driver
a_0 (m/s ²)	1.25	1.25	1.25	0.50
b_0 (m/s ²)	1.80	1.80	1.80	1.50
$T_{F,0}$ (s)	1.45	1.20	1.15	1.50
s_s (m)	3	3	3	3
$v_{des,0}$ (km/h)	100	100	100	
v_{crit} (km/h)	60	60	60	60
c_0 (s ⁻¹)	0.00042	0.00042	0.00042	0.00042
θ (m/s ²)	9.81	9.81	9.81	9.81
γ (-)	1.15	1.15	1.15	1.15
δ^* (-)	0.92	0.97	1.03	1.00
σ_δ (-)	0.03	0.10	0.10	0.00
$v_{des,t}^*$ (km/h)				85
$\sigma_{vdes,t}$ (km/h)				2.5
$T_{min,0}$ (s)	0.56	0.56	0.56	0.56
τ (s)	25	25	25	25
x_0 (m)	200	200	200	200
v_{gain} (km/h)	70	50	50	70
d_{free}^{ij} (-)	0.365	0.365	0.365	0.365
d_{sync}^{ij} (-)	0.577	0.577	0.577	0.577
d_{coop}^{ij} (-)	0.788	0.788	0.788	0.788

Table 3.2: Traffic composition

Vehicle type	Car	Car	Car	Truck
	Car driver 1	Car driver 2	Car driver 3	Truck driver
Shoulder lane	90%	0%	0%	10%
Center lane	0%	95%	0%	5%
Median lane	0%	0%	100%	0%

3.5.2 Sensitivity analysis

To perform the sensitivity analysis, we defined a base scenario and several alternative scenarios. In the base scenario, the model inputs are those defined in Section 3.5.1. In the alternative scenarios, the model inputs are the same as in the base scenario, except for one of them, which is changed by a certain percentage. The modified inputs are the vertical curve length (L), the maximum gradient compensation rate (c_0), the congestion factor on target time headway (γ), and the sensitivity to the difference between gradient and compensated gradient (θ) (see Table 3.3). Note that we also defined a reference scenario, in which it is assumed that the sag vertical curve has no influence whatsoever on the acceleration behavior of drivers (i.e., $f_g(t) = 0$ at any time t). That hypothetical behavior is modeled by setting the value of the maximum gradient compensation rate parameter to a very high value ($c_0 = 999 \text{ s}^{-1}$), leaving the other model inputs as in the base scenario. The reference scenario is used as input to analyze the traffic data obtained from simulation (see Section 3.5.3). Five simulation replications were carried out for each scenario.

Table 3.3: Sensitivity analysis: model inputs in each scenario

	L (m)	c_0 (s^{-1})	γ (-)	θ (m/s^2)
Base scenario	600	0.00042	1.15	9.81
Scenarios with modified L	$0.90 \cdot 600$	0.00042	1.15	9.81
	$0.95 \cdot 600$	0.00042	1.15	9.81
	$1.05 \cdot 600$	0.00042	1.15	9.81
	$1.10 \cdot 600$	0.00042	1.15	9.81
Scenarios with modified c_0	600	$0.90 \cdot 0.00042$	1.15	9.81
	600	$0.95 \cdot 0.00042$	1.15	9.81
	600	$1.05 \cdot 0.00042$	1.15	9.81
	600	$1.10 \cdot 0.00042$	1.15	9.81
Scenarios with modified γ	600	0.00042	$0.90 \cdot 1.15$	9.81
	600	0.00042	$0.95 \cdot 1.15$	9.81
	600	0.00042	$1.05 \cdot 1.15$	9.81
	600	0.00042	$1.10 \cdot 1.15$	9.81
Scenarios with modified θ	600	0.00042	1.15	$0.90 \cdot 9.81$
	600	0.00042	1.15	$0.95 \cdot 9.81$
	600	0.00042	1.15	$1.05 \cdot 9.81$
	600	0.00042	1.15	$1.10 \cdot 9.81$

3.5.3 Indicators

The key characteristics of the traffic flow patterns obtained from simulation were determined by analyzing the flow and speed data collected by the virtual loop detectors. To

analyze the data, we used three main indicators: a) time-to-breakdown (TTBD); b) average exit flow after breakdown (s_{cong}); and c) average vehicle delay after breakdown (AVD_{cong}).

The time-to-breakdown (TTBD) is the time when traffic becomes congested at the end of the sag vertical curve on a lane and this is followed by the formation of persistent congestion on the other lanes as well (see, for example, Figs. 3.2 and 3.3). As mentioned in Section 3.4.1, we consider that traffic becomes congested when the traffic speed goes below a certain threshold (v_{crit}). The speeds used to calculate the TTBD are those measured by the three virtual detectors (one on each lane) located at $x = 8.2$ km (i.e., near the end of the sag vertical curve).

The average total exit flow after breakdown (s_{cong}) is the average of the total exit flows measured from the time when traffic breaks down (i.e., TTBD) to the end of the simulation period. The total exit flow in a given aggregation period is defined as the sum of the flows measured by the three virtual detectors (one on each lane) located at $x = 9.9$ km (i.e., near the freeway exit point) during that period. The s_{cong} can be used as an estimate of the queue discharge capacity of the bottleneck.

The average vehicle delay after breakdown (AVD_{cong}) is calculated as follows:

$$\text{AVD}_{\text{cong}} = \frac{\text{TTS}_{\text{cong}} - \text{TTS}_{\text{ref}}}{N_{d,\text{cong}}} \quad (3.8)$$

where: TTS_{cong} is the total time spent by vehicles in the freeway stretch between the time when traffic breaks down (i.e., TTBD) and the end of the simulation period in a given scenario; TTS_{ref} is the total time spent by vehicles in the freeway stretch in the reference scenario during the same time period used to calculate TTS_{cong} ; and $N_{d,\text{cong}}$ is the number of vehicles that enter the freeway during the time period used to calculate TTS_{cong} . Note that $N_{d,\text{cong}}$ is the same in the target scenario and the reference scenario for each simulation replication.

In Eq. 3.8, the total time spent is calculated based on the total demand and exit flows using the method described in Papageorgiou et al. (2003). The total demand flow is defined as the sum of the flows measured by the three virtual detectors (one on each lane) located at $x = 0.1$ km. As mentioned above, the total exit flow is defined as the sum of the flows measured by the three virtual detectors (one on each lane) located at $x = 9.9$ km.

Finally, note that the values of the three indicators described above are expected to be different in different replications of the same scenario, because some parameters of the car-following and lane change models are stochastic (see Section 3.5.1). Therefore, the behavior of driver-vehicle units is not exactly the same in all replications of the same scenario, which influences both the demand profile and the capacity of the bottleneck.

3.6 Results

This section presents the results of the simulation study aimed at testing the face validity of the traffic model. Section 3.6.1 analyzes the main characteristics of the traffic flow patterns obtained from simulation. Section 3.6.2 presents the results of the sensitivity analysis.

3.6.1 Simulated traffic flow patterns

In Figs. 3.2 and 3.3, we can observe a typical example of the traffic flow patterns generated by the model. Those figures show flow and speed data corresponding to one of the base scenario simulation replications. In that simulation replication, traffic breaks down at $t = 84$ min (see Figs. 3.2 and 3.3(a)), when total demand is 5,200 veh/h. The bottleneck is the end of the sag vertical curve (see Fig. 3.2).

The process of congestion formation is as follows: 1) traffic becomes congested on the median lane (see Figs. 3.2 and 3.3(a)); 2) some drivers move out of the median lane, which causes a temporary increase in flow on the center and shoulder lanes (see Fig. 3.3(b)); and 3) the capacity of the center and shoulder lanes is exceeded and traffic becomes congested there as well (see Figs. 3.2 and 3.3(a)). Note that the migration of vehicles from the median lane to the other lanes causes a temporary decrease in flow on the median lane (see Fig. 3.3(b)). That decrease in flow causes the queue on that lane to dissolve; however, the decrease in flow is only temporary, so traffic becomes congested on the median lane again a few minutes later (see Figs. 3.2 and 3.3(a)).

The occurrence of congestion on a lane results in decreased lane flow (see Fig. 3.3(b)). The average total exit flow after breakdown (which is an estimation of the queue discharge capacity of the bottleneck) is about 5,000 veh/h, i.e., 4% lower than the total demand (see Fig. 3.3(c)). As a result, after traffic breaks down, a queue of vehicles forms upstream of the bottleneck (see Fig. 3.2), which leads to increased travel times (the average vehicle delay after breakdown is 13.6 s).

In conclusion, the traffic flow patterns obtained from simulation are similar to the patterns observed in real freeways. The traffic model is capable of reproducing the main phenomena that cause the formation of congestion at sags. The model generates a capacity bottleneck and reproduces its location quite accurately (i.e., end of the sag vertical curve), although the simulated free flow capacity (5,200 veh/h) is lower than in empirical traffic data from the Yamato sag (5,400 veh/h) (Patire & Cassidy, 2011). In addition, the model reproduces the congestion-induced capacity drop, although the magnitude of the drop (4%) is lower than in empirical observations (11%) (Patire & Cassidy, 2011). The model also reproduces the process of congestion formation (Koshi et al., 1992; Hatakenaka et al., 2006; Patire & Cassidy, 2011): congestion starts on the median lane and spreads to the other lanes as a result of lane changes. However, the

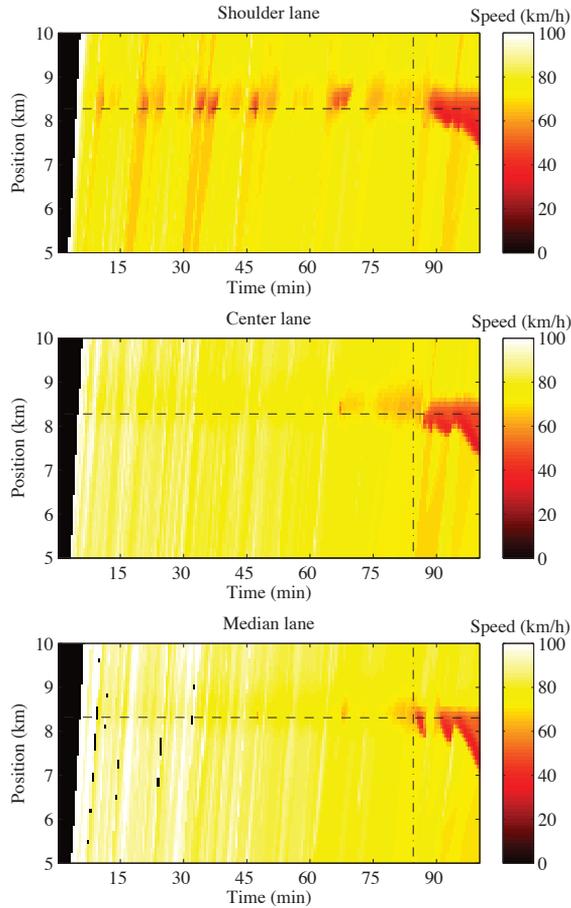


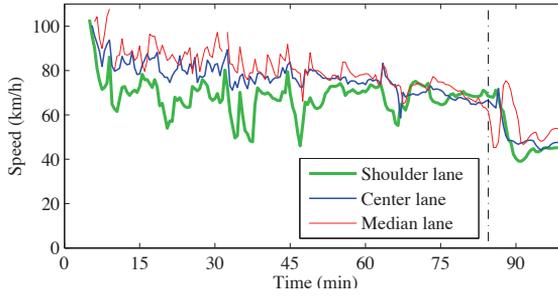
Figure 3.2: Speed contour plots of the shoulder lane, the center lane and the median lane in the base scenario (first simulation replication). The *dash-dotted* lines show the time-to-breakdown (TTBD), and the *dashed* lines indicate the location where the sag vertical curve ends.

frequency of lane changes from congested to uncongested lanes may be higher than observed in empirical data¹.

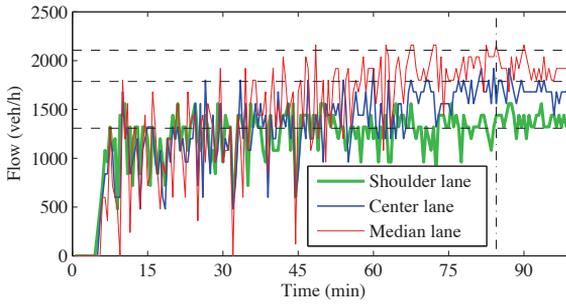
3.6.2 Sensitivity analysis

The results of the sensitivity analysis show that, in all replications of all scenarios: i) traffic breaks down at the sag, which indicates that the capacity of the sag is lower

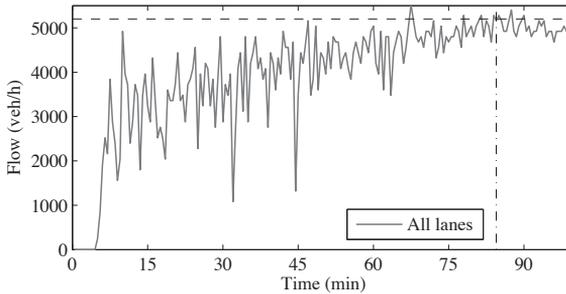
¹Note not included in the original publication: In addition, a recent study shows that the model also reproduces the oscillating nature of congested traffic at sags (Papacharalampous et al., 2015a).



(a) Speed over time per lane.



(b) Flow over time per lane.



(c) Total flow over time.

Figure 3.3: Speeds and flows over time at location $x = 8.2$ km (i.e., near the end of the sag vertical curve) in the base scenario (first simulation replication). The *dash-dotted* lines show the time-to-breakdown (TTBD). The *dashed* lines show the average demand after $t = 75$ min.

than that of the constant-gradient downhill section; ii) the head of the queue stays around the end of the vertical curve; and iii) congestion causes considerable travel time delays. However, there are notable differences between scenarios, which means that changing the values of L , c_0 , θ and γ has an impact on the characteristics of the

simulated traffic.

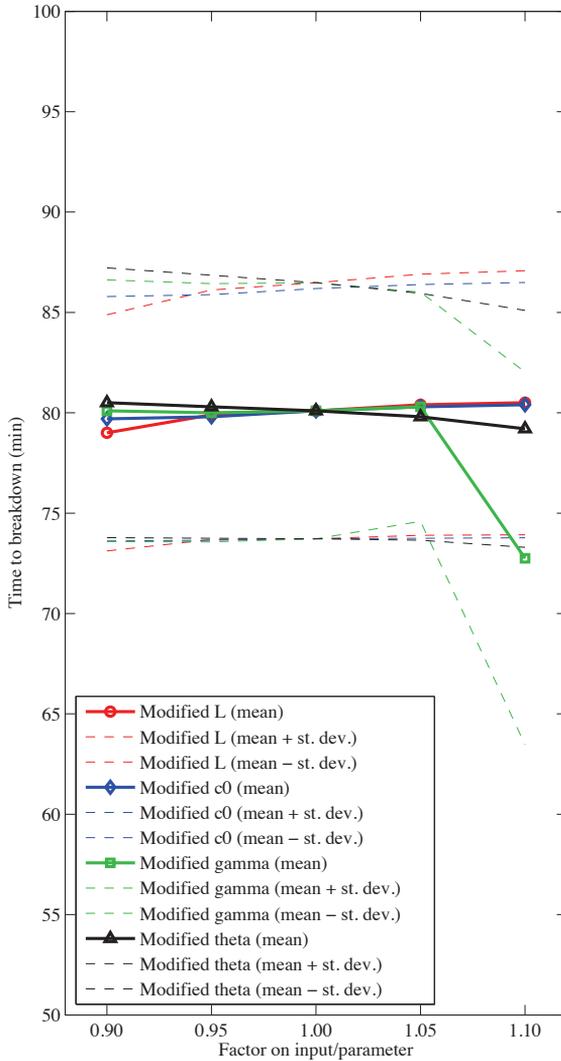


Figure 3.4: Sensitivity analysis: time-to-breakdown in each scenario (mean and mean plus/minus one standard deviation of five simulation replications).

For instance, increasing the length of the vertical curve (L) generally leads to a slightly longer time-to-breakdown (i.e., traffic breaks down a bit later, as shown in Fig. 3.4) and a higher queue discharge rate (Fig. 3.5). Consequently, it also results in a lower average vehicle delay after breakdown (Fig. 3.6). Increasing the maximum gradient compensation rate (c_0) has a similar effect to that of increasing the length of the vertical curve (see Figs. 3.4, 3.5 and 3.6). Instead, increasing the parameter θ (i.e., the sensitivity to

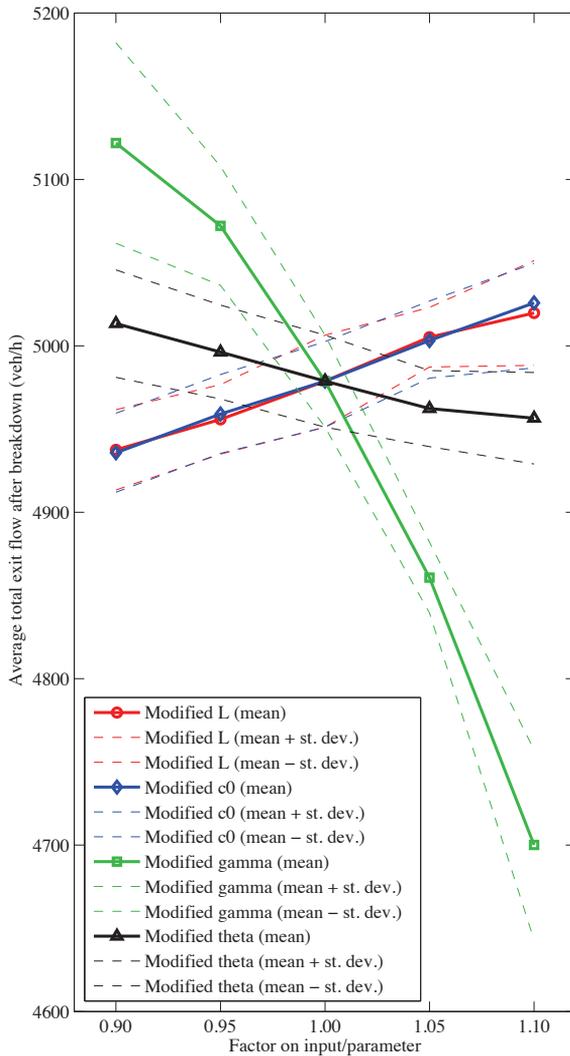


Figure 3.5: Sensitivity analysis: average total exit flow after breakdown in each scenario (mean and mean plus/minus one standard deviation of five simulation replications).

the difference between gradient and compensated gradient) has the opposite effect: a higher θ generally results in a slightly shorter time-to-breakdown and a lower queue discharge rate, which causes an increase in AVD_{cong} . Those findings can be explained as follows. A longer L (hence a more gentle vertical curve), a greater c_0 (hence a faster compensation of the increase in freeway gradient by drivers) and a lower sensitivity to the difference between gradient and compensated gradient (θ) reduce the negative

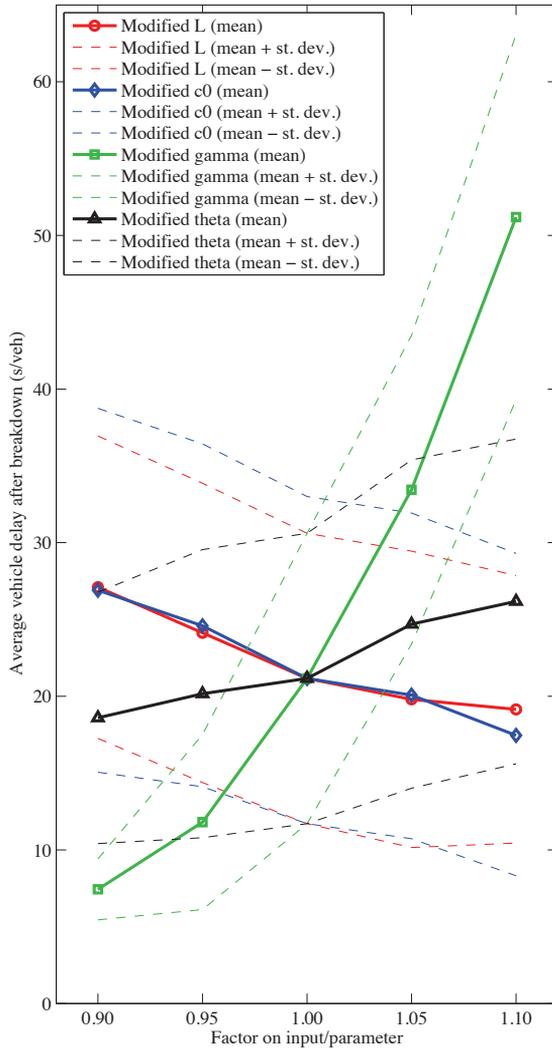


Figure 3.6: Sensitivity analysis: average vehicle delay after breakdown in each scenario (mean and mean plus/minus one standard deviation of five simulation replications).

effect that the increase in freeway gradient has on vehicle acceleration (see Eqs. 3.5-3.7), which is the main cause of the changes in car-following behavior that reduce the capacity of the sag (see Section 3.2). As a result, congestion starts later in time and the queue discharge rate increases. Therefore, the average queue length within the simulation period decreases and the traffic speed within the queue increases, which leads to decreased AVD_{cong} .

The value of the congestion factor on target time headway (γ) influences traffic flow dynamics in a different way than the value of L , c_0 and θ . First, as shown in Fig. 3.4, changing the value of parameter γ by -10, -5 or +5% in comparison to the base scenario does not have any clear effect on the time-to-breakdown (Fig. 3.4). However, if the value of parameter γ is increased by +10%, congestion starts considerably earlier (around 7 min earlier on average). Second, increasing the value of parameter γ generally leads to a lower queue discharge rate (see Fig. 3.5). As a consequence of the decreased queue discharge rate (and the unchanged or shorter time-to-breakdown), increasing the value of parameter γ results in higher average vehicle delay after breakdown (see Fig. 3.6). Those findings can be explained as follows. A higher γ means longer headways in congestion; therefore, increasing the value of γ reduces the queue discharge capacity of the sag, which results in increased average vehicle delay after breakdown. However, increasing the value of parameter γ has little influence on the time when congestion starts. The average time-to-breakdown is only significantly different when γ is 10% higher than in the base scenario. The reason why this occurs is as follows. In the other scenarios, some slow-speed regions appear near the end of the vertical curve before traffic breaks down (see, for instance, Fig. 3.2). Those slow-speed regions dissolve before they cause a full traffic breakdown because the inflow is lower than the outflow. However, increasing the value of parameter γ decreases the outflow from those low-speed regions, which in some cases prevents their dissolution. As a result, the slow-speed regions cause a full breakdown earlier than in the scenarios with lower value of γ . It is important to remark that the magnitude of the decrease in queue discharge rate caused by increasing the value of parameter γ is more pronounced than the magnitude of the decrease in queue discharge rate caused by increasing the value of θ or decreasing the value of L and c_0 (see Fig. 3.5). Consequently, the magnitude of the increase in AVD_{cong} caused by increasing the value of parameter γ is also more pronounced than the magnitude of the increase in AVD_{cong} caused by increasing parameter θ or decreasing parameters L and c_0 (see Fig. 3.6).

Finally, note that the time-to-breakdown differs substantially between simulation replications belonging to the same scenario, whereas the variation of the average exit flow after breakdown (s_{cong}) is much lower. As seen in Figs. 3.4 and 3.5, the average standard deviation of the time-to-breakdown in all scenarios is around 8% of the mean (7 min), whereas the average standard deviation of s_{cong} in all scenarios is around 0.5% of the mean (30 veh/h). From this we conclude that the variation of the AVD_{cong} within each scenario is mainly due to the variation in the time-to-breakdown.

To sum up, the results of the sensitivity analysis show that the traffic model is capable of reproducing the main characteristics of traffic flow dynamics at sags even if some model inputs are modified to some degree. This indicates that the model is quite robust. However, modifying those inputs does have an influence on the characteristics of the simulated traffic, which stresses the need to calibrate and validate the traffic model. Changing the value of L , c_0 , θ and γ influences traffic as follows. Congestion tends to start a bit earlier and be more severe if: i) the sag vertical curve is sharper (shorter

L); ii) drivers compensate more slowly for the increase in freeway gradient (lower c_0); iii) vehicle acceleration is more sensitive to the difference between gradient and compensated gradient (higher θ); or iv) drivers keep longer time headways in congestion (higher γ). It is important to remark that changing the value of parameter γ has a larger effect on the characteristics of the simulated traffic (particularly on the queue discharge capacity) than changing the values of L , c_0 and θ by the same percentage (at least in the scenarios included in the analysis). This indicates that parameter γ should be given especial attention during the model calibration/validation process. Finally, the results of the sensitivity analysis also show that the stochasticity incorporated into the traffic model (see Section 3.5.1) has a strong influence on the time when congestion starts, whereas its influence on the queue discharge capacity of the bottleneck is less significant.

3.7 Conclusions

This paper presented a microscopic model that reproduces traffic flow dynamics at sags. The traffic model consists of two sub-models: i) a car-following model; and ii) a lane change model. The main novelty is the car-following modeling approach. We propose a new car-following model that is based on assumptions about the way in which sag vertical curves affect longitudinal driving behavior that are more realistic than those of most existing models. Our car-following model is based on a similar principle to that of the model presented by Oguchi and Konuma (2009), which assumes that drivers compensate gradually for the increase in resistance force caused by an increase in freeway gradient. However, our car-following model has a more generic formulation.

The face-validity of the traffic model was tested by means of a simulation study. The traffic flow patterns obtained from simulation are similar to the patterns observed on real freeways. More specifically, the proposed traffic model is capable of reproducing the main traffic phenomena that cause the formation of congestion at sags, namely: a) decrease in vehicle acceleration and local changes in car-following behavior; b) low capacity of sags compared to normal sections (the bottleneck being the end of the vertical curve); c) congestion-induced capacity drop; d) queue dynamics; e) uneven distribution of flow across lanes in high demand conditions; and f) migration of vehicles from congested to uncongested lanes.

A sensitivity analysis indicates that the traffic model is sufficiently robust to reproduce those phenomena even if some model inputs are modified to some degree. However, modifying those inputs does have an influence on the characteristics of the simulated traffic (particularly on the queue discharge capacity of the bottleneck), which highlights the need to calibrate and validate the proposed traffic model. Parameter γ should be given especial attention in the calibration/validation process. The results of the sensitivity analysis also indicate potential ways to improve traffic flow efficiency at

sags. For instance, the severity of congestion seems to be lower if the vertical curve is more gentle and if drivers are able to compensate faster for the negative effect that the increase in freeway gradient has on vehicle acceleration.

The calibration and validation of the traffic model will require a quantitative comparison of the model output with empirical traffic data. Vehicle trajectory data from the Yamato sag (Tomei Expressway, Japan) are available (see Goñi-Ros, et al. (2014a)). However, data from additional sites will also be necessary. If possible, those additional sites should be from other countries and include sags with different vertical profiles and different number of lanes. After calibration and validation, the model could be used to evaluate the effectiveness of possible control measures to mitigate congestion at sags.

Chapter 4

Optimization of traffic flow at freeway sags by controlling the acceleration of some vehicles equipped with in-car systems

Goñi-Ros, B., Knoop, V.L., Takahashi, T., Sakata, I., van Arem, B., Hoogendoorn, S.P. (Submitted). Optimization of traffic flow at freeway sags by controlling the acceleration of vehicles equipped with in-car systems. *Transportation Research Part C: Emerging Technologies*.

Abstract

Sags are bottlenecks in freeway networks. According to previous research, the main cause is that most drivers do not accelerate enough at sags. Consequently, they keep longer headways than expected given their speed, which leads to congestion in high demand conditions. Nowadays, there is growing interest in the development of traffic control measures for sags based on the use of in-car systems. This paper determines the optimal acceleration behavior of vehicles equipped with in-car systems at sags and the related effects on traffic flow, thereby laying the theoretical foundation for developing effective traffic management applications. We formulate an optimal control problem in which a centralized controller regulates the acceleration of some vehicles of a traffic stream moving along a single-lane freeway stretch with a sag. The control objective is to minimize total travel time. The problem is solved for scenarios with different numbers of controlled vehicles and positions in the stream. The results show that the optimal behavior of controlled vehicles involves performing a

deceleration-acceleration-deceleration-acceleration (DADA) maneuver in the sag area. This maneuver induces the first vehicles located behind the controlled vehicle to accelerate fast along the vertical curve. As a result, traffic speed and flow at the end of the sag (bottleneck) increase for a limited time. The maneuver also triggers a stop-and-go wave that temporarily limits the inflow to the sag, slowing down the formation of congestion at the bottleneck. Moreover, in some cases controlled vehicles perform one or more deceleration-acceleration maneuvers upstream of the sag. This additional strategy is used to manage congestion so that the inflow is regulated more effectively. These findings are relevant for the development of new traffic management measures for sags, particularly cooperative adaptive cruise control applications. Also, our findings prove the usefulness of the proposed optimization method as a tool for control measure development.

4.1 Introduction

Sags (or sag vertical curves) are freeway sections along which the gradient increases gradually in the direction of traffic. The capacity of sags is generally lower than that of freeway sections with other vertical profiles (Xing et al., 2010). Therefore, sags constitute bottlenecks in freeway networks. Indeed, traffic often becomes congested at sags in high demand conditions (Koshi et al., 1992; Brilon & Bressler, 2004; Patire & Cassidy, 2011). In some countries, such as Japan, sags are one of the most common types of freeway bottleneck (Xing et al., 2014; Hatakenaka et al., 2006). The scientific literature suggests that the main cause of traffic congestion at sags is that most drivers do not accelerate enough as they move along the vertical curve. Generally, drivers do not compensate instantaneously for the increase in resistance force resulting from the increase in gradient, which limits the acceleration of their vehicles (Yoshizawa et al., 2012). As a consequence, most drivers keep significantly longer distance headways than expected given their speed (Koshi, 2003; Goñi-Ros et al., 2013a). This leads to periodic formation of stop-and-go waves when traffic demand is sufficiently high (Koshi et al., 1992; Patire & Cassidy, 2011; Goñi-Ros, et al., 2014a). The bottleneck is generally the end of the vertical curve (Brilon & Bressler, 2004; Patire & Cassidy, 2011).

During the last two decades, various traffic management measures have been proposed for mitigating congestion at sags. In general, the goals of those measures are: *a*) to increase the free-flow capacity of sags; *b*) to prevent the formation of congestion at sags in nearly-saturated conditions; and/or *c*) to increase the queue discharge capacity of sags. The proposed measures use very diverse strategies to achieve those goals, such as improving the ability of drivers to compensate for the increased grade resistance force at sags (Ozaki, 2003), limiting the inflow to the vertical curve (Goñi-Ros et al., 2014b), and encouraging drivers to accelerate fast after leaving congestion at sags (Sato et al., 2009). Most proposed measures, particularly those which are in more advanced development stages, use variable message signs (VMS) as actuators (Xing

et al., 2014; Goñi-Ros, et al., 2014b; Sato et al., 2009). However, in recent years there has been growing interest in the development of traffic management measures that use in-car systems as actuators. The systems used for that purpose are mainly advisory systems (Hatakenaka et al., 2006) or basic/advanced adaptive cruise control (ACC) systems (Ozaki, 2003; Kesting et al., 2006; Papacharalampous et al., 2015b). Although traffic management measures based on the use of in-vehicle systems have great potential, they are mostly in early phases of development. We argue that, at this stage, it is important to determine how equipped vehicles should behave at sags under various circumstances in order to generate the greatest possible reduction in congestion. That would clarify what are the most effective strategies to mitigate congestion at sags via in-car systems and what are the mechanisms by which those strategies reduce congestion, thereby laying the theoretical foundation for developing effective traffic management applications.

The main goal of this paper is to identify the optimal acceleration behavior of vehicles equipped with in-car systems at sags and the related effects on traffic flow dynamics assuming penetration rates that are realistic for the coming years (i.e., limited numbers of equipped vehicles). To that end, we formulate an optimal control problem in which a centralized controller regulates the acceleration of some vehicles belonging to a traffic stream that moves along a single-lane freeway stretch with a sag. The objective of the controller is to minimize the total travel time of all vehicles. The problem is solved for various scenarios defined by the number of controlled vehicles and their positions in the stream. By analyzing the results, we identify the main strategies that vehicles equipped with in-car systems should use at sags to reduce congestion to the greatest possible degree. Therefore, the main contribution of this paper is twofold. First, we present a new optimization-based method for identifying the optimal way to manage traffic at a certain type of bottleneck. Second, we determine what are the optimal strategies to manage traffic at sags via in-car systems, providing a detailed description of their principles. Development of specific traffic control measures based on those strategies is left for future work.

The rest of this paper is structured as follows. Section 4.2 presents the optimal control framework for traffic flow optimization. Section 4.3 presents the setup of the experiments that were carried out to identify the optimal acceleration behavior of vehicles equipped with in-car systems at sags. Section 4.4 reports the results of the experiments, focusing on the main strategies used by the controlled vehicles and their effects on traffic flow. Finally, Section 4.5 presents the conclusions of this study.

4.2 Optimal control framework

This section presents an optimal control framework aimed to determine the optimal trajectories of a set of vehicles as they move along a freeway stretch with a sag. Section 4.2.1 specifies the elements of the system to be controlled. Section 4.2.2 states the assumptions of the optimal control framework. Section 4.2.3 presents the formulation of the control problem.

4.2.1 Traffic system elements

The system to be controlled consists of a stream of vehicles moving along a freeway stretch. In total, the stream contains n vehicles. Each of them is assigned a number i that corresponds to its position in the stream ($i = 1, 2, \dots, n$). The set that contains all vehicle numbers i is denoted by N . There are two types of vehicles in the stream: *a*) non-controlled vehicles; and *b*) controlled vehicles. The total number of controlled vehicles is denoted by m , and the subset of N that contains the numbers i of all controlled vehicles is denoted by M . Each controlled vehicle is assigned a number j that corresponds to its position in relation to the other controlled vehicles ($j = 1, 2, \dots, m$). For instance, in the vehicle stream shown in Figure 4.1, $n = 8$, $N = \{1, 2, 3, 4, 5, 6, 7, 8\}$, $m = 2$ and $M = \{3, 7\}$.

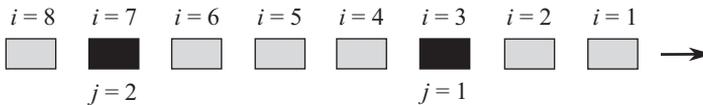


Figure 4.1: Example of vehicle stream. Gray rectangles represent non-controlled vehicles and black rectangles represent controlled vehicles. The arrow shows the direction of traffic.

The freeway stretch contains a sag vertical curve. The relation between freeway gradient and location along the freeway is known. For the sake of simplicity, we assume that the freeway stretch has one lane (i.e., overtaking is not possible) and there are no on-ramps, off-ramps, merges or diverges.

4.2.2 Assumptions

The optimal control framework is based on four assumptions:

1. A stream of vehicles moving along a freeway can be considered a discrete-time dynamical system with constant-acceleration time steps, provided that the time step length is sufficiently small.

2. The acceleration behavior of the drivers can be correctly described by the car-following model presented in Goñi-Ros et al. (2016). That model assumes that drivers accelerate with the objective of reaching/keeping their desired speed (in unconstrained driving conditions) or desired distance headway (in constrained conditions). The model also assumes that drivers operate their vehicle in such a way that they compensate gradually (not instantaneously) for the increased grade resistance force at sags.
3. Within a certain freeway section, the acceleration of some vehicles of the stream can be regulated by a centralized controller that can perfectly predict the state of the traffic stream at any time step.
4. The controller does not impose at any time step a higher acceleration than the one that would result from human driving given the current traffic conditions and position along the road.

Note that, given the objectives of this research (see Section 4.1), it is not necessary to specify how the controller acts on vehicle acceleration or how the controller determines the acceleration that would be realized by a human driver under certain circumstances.

4.2.3 Control problem formulation

State variables

The state variables of a dynamical system can be defined as the minimum set of variables that contain enough information about the history of the system to determine its future behavior. In our case, the system is a stream of vehicles moving along a freeway stretch (see Section 4.2.1) and the state variables are the variables needed to determine the future trajectory of every vehicle in the space-time plane.

As stated in Section 4.2.2, we assume that: *a*) the trajectories of human-driven vehicles are determined by the car-following model presented in Goñi-Ros et al. (2016); *b*) the trajectories of controlled vehicles are regulated by the controller, but the controller uses the car-following model to impose upper bounds on the acceleration of those vehicles. The car-following model presented in Goñi-Ros et al. (2016) calculates the acceleration of a vehicle at any time step based on the following variables: position and speed of the target vehicle, position and speed of the preceding vehicle, freeway gradient at the position of the target vehicle, and amount of freeway gradient compensated by the driver of the target vehicle. Note that freeway gradient is a variable that depends directly on vehicle position.

Therefore, the state variables of our system are: *a*) longitudinal position of (the rear-bumper of) all vehicles along the freeway ($r_i, \forall i$); *b*) speed of all vehicles ($v_i, \forall i$); and

c) amount of freeway gradient compensated by the drivers of all vehicles ($G_{\text{com},i}, \forall i$). More specifically, the state at simulation time step τ is defined by the following matrix:

$$\mathbf{x}(\tau) = \begin{bmatrix} r_1(\tau) & r_2(\tau) & \dots & r_n(\tau) \\ v_1(\tau) & v_2(\tau) & \dots & v_n(\tau) \\ G_{\text{com},1}(\tau) & G_{\text{com},2}(\tau) & \dots & G_{\text{com},n}(\tau) \end{bmatrix} \quad (4.1)$$

Control variables

The control variables can be defined as the variables that can be manipulated by the controller in order to influence the dynamics of the system. In accordance with the third and fourth assumptions stated in Section 4.2.2, in our case we chose the control variables to be the maximum accelerations of all controlled vehicles ($u_j, \forall j$). Therefore, the control input at control time step κ is defined by the following vector:

$$\mathbf{u}(\kappa) = [u_1(\kappa) \quad u_2(\kappa) \quad \dots \quad u_m(\kappa)] \quad (4.2)$$

Different counters are used for simulation time step and control time step (τ and κ) because the control time step length (T_c) can be assigned a different value than the simulation time step length (T_s), as long as T_c is a multiple of T_s and the multiplier is a natural number ($T_c = \eta \cdot T_s$, where $\eta \in \mathbb{N}^+$). If $T_c > T_s$, the control input stays constant over multiple simulation time steps.

State dynamics

Figure 4.2 shows an overview of the inputs that are necessary to update the values of the state variables (i.e., r_i , v_i and $G_{\text{com},i}, \forall i$) at every simulation time step. Vehicle positions and speeds change over time according to the following equations of motion:

$$r_i(\tau + 1) = r_i(\tau) + v_i(\tau) \cdot T_s + \frac{a_i(\tau)}{2} \cdot T_s^2 \quad (4.3)$$

$$v_i(\tau + 1) = v_i(\tau) + a_i(\tau) \cdot T_s \quad (4.4)$$

where $a_i(\tau)$ denotes the acceleration of vehicle i at simulation time step τ .

The compensated gradient changes over time as follows (see also Goñi-Ros et al. (2016)): a) drivers compensate for an increase in freeway gradient with a maximum compensation rate denoted by λ ; b) once drivers have fully compensated for an increase in gradient, G_{com} equals the actual gradient (G); and c) drivers fully compensate for any decrease in gradient instantaneously.

$$G_{\text{com},i}(\tau + 1) = \begin{cases} G(r_i(\tau + 1)) & \text{if } G(r_i(\tau + 1)) \leq G_{\text{com},i}(\tau) + \lambda_i \cdot T_s \\ G_{\text{com},i}(\tau) + \lambda_i \cdot T_s & \text{if } G(r_i(\tau + 1)) > G_{\text{com},i}(\tau) + \lambda_i \cdot T_s \end{cases} \quad (4.5)$$

Vehicle accelerations (a_i) are calculated as follows. For non-controlled vehicles, the acceleration of vehicle i at time step τ ($a_i(\tau)$) is equal to the acceleration given by the

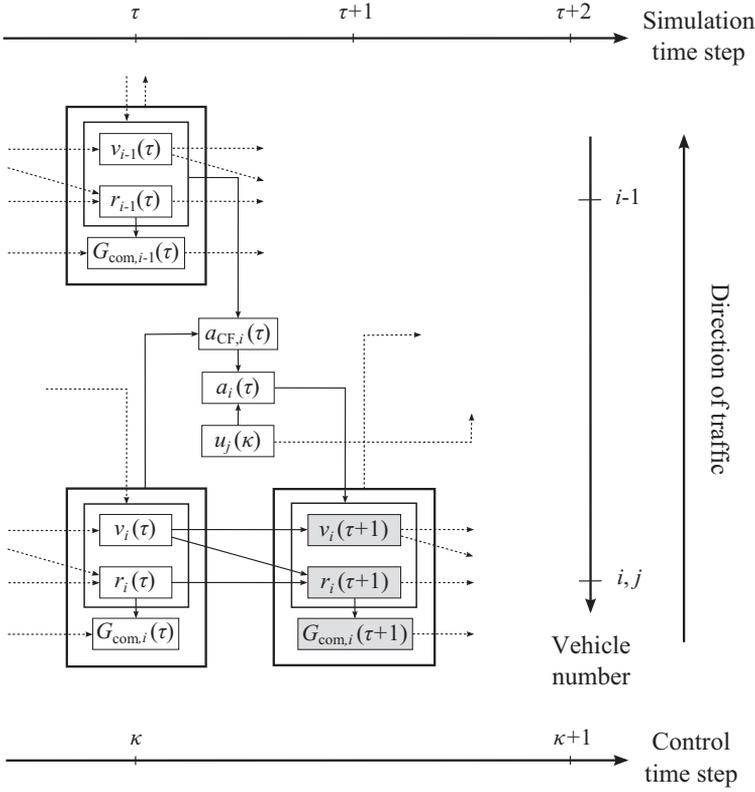


Figure 4.2: Overview of the inputs needed to calculate the state dynamics.

car-following model at that time step ($a_{CF,i}^h(\tau)$). For controlled vehicles, $a_i(\tau)$ is equal to the acceleration determined by the controller for time step τ ($a_{c,i}(\tau)$), which in turn is defined as the minimum of the control input ($u_j(\kappa)$) and the acceleration given by the car-following model ($a_{CF,i}^c(\tau)$). By defining $a_{c,i}$ in this way, we ensure that the fourth assumption stated in Section 4.2.2 holds without having to explicitly specify a state constraint in the control problem. That simplifies the optimization procedure. Thus accelerations are calculated as:

$$a_i(\tau) = \begin{cases} a_{CF,i}^h(\tau) & \text{if } i \notin M \\ a_{c,i}(\tau) & \text{if } i \in M \end{cases} \quad (4.6)$$

where:

$$a_{c,i}(\tau) = \begin{cases} \min(u_j(\kappa), a_{CF,i}^c(\tau)) & \text{if } r_i(\tau) \in [r_0^c, r_f^c] \\ a_{CF,i}^c(\tau) & \text{if } r_i(\tau) \notin [r_0^c, r_f^c] \end{cases} \quad (4.7)$$

In Equation 4.7: r_0^c and r_f^c denote the points that delimit the area where acceleration control is applicable; i and j correspond to the same vehicle; and κ is such that $\tau \cdot T_s \in [\kappa \cdot T_c, (\kappa + 1) \cdot T_c)$.

The car-following model is the same for non-controlled vehicles and controlled vehicles, thus $a_{CF,i}^h$ and $a_{CF,i}^c$ are calculated in the same way. However, there are two important differences between those two variables. First, $a_{CF,i}^h(\tau)$ is always the actual acceleration of vehicle i at simulation time step τ , whereas $a_{CF,i}^c(\tau)$ is only the realized acceleration if the control input is greater or not applicable (see Equations 4.6 and 4.7). Second, $a_{CF,i}^h(\tau)$ is always calculated based on the state of the vehicle stream at simulation time step τ , but that is not true for $a_{CF,i}^c(\tau)$. If $T_c > T_s$, then $a_{CF,i}^c(\tau)$ is calculated based on the current state only at simulation time steps when $\tau \cdot T_s = \kappa \cdot T_c$ holds (i.e., at the beginning of every control time step), or when the realized accelerations (a_i) at all previous simulation time steps within the current control step were equal to $a_{CF,i}^c(\tau)$. At all other simulation time steps, $a_{CF,i}^c(\tau)$ is calculated based on the state that would have been obtained if the realized vehicle accelerations (a_i) at all previous simulation time steps within the same control step had been equal to $a_{CF,i}^c(\tau)$ (for example, see Figure 4.5(a)). This formulation makes $a_{CF,i}^c$ independent of u_j within every control step, which simplifies the optimization procedure.

The car-following model presented in Goñi-Ros et al. (2016) is used to calculate $a_{CF,i}^h$ and $a_{CF,i}^c$. Here we only present the main features of that model; we refer to Goñi-Ros et al. (2016) for a discussion of its properties and validity. Let $a_{CF,i}$ be a variable that corresponds to $a_{CF,i}^h$ if vehicle i is non-controlled and to $a_{CF,i}^c$ if vehicle i is controlled. Then, the car-following model calculates vehicle accelerations as follows:

$$a_{CF,i}(\tau) = \max(a_{des,i}(\tau) + \delta_i(\tau), a_{min,i}, -\frac{v_i(\tau)}{T_s}) \quad (4.8)$$

In Equation 4.8, $a_{min,i}$ and $-v_i(\tau)/T_s$ define lower bounds that prevent unrealistically strong decelerations and negative speeds, respectively. The influence of traffic conditions and freeway gradient on longitudinal driving behavior is mainly described by the expression $a_{des,i}(\tau) + \delta_i(\tau)$. The desired acceleration is calculated as follows:

$$a_{des,i}(\tau) = \alpha_i \cdot \min \left[1 - \left(\frac{v_i(\tau)}{v_{des,i}} \right)^4, 1 - \left(\frac{s_{des,i}(v_i(\tau), \Delta v_i(\tau))}{s_i(\tau)} \right)^2 \right] \quad (4.9)$$

where s_i denotes net distance headway, and the relative speed (Δv_i) and desired net distance headway ($s_{des,i}$) are defined as:

$$\Delta v_i(\tau) = v_i(\tau) - v_{i-1}(\tau) \quad (4.10)$$

$$s_{des,i}(v_i(\tau), \Delta v_i(\tau)) = s_{s,i} + v_i(\tau) \cdot H_i + \frac{v_i(\tau) \cdot \Delta v_i(\tau)}{2 \cdot \sqrt{\alpha_i \beta_i}} \quad (4.11)$$

The parameters in Equations 4.9-4.11 are: desired speed ($v_{des,i}$), net distance headway at standstill ($s_{s,i}$), safe time headway (H_i), maximum acceleration (α_i) and maximum comfortable deceleration (β_i).

The term δ_i in Equation 4.8 accounts for the influence of the freeway's vertical alignment on longitudinal driving behavior. At a given simulation time step, $\delta_i(\tau)$ is equal

to the difference between the gradient in the freeway location where the vehicle is at that time step and the amount of gradient compensated by the driver until that time step, multiplied by a sensitivity parameter (θ_i):

$$\delta_i(\tau) = -\theta_i \cdot (G(r_i(\tau)) - G_{\text{com},i}(\tau)) \quad (4.12)$$

Initial state and admissible control region

The initial state of the stream of vehicles ($\mathbf{x}(0)$) is known and equal to \mathbf{x}_0 . The admissible control region \mathcal{U} contains all maximum acceleration (u_j) values that can be chosen by the controller for every controlled vehicle j . In our case, the set of admissible maximum acceleration values is the same for all controlled vehicles and for all control time steps. That set contains all real numbers between a lower bound denoted by u_{\min} and an upper bound denoted by u_{\max} . Therefore, \mathcal{U} is such that:

$$u_j(\kappa) \in [u_{\min}, u_{\max}], \forall j \in [1, m], \forall \kappa \in [0, \frac{T}{T_c}] \quad (4.13)$$

Cost function

The cost function (J) is defined as the total travel time of all vehicles from their initial positions to the arrival point R :

$$J(\mathbf{x}(0), \mathbf{x}(1), \dots, \mathbf{x}(\frac{T}{T_s}), \mathbf{u}(0), \mathbf{u}(1), \dots, \mathbf{u}(\frac{T}{T_c})) = \sum_{i=1}^n (T_s \cdot \tau_{R,i} + \Delta t_i) \quad (4.14)$$

where $\tau_{R,i}$ denotes the last simulation time step at which vehicle i is located upstream of point R :

$$\tau_{R,i} = \max(\tau \mid r_i(\tau) \leq R) \quad (4.15)$$

and Δt_i denotes the additional time required by vehicle i to move from its position at simulation time step $\tau_{R,i}$ to position R , which is calculated by solving the following quadratic equation:

$$\frac{a_i(\tau_{R,i})}{2} \cdot (\Delta t_i)^2 + v_i(\tau_{R,i}) \cdot \Delta t_i + (r_i(\tau_{R,i}) - R) = 0 \quad (4.16)$$

Optimal control problem

The discrete-time optimal control problem, which is non-linear and non-convex, can be formulated as the following mathematical program:

Find $\mathbf{u}^*(0), \mathbf{u}^*(1), \dots, \mathbf{u}^*(\frac{T}{T_c})$

that minimize $J(\mathbf{x}(0), \mathbf{x}(1), \dots, \mathbf{x}(\frac{T}{T_s}), \mathbf{u}(0), \mathbf{u}(1), \dots, \mathbf{u}(\frac{T}{T_c}))$

subject to an initial state condition, the control input constraints and the state dynamics equations:

$$\mathbf{x}(0) = \mathbf{x}_0 \quad (4.17)$$

$$\mathbf{u}(\kappa) \in \mathcal{U}, \text{ for } \kappa = 0, 1, 2, \dots, \frac{T}{T_c} \quad (4.18)$$

$$\mathbf{x}(\tau + 1) = \mathbf{f}(\mathbf{x}(\tau), \mathbf{u}(\kappa)), \text{ for } \tau = 0, 1, 2, \dots, \frac{T}{T_s} \quad (4.19)$$

where κ is such that $\tau \cdot T_s \in [\kappa \cdot T_c, (\kappa + 1) \cdot T_c)$.

The state matrix ($\mathbf{x}(\tau)$) and the control input vector ($\mathbf{u}(\kappa)$) are defined in Equations 4.1 and 4.2, respectively. The objective function (J) is calculated by using Equations 4.14-4.16. The admissible control region (\mathcal{U}) is defined in Equation 4.13. The vehicle stream state dynamics are determined by Equations 4.3-4.12.

4.3 Experimental setup

We carried out a series of optimization experiments that entailed solving the discrete-time optimal control problem presented in the previous section for various scenarios. Section 4.3.1 specifies the objectives of the experiments, Section 4.3.2 defines the scenarios, Section 4.3.3 describes the method used to solve the problem, and Section 4.3.4 defines the indicator used to evaluate the effectiveness of the optimal controller.

4.3.1 Objectives

The objectives of the optimization experiments were as follows:

1. To determine the optimal acceleration behavior of controlled vehicles as they move along a one-lane freeway with a sag in nearly-saturated traffic conditions, assuming low penetration rates.
2. To determine the effects that the optimal acceleration behavior of controlled vehicles has on traffic flow dynamics, and the reasons why the total travel time decreases as a result of those effects.

4.3.2 Scenarios

We defined eight scenarios in which the traffic stream contains 300 vehicles ($n = 300$). The scenarios differ only in two inputs: number of controlled vehicles and positions of those vehicles in the stream. All other inputs are the same in all scenarios.

Distinctive inputs

The scenarios differ in the number of controlled vehicles (m) and the positions of those vehicles in the stream (set M). To define the scenarios, we stipulated that the number of controlled vehicles should be 0, 1, 2 or 3, and the positions in the stream should be $\frac{n}{4}$, $\frac{2n}{4}$ and/or $\frac{3n}{4}$. Therefore, we assumed that the number of controlled vehicles is limited and their positions in the stream are relatively far apart. A scenario was defined for every possible configuration of set M , which makes a total of eight scenarios, including seven control scenarios and a no-control scenario (scenario with $m = 0$).

Common inputs

The simulation time step length (T_s) is 0.5 s and the control time step length (T_c) is 8 s. The total simulation period length (T) is 800 s. Note that T is a multiple of T_s and T_c , and it is large enough to ensure that all vehicles pass point R before the last simulation time step.

The freeway stretch has one lane and can be divided in three consecutive sections on the basis of its vertical profile (see Figure 4.3): a) constant-gradient downhill section; b) sag vertical curve; and c) constant-gradient uphill section. The vertical curve starts in point r_d and ends in point r_u . The gradient upstream of point r_d (which is denoted by G_d) is constant and negative, whereas the gradient downstream of point r_u (G_u) is constant and positive. Between points r_d and r_u , the freeway slope increases linearly over distance (Figure 4.3). Table 4.1 shows the values of parameters r_d , r_u , G_d and G_u , as well as the limits of the freeway section within which the control input is applicable (r_0^c and r_f^c) and the location of the arrival point used to calculate travel times (R). Note that point R is far enough from the sag vertical curve so as to ensure that traffic is in stationary conditions at that location.

All vehicle-driver units are 4 m long and are assigned the same value for every parameter of the car-following model (Table 4.1). Hence, the car-following behavior of all units is based on the same rules. This setup reduces the complexity of traffic dynamics and makes it easier to compare different scenarios.

At time zero, the stream of vehicles is in homogeneous and stationary conditions, the initial speed of all vehicle-driver units is equal to the desired speed ($v_{des,i}$), and the first vehicle of the stream is located at point r_0^c . It is assumed that the traffic density at time zero is the critical density of the constant-gradient downhill section. Therefore, since

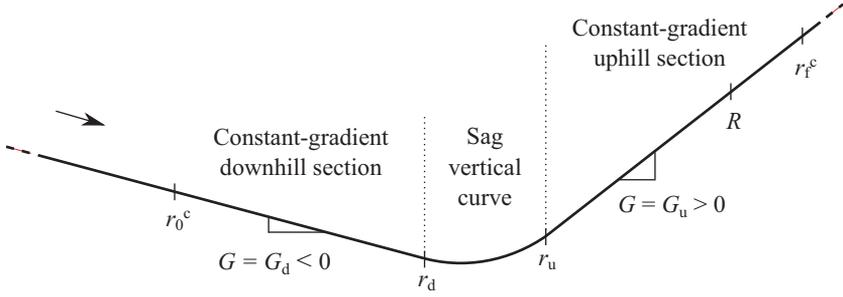


Figure 4.3: Vertical profile of the freeway stretch.

Table 4.1: Values of the parameters that describe the traffic system.

Network parameters		CF model parameters	
Parameter	Value	Parameter	Value
G_d (%)	-0.5	$v_{des,i}, \forall i$ (km/h)	120
G_u (%)	+2.5	$\alpha_i, \forall i$ (m/s ²)	1.4
r_d (m)	1000	$\beta_i, \forall i$ (m/s ²)	2.1
r_u (m)	1600	$H_i, \forall i$ (s)	1.20
R (m)	5000	$s_{s,i}, \forall i$ (m)	3
r_0^c (m)	-2000	$\theta_i, \forall i$ (dimensionless)	22
r_f^c (m)	7000	$\lambda_i, \forall i$ (s ⁻¹)	0.0004
		$a_{min,i}, \forall i$ (m/s ²)	-8

the values of parameters $s_{s,i}$, $v_{des,i}$ and H_i are the same for all vehicles i , the initial position of every vehicle i is defined as:

$$r_i(0) = r_0^c - (i-1) \cdot (s_{s,i} + v_{des,i} \cdot H_i) \quad (4.20)$$

Since $r_i(0) < r_d, \forall i$, the freeway gradient at the initial location of all vehicles is equal to the slope of the constant-gradient downhill section (G_d). Initially, the compensated gradient is equal to G_d for all vehicles, so the freeway gradient has no influence on vehicle acceleration ($\delta_i(0) = 0, \forall i$).

The lower bound (u_{min}) and the upper bound (u_{max}) that define the admissible control region are set to -0.5 m/s² and 1.4 m/s², respectively. The lower bound corresponds to a moderate deceleration rate, which prevents too strong decelerations. The upper bound is equal to the value of the car-following model parameter α (see Table 4.1), which determines the maximum vehicle acceleration.

4.3.3 Solution method

We developed a MATLAB program that solves the optimal control problem for all scenarios. The core of the program is the optimization function `fmincon`, which uses se-

quential quadratic programming to iteratively solve nonlinear optimization problems. The `fmincon` function was set to use the `sqp` solution algorithm. Note that the algorithm has a limit of cost function evaluations and requires an initial solution guess as input. In our case, we define the initial solution guess as follows. The control input is equal to the maximum acceleration parameter (see Table 4.1) at all control time steps for all controlled vehicles:

$$\mathbf{u}^g(\kappa) = [\alpha_1(\kappa) \quad \alpha_2(\kappa) \quad \dots \quad \alpha_m(\kappa)], \text{ for } \kappa = 1, 2, \dots, \frac{T}{T_c} \quad (4.21)$$

With that initial solution guess, the control input in the first optimization iteration is not lower than the acceleration given by the car-following model at any control step; therefore, the trajectories of controlled vehicles are the same as in the no-control scenario (see Equations 4.6-4.12).

4.3.4 Performance Indicator

The effectiveness of the optimal controller in every scenario is evaluated by means of the average vehicle delay (AVD). The AVD in a given scenario is calculated as follows:

$$\text{AVD} = \frac{\text{TTT} - \text{TTT}_{\text{ref}}}{n} \quad (4.22)$$

where TTT is the total travel time in that scenario, and TTS_{ref} is the total travel time in the reference scenario. The reference scenario is a hypothetical one in which there are no controlled vehicles and the gradient is constant along the whole freeway stretch.

4.4 Results

This section presents the results of the optimization experiments. Section 4.4.1 provides an overview of the results in all scenarios. Section 4.4.2 describes the primary strategy applied by the controlled vehicles, including the characteristics of the maneuver and its effects on traffic flow dynamics and total travel time. Section 4.4.3 describes what we called the supporting strategy.

4.4.1 Overview of the optimization results in all scenarios

Optimality of the solutions

The solver finds an optimal solution before reaching the maximum number of iterations for four control scenarios, namely the scenarios with M equal to $\{75\}$, $\{150\}$, $\{225\}$ and $\{75, 225\}$. Note that, given the characteristics of the optimal control problem and the solution algorithm, we cannot guarantee that the solutions found for those scenarios

correspond to global optima; they can only be considered local optima. For the other three control scenarios, the solver does not find an optimal solution before reaching the maximum number of iterations. However, for all those scenarios, the value of the cost function in the last iterations is very stable, which indicates that the solution given as output in the last iteration is probably close to the optimal solution (local optimum). Interestingly, controlled vehicles behave in a similar way in all scenarios, which suggests that the solutions may be (close to) global optima.

Main strategies applied by the controller

The results show that the optimal acceleration behavior of controlled vehicles is defined by two main strategies, which we call primary strategy and supporting strategy. The characteristics of those two strategies and their effects on traffic flow dynamics will be described in detail in Sections 4.4.2 and 4.4.3, respectively. Here we briefly describe their main principles. The *primary* strategy consists in performing a deceleration-acceleration-deceleration-acceleration (DADA) maneuver in the sag area. Remarkably, that strategy is used by all controlled vehicles in all scenarios. In all cases, the primary strategy limits the inflow to the sag and increases its outflow for a particular time interval, which produces considerable total travel time savings. The *supporting* strategy consists in performing one or more deceleration-acceleration maneuvers upstream of the vertical curve. That strategy is only used by some controlled vehicles in some scenarios. The role of the supporting strategy is to manage congestion upstream of the sag in such a way that the inflow to the vertical curve is regulated more effectively, which increases the effectiveness of the primary strategy.

Reduction in average vehicle delay (AVD)

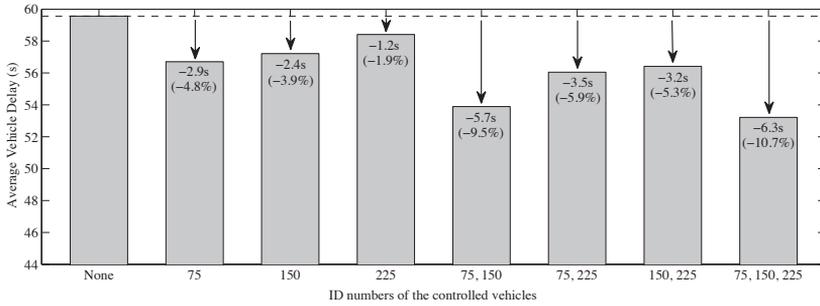
Figure 4.4(a) shows the average vehicle delay (AVD) in all scenarios. As seen in that figure, the optimal controller reduces the AVD by 2 to 11% (depending on the control scenario) in comparison with the no-control scenario. In our experiments, the greatest reduction in AVD is achieved in the scenarios with more controlled vehicles (Figure 4.4(a)). The reason is that every controlled vehicle causes a temporary period of increased sag outflow. Those periods do not overlap with each other. Therefore, the more controlled vehicles, the longer the total period of high outflow and, hence, the lower the total travel time. Note, however, that that rule might not apply if the controlled vehicles were located closer to each other and/or the number of controlled vehicles was higher. Under those circumstances, the periods of increased sag outflow caused by the controlled vehicles might overlap; hence having more vehicles behave as described in the previous paragraph might not lead to significant extra AVD savings.

If we compare scenarios with the same number of vehicles, the greatest reduction in AVD is observed in the scenarios in which the controlled vehicles are closer to the first vehicle of the stream (see Figure 4.4(a)). The reason is as follows. In all control

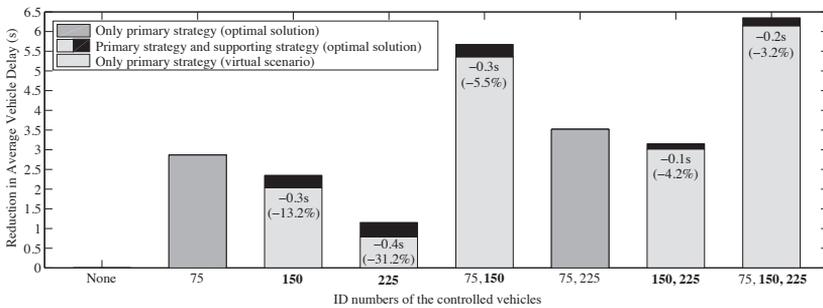
scenarios, the controlled vehicles' behavior has a similar impact on the time headways of the following vehicles at point *R*. Therefore, the contribution to total travel time of those following vehicles decreases to a greater extent in comparison with the no-control scenario if they are closer to the first vehicle of the stream, because then they influence the travel time of a greater number of vehicles (see Appendix).

Influence of each strategy on AVD reduction

Let us now compare the AVD in every scenario in which some or all controlled vehicles use the supporting strategy with the AVD in corresponding *virtual* scenarios in which



(a) Average vehicle delay (AVD) in every scenario. The numbers inside the bars indicate the absolute and relative difference in comparison with the no-control scenario.



(b) Reduction in average vehicle delay (AVD) in every control scenario in comparison with the no-control scenario. In the horizontal axis, numbers in normal font indicate controlled vehicles that use only the primary strategy, whereas numbers in bold indicate controlled vehicles that use both the primary strategy and the supporting strategy. For the control scenarios in which controlled vehicles use the supporting control strategy, the figure also shows the AVD reduction in the corresponding virtual scenarios (in which the supporting strategy is not applied). For those scenarios, the numbers inside the bars indicate the absolute and relative difference in AVD reduction between the virtual scenario (without supporting strategy) and the control scenario with the optimal solution (which includes the supporting strategy).

Figure 4.4: Average vehicle delay in every scenario and differences in comparison with the no-control scenario.

the supporting strategy is omitted. The behavior of controlled vehicles in the virtual scenarios is defined on the basis of the solutions obtained for the control scenarios: in the virtual scenarios, controlled vehicles perform a DADA maneuver in the sag area that has exactly the same characteristics as in the control scenario, but they do not perform any additional maneuver upstream of the sag. As shown in Figure 4.4(b), in the scenarios in which some controlled vehicles use the supporting strategy, the decrease in AVD in comparison with the no-control scenario would be 0.1 to 0.4 s lower if those vehicles did not use that strategy. Except for the control scenario with $M = \{225\}$, that represents a small (although significant) decrease in AVD reduction (3-13% depending on the scenario). We conclude that making some controlled vehicles apply the supporting strategy in combination with the primary strategy can reduce the AVD to a slightly greater extent than if those vehicles use only the primary strategy. However, in all scenarios, the primary strategy is the one that is more responsible for the reduction in average vehicle delay (see Figure 4.4(b)).

4.4.2 Primary strategy: deceleration - acceleration - deceleration - acceleration (DADA) maneuver in the sag area

This section describes the primary strategy applied by controlled vehicles, including the main characteristics of the maneuver as well as the typical effects of the maneuver on traffic flow and total travel time, respectively. For illustration purpose, throughout this section we show the results of the optimization experiment corresponding to the scenario with $M = \{75\}$, which constitute a good example of the typical acceleration behavior of controlled vehicles in the sag area and the resulting effects.

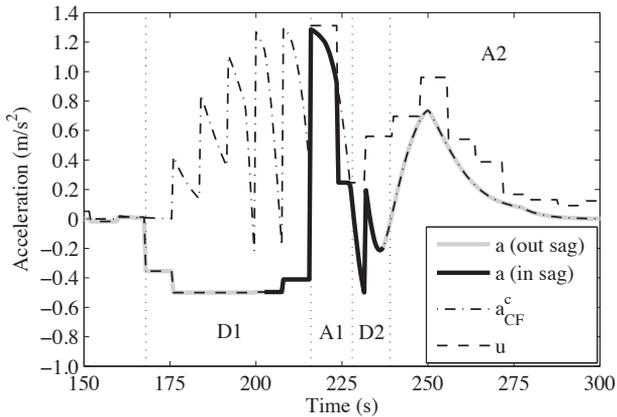
Characteristics of the DADA maneuver

In all scenarios, all controlled vehicles behave similarly when they reach the sag area. They perform a maneuver consisting of four phases (see, for example, Figures 4.5 and 4.6): first deceleration phase, first positive-acceleration phase, second deceleration phase and second positive-acceleration phase.

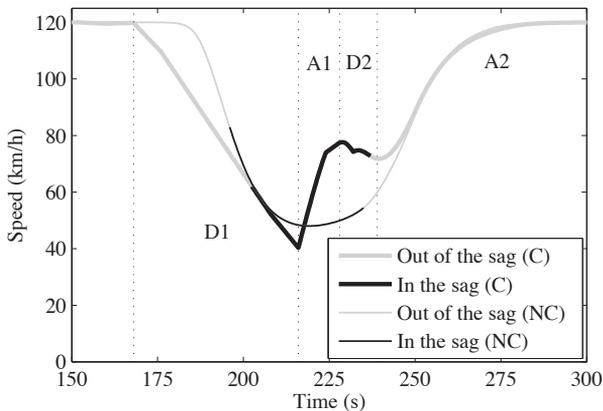
The *first deceleration phase* (D1) begins upstream of the sag or right after entering it (depending on the case). In this phase, controlled vehicles decelerate moderately (normally at around the minimum acceleration rate allowed by the controller, u_{\min}), even though they are not obliged to decelerate that fast in order to adjust to the behavior of the leader. During this phase, the distance headway of controlled vehicles increases considerably.

The *first positive-acceleration phase* (A1) begins somewhere halfway through the vertical curve. During this phase, controlled vehicles accelerate fast (with maximum acceleration rates up to 1 m/s² or higher). Fast acceleration is possible because the distance between controlled vehicles and their predecessor is quite long. The distance gap decreases quickly during this phase.

When controlled vehicles reach the last part of the vertical curve, they undergo a *second deceleration phase* (D2). At this point, their distance headway has become short enough to force them to decelerate a bit in order to adjust to the behavior of their leader. However, during this phase, the distance headway of controlled vehicles continues to decrease, because their predecessor moves at a lower speed.



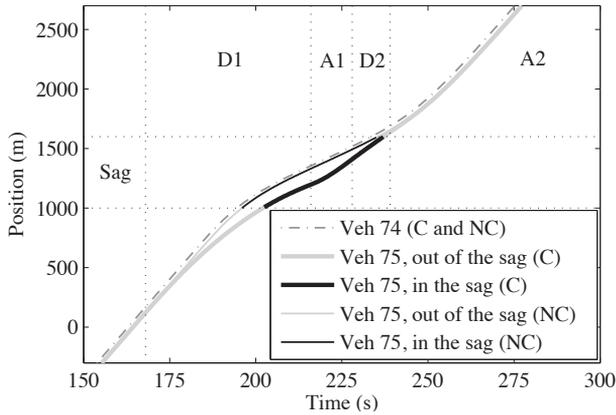
(a) Acceleration of vehicle 75 over time in the scenario with $M = \{75\}$ (a is the realized acceleration, u denotes the control input, and a_{CF}^c denotes the acceleration given by the car-following model). Note that a increases and decreases rapidly between instants 232 and 236 s; that is just a by-product of the interaction between u and a_{CF}^c at the previous control time step (the average acceleration rate in phase D2 is -0.13 m/s^2). This phenomenon is also observed in other scenarios.



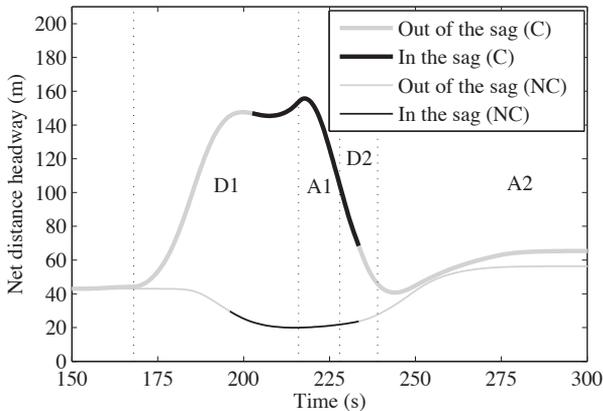
(b) Speed of vehicle 75 over time in the scenario with $M = \{75\}$ (C) and the no-control scenario (NC).

Figure 4.5: Deceleration-acceleration-deceleration-acceleration (DADA) maneuver of the controlled vehicle in the scenario with $M = \{75\}$: acceleration and speed profiles.

Eventually, controlled vehicles catch up with their leader. This generally occurs around the end of the vertical curve. At that point, the *second positive-acceleration phase* (A2) begins. In this phase, controlled vehicles simply accelerate to the desired speed. Acceleration is lower than in phase A1. Note that, after exiting the sag, the trajectories of controlled vehicles are almost the same as in the no-control scenario, although in some cases the distance headway is slightly longer.



(a) Positions of vehicle 75 and its leader over time in the scenario with $M = \{75\}$ (C) and the no-control scenario (NC).



(b) Net distance headway of vehicle 75 over time in the scenario with $M = \{75\}$ (C) and the no-control scenario (NC).

Figure 4.6: Deceleration-acceleration-deceleration-acceleration (DADA) maneuver of the controlled vehicle in the scenario with $M = \{75\}$: position and net distance headway over time

Effects on traffic flow dynamics in the sag area

In all cases, DADA maneuvers have three main effects on traffic flow dynamics. Firstly, the first positive-acceleration phase (A1) of the maneuver induces the first vehicles located behind the controlled vehicle (up to 85 vehicles in some cases) to accelerate fast along the sag (see, for example, Figure 4.7). As a result, traffic speed at the end of the sag (bottleneck) increases and stays moderately high (70-90 km/h) for a particular period (2-3 min), contrary to what happens in the no-control scenario (compare Figures 4.8(a) and 4.8(b)). That has important implications for the sag outflow. The car-following model assumes that the desired time headway of drivers decreases if the speed increases. Hence, during the period of high traffic speeds, time headways at the end of the vertical curve (bottleneck) are considerably shorter than in the no-control scenario, which implies that the sag outflow is higher (generally, around 5% higher).

Secondly, the first deceleration phase (D1) of the maneuver triggers a stop-and-go wave on the first part of the sag (see Figure 4.8(b)). That stop-and-go wave temporarily limits the inflow to the vertical curve. The reason why that is beneficial is as follows. The positive effects that phase A1 has on traffic speed and flow at the bottleneck are only temporary. Shortly after a DADA maneuver, traffic speed at the end of the sag begins to decrease (due, mainly, to the limiting effect that the vertical curve has on vehicle acceleration). As a result, traffic eventually becomes congested at that location (see Figure 4.8(b)), at which point the outflow from the sag becomes similar to that in the no-control scenario. Limiting the inflow to the vertical curve slows down the process

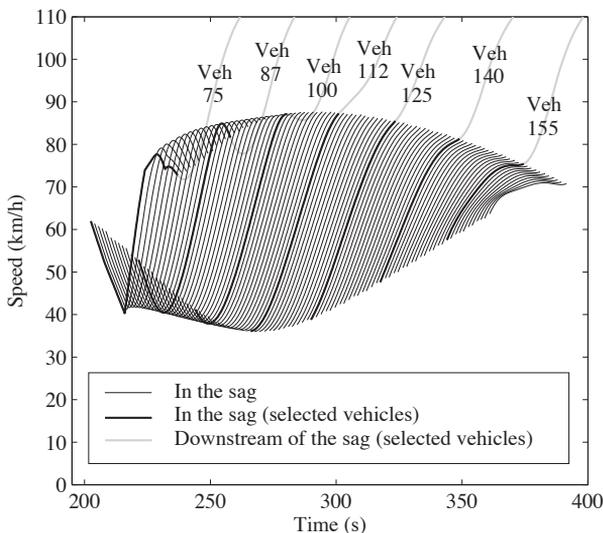
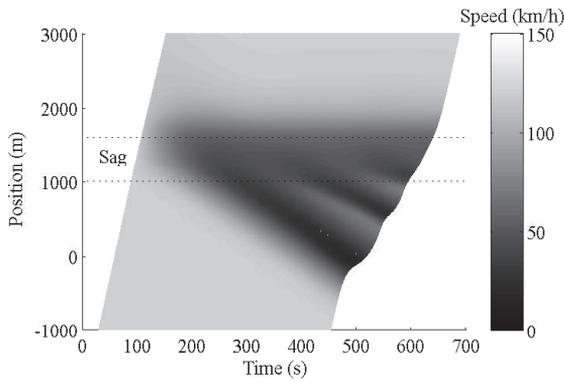


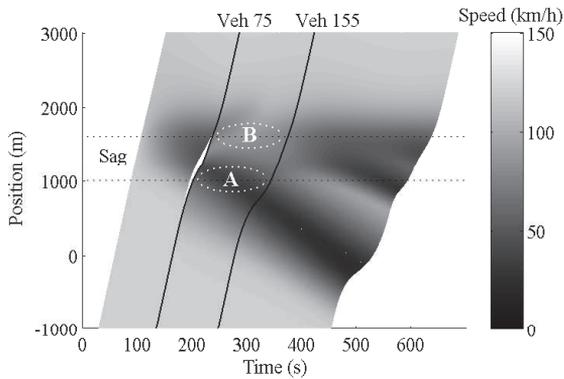
Figure 4.7: Effect of the DADA maneuver of the controlled vehicle on the behavior of the first 90 following vehicles in the control scenario with $M = \{75\}$.

of formation of congestion at the end of the sag (Goñi-Ros, et al., 2014b), hence high levels of sag outflow can be maintained for a longer period of time.

Finally, the second deceleration phase (D2) of the maneuver causes a small speed disturbance on the last part of the sag, which fades away as it propagates downstream (see Figure 4.7). The reason why that is beneficial is as follows. The occurrence of the speed disturbance makes the traffic speed at the end of the vertical curve increase for a short period of time right after the controlled vehicle passes that location, thus delaying the formation of congestion at the bottleneck (see Figure 4.7). In addition, as explained in the next section, the speed disturbance influences the time headways of the first group of vehicles located behind the controlled vehicle, which has relevant effects on total travel time.



(a) No-control scenario.

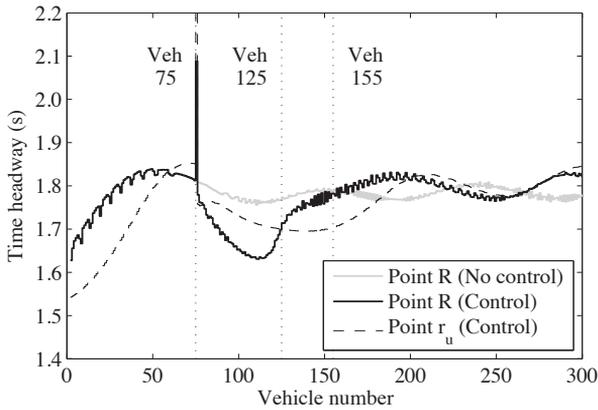


(b) Scenario with $M = \{75\}$. *A* is a region of moderately low traffic speed and limited flow around the beginning of the sag; *B* is a region of moderately high traffic speed and high flow around the end of the sag.

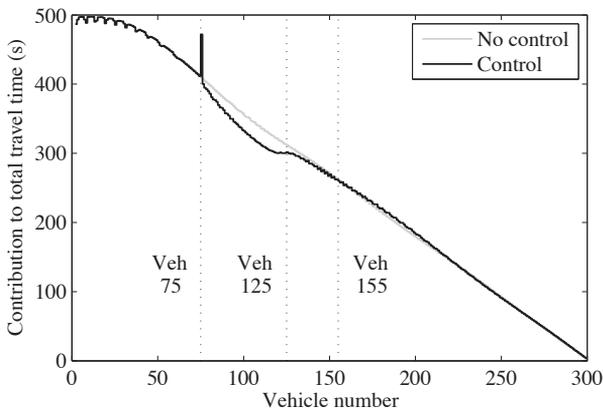
Figure 4.8: Speed contour plots of the no-control scenario and the control scenario with $M = \{75\}$.

Effects on total travel time

In all control scenarios, the vehicles that cause the decrease in total travel time in comparison with the no-control scenario are primarily the ones that exit the sag during the periods of high outflow induced by controlled vehicles. As explained in the Appendix, that is deduced from the fact that those vehicles keep shorter time headways at the arrival point (R) than in the no-control scenario, whereas most other vehicles do not (see, for example, Figure 4.9(a)).



(a) Time headway of every vehicle at points r_u and R . Time headways at point R oscillate due to differences in speed profile between consecutive vehicles as they accelerate to the desired speed after exiting the sag.



(b) Contribution to total travel time (ψ_i) of vehicles 2 to 300 (see Appendix). The area below each line corresponds to the total travel time in each scenario minus the contribution of vehicle 1 (which is the same in both scenarios).

Figure 4.9: Effect of the DADA maneuver of the controlled vehicle on the total travel time in the control scenario with $M = \{75\}$.

Furthermore, among the vehicles that exit the sag during the periods of high outflow, the ones that reduce the total travel time to a greater extent are generally the first 40-50 vehicles located behind each controlled vehicle (Figure 4.9(b)). Those vehicles show the greatest differences in time headway at point R in comparison with the no-control scenario (Figure 4.9(a)). The main reason is that their time headways decrease after exiting the sag (Figure 4.9(a)), whereas the time headways of the next group of vehicles increase. That occurs because the first 40-50 vehicles located behind a controlled vehicle accelerate to the desired speed a little faster than their immediate predecessors, since they are affected by the speed disturbance caused by the second deceleration phase of the controlled vehicle (Figure 4.7). Instead, the next vehicles accelerate to the desired speed a bit more slowly than their immediate predecessors, because they exit the sag at lower speeds (due to the vertical curvature) and they are not affected by the aforementioned speed disturbance (Figure 4.7).

We conclude that an important reason why DADA maneuvers include a second deceleration phase (D2) is that the resulting speed disturbance makes the first group of following vehicles reduce their time headways after exiting the vertical curve. The vehicles that are not affected by the disturbance increase their headways after exiting the sag. Therefore, the second deceleration phase increases the effectiveness of controlled vehicles in reducing total travel time.

It is worth noting that, in all control scenarios, the time headway of some vehicles at point R (including the controlled vehicle, in some cases) is longer than in the no-control scenario (see, for example, Figure 4.9(a)). Overall, however, the increase in the contributions to total travel time of those vehicles is compensated by the decrease in the contributions of the other vehicles (Figure 4.9(b)).

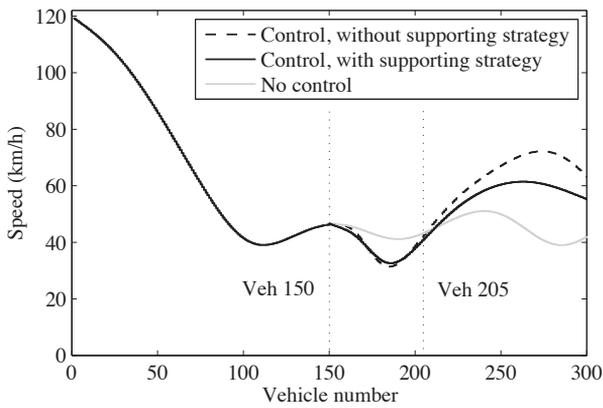
4.4.3 Supporting strategy: deceleration - acceleration maneuvers upstream of the sag area

In some scenarios, some controlled vehicles apply an additional strategy besides performing a DADA maneuver in the sag area, which we call *supporting* strategy. The supporting strategy consists in performing one or more deceleration-acceleration maneuvers upstream of the sag, catching up with the preceding vehicle before entering the vertical curve. The characteristics of those maneuvers are very situation-specific: in different scenarios, they start in different locations, they are quicker or slower, they are performed consecutively or separately, and they last for different periods of time.

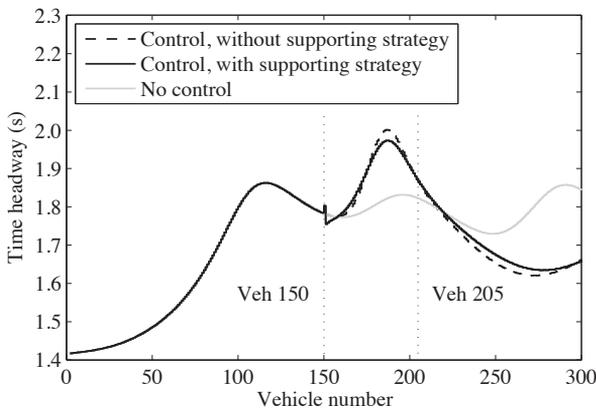
However, their overall effect on traffic flow dynamics is similar. Essentially, those maneuvers change the location and severity of congestion upstream of the vertical curve in such a way that, after the stop-and-go wave caused by the primary strategy moves away from the sag, the inflow to the bottleneck stays somewhat lower than if the supporting strategy was not applied (see, for instance, Figure 4.10(b)). As a result, the primary strategy is able to produce higher traffic speeds and flows at the end of the sag

after the above-mentioned stop-and-go wave moves away from the vertical curve (see Figures 4.11(a) and 4.11(b)), which brings about some extra total travel time savings (see Figure 4.4(b)).

It is important to highlight that the supporting strategy is not applied alone in any scenario. In fact, as explained in Section 4.4.1, if we compare the average vehicle delay (AVD) in every scenario in which some or all controlled vehicles use the supporting strategy with the AVD in corresponding virtual scenarios in which the supporting strategy is omitted, we can observe that the primary strategy is the one that contributes the most to reduce the total travel time (see Figure 4.4(b)).



(a) Speed of every vehicle at point $r_d + 150$ m.



(b) Time headway of every vehicle at point $r_d + 150$ m.

Figure 4.10: Speed and time headway of every vehicle on the first part of the sag in the no-control scenario, in the control scenario with $M = \{150\}$ (optimal solution) and in the virtual scenario without supporting strategy corresponding to that control scenario.

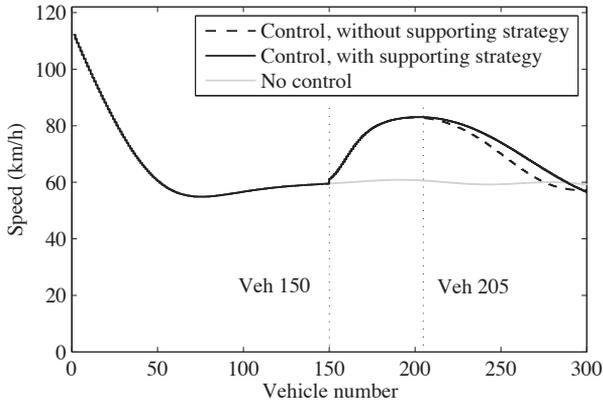
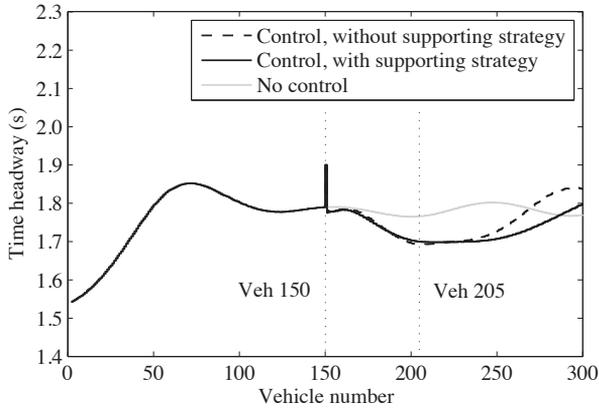
(a) Speed of every vehicle at the end of the sag (r_u).(b) Time headway of every vehicle at the end of the sag (r_u).

Figure 4.11: Speed and time headway of every vehicle at the end of the sag in the no-control scenario, in the control scenario with $M = \{150\}$ (optimal solution) and in the virtual scenario without supporting strategy corresponding to that control scenario.

4.5 Conclusions

The goal of this paper was to identify the main strategies that define the optimal acceleration behavior of vehicles equipped with in-car systems at sags and their effects on traffic flow, considering as objective the minimization of total travel time. To this end, we formulated an optimal control problem in which a centralized controller regulates the acceleration of some vehicles belonging to a traffic stream that moves along a one-lane freeway stretch with a sag. The problem was solved for various scenarios defined by the number of controlled vehicles and their positions in the stream, assuming

penetration rates that are realistic for the coming years (not higher than 1%).

The results show that traffic management measures based on the use of in-car systems have the potential to considerably reduce congestion at sags, even if the number of equipped vehicles is low. The optimal acceleration behavior of equipped vehicles is defined by two strategies, which we called primary strategy and supporting strategy. The *primary* strategy is applied in all cases and is the one that contributes the most to reduce the total travel time. It involves performing a specific type of deceleration-acceleration-deceleration-acceleration (DADA) maneuver in the sag area. This maneuver induces the first group of following vehicles to accelerate fast along the sag. As a result, traffic speeds at the end of the sag (bottleneck) increase for a particular time interval. During this period, the sag outflow increases up to 5% in comparison with the no-control scenario, which produces substantial travel time savings for all vehicles located behind the equipped vehicle. Interestingly, every DADA maneuver triggers a stop-and-go wave on the first part of the sag that temporarily limits the inflow to the vertical curve. Limiting the inflow is beneficial because it slows down the process of formation of congestion at the end of the sag (bottleneck), hence high levels of sag outflow can be maintained for a longer period of time. The *supporting* strategy consists in performing one or more deceleration-acceleration maneuvers upstream of the sag. Controlled vehicles use the supporting strategy to manage congestion upstream of the sag in such a way that the inflow to the vertical curve is regulated more effectively, thus enhancing the effectiveness of the primary strategy.

Our findings provide valuable insight into how congestion can be reduced at sags by means of traffic management measures using in-car systems. More specifically, they highlight the relevance of limiting the inflow to sags—in line with Goñi-Ros, et al. (2014b)—and of motivating drivers to accelerate fast along the vertical curve; in addition, they indicate ways to do that by regulating the acceleration of vehicles equipped with in-car systems. Moreover, our findings prove the usefulness of the proposed optimization method as a tool for control measure development. We conclude that this method can be used to identify effective traffic management strategies for other types of bottlenecks besides sags.

Further research is necessary to determine whether the traffic management strategies identified in this paper would also be the most effective in other scenarios (such as scenarios with multi-lane freeways, different positions of the equipped vehicles in the stream, higher penetration rates and/or lower traffic demand). In order to do this, additional optimization experiments need to be carried out. Also, further research is necessary to translate the identified strategies into effective traffic control measures. Mainly, these measures should induce vehicles equipped with in-car systems (e.g., cooperative ACC systems) to perform a DADA maneuver at sags in high demand conditions. Additional strategies may be necessary to facilitate the DADA maneuvers and to enhance their effectiveness in mitigating congestion.

4.6 Appendix

In Section 4.2.3, total travel time (TTT) was defined as the sum of the travel times of all vehicles i from their initial positions to point R (which are denoted by TT_i):

$$TTT = \sum_{i=1}^n TT_i \quad (4.23)$$

The time headway of vehicle i at point R can be defined as the travel time of vehicle i minus the travel time of its predecessor:

$$h_{i,R} = TT_i - TT_{i-1}, \quad \text{for } i = 2, \dots, n \quad (4.24)$$

From Equation 4.24, it can be derived that the travel time of vehicle $i = 2, \dots, n$ is equal to the travel time of vehicle 1 plus the headways of vehicles 2 to i :

$$TT_i = \begin{cases} TT_1 & \text{if } i = 1 \\ TT_1 + \sum_{v=2}^i h_{v,R} & \text{if } i = 2, \dots, n \end{cases} \quad (4.25)$$

By substituting Equation 4.25 into Equation 4.23 and rearranging the terms, the total travel time can be expressed as a function of the travel time of vehicle 1 and the headways of vehicles 2 to n at point R :

$$TTT = n \cdot TT_1 + \sum_{i=2}^n ((n-i+1) \cdot h_{i,R}) \quad (4.26)$$

From Equation 4.26, it follows that the contribution of each vehicle to total travel time (Ψ_i) is:

$$\Psi_i = \begin{cases} n \cdot TT_1 & \text{if } i = 1 \\ (n-i+1) \cdot h_{i,R} & \text{if } i = 2, \dots, n \end{cases} \quad (4.27)$$

The sum of the contributions to total travel time of all vehicles equals the total travel time:

$$TTT = \sum_{i=1}^n \Psi_i \quad (4.28)$$

We conclude that the vehicles that reduce the total travel time the most in a given control scenario are the ones whose Ψ_i decreases to a greater extent in comparison with the no-control scenario. From Equation 4.27, it follows that Ψ_i decreases more for vehicles whose time headways at point R decrease to a greater extent and are located closer to the first vehicle of the stream.

Chapter 5

Mainstream traffic flow control at sags

Goñi-Ros, B., Knoop, V.L., van Arem, B., Hoogendoorn, S.P. (2014). Mainstream Traffic Flow Control at Sags. *Transportation Research Record: Journal of the Transportation Research Board*, No. 2470, 57–64. © Transportation Research Board of the National Academies.

Abstract

Sags are freeway sections along which the gradient changes significantly from downward to upward. The capacity of sags is considerably lower than the capacity of normal sections. Consequently, sags are often freeway bottlenecks. Recently, several control measures have been proposed to improve traffic flow efficiency at sags. Those measures generally aim to increase the capacity of the bottleneck, to prevent traffic flow perturbations in nearly saturated conditions, or both. This paper presents an alternative type of measure based on the concept of mainstream traffic flow control. The proposed control measure regulates traffic density at the bottleneck area to keep it below the critical density and hence prevent traffic from breaking down while maximizing outflow. Density is regulated by means of a variable speed limit section that regulates the inflow to the bottleneck. Speed limits are selected on the basis of a feedback control law. The authors evaluate the effectiveness of the proposed control strategy by means of a simple case study by using microscopic traffic simulation. The results show a significant increase in bottleneck outflow, particularly during periods of high demand, which leads to a considerable decrease in total delay. This finding suggests that mainstream traffic flow control strategies that use variable speed limits have the potential to improve substantially the performance of freeway networks containing sags.

5.1 Introduction

Sags are freeway sections along which the gradient changes significantly from downward to upward in the direction of traffic (Furuichi et al., 2003). The capacity of sags is considerably lower than that of normal freeway sections (Koshi et al., 1992; Xing et al., 2010). In general, a bottleneck is located 0.5 to 1 km downstream of the bottom of the sag (Brilon & Bressler, 2004). As a consequence of the reduced capacity, traffic often breaks down at sags in conditions of high demand. The formation of congestion results in a further decrease in bottleneck capacity (Koshi et al., 1992). Recently, various control measures have been proposed to improve traffic flow efficiency at sags. Generally, those measures aim to increase the capacity of the bottleneck, to prevent traffic flow perturbations in nearly saturated conditions, or both.

This paper presents an alternative type of control measure and evaluates its potential effectiveness, performing a proof of principle. The proposed measure is based on the concept of mainstream traffic flow control (Carlson et al., 2011). The traffic density at the bottleneck area is regulated to keep it below the critical density and hence prevent traffic from breaking down. The capacity drop caused by congestion does not occur, so the outflow from the bottleneck can be higher. The density at the bottleneck area is regulated by means of a variable speed limit section that regulates the inflow to the bottleneck. Speed limits are selected on the basis of a proportional-feedback control law. The effectiveness of the proposed control measure is evaluated by means of a simple case study by using microscopic traffic simulation. Traffic flow is simulated in a single-lane freeway stretch containing a sag, with and without implementing the control strategy. The results show that the control measure increases the bottleneck outflow significantly (particularly in periods of extremely high demand), and this increase in flow leads to a considerable decrease in total delay (TD). This finding suggests that mainstream traffic flow control strategies using variable speed limits can considerably improve traffic flow efficiency in freeway networks containing sags.

The next section contains a literature review on the characteristics of traffic flow at sags and on types of control measures for mitigating congestion at that type of bottleneck. Then, the proposed control strategy and the method used to evaluate its effectiveness are described. Next, the results of the evaluation (including a sensitivity analysis) are presented. The final section presents the conclusions of this study and some suggestions for future research.

5.2 Background

5.2.1 Sags as freeway bottlenecks

Bottlenecks are freeway sections that have a lower capacity than the immediate upstream section. Generally, the causes of that lower capacity are spatial inhomogeneities (e.g., lane drops, ramps, tunnels), traffic conditions (e.g., slow vehicles, accidents), and environmental conditions (e.g., adverse weather) (Helbing, 2001; Schönhof & Helbing, 2007). When traffic demand exceeds the capacity of a bottleneck, congestion forms upstream of the bottleneck. The capacity of a bottleneck depends on the traffic state: the capacity in congested traffic conditions (queue discharge capacity) is generally lower than the capacity in uncongested traffic conditions (free-flow capacity). The difference, called capacity drop, ranges from 3% to 20% according to different studies (Hall & Agyemang-Duah, 1991; Cassidy & Bertini, 1999; Srivastava & Geroliminis, 2013).

Several empirical studies show that the capacity of sags can be significantly lower than the capacity of flat sections (Koshi et al., 1992; Xing et al., 2010). In general, the bottleneck is located 0.5 to 1 km downstream of the bottom of the sag (Brilon & Bressler, 2004). Xing et al. (2010) present empirical measurements of free-flow capacities and queue discharge capacities of various sag sections of Japanese freeways. Most of the measurements were taken on holidays, when traffic demand consists mainly of passenger cars and the percentage of heavy vehicles is relatively low. According to the data presented in that study, the average free-flow capacity is 3,150 vehicles per hour (vph) at two-lane sags and 5,340 vph at three-lane sags. The average queue discharge capacity is 2,780 vph at two-lane sags and 4,600 vph at three-lane sags, which means that the capacity drop is 12% and 14%, respectively. Similar capacity estimates have been reported by other authors (Koshi et al., 1992; Patire & Cassidy, 2011).

In a comparison of the capacities of sags with those of flat sections, one observes that the free-flow capacity and the queue discharge capacity of sags are considerably lower. At flat sections, free-flow capacities are generally around 4,000 passenger car units per hour (pcu/h) (for two lanes) and 6,000 pcu/h (for three lanes) (Koshi et al., 1992). If a 10% capacity drop is assumed, queue discharge capacities for flat sections of 3,600 pcu/h (for two lanes) and 5,400 pcu/h (for three lanes) are obtained. Therefore, the free-flow capacity and the queue discharge capacity of two-lane freeways are around 20% lower at sags than at flat sections (10% to 15% lower in three-lane freeways).

The main cause of capacity reduction at sags seems to be related to the impact of the increase in freeway gradient on the behavior of drivers. Several empirical studies show that two important changes in longitudinal driving behavior occur when vehicles go through a sag. First, drivers tend to reduce speed (Furuichi et al., 2003; Brilon & Bressler, 2004). Second, drivers tend to keep longer distance headways than expected

given their speed (Koshi, 2003; Goñi-Ros, et al., 2014a). These local changes in driving behavior seem to be caused by the inability of drivers to accelerate sufficiently and compensate for the increase in the resistance force resulting from the increase in slope (Yoshizawa et al., 2012).

5.2.2 Control measures to mitigate congestion at sags

In the last two decades, several measures have been proposed to prevent or delay the formation of congestion at sags and to reduce its severity. In general, those measures can be sorted into three categories: measures that aim (a) to increase the free-flow capacity of sag bottlenecks, (b) to prevent traffic flow perturbations at sag bottlenecks in nearly saturated conditions, and (c) to increase the queue discharge capacity of active sag bottlenecks. An example of a measure from the first category is equipping vehicles with adaptive cruise control systems, which perform the acceleration task more efficiently than human drivers (Ozaki, 2003). Another example is distribution of the traffic flow more evenly across lanes to use the bottleneck capacity more efficiently (Hatakenaka et al., 2006; Xing et al., 2010). The second category comprises measures such as preventing the formation of long vehicle platoons (Hatakenaka et al., 2006) and discouraging drivers from performing lane changes to the busiest lanes (Hatakenaka et al., 2006; Patire & Cassidy, 2011). The third category comprises measures such as giving information to drivers about the location of the head of the queue and encouraging them to recover speed after leaving congestion (Xing et al., 2007; Sato et al., 2009). In addition, control measures belonging to the above-mentioned categories have been proposed for other types of bottlenecks besides sags [e.g., on-ramps (Laval et al., 2007) and weaving sections (Daganzo, 2002)]. The potential effectiveness of most of those measures has been demonstrated by means of empirical data analysis or simulation.

However, an additional category of measures could improve traffic flow efficiency at sags, but it has received little attention in the recent literature, namely measures of mainstream traffic flow control. In mainstream traffic flow control, the inflow to a bottleneck is regulated by creating a controlled section upstream. The traffic density at the bottleneck area is kept below the critical density. Consequently, even if demand gets extremely high, traffic does not break down at the bottleneck; the capacity drop does not occur, so the outflow from the bottleneck can be higher than its queue discharge capacity. Mainstream traffic flow control is a concept that was first applied in the 1950s (Greensberg & Daou, 1960). Recently, it has been presented as an effective measure to mitigate congestion at on-ramp bottlenecks (Carlson et al., 2011). The current authors argue that mainstream traffic flow control can also be used to improve traffic flow efficiency at sags, either by itself or in combination with other types of measures. This control concept can result in relevant improvements in traffic flow efficiency only if the queue discharge capacity of the bottleneck is significantly lower than the queue discharge rate of the controlled section. This relationship is usually the case

with sag bottlenecks (Koshi et al., 1992; Xing et al., 2010; Patire & Cassidy, 2011).

5.3 Control strategy

This section describes the characteristics of a strategy for mainstream traffic flow control aimed at mitigating congestion at sags. The control goal is to minimize the total time spent by vehicles in the network over a certain period. If one assumes that the flow entering the network cannot be influenced by any control measure, then minimizing the total time spent is equivalent to maximizing the time-weighted sum of exit flows (Papageorgiou et al., 2003). For the sake of simplicity, the authors consider a simple network consisting of a freeway stretch with a sag (bottleneck) without any on- or off-ramps. Hence, the network designated for control has a single entry point and a single exit point. However, the control strategy described in this section could be generalized to more complex networks, possibly in combination with other control measures.

5.3.1 Control concept: mainstream traffic flow control

The outflow from a sag bottleneck (q_b) is lower than or equal to its capacity ($q_{b,\max}$) regardless of the traffic demand. Therefore, with no other bottleneck in or downstream of the network, the network exit flow (s) is mainly constrained by the capacity of the sag bottleneck:

$$s \approx q_b \leq q_{b,\max} \quad (5.1)$$

As noted earlier, the capacity of a bottleneck depends on the traffic state: the queue discharge capacity of the bottleneck ($q_{b,\max}^c$) is lower than its free-flow capacity ($q_{b,\max}^f$):

$$q_{b,\max} = \begin{cases} q_{b,\max}^f & \text{in uncongested traffic conditions} \\ q_{b,\max}^c & \text{in congested traffic conditions} \end{cases} \quad (5.2)$$

where

$$q_{b,\max}^c < q_{b,\max}^f \quad (5.3)$$

Because network exit flows (s) can be higher if traffic flow at the bottleneck is uncongested than if it is congested, a way to maximize the time-weighted sum of exit flows in the network (control goal) is to prevent traffic from breaking down at the sag bottleneck area. To that end, the authors propose a control strategy that is based on the concept of mainstream traffic flow control. The control strategy aims to regulate the traffic inflow to the sag bottleneck ($q_{b,\text{in}}$) to achieve a desired traffic state at the bottleneck that maximizes outflow. The inflow to the sag bottleneck is regulated by means of a controlled section upstream of the bottleneck (Figure 5.1). On that controlled section, the speed limit is variable. Speed limits are set by the controller on the basis of measurements of the traffic conditions (density) at the bottleneck. As a result

of the fundamental relation between traffic speed and flow, the outflow from the controlled section (q_c) depends on the speed limit (if one assumes that drivers comply with it). The inflow to the bottleneck is approximately equal to the outflow from the controlled section ($q_{b,in} \approx q_c$). By applying an appropriate speed limit on the controlled section, the inflow to the bottleneck can be kept slightly below its free-flow capacity ($q_c \approx q_{b,in} < q_{b,max}^f$). Therefore, even in conditions of high demand, the density at the bottleneck does not go above the critical density and traffic does not break down at the bottleneck area (Figure 5.1). Yet congestion is not completely prevented: traffic flow becomes congested on the controlled section and upstream of it. However, if an appropriate speed limit is applied, the outflow from the controlled section can be higher than the queue discharge capacity of the bottleneck ($q_c < q_{b,max}^c$). As a result, higher exit flows (s) can be obtained than if traffic flow becomes congested at the bottleneck area (Figure 5.1). This condition should result in a higher time-weighted sum of exit flows and a lower total time spent.

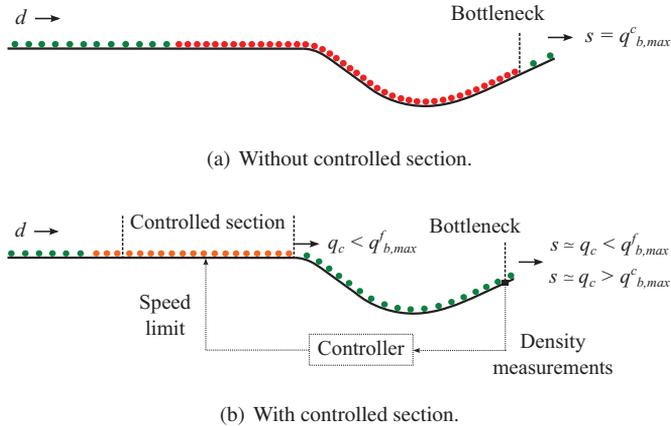


Figure 5.1: Flows within network in two scenarios: without controlled section and with controlled section ($d =$ demand flow).

5.3.2 Control law: proportional feedback

The controller determines the speed limits to be applied on the controlled section by means of a proportional feedback control law that is similar in nature to the one used by the ramp-metering control algorithm ALINEA (Papageorgiou et al., 1991). The control law requires (a) a target traffic density at the sag bottleneck area and (b) real-time measurements of the density at the sag bottleneck area. As explained earlier, the target density should be slightly lower than the critical density of the bottleneck. The density at the bottleneck is measured in real time by means of loop detectors. The control law determines the speed limit to be applied on the controlled section (v_{lim}) based on the difference between the target density ($\rho_{b,0}$) and the measured density

(ρ_b). The speed limit is reevaluated each time that the controller receives a new density measurement; hence, the control time step period (T_c) is equal to the sampling period of the detector (T_s). However, a delay ($r \cdot T_c$) occurs between the time the detector time sampling period finishes and the time the new speed limit is actually applied on the control section.

$$v_{\text{lim}}(k) = v_{\text{lim},0} + K_p \cdot (\rho_{b,0} - \rho_b(k-r)) \quad (5.4)$$

where k = control time step index, K_p = proportional gain, r = control time step delay, and $v_{\text{lim},0}$ = target speed limit when $\rho_{b,0}$ is equal to $\rho_b(k-r)$.

Furthermore, three constraints are imposed on the speed limits displayed on the message signs to make it easier for drivers to comply with them. First, the speed limit is always rounded to a value multiple of 10. Second, the speed limit cannot be lower than a minimum threshold ($v_{\text{lim},\text{min}}$). Third, the change in speed limit between two consecutive control steps cannot be higher than a maximum change rate (Δv_{lim}).

By means of the feedback control law described earlier, the controller should be able to regulate dynamically the speed limit on the controlled section so that the outflow from the bottleneck is maximized. In stationary high-demand conditions, the controller maintains the density (ρ_b) near the target value ($\rho_{b,0}$) to prevent traffic from breaking down at the bottleneck. Furthermore, the controller should be able to react immediately to density deviations. If the measured density is significantly lower than the target density (e.g., because the demand is low), the controller will choose to apply a high speed limit (or even the regular speed limit) to maximize the inflow to the bottleneck. If the measured density is higher than the target density (e.g., because traffic has broken down at the bottleneck), the controller will choose to apply a lower speed limit to reduce the density at the bottleneck to the target value. The latter process is extremely important because traffic flow in nearly saturated conditions can easily destabilize and become congested, and the controller must be able to react to that possibility. Finally, the controller reacts to density deviations with a certain delay. This delay is the result of the control delay ($r \cdot T_c$) but also of the time needed by drivers to cover the distance between the controlled section and the bottleneck.

5.4 Method of performance evaluation

A case study was used to evaluate the performance of the proposed control strategy. A model of longitudinal driving behavior that takes into account the influence of changes in gradient on vehicle acceleration simulated traffic flow on a sag in two scenarios (Goñi-Ros et al., 2013b): (a) a no-control scenario (no control measures are implemented) and (b) a control scenario (the proposed control measure is operative). The performance of the control strategy was assessed by comparing the TD experienced by drivers in the two scenarios.

5.4.1 Model of longitudinal driving behavior

The model of longitudinal driving behavior determines the acceleration of every vehicle at each simulation time step. Vehicle acceleration is assumed to stay constant over the period $[t, t + \Delta t]$, where Δt is the simulation step period. The model determines vehicle acceleration (\dot{v}) by means of a two-term additive function:

$$\dot{v}(t) = f_r(t) + f_g(t) \quad (5.5)$$

where f_r describes regular car-following behavior and f_g is the freeway gradient term. The term f_r accounts for the influence of vehicle speed (v), relative speed to the leading vehicle (Δv), and net distance headway (s) on vehicle acceleration. The formulation of f_r is based on the IDM+ model (Schakel et al., 2012):

$$f_r(t) = a \cdot \min \left[1 - \left(\frac{v(t)}{v_{\text{des}}(x(t), t)} \right)^4, 1 - \left(\frac{s_{\text{des}}(v(t), \Delta v(t))}{s(t)} \right)^2 \right] \quad (5.6)$$

where

$$s_{\text{des}}(v(t), \Delta v(t)) = s_0 + v(t) \cdot \tau(v(t)) + \frac{v(t) \cdot \Delta v(t)}{2\sqrt{ab}} \quad (5.7)$$

and s_{des} = dynamic desired net distance headway, v_{des} = desired speed, x = position along the highway, t = time, a = maximum acceleration, b = maximum comfortable deceleration, s_0 = net distance headway at standstill, and τ = safe time headway.

The desired speed depends on the position along the freeway x and on time t because some freeway sections may have variable speed limits. Also, the value of τ depends on the traffic state. In congested traffic conditions (i.e., below the critical speed v_{crit}), the value of τ is higher than in uncongested conditions:

$$\tau(v(t)) = \begin{cases} \tau_f & \text{if } v(t) \geq v_{\text{crit}} \\ \gamma \cdot \tau_f & \text{if } v(t) < v_{\text{crit}} \end{cases} \quad (5.8)$$

where τ_f is the safe time headway in uncongested conditions and γ is a factor greater than 1.

The second term (f_g) in Equation 5.5 accounts for the influence of changes in freeway gradient on vehicle acceleration. At a given time t , this influence is gravity acceleration ($g = 9.81 \text{ m/s}^2$) multiplied by the difference between the gradient at the location of the vehicle at that time ($G(x(t))$) and the gradient compensated by the driver until that time ($G_c(t)$):

$$f_g(t) = -g \cdot [G(x(t)) - G_c(t)] \quad (5.9)$$

The compensated gradient (G_c) is a variable that accounts for drivers' limited ability to accelerate on freeway sections where the slope increases. The authors assume that drivers compensate for positive changes in slope linearly over time (with a maximum

rate of gradient compensation defined by parameter c). Furthermore, the authors assume that drivers can fully compensate for negative changes in gradient.

$$G_c(t) = \begin{cases} G(x(t)) & \text{if } G(x(t)) \leq G(t_c) + c \cdot (t - t_c) \\ G(t_c) + c \cdot (t - t_c) & \text{if } G(x(t)) > G(t_c) + c \cdot (t - t_c) \end{cases} \quad (5.10)$$

where

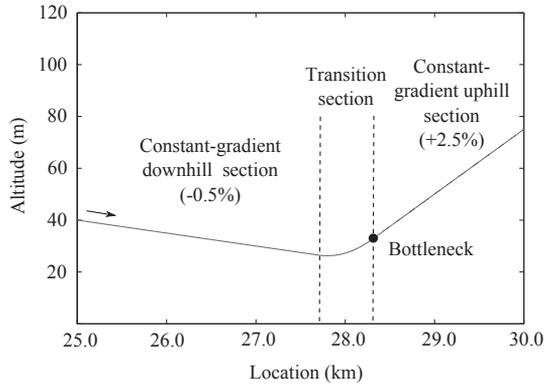
$$t_c = \max [t \mid G_c(t) = G(x(t))] \quad (5.11)$$

If the rate at which the freeway slope increases over time is lower than the driver's maximum gradient compensation rate (c), then $G_c(t) = G(t)$ for all t . Therefore, $f_g(t) = 0$ for all t , which means that vehicle acceleration is not affected by the increase in gradient. However, if the rate at which the freeway slope increases over time is higher than the driver's maximum rate of gradient compensation (c), then $G_c(t) < G(t)$ for a certain period. During that period, G_c increases linearly over time, but f_g is negative, which limits vehicle acceleration. This limitation in vehicle acceleration seems to be the main cause of local changes in longitudinal driving behavior that reduce the capacity of sags (Yoshizawa et al., 2012). The model of longitudinal driving behavior generates the main bottleneck of sags at the end of the transition section (Figure 5.2) because the maximum difference between $G_c(t)$ and $G(t)$ occurs at that location. This situation is in line with empirical observations (Koshi et al., 1992; Brilon & Bressler, 2004). In addition, the model of longitudinal driving behavior is face valid, but it has not yet been calibrated (Goñi-Ros et al., 2013b).

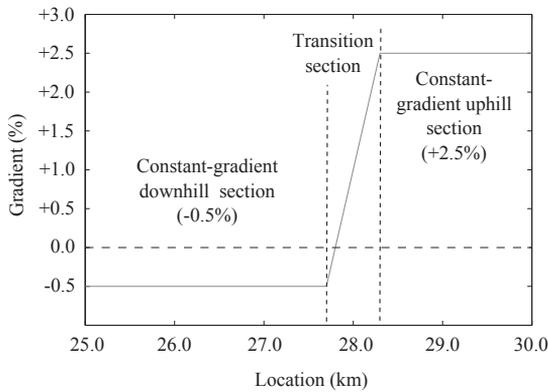
5.4.2 Simulation settings

Network characteristics

The simulated network is a freeway stretch 30 km long that contains a sag. The stretch has a constant-gradient downhill section that goes from location $x = 0$ to $x = 27.7$ km, a transition section that goes from $x = 27.7$ km to $x = 28.3$ km, and a constant-gradient uphill section that goes from $x = 28.3$ km to $x = 30.0$ km (Figure 5.2). On the transition section, the freeway slope increases linearly over distance. The long length of the freeway stretch ensures that the flow entering the network is not influenced by the traffic conditions at the sag bottleneck area. The regular speed limit is 120 km/h. The network has only one lane (with no overtaking possibilities) and no ramps or horizontal curves. Four detectors are in the network, and they are used to monitor traffic conditions at key locations: (a) network entry area ($x = 0.3$ km), (b) the area where the controlled section is located in the control scenario ($x = 27.0$ km), (c) the sag bottleneck area ($x = 28.3$ km), and (d) the network exit area ($x = 29.9$ km).



(a) Altitude versus location.



(b) Gradient versus location.

Figure 5.2: Vertical alignment of network (from $x = 25.0$ km to $x = 30.0$ km).

Traffic demand

The simulation period is 10,000 s. At $t = 0$, no vehicles are in the network. Network loading starts in the first simulation time step. The demand profile (i.e., flow at $x = 0$ over time) contains three periods that are relevant for testing the proposed control strategy. First, from $t = 0$ to $t = 2,000$ s, demand increases and goes above the capacity of the sag bottleneck. Second, from $t = 2,000$ s to $t = 3,000$ s, demand decreases considerably. Third, from $t = 3,000$ s to $t = 7,000$ s, demand increases again, goes above the capacity of the sag bottleneck, and stays at that level. The controller should be able to control traffic adequately in periods of high and low demand, and it should be able to react adequately to demand fluctuations. From $t = 9,000$ s to $t = 10,000$ s, demand is 0. This period of no demand is necessary to ensure that all vehicles exit the network before the end of the simulation period and to allow comparison of network

performance in different scenarios. Figure 5.3 shows the demand profile, including the flows measured by the detector at $x = 0.3$ km, during the whole simulation period.

Longitudinal driving behavior

For the sake of simplicity, homogeneous vehicle and driver characteristics are assumed. All vehicles are 4 m long. The parameters of the model of longitudinal driving behavior are shown in Table 5.1.

Table 5.1: Parameter values.

Long. driving behavior model		Controller	
Parameter	Value	Parameter	Value
v_{des} (km/h)	120	T_s (s)	30
a (m/s^2)	1.45	T_c (s)	30
b (m/s^2)	2.10	$v_{lim,0}$ (km/h)	60
τ_f (s)	1.20	K_p ($km^2/h/veh$)	4.8
s_0 (m)	3	$\rho_{b,0}$ (veh/km)	18.0
v_{crit} (km/h)	65	r (-)	2
γ (-)	1.15	$v_{lim,min}$ (km/h)	20
c (s^{-1})	0.0001	Δv_{lim} (km/h)	20
Δt (s)	0.5		

Control

In the control scenario, a controlled section is added to the network. In that section, the speed limit is variable and is displayed on message signs. The controlled section is 1.0 km long. That length gives drivers sufficient time to adapt to the speed limit before leaving the controlled section. The controlled section is between $x = 26.3$ km and $x = 27.3$ km. The downstream end of the controlled section is 1.0 km upstream of the end of the transition section (i.e., the bottleneck) to ensure that drivers have sufficient time to accelerate and that the traffic speed at the bottleneck is not influenced by the speed limit on the controlled section. Three message signs are located at different points of the controlled section: (a) upstream end ($x = 26.3$ km), (b) center point ($x = 26.8$ km), and (c) downstream end ($x = 27.3$ km). Only the first two message signs display the variable speed limits (v_{lim}). The sign at the downstream end of the controlled section always displays the freeway's regular speed limit. The variable speed limits are selected on the basis of the feedback control law described earlier. The controller uses density measurements from the detector at the bottleneck ($x = 28.3$ km) as input. The values of the control parameters, shown in Table 5.1, were selected after the controller performance for different sets of values was analyzed. No optimization method was used to tune the controller.

In the control scenario, the model of longitudinal driving behavior is extended on the basis of two assumptions: (a) a driver notices the message signs displaying the variable speed limits when the distance between the driver and the sign is 300 m or less, and (b) longitudinal driving behavior after a driver notices a message sign can be adequately reproduced by changing the value of the desired speed parameter (v_{des}) to the displayed speed limit (also assumed is that all drivers fully comply with speed limits) while keeping the remaining parameter values unchanged. A change in the desired speed parameter does not result in an instantaneous change in vehicle speed.

5.4.3 Performance indicator: TD

The performance of the proposed control strategy is evaluated by comparing the TD in the no-control scenario with that in the control scenario. The TD in a given scenario is defined as

$$TD = TTS - TTS_{ref} \quad (5.12)$$

where TTS is the total time spent in that scenario and TTS_{ref} is the total time spent in the reference scenario. The total time spent is calculated on the basis of both demand flows and exit flows (Papageorgiou et al., 2003). The demand flows are the flows measured by the detector located at $x = 0.3$ km; the exit flows are the flows measured by the detector located at $x = 29.9$ km. The reference scenario is a hypothetical one in which the freeway vertical alignment is assumed to have no influence on the acceleration behavior of drivers; hence, the sag is not a bottleneck. This situation is modeled by setting the value of the maximum gradient compensation rate parameter c to an extremely high value: $c = 999 \text{ s}^{-1}$.

5.5 Results

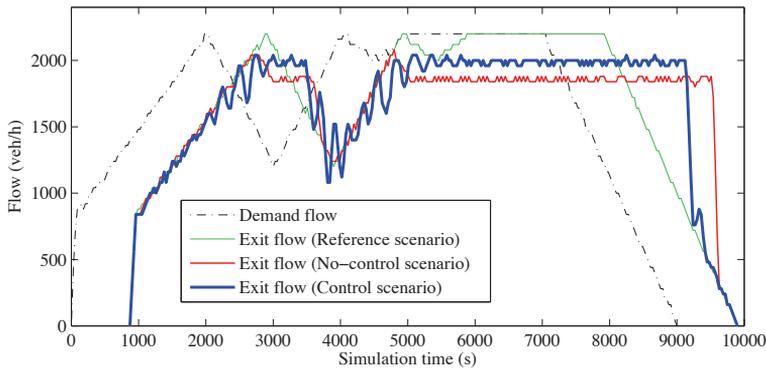
5.5.1 Reference scenario

In the reference scenario, traffic flow remains uncongested everywhere in the network during the entire simulation period. Thus, the exit flow profile over time is similar to the demand flow profile, with an offset of around 900 s (Figure 5.3(a)). The total time spent is 1,035 vehicle hours (veh-h).

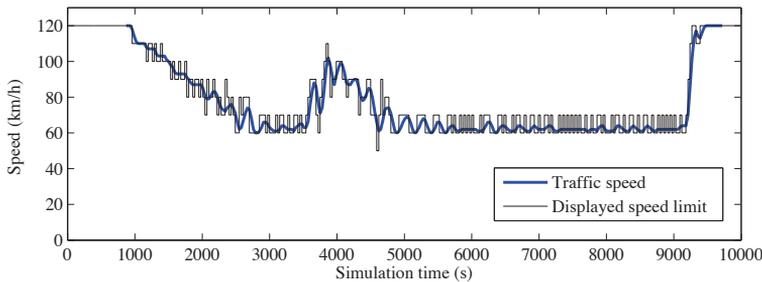
5.5.2 No-control scenario

In the no-control scenario, traffic breaks down at the sag bottleneck when the inflow goes above 2,050 vph (which can be considered the free-flow capacity of the bottleneck). When traffic breaks down, the outflow from the bottleneck decreases to around 1,855 vph (which can be considered the queue discharge capacity) and reduces the

network exit flow to 1,855 vph as well (Figure 5.3(a)). During the simulation period, traffic breaks down twice. After the first breakdown, the demand flow decreases considerably and allows the first queue to dissolve. Afterward, the demand flow again increases above the free-flow capacity of the bottleneck, causing a second breakdown (Figure 5.3(a)). In both cases, because the demand flow is higher than the exit flow, the number of vehicles within the network increases. This accumulation of vehicles results in a higher TTS than the TTS_{ref} . The TTS in the no-control scenario is 1,237 veh-h, so the TD is 202 veh-h.



(a) Demand and exit flows over time in all scenarios.



(b) Speed limit and traffic speed over time at location $x = 27.0$ km (i.e., within the controlled section) in the control scenario.

Figure 5.3: Simulation results.

5.5.3 Control scenario

In the control scenario, the outflow from the controlled section is regulated so that it does not go above the free-flow capacity of the bottleneck. Because of that limit, traffic does not break down at the bottleneck during the entire simulation period. In conditions of high demand, congestion forms in the controlled section; however, the outflow from the controlled section is higher (around 1,985 vph) than the queue discharge capacity

of the bottleneck (which is around 1,855 vph) (Figure 5.3(a)). As a result, in periods of high demand, network exit flows are around 1,985 vph (i.e., 7% higher than in the no-control scenario) (Figure 5.3(a)). Therefore, fewer vehicles accumulate in the network, and the result is a considerably lower TD. In the control scenario, the TTS is 1,177 veh-h (5% lower than in the no-control scenario), so the TD is 142 veh-h (30% lower than in the no-control scenario).

The controller is able to react adequately to fluctuations in demand. Demand flows reach high levels before $t = 2,000$ s (Figure 5.3(a)). When density at the bottleneck gets close to the target density, the controller sets a speed limit of 60 to 70 km/h in the controlled section (around $t = 2,700$ s, Figure 5.3(b)). Between $t = 2,000$ s and $t = 3,000$ s, demand significantly decreases (Figure 5.3(a)), and this decrease results in low densities at the bottleneck. When such low densities are measured, the controller increases the speed limit in the controlled section (Figure 5.3(b)). The reason for the speed limit increase is that demand is too low to cause traffic to break down at the bottleneck, so the need to restrict inflow is lower. Afterward, between $t = 3,000$ s and $t = 4,000$ s, the demand increases again (Figure 5.3(a)). The controller responds by decreasing the speed limit in the controlled section to 60 to 70 km/h again (Figure 5.3(b)) to prevent traffic from breaking down at the bottleneck. Because of the proportional structure of the controller, demand fluctuations result in speed limit oscillations (Figure 5.3(b)). However, in this case study, oscillations seem to dampen out with time, so the system does not become unstable.

5.6 Sensitivity analysis

The authors selected the values of the controller parameters (Table 5.1) to ensure high controller performance under the assumption that drivers behave in accordance with the model of longitudinal driving behavior. However, the authors also analyzed the performance of the controller by assuming that drivers do not behave exactly as described by that model. More specifically, the authors investigated the sensitivity of the controller performance to two key parameters of the model of longitudinal driving behavior that have a significant influence on the capacity of the sag bottleneck: the maximum gradient compensation rate, c , and the congestion factor on safe time headway, γ . First, the authors evaluated the performance of the controller by assuming a lower and a higher value for parameter c (i.e., 0.00005 s^{-1} and 0.00015 s^{-1} , respectively) while the other parameters remained unchanged. Second, they evaluated the performance of the controller by assuming a lower and a higher value for parameter γ (i.e., 1.12 and 1.18, respectively) while the other parameters remained unchanged.

The results indicate that the reduction in TD resulting from the implementation of the proposed control strategy significantly depends on the value of parameter c . If $c = 0.00010 \text{ s}^{-1}$ (default value), the TD in the control scenario is 30% lower than in the no-control scenario. If $c = 0.00005 \text{ s}^{-1}$, that percentage is 36%, whereas, if

$c = 0.00015 \text{ s}^{-1}$, that percentage is 23% (Table 5.2). The main reason for those differences is that a higher (or lower) value of c results in a higher (or lower) queue discharge capacity of the sag bottleneck; hence, it also results in higher (or lower) exit flows in the no-control scenario. In contrast, in the control scenario, exit flows are almost the same regardless of the value of c . Therefore, the controller reduces TD to a larger extent if the value of c is lower.

The reduction in TD resulting from the implementation of the controller does not significantly depend on the value of parameter γ . If $\gamma = 1.15$ (default value), the TD in the control scenario is 30% lower than in the no-control scenario. If $\gamma = 1.12$, that percentage is 31%, whereas, if $\gamma = 1.18$, that percentage is 29% (Table 5.2). The main reason that the percentages are similar is that a higher (or lower) value of γ results in a lower (or higher) queue discharge capacity of both the sag bottleneck and the controlled section. Therefore, a higher (or lower) value of γ results in lower (or higher) exit flows in both the no-control scenario and the control scenario.

To conclude, the sensitivity analysis shows that the results of evaluation of the controller performance depend on the specification of the model of longitudinal driving behavior. However, the sensitivity analysis also shows that the controller is able to reduce TD significantly even after the values of key model parameters are changed.

Table 5.2: Performance of controller, including sensitivity analysis

	Value, by scenario				
	Reference	Varied c		Varied γ	
		A	B	C	D
Model parameter					
$c \text{ (s}^{-1}\text{)}$	0.00010	0.00005	0.00015	0.00010	0.00010
$\gamma \text{ (unitless)}$	1.15	1.15	1.15	1.12	1.18
TD (veh-h)					
No-control scenario	202	227	177	157	244
Control scenario	142	145	137	108	173
Difference (veh-h)					
Absolute	-60	-82	-40	-49	-71
Relative	-29.7	-36.1	-22.6	-31.2	-29.0

5.7 Conclusions

The capacity of sags is considerably lower than the capacity of normal freeway sections. Consequently, sags are often bottlenecks in freeway networks. This paper presented a new control strategy to mitigate congestion at sags on the basis of the concept of mainstream traffic flow control. By limiting the traffic speed (and hence the flow) in a controlled section upstream of the bottleneck, the proposed strategy regulates the

density at the bottleneck area to keep it slightly below the critical density and hence prevents traffic from breaking down. The capacity drop attributable to congestion does not occur, so the outflow from the bottleneck can be higher. The speed limit on the controlled section is set by using a proportional feedback control law. The performance of the proposed control strategy was evaluated by means of a simple case study by using microscopic traffic simulation. The results show a considerable improvement in the efficiency of traffic flow. In periods of high demand, the flow exiting the network is around 7% higher in the control scenario than in the no-control scenario, and this higher exiting flow reduces the TD by around 30%. A sensitivity analysis shows that the controller is able to reduce TD considerably, even if different values for key parameters of the model of longitudinal driving behavior are assumed. Despite the simplicity of the case study, the findings here show for the first time that mainstream traffic flow control strategies using variable speed limits have the potential to improve traffic flow efficiency considerably in freeway networks containing sags.

Further research is necessary to make a more thorough evaluation of the performance of the proposed control strategy. Such evaluation requires extending the case study to include a multilane network and heterogeneous traffic. In addition, the model of longitudinal driving behavior should take into account the level of compliance of drivers with variable speed limits, because driver compliance may have a strong influence on the performance of the control strategy. In addition, the model should be calibrated and validated. Further research should be performed to refine the controller design and improve its performance. For example, the oscillatory behavior of the controller could be mitigated by using an alternative type of control law (e.g., proportional-integral feedback). Furthermore, other means to regulate speed in the controlled section could be tested. An alternative to displaying variable speed limits on message signs could be to regulate the speed of vehicles equipped with cooperative adaptive cruise control systems (via infrastructure-to-vehicle communication). Finally, the controller design could be extended to make it operational in more-complex networks (e.g., networks with ramps or other types of bottlenecks). This extension may require combining the control strategy presented in this paper with other types of control measures.

Chapter 6

Conclusions

6.1 Main findings

6.1.1 Causes of stop-and-go waves at sags

At sags, the most frequently observed congested traffic patterns are moving localized clusters and oscillating congested traffic, both of which are patterns characterized by the presence of stop-and-go waves. However, the primary cause of stop-and-go waves at sags has not been identified yet. There is evidence that the car-following behavior of most drivers changes at sags (drivers tend to keep longer headways than expected given their speed). Some researchers state that stop-and-go waves form mainly as a result of the cumulative effect of these local changes in car-following behavior. On the other hand, other researchers argue that disruptive lane changes may be the primary cause (as is usually the case in other types of bottlenecks).

This thesis analyzes 2284 vehicle trajectories at a sag of a three-lane Japanese freeway and shows that, in most cases (16 out of 18), the factor triggering the formation and growth of stop-and-go waves is related to car-following phenomena. Lane changes are a significantly less frequent triggering factor (2 cases out of 18). These car-following phenomena may be related to the fact that drivers keep longer headways than expected given their speed when they drive through sags (as suggested by other researchers). However, it is not possible to validate this hypothesis based on the data analyzed in this thesis, because the resolution of the trajectories is not high enough.

6.1.2 Microscopic modeling of traffic flow at sags

The ability to simulate traffic at sags in a realistic way is crucial for the development and evaluation of dynamic traffic management (DTM) measures for this type of bottlenecks. The findings reported in Chapter 2 indicate that, in order to accurately simulate traffic flow at sags on a microscopic level, it is necessary to use a car-following model

that takes into account the influence of the freeway's vertical profile. A review of the literature shows that most existing car-following models of that type are based on assumptions about the way drivers compensate for the increased grade resistance force at sags that are inconsistent with empirical observations. That hampers their validity. The only car-following model that is based on accurate assumptions is the one proposed by Oguchi and Konuma (2009), which assumes that drivers compensate *gradually* for the increase in grade resistance force at sags. However, an important drawback of this model is that it is not sufficiently generic, since the values of some parameters depend on the vertical profile of the sag.

This thesis presents a new car-following model that is based on a similar principle to that of the model presented by Oguchi and Konuma (2009) but has a more generic formulation. Simulation experiments show that a microscopic traffic model consisting of the proposed car-following model and a regular lane change model (LMRS) reproduces the main phenomena that characterize traffic dynamics at sags. In particular, the model generates a bottleneck caused by the increase in gradient and reproduces its location very accurately. The model also reproduces the oscillating nature of congested traffic at sags. A sensitivity analysis indicates potential ways to alleviate traffic congestion at sags. For instance, it is found that congestion is less severe if drivers: *a*) compensate faster for the increased grade resistance force at sags; and *b*) have shorter desired time headways in congestion.

6.1.3 Optimal acceleration behavior of vehicles equipped with in-car systems at sags

In recent years there has been a growing interest in the development of DTM measures for sags that use in-vehicle systems as actuators. At present, however, it is not known how vehicles equipped with those systems should behave at sags in order to minimize congestion. The traffic flow optimization experiments carried out in this thesis (which consider scenarios with a single-lane freeway, nearly-saturated traffic conditions and low penetration rates of equipped vehicles) indicate that the optimal acceleration behavior of equipped vehicles at sags is defined by two separate strategies, which are given the names of primary and supporting strategies.

The *primary* strategy is the one that contributes the most to reducing total delay. It involves performing a deceleration-acceleration-deceleration-acceleration (DADA) maneuver in the sag area. This maneuver induces the first group of following vehicles (up to 85 vehicles, in some cases) to accelerate fast along the sag, reaching the end of the curve (bottleneck) at moderately high speeds. As a result, for a particular time interval (2-3 min), the outflow from the sag is significantly higher (up to 5% higher) than the regular queue discharge rate. This produces substantial delay savings for all vehicles located behind the equipped vehicle. The maneuver also triggers a stop-and-go wave on the first part of the sag that temporarily limits the inflow into the vertical curve.

Limiting the inflow is beneficial because it slows down the formation of congestion at the bottleneck.

The *supporting* strategy consists in performing one or more deceleration-acceleration maneuvers upstream of the sag. This strategy is used in some cases to manage congestion upstream of the vertical curve in such a way that the inflow into the sag is regulated more effectively. This increases slightly the effectiveness of the primary strategy.

6.1.4 Mainstream traffic flow control using variable speed limits at sags

The findings reported in Chapter 4 underline the relevance of limiting the inflow into sags as an approach to mitigate traffic congestion at this type of bottlenecks. Previous research shows that DTM measures based on the concept of mainstream traffic flow control (MTFC), which essentially regulate the density at a bottleneck by limiting the inflow, can significantly alleviate traffic congestion. However, no automatic MTFC measures have been specifically developed for sags yet.

This dissertation presents a new MTFC measure that applies variable speed limits upstream of the sag in order to regulate the density at the end of the vertical curve (bottleneck). Density is kept slightly below the critical density, hence preventing traffic from breaking down while maximizing outflow. Speed limits are set on the basis of a simple feedback control law. Simulation experiments using the traffic model developed in this thesis (and assuming that all drivers fully comply with speed limits) show that the proposed MTFC measure: *a*) is able to adapt to fluctuations in demand; and *b*) increases the sag outflow by around 7% in periods of high traffic demand. This leads to a substantial decrease in total delay. Furthermore, a sensitivity analysis shows that the proposed measure reduces the total delay considerably even if different values for key parameters of the car-following model are assumed.

6.2 Conclusions

Previous research shows that, at sags, most drivers operate the vehicle controls in a way that does not allow them to compensate instantaneously for the increase in resistance force that results from the increase in gradient. The non-compensated resistance force limits the acceleration of their vehicles, which leads to local changes in car-following behavior. More specifically, drivers tend to keep longer headways than expected given their speed. The findings reported in the dissertation support the hypothesis that congestion (in the form of stop-and-go waves) occurs at sags primarily as a result of the cumulative effect of these local changes in car-following behavior rather than as a result of lane-changing maneuvers. Note, however, that it is not possible to conclude this absolutely due to the reduced size and scope of the data set analyzed in the thesis. The

findings give us a better understanding of the causes of congestion at sags. Moreover, the discovery that lane-changing maneuvers are not the primary cause of stop-and-go waves at the studied sag, which is in line with observations made by Zheng et al. (2011) at a different site, is very relevant because it implies that stop-and-go waves probably form at sags for a different reason than in most other types of freeway sections.

The ability to simulate traffic at sags in a realistic way is essential for the development of DTM measures for this type of bottleneck. The conclusion that congestion at sags is most likely caused by local changes in car-following behavior implies that accurate modeling of this dimension of driving behavior is crucial for modeling traffic at sags on a microscopic level. However, a literature review revealed that most existing car-following models do not take into account the influence of the freeway's vertical profile, and those that do, do so mostly on the basis of incorrect assumptions, which hinders their validity. This means that existing car-following models can generally not reproduce traffic dynamics at sags in an accurate manner, whereas there is a need for models that can do so. To address this issue, a new car-following model was developed in this dissertation, which generalizes the Intelligent Driver Model (IDM) on the basis of two assumptions: *a*) drivers compensate linearly over time for the increase in grade resistance force at sags; *b*) the compensation rate is generally lower than the rate at which the grade resistance force increases as the vehicle moves along the vertical curve. These assumptions are in line with empirical observations and imply that drivers do not attain their desired acceleration when they go through sags. The simulation results reported in this thesis show that a microscopic traffic model composed of the proposed car-following model and a regular lane change model (LMRS) is able to reproduce the main phenomena that characterize traffic dynamics at sags, including the oscillatory nature of congested traffic and the location of the bottleneck (end of the sag). Therefore, it is concluded that the proposed modeling approach is face-valid and its underlying theoretical assumptions are correct. Nevertheless, the model still needs to be calibrated and validated. In general, the modeling of traffic on freeways with vertical curves remains an underdeveloped field of study that needs more attention.

As mentioned above, the ability to accurately model traffic flow at sags is crucial for the development of specific DTM measures for this type of bottleneck. This dissertation identifies various potentially effective traffic management strategies by analyzing the properties of the newly developed car-following model and by performing optimization experiments using an adapted version of that model. A mathematical optimization method was used for the first time for identifying traffic management strategies for sags, and the results prove the usefulness of this method. It is concluded that the same optimization method (or an extended version of it) could be used for identifying traffic management strategies for other types of bottlenecks. The findings reported in this thesis suggest four main strategies to reduce the severity of congestion at sags: *a*) improving the ability of drivers to compensate for the increase in grade resistance force at sags; *b*) inducing drivers to keep shorter headways when they are queuing; *c*) inducing drivers to accelerate fast along sags and reach the end of the vertical curve at

high speeds; and *d*) limiting the inflow into sags. The main effect of all these strategies is an increased outflow from the sag in high demand conditions, which reduces the total delay caused by the bottleneck. DTM measures based on strategies *a* and *b* have been presented in recent years, but there are no modern applications of strategies *c* and *d*. This thesis proposes two possible ways to implement the latter two strategies in practice.

The first way consists in making specific vehicles equipped with in-car systems (for example, ACC systems that communicate with the freeway infrastructure) perform a deceleration-acceleration-deceleration-acceleration (DADA) maneuver at sags in conditions when traffic demand approaches the bottleneck capacity. Simulation results reported in this thesis indicate that, if done correctly, the DADA maneuver of each vehicle can induce the first group of following vehicles to accelerate fast along the sag (strategy *c*) and trigger a stop-and-go wave that temporarily limits the inflow into the vertical curve (strategy *d*). As a result, after each DADA maneuver, traffic stays uncongested at the end of the sag (bottleneck) for a certain time period (2-3 min). During that period, the sag outflow is higher (up to 5% higher) than the queue discharge capacity. This leads to substantial delay savings in comparison with a no-control scenario, especially if the distance between controlled vehicles is optimized so as to ensure that sag outflow stays high for a total period as long as possible (for example, in the scenarios analyzed in the thesis, the ideal situation would be to have around 80 non-equipped vehicles between each pair of equipped vehicles). Note that this type of measure does not require high penetration rates of equipped vehicles, which facilitates its implementation. However, various issues need to be addressed before this type of measure can be implemented in real freeways. The simulations presented in this dissertation consider scenarios with a single-lane freeway; one of the main questions is whether this type of measure would be effective in multi-lane freeways, where vehicles may change lane changes in front of equipped vehicles while these are performing a DADA maneuver. Actually, for real-life applications, additional measures may be necessary to facilitate the DADA maneuvers and to enhance their effectiveness in mitigating congestion.

The second way consists in regulating the inflow into sags (strategy *d*) by modifying the speed limits on a freeway section located immediately upstream of the vertical curve on the basis of a feedback control law. Inflow should be regulated in such a way that the traffic density at the end of the sag (bottleneck) stays slightly below the critical density, hence traffic flow remains uncongested. This type of measure is based on a relatively simple principle. Simulation results reported in this dissertation show that it can substantially increase the sag outflow during periods of high demand (the increase can be of up to 7%), thereby producing considerable delay savings. It is concluded that this type of measure can substantially reduce congestion in freeway networks with sags. However, some issues need to be addressed before this type of measure can be implemented in practice, such as improving the stability of the controller and finding ways to ensure that drivers comply with the speed limits.

6.3 Implications for practice

The findings and conclusions of this thesis have important implications for practice, since they open the possibility of improving microscopic simulation tools and developing more effective measures to manage traffic at sags.

The findings on the causes of congestion at sags provide a theoretical foundation that can facilitate the understanding and modeling of traffic dynamics at this type of free-way sections. The thesis presented a new microscopic traffic model that was developed on the basis of these findings, and showed that the model correctly reproduces traffic dynamics at sags. The proposed modeling approach allows for an improved representation of traffic flow in microscopic simulation tools. This is expected to improve the predictive power of these tools and increase their accuracy in evaluating the effects of DTM measures or infrastructure adjustments. Policy makers, road authorities and traffic engineering companies often require this type of evaluations.

The findings on strategies to reduce congestion at sags provide a foundation for the development of new DTM measures. Particularly, the thesis identified two traffic management strategies (inducing drivers to accelerate fast along sags, and limiting the inflow into the bottleneck) that are currently not being applied in practice. Moreover, the thesis proposed two specific DTM concepts based on these strategies: *a*) making vehicles equipped with in-car systems perform a specific acceleration-related maneuver at sags in order to induce the following vehicles to accelerate fast and reach the bottleneck at high speeds; and *b*) using variable speed limits to regulate the inflow into sags and prevent congestion at the bottleneck. The findings reported in the dissertation indicate that both DTM concepts have the potential to substantially reduce congestion at sags. Furthermore, on a more general level, they confirm the high potential of DTM measures based on the use of in-car systems, particularly advanced adaptive cruise control (ACC) systems, and DTM measures based on the concept of mainstream traffic flow control. Therefore, it is expected that industry will explore the potential market opportunities related to these new technologies and on this basis will develop them further.

6.4 Recommendations for future research

Determining the reasons why the car-following behavior of drivers changes at sags is crucial for developing a better understanding of the causes of congestion at these type of bottlenecks. Previous research indicates that the main reason is related to the characteristics of vehicle control behavior¹ at sags. However, these are not well known

¹Here *vehicle control behavior* means the way drivers operate the vehicle controls (e.g., pedals and gear box) in response to the roadway characteristics, traffic conditions and environmental conditions.

yet. Hence, we recommend to investigate more thoroughly how drivers operate the vehicle controls at sags in various circumstances and how and to what extent this influences longitudinal driving behavior. A possible approach is to analyze real vehicle control operation data together with the corresponding vehicle trajectory and vertical alignment data. Nevertheless, to the best of the author's knowledge, this type of data is not available. An alternative could be to use driving simulator data (Yoshizawa et al., 2012), although the reliability of this type of data is questionable. Otherwise, real driving data could be collected by carrying out experiments with test vehicles. If possible, these experiments should be performed with different vehicles, with different drivers, in different traffic conditions, and on sags with different vertical profiles. Another relevant research topic is whether, how and to what extent the freeway's vertical profile influences lateral driving behavior. Currently, very little is known about this topic. A possible research approach is to analyze high-resolution trajectory data.

This dissertation presented a new microscopic traffic model and demonstrated its face-validity. However, the proposed traffic model should be calibrated and validated, which requires a quantitative comparison of the model output with empirical trajectory data. Vehicle trajectories from the Yamato sag (Tomei Expressway, Japan) are available, but their resolution is not high enough for that purpose. Therefore, high-resolution data (including lane changes) from that site and/or other sites will be necessary to calibrate and validate the traffic model. Preferably, the additional sites should be from other countries and include sags with different vertical profiles and different number of lanes. Also, the trajectory data should correspond to periods with different traffic conditions, including free flow, onset of congestion and fully congested conditions. Furthermore, it is recommended to perform a thorough analysis of the mathematical properties of the traffic model and, particularly, of its car-following model. Such analysis is essential to gain more insight on how the model works and to identify potential limitations. Also, it could indicate ways to improve the model performance further. Finally, it would be very interesting to compare the performance of this model with that of existing models. Another relevant research topic is the development of macroscopic approaches to model traffic at sags. Here, one of the first necessary steps is estimating the fundamental relation between traffic flow, density and speed. The shape of this relation at sags is not well known, although some preliminary results have been presented in the literature (Goñi-Ros et al., 2013a).

The thesis identified the two strategies that define the optimal acceleration behavior of vehicles equipped with in-car systems at sags in a series of scenarios. These findings are very relevant for the development of new DTM measures for sags. However, more research is needed to determine whether the identified strategies would also be optimal in a wider range of scenarios or other strategies would be more appropriate in other conditions. Relevant scenarios that need to be studied include scenarios with multiple lanes, higher penetration rates and lower traffic demand. To this end, additional optimization experiments are necessary. Note that performing optimization experiments for scenarios with multiple lanes will require modifying the formulation of the optimal

control problem in order to include the possibility that vehicles change lanes and to specify the lateral driving behavior of drivers. Moreover, further research is needed to analyze the working principles of the identified strategies more in detail and to determine how they can be translated into effective DTM measures. In this respect, it is important to consider the drivers' acceptance of and interaction with the in-car systems, since this will have a big influence on the effectiveness of the DTM measures. Currently, ACC applications using infrastructure-to-vehicle or vehicle-to-vehicle communication based on the strategies identified in this thesis are under development. Preliminary work shows promising results (Papacharalampous et al., 2015b).

This thesis presented a new DTM measure that uses variable speed limits to regulate the traffic density at the end of a sag (bottleneck) and prevent congestion at that location. It is shown that the proposed measure can considerably reduce total delay. However, there is room to improve and extend the controller design and, hence, to increase the performance of this type of measure. For example, the oscillatory behavior of the controller could be mitigated by using an alternative type of control law (e.g., proportional-integral feedback) or by using traffic state predictions. Also, the controller design can be extended to make it operational in more complex freeway networks, such as networks with ramps or other types of bottlenecks. This extension may require combining the proposed measure with other types of control measures (e.g., ramp metering). Furthermore, it would be interesting to test other means to regulate speed in the controlled section. An alternative or an addition to displaying variable speed limits on message signs could be to regulate the speed of vehicles equipped with advanced adaptive cruise control systems via infrastructure-to-vehicle communication (Wang et al., 2015). Also, it would be interesting to study ways to ensure that the variable speed limits do not cause the formation of long-lasting stop-and-go waves. Finally, it is recommended to perform a more extensive evaluation of the performance of the proposed DTM measure. Such evaluation requires performing simulations for scenarios with multiple lanes and heterogeneous traffic. In addition, the car-following model should be adjusted to take into account the level of compliance of drivers with speed limits.

References

- Ahn, S., & Cassidy, M. J. (2007). Freeway traffic oscillations and vehicle lane-change maneuvers. In R. E. Allsop, M. G. H. Bell, B. G. Heydecker (Eds.), *Proceedings of the 17th International Symposium of Transportation and Traffic Theory* (p. 691–710). Amsterdam: Elsevier.
- Bando, M., Hasebe, K., Nakayama, A., Shibata, A., & Sugiyama, Y. (1995). Dynamical model of traffic congestion and numerical simulation. *Physical Review E: Statistical, Nonlinear, Biological, and Soft Matter Physics*, 51(2), 1035–1042.
- Banks, J. H. (2000). Are Minimization of Delay and Minimization of Freeway Congestion Compatible Ramp Metering Objectives? *Transportation Research Record: Journal of the Transportation Research Board*, 1727, 112–119.
- Bexelius, S. (1968). An extended model for car-following *Transportation Research*, 2(1), 13–21.
- Brilon, W., & Bressler, A. (2004). Traffic Flow on Freeway Upgrades. *Transportation Research Record: Journal of the Transportation Research Board*, 1883, 112–121.
- Carlson, R. C., Papamichail, I., & Papageorgiou, M. (2011). Local Feedback-Based Mainstream Traffic Flow Control on Motorways Using Variable Speed Limits. *IEEE Transactions on Intelligent Transportation Systems*, 12(4), 1261–1276.
- Carlson, R. C., Papamichail, I., & Papageorgiou, M. (2013). Comparison of Local Feedback Controllers for the Mainstream Traffic Flow on Freeways Using Variable Speed Limits. *Journal of Intelligent Transportation Systems: Technology, Planning, and Operations*, 17(4), 268–281.
- Carlson, R. C., Papamichail, I., Papageorgiou, M., & Messmer, A. (2010). Optimal mainstream traffic flow control of large-scale motorway networks. *Transportation Research Part C: Emerging Technologies*, 18(2), 193–212.
- Cassidy, M. J., & Bertini, R. L. (1999). Some traffic features at freeway bottlenecks. *Transportation Research Part B: Methodological*, 33(1), 25–42.
- Chandler, R. E., Herman, R., & Montroll, E. W. (2016). Traffic Dynamics: Studies in Car Following. *Operations Research*, 6(2), 165–184.
- Daganzo, C. F. (1994). The cell transmission model: A dynamic representation of highway traffic consistent with the hydrodynamic theory. *Transportation Research Part B: Methodological*, 28(4), 269–287.
- Daganzo, C. F. (1997). *Fundamentals of Transportation and Traffic Operations* (1st Edition). Oxford: Pergamon.

- Daganzo, C. F. (2002). A behavioral theory of multi-lane traffic flow. Part I: Long homogeneous freeway sections. *Transportation Research Part B: Methodological*, 36(2), 131–158.
- Findley, D. J., Schroeder, B. J., Cunningham, C. M., & Brown, T. H. (2015). *Highway Engineering: Planning, Design, and Operations* (1st Edition). Oxford: Butterworth-Heinemann.
- Furuichi, T., Yamamoto, S., Kotani, M., & Iwasaki, M. (2003). Characteristics of Spatial Speed Change at Motorway Sag Sections and Capacity Bottlenecks. Presented at *82nd Annual Meeting of the Transportation Research Board*, Washington, D. C.
- Gipps, P. G. (1981). A behavioural car-following model for computer simulation. *Transportation Research Part B: Methodological*, 15(2), 105–111.
- Gipps, P. G. (1986). A model for the structure of lane-changing decisions. *Transportation Research Part B: Methodological*, 20(5), 403–414.
- Goñi-Ros, B., Knoop, V. L., van Arem, B., & Hoogendoorn, S. P. (2012). Reducing congestion at uphill freeway sections by means of a Gradient Compensation System. Presented at *2012 IEEE Intelligent Vehicles Symposium*, Alcalá de Henares, Spain.
- Goñi-Ros, B., Knoop, V. L., van Arem, B., & Hoogendoorn, S. P. (2013a). Car-following Behavior at Sags and its Impacts on Traffic Flow. Presented at *92nd Annual Meeting of the Transportation Research Board*, Washington, D. C.
- Goñi-Ros, B., Knoop, V. L., van Arem, B., & Hoogendoorn, S. P. (2013b). Modeling driving behavior and traffic flow at sags. Presented at *20th ITS World Congress*, Tokyo, Japan.
- Goñi-Ros, B., Knoop, V. L., van Arem, B., & Hoogendoorn, S. P. (2014a). Empirical analysis of the causes of stop-and-go waves at sags. *IET Intelligent Transport Systems*, 8(5), 499–506.
- Goñi-Ros, B., Knoop, V. L., van Arem, B., & Hoogendoorn, S. P. (2014b). Mainstream Traffic Flow Control at Sags. *Transportation Research Record: Journal of the Transportation Research Board*, 2470, 57–64.
- Goñi-Ros, B., Knoop, V. L., Shiomi, Y., Takahashi, T., van Arem, B., & Hoogendoorn, S. P. (2016). Modeling Traffic at Sags. *International Journal of Intelligent Transportation Systems Research*, 14(1), 64–74.
- Greensberg, H., & Daou, A. (1960). The Control of Traffic Flow to Increase the Flow. *Operations Research*, 8(4), 524–532.
- Greenshields, B. D., Bibbins, J. R., Channing, W. S., & Miller, H. H. (1935). A study of traffic capacity. In R. W. Crum (Ed.), *Proceedings of the 14th Annual Meeting of the Highway Research Board* (p. 448–477). Washington, D.C.: Highway Research Board.
- Hall, F. L., & Agyemang-Duah, K. (1991). Freeway Capacity Drop and the Definition of Capacity. *Transportation Research Record: Journal of the Transportation Research Board*, 1320, 91–98.
- Hatakenaka, H., Hirasawa, T., Yamada, K., Yamada, H., Katayama, Y., & Maeda, M.

- (2006). Development of AHS for Traffic Congestion in Sag Sections. Presented at *13th ITS World Congress*, London, U. K.
- Hegyi, A., Bellemans, T., & De Schutter, B. (2009). Freeway traffic management and control. In R. A. Meyers (Ed.), *Encyclopedia of Complexity and Systems Science* (p. 3943–3964). New York: Springer.
- Hegyi, A., De Schutter, B., & Hellendoorn, J. (2005). Optimal Coordination of Variable Speed Limits to Suppress Shock Waves. *IEEE Transactions on Intelligent Transportation Systems*, 6(1), 102–112.
- Hegyi, A., & Hoogendoorn, S. P. (2010). Dynamic speed limit control to resolve shock waves on freeways - Field test results of the SPECIALIST algorithm. Presented at *13th International IEEE Conference on Intelligent Transportation Systems*, Madeira Island, Portugal.
- Hegyi, A., Hoogendoorn, S. P., Schreuder, M., Stoelhorst, H., & Viti, F. (2008). SPECIALIST: A dynamic speed limit control algorithm based on shock wave theory. Presented at *11th International IEEE Conference on Intelligent Transportation Systems*, Beijing, China.
- Helbing, D. (2001). Traffic and related self-driven many-particle systems. *Reviews of Modern Physics*, 73(4), 1067–1141.
- Helbing, D., Treiber, M., Kesting, A., & Schönhof, M. (2009). Theoretical vs. empirical classification and prediction of congested traffic states. *The European Physical Journal B: Condensed Matter and Complex Systems*, 69, 583–598.
- Hoogendoorn, R. G., Hoogendoorn, S. P., Brookhuis, K. A., & Daamen, W. (2011). Longitudinal Driving Behavior under Adverse Conditions: A Close Look at Psycho-spacing Models. *Procedia - Social and Behavioral Sciences*, 20, 536–546.
- Jayakrishnan, R., Mahmassani, H. S., & Hu, T. (1994). An evaluation tool for advanced traffic information and management systems in urban networks. *Transportation Research Part C: Emerging Technologies*, 2(3), 129–147.
- Kesting, A., Treiber, M., & Helbing, D. (2007). General Lane-Changing Model MOBIL for Car-Following Models. *Transportation Research Record: Journal of the Transportation Research Board*, 1999, 86–94.
- Kesting, A., Treiber, M., Lämmer, S., Schönhof, M., & Helbing, D. (2006). Decentralized approaches to adaptive traffic control and an extended level of service concept. Presented at *17th International Conference on the Application of Computer Science and Mathematics in Architecture and Civil Engineering*, Weimar, Germany.
- Keyvan-Ekbatani, M., Knoop, V. L., & Daamen, W. (2015). Categorization of the lane change decision process on freeways. *Transportation Research Part C: Emerging Technologies* (available on-line).
- Knoop, V. L., Duret, A., Buisson, C., & van Arem, B. (2010). Lane distribution of traffic near merging zones influence of variable speed limits. Presented at *13th International IEEE Conference on Intelligent Transportation Systems*, Madeira Island, Portugal.

- Komada, K., Masukura, S., & Nagatani, T. (2009). Effect of gravitational force upon traffic flow with gradients. *Physica A: Statistical Mechanics and its Applications*, 388(14), 2880–2894.
- Koshi, M. (2003). An Interpretation of a Traffic Engineer on Vehicular Traffic Flow. In M. Fukui, Y. Sugiyama, M. Schreckenberg, & D. E. Wolf (Eds.), *Traffic and Granular Flow'01* (p. 199–210). Berlin: Springer.
- Koshi, M., Iwasaki, M., & Okhura, I. (1981). Some findings and an overview on vehicular flow characteristics. In *Proceedings of the 8th International Symposium of Transportation and Traffic Theory* (p. 403–426). Toronto: University of Toronto Press.
- Koshi, M., Kuwahara, M., & Akahane, H. (1992). Capacity of Sags and Tunnels on Japanese Motorways. *ITE Journal*, 62(5), 17–22.
- Kotsialos, A., Papageorgiou, M., Diakaki, C., Pavlis, Y., & Middelham, F. (2002). Traffic Flow Modeling of Large-Scale Motorway Networks Using the Macroscopic Modeling Tool METANET. *IEEE Transactions on Intelligent Transportation Systems*, 3(4), 282–292.
- Kraan, M., van der Zijpp, N., Tutert, B., Vonk, T., & van Megen, D. (1999). Evaluating Networkwide Effects of Variable Message Signs in the Netherlands. *Transportation Research Record: Journal of the Transportation Research Board*, 1689, 60–67.
- Laval, J. A. (2009). Effects of geometric design on freeway capacity: Impacts of truck lane restrictions. *Transportation Research Part B: Methodological*, 43(6), 720–728.
- Laval, J. A., Cassidy, M. J., & Daganzo, C. F. (2007). Impacts of Lane Changes at Merge Bottlenecks: A Theory and Strategies to Maximize Capacity. In A. Schadschneider, T. Pöschel, R. Kühne, M. Schreckenberg, D. E. Wolf (Eds.), *Traffic and Granular Flow'05*. (p. 577–586). Berlin: Springer.
- Laval, J. A., & Leclercq, L. (2010). A mechanism to describe the formation and propagation of stop-and-go waves in congested freeway traffic. *Philosophical Transactions of the Royal Society A: Mathematical Physical and Engineering Sciences*, 368, 4519–4541.
- Laval, J. A., & Leclercq, L. (2013). The Hamilton-Jacobi partial differential equation and the three representations of traffic flow. *Transportation Research Part B: Methodological*, 52, 17–30.
- Lebacque, J. P. (1996). The Godunov scheme and what it means for first order traffic flow models. In J. B. Lesort (Ed.), *Proceedings of the 13th International Symposium on Transportation and Traffic Theory* (p. 647–678). New York: Elsevier.
- Leclercq, L., Laval, J. A., & Chiabaut, N. (2011). Capacity Drops at Merges: an endogenous model. *Procedia - Social and Behavioral Sciences*, 17, 12–26.
- Levinson, D., & Zhang, L. (2006). Ramp meters on trial: Evidence from the Twin Cities metering holiday. *Transportation Research Part A: Policy and Practice*, 40(10), 810–828.
- Lighthill, M. J., & Whitham, G. B. (1955). On Kinematic Waves. II. A Theory of

- Traffic Flow on Long Crowded Roads. *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 229, 317–345.
- Ludmann, J. (1998). *Beeinflussung des Verkehrsablaufs auf Strassen: Analyse mit dem fahrzeugorientierten Verkehrsflusssimulationsprogramm PELOPS*. Schriftenreihe, Institut für Kraftfahrwesen, Aachen.
- Masher, D. P., Ross, D. W., Wong, P. J., Tuan, P. L., Zeidler, H. M., & Peracek, S. (1975). *Guidelines for design and operation of ramp control systems*. Stanford Research Institute, Menlo Park, CA.
- Moore, D. S., McCabe, G. P., & Craig, B. (2011). *Introduction to the Practice of Statistics* (7th International Edition). London: W. H. Freeman & Co.
- Nagel, K., & Schreckenberg, M. (1992). A cellular automaton model for freeway traffic. *Journal de Physique I*, 2(12), 2221–2229.
- Newell, G. F. (1993). A simplified theory of kinematic waves in highway traffic. *Transportation Research Part B: Methodological*, 27(4), 281–313.
- Oguchi, T., & Konuma, R. (2009). Comparative study of car-following models for describing breakdown phenomena at sags. Presented at *16th ITS World Congress*, Stockholm, Sweden.
- Okamura, H., Watanabe, S., & Watanabe, T. (2000). An Empirical Study on the Capacity of Bottlenecks on the Basic Suburban Expressway Sections in Japan. Presented at *4th International Symposium on Highway Capacity*, Maui, Hawaii.
- Ossen, S. (2008). Longitudinal Driving Behavior: Theory and Empirics. *TRAIL Thesis Series, T2008/8*. Delft: TRAIL Research School.
- Ozaki, H. (2003). Modeling of Vehicular Behavior from Road Traffic Engineering Perspectives. In M. Fukui, Y. Sugiyama, M. Schreckenberg, & D. E. Wolf (Eds.), *Traffic and Granular Flow '01* (p. 281–292). Berlin: Springer.
- Papacharalampous, A. E., Wang, M., Knoop, V. L., Goñi-Ros, B., Schakel, W. J., van Arem, B., & Hoogendoorn, S. P. (2015a). *Reducing congestion at sags by means of Cooperative Advanced Driver Assistance Systems – Stage 6.5: Effectiveness of C2X controller using multilane traffic flow simulation* (Internal Report). Delft University of Technology.
- Papacharalampous, A. E., Wang, M., Knoop, V. L., Goñi-Ros, B., Takahashi, T., Sakata, I., van Arem, B., & Hoogendoorn, S. P. (2015b). Mitigating Congestion at Sags with Adaptive Cruise Control Systems. Presented at *18th International IEEE Conference on Intelligent Transportation Systems*, Las Palmas de Gran Canaria, Spain.
- Papageorgiou, M., Diakaki, C., Dinopoulou, V., Kotsialos, A., & Wang, Y. (2003). Review of road traffic control strategies. *Proceedings of the IEEE*, 91(12), 2043–2067.
- Papageorgiou, M., Hadj-Salem, H., & Blosseville, J. M. (1991). ALINEA: A local feedback control law for on-ramp metering. *Transportation Research Record: Journal of the Transportation Research Board*, 1320, 58–64.
- Papageorgiou, M., Hadj-Salem, H., & Middleham, F. (1997). ALINEA local ramp metering: summary of field results. *Transportation Research Record: Journal*

- of the Transportation Research Board, 1603, 90–98.*
- Papageorgiou, M., & Messmer, A. (1991). Dynamic network traffic assignment and route guidance via feedback regulation. *Transportation Research Record: Journal of the Transportation Research Board, 1306, 49–58.*
- Papamichail, I., Papageorgiou, M., Vong, V., & Gaffney, J. (2010). Heuristic Ramp-Metering Coordination Strategy Implemented at Monash Freeway, Australia. *Transportation Research Record: Journal of the Transportation Research Board, 2178, 10–20.*
- Patire, A. D. (2010). *Observations of lane changing patterns on an uphill expressway* (Ph.D. Thesis). University of California, Berkeley.
- Patire, A. D., & Cassidy, M. J. (2011). Lane changing patterns of bane and benefit: Observations of an uphill expressway. *Transportation Research Part B: Methodological, 45(4), 656–666.*
- Payne, H. J. (1971). Models of freeway traffic and control. *Mathematical Models of Public Systems, 28, 51–61.*
- Richards, P. I. (1956). Shock Waves on the Highway. *Operations Research, 4(1), 42–51.*
- Rijkswaterstaat (Dutch Ministry of Infrastructure and the Environment) (2003). *Handbook of Sustainable Traffic Management*. Rijkswaterstaat.
- Sato, H., Xing, J., Tanaka, S., & Watauchi, T. (2009). An Automatic Traffic Congestion Mitigation System by Providing Real Time Information on Head of Queue. Presented at *16th ITS World Congress*, Stockholm, Sweden.
- Schakel, W. J., Knoop, V. L., & van Arem, B. (2012). Integrated Lane Change Model with Relaxation and Synchronization. *Transportation Research Record: Journal of the Transportation Research Board, 2316, 47–57.*
- Schakel, W. J., & van Arem, B. (2014). Improving Traffic Flow Efficiency by In-Car Advice on Lane, Speed, and Headway. *IEEE Transactions on Intelligent Transportation Systems, 15(4), 1597–1606.*
- Schönhof, M., & Helbing, D. (2007). Empirical Features of Congested Traffic States and Their Implications for Traffic Modeling. *Transportation Science, 41(2), 135–166.*
- Schreiter, T., van Lint, J. W. C., Yuan, Y., & Hoogendoorn, S. P. (2010). Propagation Wave Speed Estimation of Freeway Traffic with Image Processing Tools. Presented at *89th Annual Meeting of the Transportation Research Board*, Washington, D. C.
- Smilowitz, K., Daganzo, C. F., Cassidy, M. J., & Bertini, R. L. (1999). Some observations of highway traffic in long queues. *Transportation Research Record: Journal of the Transportation Research Board, 1678, 225–233.*
- Smulders, S. (1990). Control of freeway traffic flow by variable speed signs. *Transportation Research Part B: Methodological, 24(2), 111–132.*
- Srivastava, A., & Geroliminis, N. (2013). Empirical observations of capacity drop in freeway merges with ramp control and integration in a first-order model. *Transportation Research Part C: Emerging Technologies, 30, 161–177.*

- Sugiyama, Y., Nakayama, A., Fukui, M., Hasebe, K., Kikuchi, M., Nishinari, K., Tadaki, S., & Yukawa, S. (2005). Observation, theory and experiment for free-way traffic as physics of many-body system. In M. Schreckenberg, P. Bovy, S. P. Hoogendoorn, & D. E. Wolf (Eds.), *Traffic and Granular Flow'03* (p. 45–59). Berlin: Springer.
- Tilch, B., & Helbing, D. (2000). Evaluation of Single Vehicle Data in Dependence of the Vehicle-Type, Lane, and Site. In D. Helbing, H. J. Herrmann, M. Schreckenberg, & D. E. Wolf (Eds.), *Traffic and Granular Flow 99: Social, Traffic, and Granular Dynamics* (p. 333–338). Berlin: Springer.
- Toledo, T., Koutsopoulos, H. N., & Ben-Akiva, M. (2007). Integrated driving behavior modeling. *Transportation Research Part C: Emerging Technologies*, 15(2), 96–112.
- Transportation Research Board (2010). *Highway Capacity Manual*. Transportation Research Board.
- Treiber, M., Hennecke, A., & Helbing, D. (2000). Congested traffic states in empirical observations and microscopic simulations. *Physical Review E: Statistical, Nonlinear, and Soft Matter Physics*, 62(2), 1805–1824.
- Treiber, M., Kesting, A., & Helbing, D. (2006). Delays, inaccuracies and anticipation in microscopic traffic models. *Physica A: Statistical Mechanics and its Applications*, 360(1), 71–88.
- Van den Hoogen, E., & Smulders, S. (1994). Control by variable speed signs: results of the Dutch experiment. Presented at *7th International Conference on Road Traffic Monitoring and Control*, London, U. K.
- Van Wee, M., Annema, J. A., & Banister, D. (2013). *The Transport System and Transport Policy: An Introduction*. Cheltenham: Edward Elgar Publishing.
- Wang, M., Daamen, W., Hoogendoorn, S. P., & van Arem, B. (2015). Connected Variable Speed Limits Control and Vehicle Acceleration Control to Resolve Moving Jams. Presented at *94th Annual Meeting of the Transportation Research Board*, Washington, D. C.
- Wattleworth, J. A. (1967). Peak-period analysis and control of a freeway system. *Highway Research Record*, 157, 1-21.
- Weigl, S., Bogenberger, K., & Bertini, R. L. (2013). Traffic Management Effects of Variable Speed Limit System on a German Autobahn: Empirical Assessment Before and After System Implementation. *Transportation Research Record: Journal of the Transportation Research Board*, 2380, 48–60.
- Wolshon, B., & Lambert, L. (2006). Reversible Lane Systems: Synthesis of Practice. *Journal of Transportation Engineering*, 132(12), 933–944.
- Wu, N. (2006). Equilibrium of Lane Flow Distribution on Motorways. *Transportation Research Record: Journal of the Transportation Research Board*, 1965, 48–59.
- Xing, J., Muramatsu, E., & Harayama, T. (2012). Balance Lane Use with VMS to Mitigate Motorway Traffic Congestion. Presented at *19th ITS World Congress*, Vienna, Austria.
- Xing, J., Muramatsu, E., & Harayama, T. (2014). Balance Lane Use with VMS

- to Mitigate Motorway Traffic Congestion. *International Journal of Intelligent Transportation Systems Research*, 12(1), 26–35.
- Xing, J., Sagae, K., & Muramatsu, E. (2010). Balance Lane Use of Traffic to Mitigate Motorway Traffic Congestion with Roadside Variable Message Signs. Presented at *17th ITS World Congress*, Busan, South Korea.
- Xing, J., Takahashi, H., & Takeuchi, T. (2007). Increasing Bottleneck Capacity through Provision of Bottleneck Location Information. Presented at *11th World Conference on Transport Research*, Berkeley, California.
- Yeo, H., & Skabardonis, A. (2009). Understanding stop-and-go traffic in view of asymmetric traffic theory. In W. Lam, S.C. Wong, & H.K. Lo (Eds.), *Proceedings of the 18th International Symposium on Transportation and Traffic Theory* (p. 99–116). Hong Kong: Springer.
- Yokota, T., Kuwahara, M., & Ozaki, H. (1998). A Study of AHS effects on Traffic Flow at Bottlenecks. Presented at *5th ITS World Congress*, Seoul, South Korea.
- Yoshizawa, R., Shiomi, Y., Uno, N., Iida, K., & Yamaguchi, M. (2012). Analysis of Car-following Behavior on Sag and Curve Sections at Intercity Expressways with Driving Simulator. *International Journal of Intelligent Transportation Systems Research*, 10(2), 56–65.
- Yuan, K., Knoop, V. L., & Hoogendoorn, S. P. (2016). A microscopic investigation into the capacity drop: impacts of bounded acceleration and reaction time. Presented at *95th Annual Meeting of the Transportation Research Board*, Washington, D. C.
- Zheng, Z., Ahn, S., Chen, D., & Laval, J. A. (2011). Freeway traffic oscillations: Microscopic analysis of formations and propagations using Wavelet Transform. *Transportation Research Part B: Methodological*, 45(9), 1378–1388.

Summary

The main purpose of freeway networks is to provide short and reliable travel times between key locations of a region while ensuring safe operations. Frequently, however, freeway networks cannot perform this function properly due to the occurrence of traffic *congestion*. Congestion is a traffic condition characterized by high densities and low speeds that arises when too many vehicles try to transverse a freeway section with limited capacity. The occurrence of traffic congestion inherently causes delays. Furthermore, it reduces travel time reliability and traffic safety, and increases fuel consumption and air pollution.

Sag vertical curves (or *sags*) are roadway sections along which the gradient increases significantly in the direction of traffic. Previous research shows that the capacity of freeway sags is considerably lower than the capacity of freeway sections with other vertical profiles. Hence, traffic generally becomes congested at freeway sags in high demand conditions. Generally, congestion occurs first on the median lane and then spreads to the other lanes. The congestion patterns observed at freeway sags are characterized by the formation of stop-and-go waves, i.e., spatially-confined regions of low traffic speed that propagate upstream at a constant velocity. The bottleneck is generally the end of the vertical curve (see Figure I.A).

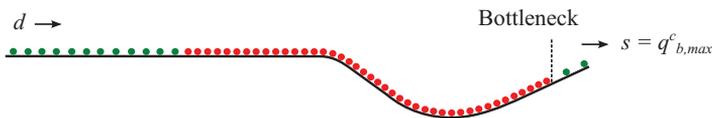


Figure I.A: Typical traffic situation on a freeway stretch with a sag in high demand conditions if no control measure is applied (d = demand flow, s = exit flow, $q_{b,max}^c$ = queue discharge capacity of the sag bottleneck; green and red circles are vehicles that are moving at high and low speed, respectively).

In some countries, sags are one of the most common types of freeway bottleneck. For example, on Japanese intercity freeways, around 60% of traffic jams occur at sags. Because of this reason, various dynamic traffic management (DTM) measures aimed at reducing congestion at sags have been developed in the past years. However, despite the progress made, several challenges still need to be addressed in order to improve our ability to mitigate congestion at this type of bottlenecks. These challenges include: *a*) gaining a better understanding of the causes of congestion at sags; *b*) developing new

mathematical models to represent traffic at sags in a more realistic way; and *c*) identifying the most effective strategies to manage traffic at sags. With regard to the latter point, current developments in dynamic traffic management (such as mainstream traffic flow control and in-car system applications) need to be considered.

The main objective of this dissertation is to develop new concepts for traffic management at sags based on a thorough understanding of the causes of congestion at this type of bottlenecks. To this end, the dissertation answers the following research questions: *a*) what is the primary cause of stop-and-go waves at sags?; *b*) how should traffic be modeled on a microscopic level in order to accurately reproduce its characteristics on freeways with sags?; *c*) how should vehicles equipped with in-car systems behave at sags in order to minimize the severity of congestion?; and *d*) can mainstream traffic flow control measures based on the use of variable speed limits significantly reduce congestion at sags?

Causes of stop-and-go waves at sags

It is well known that, in high demand conditions, traffic becomes congested at freeway sags in the form of stop-and-go waves. However, there is no consensus in the literature on what the primary cause of these waves is. There is evidence that the car-following behavior of most drivers changes at sags: drivers tend to keep longer headways than expected given their speed. The main reason for this appears to be inadequate acceleration behavior. Some researchers argue that stop-and-go waves form mainly as a result of the cumulative effect of these local changes in car-following behavior. On the other hand, other researchers suggest that disruptive lane changes may be the main cause of stop-and-go waves at sags.

This thesis contributes to the understanding of the causes of congestion at sags by analyzing empirical trajectory data from a freeway stretch with a sag and showing that, in most cases (89%), the factor triggering the formation and growth of stop-and-go waves at the study site is related to car-following phenomena, whereas lane changes are a much less frequent triggering factor (11% of cases). In view of these results, it is concluded that the main cause of stop-and-go waves at sags are not disruptive lane changes but car-following phenomena (which are probably induced by the increase in freeway gradient, as suggested by other researchers). This is not to say, however, that lane changes play no part in congestion dynamics.

Microscopic modeling of traffic flow at sags

Traffic models are essential tools for understanding traffic flow dynamics and also for developing and assessing DTM measures. The findings reported in the previous section indicate that using a car-following model that accurately captures the influence of the freeway vertical profile is crucial for modeling traffic at sags on a microscopic level.

On the other hand, there is no evidence that lane-changing behavior is influenced by the vertical profile of the freeway. Thus, regular lane change models are generally deemed suitable for modeling traffic at sags on a microscopic level.

This thesis contributes to the theoretical understanding and mathematical description of microscopic driving behavior at sags by proposing a new model that describes car-following behavior at this type of freeway sections more accurately than existing models. The proposed car-following model generalizes the Intelligent Driver Model (IDM) by adding a term that describes the influence of the freeway vertical profile on vehicle acceleration based on two assumptions (which are supported by empirical evidence): *a*) drivers compensate linearly over time for the increase in resistance force caused by an increase in slope; and *b*) the compensation rate (which is defined by a parameter) is usually lower than the rate at which the resistance force increases as a vehicle moves through a sag.

The thesis shows that a microscopic traffic model composed of the new car-following model and a regular lane change model (LMRS) is able to reproduce the main phenomena that characterize traffic flow at sags. In particular, the model generates a bottleneck caused by the vertical curve and accurately reproduces its location. In view of these results, it is concluded that the traffic model (and particularly its car-following model) is face-valid and its underlying assumptions are plausible. Furthermore, a sensitivity analysis reveals potential ways to mitigate congestion at sags. Specifically, it is found that congestion is less severe if drivers: *a*) compensate faster for the increase in grade resistance force at sags; and *b*) have shorter desired headways when they are queuing.

Optimal acceleration behavior of vehicles equipped with in-car systems at sags

Nowadays, there is growing interest in the development of DTM measures for sags based on the use of in-car systems (e.g., advisory systems or advanced adaptive cruise control systems). However, most proposed measures are in early phases of development. The thesis contributes to the development of this type of measures by determining how equipped vehicles should move at sags in order to minimize congestion. An optimal control method is used for this purpose. The optimal control problem consists in finding the acceleration profiles of a number of vehicles as they move along a single-lane freeway stretch with a sag in such a way that the total travel time of all vehicles is minimized. The problem is solved for various scenarios defined by the number of controlled vehicles and their positions in traffic.

The thesis shows that the optimal acceleration behavior of equipped vehicles involves performing a deceleration-acceleration-deceleration-acceleration (DADA) maneuver in the sag area (see Figure I.B). This maneuver induces the first group of following vehicles to accelerate fast along the vertical curve. As a result, traffic speed and flow at the end of the sag (bottleneck) increase for a particular time period, which produces substantial total travel time savings. The maneuver also triggers a stop-and-go wave

that temporarily limits the inflow into the sag, slowing down the formation of congestion at the bottleneck. Moreover, in some cases, controlled vehicles perform one or more deceleration-acceleration maneuvers upstream of the sag. This additional strategy is used to manage congestion upstream of the sag in such a way that the inflow into the vertical curve is regulated more effectively.

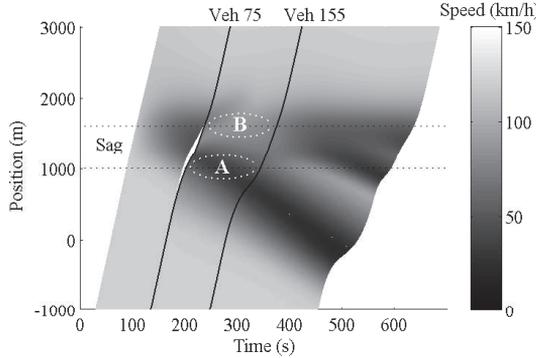


Figure I.B: Primary strategy example (scenario with one controlled vehicle, the 75th vehicle). A is a region of moderately low traffic speed and limited flow around the beginning of the sag; B is a region of moderately high traffic speed and high flow around the end of the sag.

Mainstream traffic flow control using variable speed limits at sags

The findings reported in the previous section show that limiting the inflow into sags can help to alleviate congestion at this type of bottlenecks. Furthermore, previous research indicates that DTM measures based on the concept of mainstream traffic flow control (MTFC), which essentially regulate the inflow into a bottleneck, can substantially reduce congestion at various types of bottlenecks. However, no automatic MTFC measures have been specifically developed for sags yet.

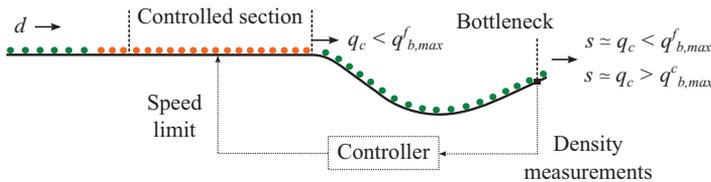


Figure I.C: Traffic situation on a freeway stretch with a sag in high demand conditions if the MTFC measure is applied (d = demand flow, s = exit flow, q_c = outflow from the controlled section, $q_{b,max}^c$ = queue discharge capacity of the sag bottleneck, $q_{b,max}^f$ = free flow capacity of the sag bottleneck; green and orange circles are vehicles that are moving at high and low speed, respectively).

This thesis presents a concept for a MTFC measure for sags (see Figure I.C). The measure applies variable speed limits upstream of the sag in order to regulate the density

at the end of the vertical curve (bottleneck). Density is kept slightly below the critical density, hence preventing traffic from becoming congested while maximizing outflow. Speed limits are set on the basis of a feedback control law. Simulation experiments show that the MTFC measure considerably increases the sag outflow in periods of high traffic demand, which leads to substantial total travel time savings. In view of these results, it is concluded that MTFC measures using variable speed limits have the potential to substantially reduce congestion in freeway networks with sags.

Conclusions

The findings of this dissertation suggest that the main cause of congestion (stop-and-go waves) at sags is that most drivers do not accelerate enough to compensate instantaneously for the increase in grade resistance at this type of freeway sections. This limits vehicle acceleration and makes drivers keep longer headways than expected given their speed, which results in a local decrease of the freeway capacity. This conclusion implies that lane changes are not the primary cause of congestion at sags.

The thesis provides an accurate mathematical description of microscopic driving behavior on freeways with sags. The main novelty is a car-following model that assumes that drivers compensate gradually (not instantaneously) for the increase in grade resistance at sags. This model facilitates improved representation of traffic in simulation tools. In addition, potentially effective traffic management strategies can be identified by analyzing the properties of the model and/or performing optimization experiments using the model. An optimal control method is used for this purpose for the first time in this thesis. In view of the results, it is concluded that the proposed method (possibly extended) could also be used to identify strategies to manage traffic at other types of bottlenecks.

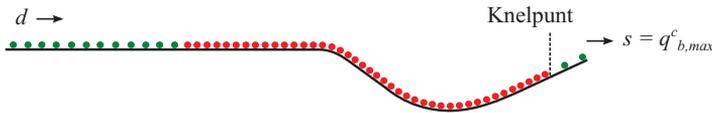
The thesis identifies four traffic management strategies for sags: *a*) improving the ability of drivers to compensate for the increase in grade resistance at sags; *b*) inducing drivers to keep shorter headways when they are queuing; *c*) inducing drivers to accelerate fast through sags and reach the end of the vertical curve at high speeds; and *d*) limiting the inflow into sags. The main goal of all these strategies is to increase the outflow from the sag in high demand conditions, which leads to a decrease in total travel time. Various DTM measures based on strategies *a* and *b* have been proposed in recent years. However, there are no modern applications of strategies *c* and *d*. The thesis proposes two possible ways for implementing the latter two strategies in practice. The first way consists in inducing vehicles equipped with in-car systems (e.g., advanced adaptive cruise control systems) to perform a specific acceleration-related maneuver at sags if traffic demand approaches the bottleneck capacity. The second way involves regulating the traffic density at the bottleneck by modifying the speed limits (and thus the flow) on a freeway section located immediately upstream of the vertical curve on the basis of a feedback control law. Simulation experiments show that both approaches have the potential to considerably reduce congestion at sags.

The following issues should be addressed in future research: *a)* determining how drivers operate the vehicle controls at sags under various circumstances and how that influences their car-following behavior; *b)* refining, calibrating and validating the proposed car-following model; *c)* determining the optimal behavior of vehicles equipped with in-car systems at sags in a wider range of scenarios (e.g., scenarios with multiple lanes, higher penetration rates, etc.); and *d)* developing implementable DTM measures based on the strategies and concepts presented in this thesis.

Samenvatting

Snelwegen hebben als voornaamste doel te zorgen voor korte en betrouwbare reistijden tussen belangrijke locaties in een regio terwijl de veiligheid gewaarborgd wordt. De snelwegen kunnen deze functie echter regelmatig niet naar behoren vervullen vanwege het optreden van *congestie*. Congestie is een verkeersconditie die gekenmerkt wordt door hoge dichtheden en lage snelheden. Dit doet zich voor wanneer te veel voertuigen over een snelweg met beperkte capaciteit proberen te rijden. Vertragingen zijn inherent aan het optreden van congestie. Congestie verlaagt bovendien de reistijdbetrouwbaarheid en verkeersveiligheid, en verhoogt het brandstofverbruik en de luchtvervuiling.

Sags zijn wegdelen waar de gradint aanmerkelijk toeneemt in de richting van het verkeer, dus het deel van een helling voorbij het diepste punt van een dal. Eerder onderzoek toont aan dat de capaciteit van sags op snelwegen aanzienlijk lager is dan de capaciteit van snelwegdelen met andere verticale profielen. Bij een hoge verkeersvraag ontstaat er daarom doorgaans congestie op sags welke onderdeel zijn van snelwegen. Congestie doet zich meestal eerst voor op de middelste rijstrook waarna het zich verspreidt naar de andere rijstroken. De congestiepatronen bij sags bestaan uit stop-and-go waves. Dit zijn ruimtelijk begrensde gebieden met lage verkeerssnelheden welke zich stroomopwaarts verplaatsen met een constante snelheid. Het knelpunt is doorgaans het eind van de verticale bocht (zie Figuur II.A).



Figuur II.A: Typische verkeerssituatie op een snelwegdeel met een sag bij een hoge verkeersvraag wanneer er geen beheersmaatregel is toegepast (d = verkeersvraag, s = uitstroom, $q_{b,max}^c$ = uitstroomcapaciteit van de sag; groene en rode cirkels zijn voertuigen die respectievelijk met een hoge en een lage snelheid voortbewegen).

In enkele landen zijn sags een van de voornaamste type knelpunten op de snelwegen. Zo doet in Japan ongeveer 60% van de files zich voor bij sags. Hierdoor zijn er de afgelopen jaren verscheidene dynamische verkeersmanagement (DVM) maatregelen ontwikkeld met het doel om de congestie bij sags te verminderen. Hoewel er vooruitgang is geboekt, zijn er nog verschillende uitdagingen om de congestie op deze typen

knelpunten verder te verminderen. Tot deze uitdagingen behoren: *a)* het verkrijgen van beter inzicht in de oorzaken van congestie op sags; *b)* het ontwikkelen van nieuwe wiskundige modellen om verkeer op sags op een meer realistische manier weer te geven; en *c)* het identificeren van de meest effectieve strategieën om verkeer bij sags te beheersen. Met betrekking tot dit laatste punt dienen huidige ontwikkelingen in dynamisch verkeersmanagement (zoals het gebruikelijke beheer van verkeersstromen en in-car toepassingen) te worden overwogen.

Het hoofddoel van deze dissertatie is het ontwikkelen van nieuwe verkeersmanagementconcepten voor sags, gebaseerd op de oorzaken van congestie bij deze typen knelpunten. Hiertoe beantwoordt deze dissertatie de volgende onderzoeksvragen: *a)* wat is de primaire oorzaak van stop-and-go waves bij sags?; *b)* hoe dient het verkeer op een microscopisch niveau gemodelleerd te worden om de karakteristieken op snelwegen met sags op een representatieve wijze te reproduceren?; *c)* hoe dienen voertuigen welke zijn uitgerust met in-car systemen zich te gedragen bij sags om de ernst van de congestie te minimaliseren?; en *d)* kunnen gebruikelijke reguleringsmaatregelen voor verkeersstromen welke zijn gebaseerd op het gebruik van variabele snelheidslimieten de congestie op sags substantieel reduceren?

Oorzaken van stop-and-go waves op sags

Het is bekend dat congestie zich op sags voordoet in de vorm van stop-en-go waves. Er is echter geen consensus in de literatuur over de primaire oorzaak van deze stop-en-go waves. Er is bewijs dat het voertuigvolggedrag van de meeste bestuurders verandert bij sags: bestuurders zijn geneigd om langere volgafstanden aan te houden dan op basis van hun snelheid kan worden verwacht. De voornaamste reden blijkt te liggen in inadequaate versnellingsgedrag. Enkele onderzoekers stellen dat stop-en-go waves hoofdzakelijk het resultaat zijn van het cumulatieve effect van deze lokale veranderingen in voertuigvolggedrag. Andere onderzoekers suggereren dat versturende rijstrookwisselingen de oorzaak kunnen zijn van stop-and-go waves op sags.

Dit proefschrift draagt bij aan het begrip van de oorzaken van congestie op sags. Hiertoe zijn empirische trajectorie-data van een snelwegdeel met een sag geanalyseerd en is laten zien dat voor dit deel in de meeste gevallen (89%) de vorming en groei van stop-en-go waves gerelateerd is aan voertuigvolgfenomenen, terwijl veranderingen van rijstrook veel minder vaak de teweegbrengende factor zijn (11% van de gevallen). In het licht van deze resultaten wordt geconcludeerd dat de hoofdoorzaak van stop-and-go waves bij sags niet versturende rijstrookwisselingen zijn maar voertuigvolgfenomenen (welke mogelijk veroorzaakt worden door de toename van de gradint van de snelweg zoals gesuggereerd door andere onderzoekers). Dit wil echter niet zeggen dat rijstrookwisselingen geen rol spelen in de dynamiek van congestie.

Microscopisch modelleren van verkeersstromen op sags

Verkeersmodellen zijn essentiële middelen om de dynamiek van verkeersstromen te begrijpen en DVM-maatregelen te ontwikkelen en bepalen. De bevindingen welke gerapporteerd zijn in de voorgaande sectie duiden er op dat het bij het op een microscopisch niveau modelleren van verkeer op sags cruciaal is om een voertuigvolgmodel te gebruiken dat rekening houdt met de invloed van het verticale profiel van de snelweg. Aan de andere kant is er geen bewijs dat het gedrag bij rijstrookwisselingen beïnvloed wordt door het verticale profiel van de snelweg. Daarom worden reguliere modellen voor rijstrookwisselingen geschikt geacht voor het modelleren van sags op een microscopisch niveau.

Dit proefschrift draagt bij aan het theoretische begrip en de wiskundige beschrijving van microscopisch rijgedrag op sags door het introduceren van een nieuw model dat het voertuigvolggedrag op deze typen snelwegdelen nauwkeuriger beschrijft dan andere modellen. Het gintrodeerde voertuigvolgmodel veralgemeniseert het Intelligent Driver Model (IDM) door een term toe te voegen die de invloed van het verticale profiel van de snelweg op de versnelling van het voertuig beschrijft. Die is gebaseerd op twee aannames (beide ondersteund door empirisch bewijs): *a*) bestuurders compenseren lineair over de tijd voor de toename in de weerstand welke veroorzaakt wordt door de toename van de helling; en *b*) de mate van compensatie (gedefinieerd door een parameter) is doorgaans lager dan de mate waarmee de weerstand toeneemt wanneer een voertuig zich over een sag verplaatst.

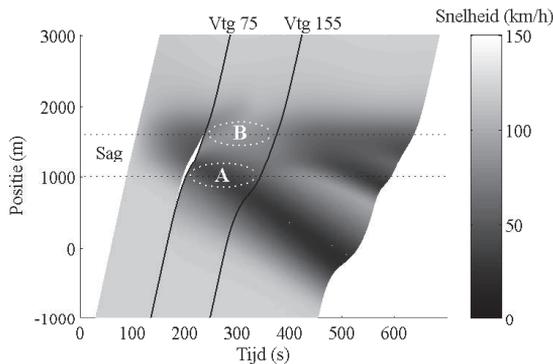
Het proefschrift laat zien dat een microscopisch verkeersmodel opgebouwd uit het nieuwe voertuigvolgmodel en een regulier rijstrookwisselingsmodel (LMRS) in staat is om de belangrijkste fenomenen van verkeersstromen op sags te reproduceren. Het model genereert een knelpunt dat wordt veroorzaakt door de verticale bocht en reproduceert de locatie op een nauwkeurige manier. In het licht van deze resultaten wordt geconcludeerd dat het verkeersstroommodel (en in bijzonder het voertuigvolgmodel) op het eerste gezicht valide lijkt te zijn en dat de onderliggende aannames plausibel zijn. Verder laat een gevoeligheidsanalyse manieren zien om congestie bij sags te mitigeren. Congestie blijkt minder extreem te zijn als bestuurders: *a*) sneller compenseren voor de toename in de hellingsweerstand op sags; en *b*) kortere gewenste volgafstanden hebben wanneer ze in de file staan.

Optimaal versnellingsgedrag van voertuigen uitgerust met in-car systemen op sags

Er is tegenwoordig groeiende interesse in de ontwikkeling van DVM-maatregelen voor sags welke gebaseerd zijn op het gebruik van in-car systemen (bijvoorbeeld, advies-systemen of geavanceerde adaptieve cruise control systemen). De meeste voorgestelde maatregelen zijn echter in vroege stadia van ontwikkeling. Dit proefschrift draagt bij

aan de ontwikkeling van deze typen maatregelen door te bepalen hoe uitgeruste voertuigen zich zouden moeten verplaatsen op sags om congestie te minimaliseren. Hiertoe is een optimal control methode gebruikt. Binnen het optimal control probleem wordt gezocht naar versnellingsprofielen voor een aantal voertuigen op een enkelbaans stuk snelweg met een sag waarbij de totale reistijd van alle voertuigen geminimaliseerd wordt. Het probleem is opgelost voor verschillende scenarios gedefinieerd door het aantal gereguleerde voertuigen en hun verkeerspositie.

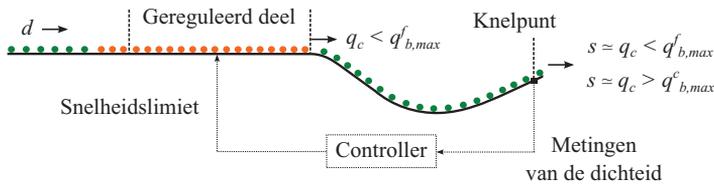
Het proefschrift laat zien dat het optimale versnellingsgedrag van uitgeruste voertuigen bestaat uit het uitvoeren van een vertraging-versnelling-vertraging-versnelling (V V V V) manoeuvre in het gebied van de sag (zie Figuur II.B). Deze manoeuvre zet de eerste groep van volgende voertuigen aan om snel te versnellen langs de verticale helling. Hierdoor nemen de verkeerssnelheid en -stroom aan het eind van de sag (knelpunt) voor een bepaalde periode toe wat leidt tot substantiele besparingen op de totale reistijd. De manoeuvre veroorzaakt ook een stop-and-go wave die de instroom van de sag tijdelijk beperkt, wat de vorming van congestie bij het knelpunt vermindert. In enkele gevallen voeren de gereguleerde voertuigen bovendien een of meerdere vertraging-versnelling manoeuvres uit stroomopwaarts van de sag. Deze additionele strategie wordt gebruikt om de congestie stroomopwaarts van de sag zo te reguleren dat de instroom in the verticale bocht effectiever geregeld is.



Figuur II.B: Voorbeeld van de basisstrategie (scenario met een gereguleerd voertuig, het 75e voertuig). *A* is een gebied met een vrij lage verkeerssnelheid en beperkte verkeersstromen rond het begin van de sag; *B* is een gebied met een vrij hoge verkeerssnelheid en grote verkeersstromen rond het eind van de sag.

Gebruikelijke verkeersstroomregulering met variabele snelheidslimieten op sags

De bevindingen in de vorige sectie laten zien dat het beperken van de instroom op sags kan helpen bij het verlichten van congestie bij dit type knelpunten. Eerder onderzoek laat bovendien zien dat DVM-maatregelen gebaseerd op de gebruikelijke verkeersstroomregulering (MTFC, het reguleren van de instroom naar een knelpunt) congestie substantieel kunnen reduceren bij verschillende typen knelpunten. Er zijn echter nog geen automatische MTFC-maatregelen ontwikkeld specifiek voor sags.



Figuur II.C: Verkeerssituatie op een snelwegdeel met een sag met een hoge verkeersvraag wanneer de MTFC maatregel is toegepast (d = verkeersvraag, s = uitstroom, q_c = uitstroom van de gereguleerde sectie, $q_{b,max}^e$ = uitstroomcapaciteit van de sag, $q_{b,max}^f$ = free-flow capaciteit van de sag; groene en oranje cirkels zijn voertuigen die respectievelijk met een hoge en een lage snelheid voortbewegen).

Dit proefschrift presenteert een concept voor een MTFC-maatregel voor sags (zie Figuur II.C). De maatregel past variabele snelheidslimieten toe stroomopwaarts van de sags om de dichtheid aan het eind van de verticale bocht (het knelpunt) te reguleren. De dichtheid wordt enigszins onder de kritieke dichtheid gehouden om congestie te voorkomen en de uitstroom te maximaliseren. De snelheidslimieten zijn gebaseerd op een feedback control regel. Simulaties laten zien dat de MTFC-maatregel de uitstroom van sags bij een hoge verkeersvraag aanzienlijk verhoogt, wat leidt tot substantiele besparingen op de totale reistijd. In het licht van deze resultaten wordt geconcludeerd dat MTFC-maatregelen welke gebruikmaken van variabele snelheidslimieten de potentie hebben om congestie op snelwegen met sags wezenlijk te reduceren.

Conclusies

De bevindingen van dit proefschrift wijzen erop dat de hoofdoorzaak van congestie (stop-and-go waves) bij sags is dat bestuurders niet genoeg versnellen om onmiddellijk te compenseren voor de toename in de hellingsweerstand op dit type snelwegdelen. Dit beperkt de voertuigversnelling en zorgt ervoor dat bestuurders langere volgafstanden aanhouden dan op basis van hun snelheid kan worden verwacht, wat resulteert in een lokale afname van de capaciteit van de snelweg. Deze conclusie impliceert dat rijstrookwisselingen niet de primaire oorzaak zijn van congestie op sags.

Dit proefschrift presenteert een accurate mathematische beschrijving van microscopisch rijgedrag op snelwegen met sags. De belangrijkste noviteit is een voertuigvolgmodel dat beschrijft dat bestuurders geleidelijk aan (en niet onmiddellijk) compenseren voor de toename in de hellingsweerstand bij sags. Dit model maakt het mogelijk de representatie van verkeer in simulatietools te verbeteren. Bovendien kunnen mogelijk effectieve verkeersmanagementstrategieën gidentificeerd worden door de karakteristieken van het model te analyseren en/of optimalisatie-experimenten uit te voeren met het model. Hiertoe is in dit proefschrift voor de eerste keer een optimal control methode gebruikt. Met het oog op de resultaten wordt geconcludeerd dat de (mogelijk uitgebreide) voorgestelde methode ook gebruikt kan worden om strategieën te identificeren om het verkeer te reguleren bij andere typen knelpunten.

Dit proefschrift identificeert vier typen verkeersmanagementstrategien voor sags: *a*) het verbeteren van de mogelijkheid van bestuurders om te compenseren voor de toename van de hellingsweerstand op sags; *b*) het aanzetten van bestuurders tot het aanhouden van kortere volgafstanden wanneer ze in de file staan; *c*) het aanzetten van bestuurders tot het snel versnellen op sags en het met hoge snelheden bereiken van het einde van de verticale bocht; en *d*) het beperken van de instroom van sags. Het hoofddoel van alle strategien is het bij een hoge verkeersvraag laten toenemen van de uitstroom van de sags, wat leidt tot een afname van de totale reistijd. In de afgelopen jaren zijn er verschillende DVM-maatregelen voorgesteld welke gebaseerd zijn op strategien *a* en *b*. Hedentendage zijn er echter geen toepassingen van strategien *c* en *d*. Dit proefschrift stelt twee mogelijke manieren voor om de laatste twee strategien in de praktijk te brengen. De eerste manier bestaat uit het aanzetten van voertuigen welke zijn uitgerust met in-car systemen (bijvoorbeeld geavanceerde adaptieve cruise control systemen) tot het uitvoeren van een specifieke versnellingsgerelateerde manoeuvre op sags wanneer de verkeersvraag de capaciteit van het knelpunt nadert. De tweede manier bestaat uit het reguleren van de verkeersdichtheid bij het knelpunt. Hiertoe worden de snelheidslimieten (en dus de verkeersstromen) op een snelwegdeel direct stroomopwaarts van de verticale bocht gewijzigd gebaseerd op een feedback control regel. Experimenten laten zien dat beide aanpakken de potentie hebben om de congestie op sags aanzienlijk te reduceren.

De volgende punten moeten worden aangepakt in toekomstig onderzoek: *a*) bepalen hoe bestuurders de beheersmaatregelen bij sags uitvoeren onder verschillende omstandigheden en hoe dat hun voertuigvolgedrag beïnvloedt; *b*) verfijnen, kalibreren en valideren van het gintrodeerde voertuigvolgmodel; *c*) bepalen van het optimale gedrag van voertuigen welke zijn uitgerust met in-car systemen op sags met meer uiteenlopende scenarios (bijvoorbeeld scenarios met meerdere rijstroken, hogere penetratiegraden, etc.); en *d*) ontwikkelen van implementeerbare DVM-maatregelen gebaseerd op de strategien en concepten gepresenteerd in dit proefschrift.

About the author

Bio



Bernat Goñi-Ros was born in Barcelona (Catalonia, Spain) in 1982. In 2005, he obtained a bachelor's degree with distinction in Environmental Science and Engineering from Autonomous University of Barcelona. In the same year, he received a national award from the Spanish Ministry of Education in recognition of his academic achievements. After three months doing volunteer work in Rio de Janeiro (Brazil), he joined CIREM, a Barcelona-based consulting firm that specializes in urban planning, where he worked as a consultant for three years. In 2008, he received a grant to spend four months as a visiting researcher in the Department of City & Regional Planning at Cornell University (New York). In 2009, he moved to the Netherlands to pursue a master's degree in Transport, Infrastructure and Logistics at Delft University of Technology. He graduated with distinction in 2011, after defending his master thesis "Cost-benefit analysis of RetBus, a new Bus Rapid Transit system in Barcelona". Immediately after that, he started his Ph.D. in the Department of Transport & Planning at Delft University of Technology. The main goal of his research project was to develop new traffic management strategies to mitigate congestion at freeway sags. The project was funded by Toyota Motor Europe and was supervised by Bart van Arem, Serge Hoogendoorn and Victor Knoop. During his doctoral studies, Bernat worked as a teaching assistant and supervised the thesis of a graduate student. In addition, he organized various seminars and courses for doctoral students. He also served as a reviewer for various international journals and conferences. Bernat is currently a postdoctoral researcher in the Department of Transport & Planning at Delft University of Technology, where he works in various research projects aimed at developing theories and models to explain and predict the dynamics of pedestrian and cyclist flows. Bernat's research interests include transportation planning, traffic flow theory and intelligent transportation systems, with a focus on urban mobility and sustainability.

Publications

Journal articles

- Goñi-Ros, B., Knoop, V.L., Shiomi, Y., Takahashi, T., van Arem, B., & Hoogendoorn, S.P. (2016) Modeling traffic at sags. *International Journal of Intelligent Transportation Systems Research*, 14(1), 64–74.
- Goñi-Ros, B., Knoop, V.L., van Arem, B., & Hoogendoorn, S.P. (2014) Mainstream Traffic Flow Control at Sags, *Transportation Research Record: Journal of the Transportation Research Board*, No. 2470, 57–64.
- Goñi-Ros, B., Knoop, V.L., van Arem, B., & Hoogendoorn, S.P. (2014) Empirical analysis of the causes of stop-and-go waves at sags, *IET Intelligent Transport Systems*, 8(5), 499–506.

The following article is currently under review:

- Goñi-Ros, B., Knoop, V.L., Takahashi, T., Sakata, I., van Arem, B., & Hoogendoorn, S.P. (Submitted) Optimization of traffic flow at freeway sags by controlling the acceleration of vehicles equipped with in-car systems. *Transportation Research Part C: Emerging Technologies*.

Peer-reviewed conference contributions

- Zhang, F., Chen, Y., Goñi-Ros, B., Gao, J., & Knoop, V.L. (2015). *Routing Strategy Including Time and Carbon Dioxide Emissions: Effects on Network Performance*. Presented at 95th Annual Meeting of the Transportation Research Board, 10-14 January 2016, Washington, D.C.
- Goñi-Ros, B., Knoop, V.L., Kitahama, K., van Arem, B., & Hoogendoorn, S.P. *Traffic Flow Optimization at Sags by Controlling the Acceleration of Some Vehicles*. Presented at 11th Conference on Traffic and Granular Flow, 28-30 October 2015, Nootdorp, the Netherlands.
- Papacharalampous, A.E., Wang, M., Knoop, V.L., Goñi-Ros, B., Takahashi, T., Sakata, I., van Arem, B., & Hoogendoorn, S.P. *Mitigating Congestion at Sags with Adaptive Cruise Control Systems*. Presented at 18th International IEEE Conference on Intelligent Transportation Systems, 15-18 September 2015, Las Palmas de Gran Canaria, Spain.
- Goñi-Ros, B., Knoop, V.L., van Arem, B., & Hoogendoorn, S.P. *Mainstream traffic flow control at sags*. Presented at 93rd Annual Meeting of the Transportation Research Board, 12-16 January 2014, Washington, D.C.

- Goñi-Ros, B., Knoop, V.L., van Arem, B., & Hoogendoorn, S.P. *Modeling driving behavior and traffic flow at sags*. Presented at 20th ITS World Congress, 14-18 October 2013, Tokyo, Japan.
- Goñi-Ros, B., Knoop, V.L., Schakel, W.J., van Arem, B., & Hoogendoorn, S.P. (2015). *A model of car-following behavior at sags*, in: Chraïbi, M., Boltes, M., Schadschneider, A., Seyfried, A. (Eds.) *Proceedings of the 10th Conference on Traffic and Granular Flow 2013* (p. 385–393). Springer: Berlin.
- Goñi-Ros, B., Knoop, V.L., van Arem, B., & Hoogendoorn, S.P. *Car-following Behavior at Sags and its Impacts on Traffic Flow*. Presented at 92nd Annual Meeting of the Transportation Research Board, 13-17 January 2013, Washington, D.C.
- Goñi-Ros, B., Knoop, V.L., van Arem, B., & Hoogendoorn, S.P. *Reducing congestion at uphill freeway sections by means of a Gradient Compensation System*. Presented at 2012 IEEE Intelligent Vehicles Symposium, 3-7 June 2012, Alcalá de Henares, Spain. (Nominated for best conference paper)

Technical reports

- Goñi-Ros, B., Knoop, V.L., Wang, M., Papacharalampous, A., van Arem, B., & Hoogendoorn, S.P. *Reducing congestion at sags by means of Cooperative Advanced Driver Assistance Systems – Final report*. Department of Transport & Planning, Delft University of Technology, the Netherlands, January 2016.
- Papacharalampous, A., Wang, M., Knoop, V.L., Goñi-Ros, B., Schakel, W.J., van Arem, B., & Hoogendoorn, S.P. *Reducing congestion at sags by means of Cooperative Advanced Driver Assistance Systems – Stage 6.5: Effectiveness of C2X controller using multilane traffic flow simulation*. Department of Transport & Planning, Delft University of Technology, the Netherlands, December 2015.
- Schakel, W.J., Papacharalampous, A., Wang, M., Knoop, V.L., Goñi-Ros, B., van Arem, B., & Hoogendoorn, S.P. *Reducing congestion at sags by means of Cooperative Advanced Driver Assistance Systems – Stage 6: Development and assessment of C2X controller (Annex to report)*. Department of Transport & Planning, Delft University of Technology, the Netherlands, November 2015.
- Goñi-Ros, B., Knoop, V.L., van Arem, B., & Hoogendoorn, S.P. *Reducing congestion at sags by means of Cooperative Advanced Driver Assistance Systems – Stage 5: Design of an optimal controller and analysis of its behavior*. Department of Transport & Planning, Delft University of Technology, the Netherlands, May 2014.
- Goñi-Ros, B., Knoop, V.L., van Arem, B., & Hoogendoorn, S.P. *Reducing congestion at sags by means of Cooperative Advanced Driver Assistance Systems –*

Stage 4: Quantitative evaluation of the effectiveness of control concepts. Department of Transport & Planning, Delft University of Technology, the Netherlands, July 2013.

- Goñi-Ros, B., Knoop, V.L., van Arem, B., & Hoogendoorn, S.P. *Reducing congestion at sags by means of Cooperative Advanced Driver Assistance Systems – Stage 3: Analysis and evaluation of control systems to reduce congestion at sags.* Department of Transport & Planning, Delft University of Technology, the Netherlands, April 2013.
- Goñi-Ros, B., Knoop, V.L., van Arem, B., & Hoogendoorn, S.P. *Reducing congestion at sags by means of Cooperative Advanced Driver Assistance Systems – Stage 2: Modeling driving behavior and traffic flow at sags.* Department of Transport & Planning, Delft University of Technology, the Netherlands, November 2012.
- Goñi-Ros, B., Knoop, V.L., van Arem, B., & Hoogendoorn, S.P. *Reducing congestion at sags by means of cooperative ACC systems – Phase 1: Theoretical framework of the causes of congestion at sags.* Department of Transport & Planning, Delft University of Technology, the Netherlands, April 2012.

TRAIL Thesis Series

The TRAIL Thesis Series is a series of the Netherlands TRAIL Research School on Transport, Infrastructure and Logistics. The following list contains the most recent dissertations in the TRAIL Thesis Series. For a complete overview of more than 150 titles see the TRAIL website: www.rsTRAIL.nl.

Goñi-Ros, B., *Traffic Flow at Sags: Theory, Modeling and Control*, T2016/2, March 2016, TRAIL Thesis Series, the Netherlands.

Khademi, E., *Effects of Pricing Strategies on Dynamic Repertoires of Activity-Travel Behaviour*, T2016/1, February 2016, TRAIL Thesis Series, the Netherlands.

Cong, Z., *Efficient Optimization Methods for Freeway Management and Control*, T2015/17, November 2015, TRAIL Thesis Series, the Netherlands.

Kersbergen, B., *Modeling and Control of Switching Max-Plus-Linear Systems: Rescheduling of railway traffic and changing gaits in legged locomotion*, T2015/16, October 2015, TRAIL Thesis Series, the Netherlands.

Brands, T., *Multi-Objective Optimisation of Multimodal Passenger Transportation Networks*, T2015/15, October 2015, TRAIL Thesis Series, the Netherlands.

Ardıç, Özgül, *Road Pricing Policy Process: The interplay between policy actors, the media and public*, T2015/14, September 2015, TRAIL Thesis Series, the Netherlands.

Xin, J., *Control and Coordination for Automated Container Terminals*, T2015/13, September 2015, TRAIL Thesis Series, the Netherlands.

Anand, N., *An Agent Based Modelling Approach for Multi-Stakeholder Analysis of City Logistics Solutions*, T2015/12, September 2015, TRAIL Thesis Series, the Netherlands.

Hurk, E. van der, *Passengers, Information, and Disruptions*, T2015/11, June 2015, TRAIL Thesis Series, the Netherlands.

Davydenko, I., *Logistics Chains in Freight Transport Modelling*, T2015/10, May 2015, TRAIL Thesis Series, the Netherlands.

Schakel, W., *Development, Simulation and Evaluation of In-car Advice on Headway, Speed and Lane*, T2015/9, May 2015, TRAIL Thesis Series, the Netherlands.

Dorsser, J.C.M. van, *Very Long Term Development of the Dutch Inland Waterway Transport System: Policy analysis, transport projections, shipping scenarios, and a new perspective on economic growth and future discounting*, T2015/8, May 2015, TRAIL Thesis Series, the Netherlands.

Hajiahmadi, M., *Optimal and Robust Switching Control Strategies: Theory, and Applications in Traffic Management*, T2015/7, April 2015, TRAIL Thesis Series, the Netherlands.

Wang, Y., *On-line Distributed Prediction and Control for a Large-scale Traffic Network*, T2015/6, March 2015, TRAIL Thesis Series, the Netherlands.

Vreeswijk, J.D., *The Dynamics of User Perception, Decision Making and Route Choice*, T2015/5, February 2015, TRAIL Thesis Series, the Netherlands.

Lu, R., *The Effects of Information and Communication Technologies on Accessibility*, T2015/4, February 2015, TRAIL Thesis Series, the Netherlands.

Ramos, G. de, *Dynamic Route Choice Modelling of the Effects of Travel Information using RP Data*, T2015/3, February 2015, TRAIL Thesis Series, the Netherlands.

Sierzchula, W.S., *Development and Early Adoption of Electric Vehicles: Understanding the tempo*, T2015/2, January 2015, TRAIL Thesis Series, the Netherlands.

Vianen, T. van, *Simulation-integrated Design of Dry Bulk Terminals*, T2015/1, January 2015, TRAIL Thesis Series, the Netherlands.

Risto, M., *Cooperative In-Vehicle Advice: A study into drivers' ability and willingness to follow tactical driver advice*, T2014/10, December 2014, TRAIL Thesis Series, the Netherlands.

Djukic, T., *Dynamic OD Demand Estimation and Prediction for Dynamic Traffic Management*, T2014/9, November 2014, TRAIL Thesis Series, the Netherlands.

Chen, C., *Task Complexity and Time Pressure: Impacts on Activity-Travel Choices*, T2014/8, November 2014, TRAIL Thesis Series, the Netherlands.

Wang, Y., *Optimal Trajectory Planning and Train Scheduling for Railway Systems*, T2014/7, November 2014, TRAIL Thesis Series, the Netherlands.

Wang, M., *Generic Model Predictive Control Framework for Advanced Driver Assistance Systems*, T2014/6, October 2014, TRAIL Thesis Series, the Netherlands.

Kecman, P., *Models for Predictive Railway Traffic Management*, T2014/5, October 2014, TRAIL Thesis Series, the Netherlands.

Davarynejad, M., *Deploying Evolutionary Metaheuristics for Global Optimization*, T2014/4, June 2014, TRAIL Thesis Series, the Netherlands.

Li, J., *Characteristics of Chinese Driver Behavior*, T2014/3, June 2014, TRAIL Thesis Series, the Netherlands.