Capacity Drop on Freeways: Traffic Dynamics, Theory and Modeling

Yuan, Kai

DOI
10.4233/uuid:d194ac2a-3176-4550-a5d1-ae231c3a44fd

Publication date
2016

Document Version
Final published version

Citation (APA)
Yuan, K. (2016). *Capacity Drop on Freeways: Traffic Dynamics, Theory and Modeling*. [Dissertation (TU Delft), Delft University of Technology]. TRAIL Research School. https://doi.org/10.4233/uuid:d194ac2a-3176-4550-a5d1-ae231c3a44fd

Capacity Drop on Freeways: Traffic Dynamics, Theory and Modeling

Kai Yuan

Delft University of Technology, 2016

This thesis is a result from a project
funded by the China scholarship council (CSC).

Cover illustration: Hanqi Huang & Zelin Lu

Capacity Drop on Freeways: Traffic Dynamics, Theory and Modeling

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 14 december 2016 om 10.00 uur

door

Kai YUAN

Master of Engineering in Traffic and Transportation Planning and Management,

Beijing Jiaotong University,

geboren te Yiyang, Jiangxi Province, China

Dit proefschrift is goedgekeurd door de:
promotor: Prof. dr. ir. Serge P. Hoogendoorn
copromotor: Dr. Victor L. Knoop

Samenstelling van de promotiecommissie :

Rector Magnificus	voorzitter
Prof. dr. ir. S.P. Hoogendoorn	Technische Universiteit Delft, promotor
Dr. V.L. Knoop	Technische Universiteit Delft, copromotor
Onafhankelijke leden:	
Dr. S. Ahn	University of Wisconsin - Madison
Dr. J. Laval	Georgia Institute of Technology
Prof. dr. ir. B. De Schutter	Technische Universiteit Delft
Prof. dr. ir. C.M.J. Tampère	Katholieke Universiteit Leuven
Prof. dr. ir. B. van Arem	Technische Universiteit Delft

TRAIL Thesis Series no. T2016/24, the Netherlands TRAIL Research School

TRAIL
P.O. Box 5017
2600 GA Delft
The Netherlands
E-mail: info@rsTRAIL.nl

ISBN 978-90-5584-212-4

Printed in the Netherlands

Dedicated to my beloved parents

献给我最亲爱的父亲母亲

Preface

It was in November 2011 that I had my first interview in the 3rd PhD workshop in Beijing for my PhD position. I had a very nice talk with Prof. Fokkema and Marianne there, which also opened a door to my first interview with Serge Hoogendoorn and Victor Knoop in December 2011. In May 2012, I am granted the CSC (China Scholarship Council) scholarship for my PhD research. Finally, in September 2012, I arrived in the Netherlands and became a PhD candidate. I would like to thank those who helped and encouraged me to go abroad. I also want to thank the China Scholarship Council for financially supporting my PhD research.

While my PhD research started in September 2012, I began working on the capacity drop phenomenon in my second PhD year. In my first PhD year, I worked on distributed traffic signal control in urban environments. After realizing that I have a very strong interest in traffic flow theory, I decided to change my research topic. I told Victor my decision. He agreed immediately. He told me that there are many debates about the cause of the capacity drop phenomenon, and suggested me to work on the capacity drop. Victor's suggestion was very attractive to me. In my eagerness to end well known debates, I set out on a journey of understanding the capacity drop.

Time flies! At present, I am proud and excited to present this book to conclude my PhD research. On completion of my PhD research, I would like to thank those who helped and accompanied me in these four years.

Serge and Victor, I am very lucky to have you as my supervisors. I want to thank you for giving me the opportunity to pursue this PhD. During the four-year study, I am so grateful for working under your expert guidances. Your constructive comments, brilliant ideas, boundless enthusiasm, tremendous support and considerable encouragement let me enjoy every moment when exploring traffic science. Working with you gave me many opportunities to interact with excellent international researchers and experts, which is amazing! Particularly, I would like to take this opportunity to thank my daily supervisor, Victor. I appreciate your infinite patience in years. It is not easy to keep your door open to discussions at any time. In the past four years, you have shared a lot of knowledge and experiences to guide me to be an independent researcher. Thanks!

Also Jorge Laval from the Georgia Institute of Technology in Atlanta has supervised a part of this research when I visited Georgia Institute of Technology. Jorge, I really

like working with you. Thank you for discussing with me almost twice a week. I have learned a lot from you during the stay. Your good humor allowed me to enjoy every minute when meeting you.

When I stayed in Atlanta, Hyun and Joy also inspired me. I am so happy to work with you in the same office - Room 207. Haiyang Zou, knowing you made the stay in Atlanta a very interesting period. Every weekend, shopping with you and your friends gave me a lot of great time.

I want to thank Ludovic Leclercq, Thomas Schreiter and Rui Jiang for being co-authors of several papers. Your comments and contributions made me confident about my works, and helped to improve the quality of the papers. Thanks a lot!

Many thanks go to Bart De Schutter, Bart van Arem, Chris Tampère, Jorge Laval and Soyoung Ahn for being my PhD committee members. I am grateful for your valuable comments on my research and the manuscript of this book.

I also want to thank colleagues at Delft University of Technology. It is an amazing journey to conduct my PhD research in the Transport and Planning department. I had a lot of fun when working with you. I appreciate every interesting discussion and all the support I received from you. I enjoyed myself in many social activities, such as travels, drinks, exchanging gifts and dinners.

During the past four years, many people helped me out when I encountered difficulties in daily life. I am grateful for your kindness.

I would like to thank Qichen - my paranymph - for your friendship and for assisting me when finishing this book. Thank Jeroen for translating the summary of my thesis into Dutch. Thank Victor for the Dutch translation of my propositions. Thank Hanqi Huang and Zelin Lu for drawing the cover illustration. Thank technical and administrative staffs in TU Delft as well as TRAIL Research School for taking care of many practical issues for my PhD defense.

Finally I want to thank my parents. 爸爸妈妈，感谢你们对我无尽的关爱，让我在异国他乡求学时，不会感到孤单。感谢你们对我的支持与鼓励，让我一直有勇气面对困难。谢谢，我爱你们！

Kai Yuan
China, November 2016

Contents

List of Figures

List of Tables

Notation

List of acronyms

CFL	:	Courant-Friedrichs-Lewy
DVC	:	Driver-Vehicle-Combination
$GSOM$	:	Generic Second-Order Model
ITS	:	Intelligent Transportation System
LWR	:	Lighthill-Whitham-Richards

List of symbols

a	:	Acceleration
a^{desire}	:	Desired acceleration
a_i^{desire}	:	Desired acceleration of vehicle i
$a_{\min}$, $a_{\max}$	:	Minimal and maximal desired acceleration
Δa	:	Acceleration difference
C	:	Free-flow (pre-queue) capacity
h	:	Time headway
$h_{\min}$	:	Minimal time headway
H	:	Sum of free-flow time headways from the second vehicle to the last vehicle
H_{cri}	:	sum of free-flow time headways from vehicle 2 to vehicle N in no-capacity-drop condition.
i	:	Object (e.g., Lagrangian cluster or vehicle) index
k	:	Time instant
l_{down}	:	Space on the outgoing link
l_{up}^p	:	Space upstream of the node on the link with priority
l_{up}^n	:	Space upstream of the node on the other link
m^p	:	Number of lanes on the link with priority
m_{down}	:	Number of lanes on the outgoing link
N	:	Cumulative number of vehicles
ΔN	:	Number of vehicles in one Lagrangian cluster
q	:	Traffic flow
q_d	:	Queue discharge rate
q_j	:	Traffic flow in a congested state
Q, Q_a	:	Flow-density fundamental diagram, the subscription a indicates the acceleration branch
q_0	:	Queue discharge rate of wide moving jams
r	:	Correlation coefficient
s	:	Spacing
s_{cri}	:	Critical spacing
s_d	:	Spacing corresponding to queue discharge rate
s_c^p	:	Critical spacing on the incoming link with priority
Δs	:	Spacing in wide moving jam in which vehicular speed are zero m/s
$s_i^{\text{extension}}$	:	Extension of spacing between vehicle i and $i-1$
$s_{i,N}^{\text{extension}}$	:	Extension of spacing between first and the last vehicle
s_{shift}	:	Distance vehicle travels during the reaction time extension
$S(v)$	:	Inverse of $V(s)$, i.e., $S(v) = V^{-1}(s)$
t	:	Time
t_r	:	Reaction time
t_{ex}	:	Reaction time extension
Δt	:	Time step

$\Delta t_1, \Delta t_2$	:	Time slot during acceleration
U	:	Uniform distribution
v	:	Vehicular speed
v_t, v_s	:	Time mean speed, space mean speed
v_j	:	Vehicular speed in a congested state
v_f	:	Free-flow speed
$v_{\max}$	:	Maximum free-flow speed
$v_j^{\max}$	:	Lowest speed in congestion leading to no capacity drop
V, V_a	:	Speed-spacing fundamental diagram, the subscription a indicates the acceleration branch
V_i^*	:	Default fundamental speed-spacing relation
w, w_a	:	Congestion wave speed, the subscription a indicates the acceleration branch in the fundamental diagram
x	:	Location
Δx_{down}	:	Distance between the node and the most upstream Lagrangian cluster on the outgoing link
α	:	Parameter indicating by how much the queue discharge rate increases as the speed in congestion grow by one unit
δ	:	Merging ratio
ρ	:	Traffic density
ρ_j	:	Traffic density in a congested state
ρ_d	:	Traffic density corresponding to queue discharge rate
ρ_{cri}	:	Critical density
$\rho_{\max}$	:	Maximum jam density
ρ_{slow}	:	Density in heavy congestion J_{slow}
ρ_{fast}	:	Density in light congestion J_{fast}
τ	:	Trip time of wave propagating upstream in Newell's first order car-following model
γ	:	Parameter indicating the reaction time extension in wide moving jams
β	:	Parameters related to relation between the mean desired acceleration and the vehicular speed
σ	:	Parameters related to relation between the standard deviation of desired acceleration and the vehicular speed
$\Theta, \Phi, \tilde{\Theta}, \tilde{\Phi}$	:	Function
$\tilde{\sigma}, \tilde{t}, \tilde{w}, \tilde{N}, \tilde{\rho}_j, \tilde{v}_f, \tilde{q}, \tilde{\beta}, \tilde{x}$	:	Dimensionless parameters

Chapter 1

Introduction

1.1 Background and motivation

Due to the increasing number of vehicles on freeways, traffic congestion has been a daily problem in urbanized areas all over the word. Congestion causes traffic delays, decreased traffic safety, air and noise pollution. All of those adverse effects mean large financial losses. Hence, for decades, many efforts have been devoted to mitigating traffic jams and increasing the mobility on roads.

Dynamic traffic management (e.g., on-ramp metering, route guidance and dynamic speed limits), which is also called active traffic management, can be one of the most cost-efficient approaches to address the congestion problem. It is expected to be efficient and can have immediate effects on mitigating congestion. It aims to improve traffic network performance by controlling traffic flows, according to current and sometimes expected traffic conditions on roads. To achieve this goal, traffic flow theory, i.e., knowledge of fundamental freeway traffic flow characteristics (e.g., traffic density, time mean and space-mean speed) and relevant analytical approaches (e.g., shock wave analysis and traffic flow modeling) are essential requirements (May, 1990). The traffic flow theory is based on empirical observations on roads.

One of important phenomena for the magnitude of traffic delay is the capacity drop. It describes the fact that once congestion occurs, the queue discharge rate is generally substantially lower than the free-flow capacity. The queue discharge rate is the flow detected in the downstream of a traffic jam, which is referred to as the effective capacity in some papers. The capacity (also called pre-queue capacity or free-flow capacity) is the maximum flow that can be observed. In the last decade of the 20th century, the capacity drop phenomenon was empirically confirmed on freeways (Banks, 1991a; Hall & Agyemang-Duah, 1991). Since then, a large number of empirical observations reveal a wide range of the capacity drop magnitudes which indicate to what extent the queue discharge rate is reduced from the capacity.

The study of the capacity drop is relevant for improving freeway performance. As a

consequence of the capacity drop, the potential capacity of the road is not fully used. The magnitude of the capacity drop has a considerable influence on the total traffic delays. This is why many of the current approaches in traffic management aim to prevent or delay the capacity drop. Hence, studying the capacity drop can dramatically benefit developing traffic operation strategies for greater road mobility.

1.2 Freeway traffic flow fundamentals

This section briefly presents some fundamental concepts in the freeway traffic flow theory for understanding traffic dynamics. The general process of studying traffic dynamics in traffic flow theory is presented in Figure 1.1. Traffic data are the basis of traffic flow theory. The development of the data collection methods opens a door for a proper understanding on the traffic dynamics which are revealed in empirical observations. To better understand traffic dynamics and even further improve traffic performances, theoretical assumptions are usually proposed to explain the observed traffic phenomena. Note that convincing assumptions should originate from empirical observations (i.e., data). Finally, the assumptions are incorporated into traffic flow models. The models are run in simulations to support those assumptions by reproducing the traffic phenomena. Being able to reproduce traffic dynamics also indicates that the models can be used in practical applications. In sum, the data is the input while the supported assumptions and traffic flow models are outputs in the work for understanding traffic dynamics.

Before describing the study on the capacity drop in this thesis, this section briefly presents some relevant concepts during the process in understanding traffic dynamics. Those concepts will be used through the whole work in this thesis, directly or indirectly. This section consists of the introduction of traffic data (Section 1.2.1), traffic dynamics (Section 1.2.2) as well as traffic models (Section 1.2.3).

1.2.1 Traffic data

Traffic flow theory is based on empirical studies, which means traffic data are crucial. Generally, traffic data are categorized into cross-section data, fixed-section trajectory data, and floating-car trajectory data. This section describes these three types of data, which are analyzed in following chapters in this thesis.

Cross-sectional data is usually measured at a fixed cross-section on the road by sensors (Treiber & Kesting, 2013). The loop installed beneath the road surface is a most popular way of collecting cross-sectional data. Those data collected from loops are referred to as loop detector data in this thesis. In some countries (e.g., the Netherlands), loops are installed in pairs with a short and fixed distance (Hoogendoorn, 2007). With these double-loop measurements, speed, length, type, time headway and time gap of each

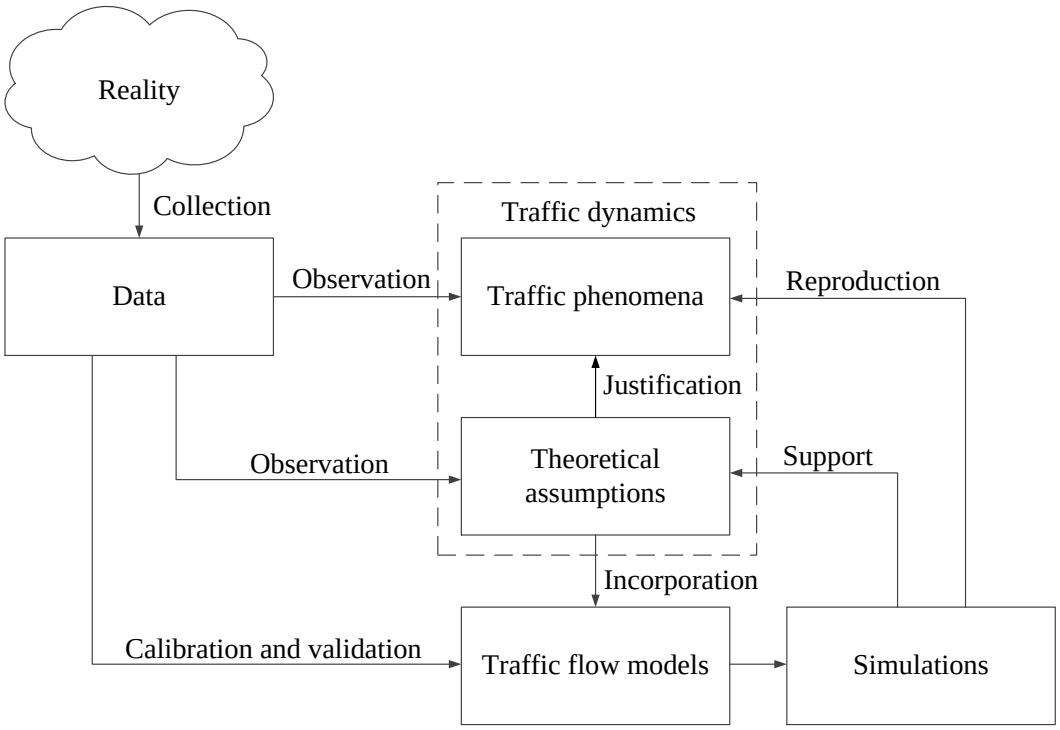


Figure 1.1: Process in understanding traffic dynamics.

vehicle can be determined. Frequently, those collected data are aggregated to give flow and mean speed by averaging the counts and vehicular speed over a fixed time period (e.g., 1 min or 5 min). This aggregation often gives the time-mean speed, which actually overestimates the instantaneous speed, and harmonic average of the collected speed as an approximation of the space-mean speed. To correctly compute traffic density, space-mean speed which is the average of all vehicular speeds on a road section, is preferred. The error between the time-mean speed and the space mean speed should not be ignored, especially in congested states where the error can be up to a factor of 4 (Knoop et al., 2009).

Vehicles can be directly observed by collecting fixed-section trajectory or floating-car trajectory data. As the name implies, trajectory data can show trajectories of all detected vehicles. Fixed-section trajectory data can be collected by installing cameras on a high infrastructure nearby a targeted road section or a helicopter above the road section. Installing dense loop detectors (e.g., every 100 m) along a road is also an alternative approach for collecting the fixed-section trajectory data. These data collections are usually quite expensive, which makes that fixed-section trajectory data are not so accessible as the loop detector data. Moreover, the road sections where fixed-section trajectory data are captured are usually much shorter than that where loops are installed. Generally, the fixed-section trajectory is called the trajectory data for simplicity. The other type of data, which also can give driving details of individual vehicle, is the floating-car trajectory data. These data are collected by using in-car devices (e.g., GPS or app in mobiles) to report vehicles' locations in a fixed time interval. Those ve-

hicles with in-car devices are called probe vehicles. Different from the trajectory data, which can give all vehicles' trajectories in detected road sections, the floating-car data generally only record information of a certain number of detected vehicles.

In this thesis, loop detector data are used to unveil features of the capacity drop phenomenon (see Chapter 2 and Chapter 3), floating-car trajectory data are used to analyze desired acceleration stochasticity of an individual driver in Chapter 6.

1.2.2 Traffic flow dynamics

This section briefly reviews some traffic flow dynamics which are important for the study on the capacity drop in this thesis. The traffic flow dynamic describes the interaction among driver-vehicle-combinations (DVC) on roads and the resulting freeway traffic states (Treiber & Kesting, 2013). The interaction generally means the motion and reaction of individual driver-vehicle unit to surroundings (e.g., acceleration or deceleration of its predecessor) during driving on freeways at a microscopic level. Its resulting traffic states include the traffic state evolution and the traffic phenomena description at macroscopic level. Treiber and Kesting (2013) argue that the time scale of traffic flow dynamics ranges from one second to a few hours, distinguishing from vehicle dynamics and transportation planning by the time scales.

The motion and interaction of individuals pertain to driver behavior. On freeways, the driver behaviors are categorized into: (a) longitudinal behaviors and (b) lateral behaviors. The longitudinal behavior are the movement on one lane along the road, including accelerating, decelerating and cruising process. The lateral behavior means the lane changing maneuver, the movement from one lane to an adjacent lane. For maneuvering or control tasks, Minderhoud (1999) categorizes the driver behavior discerning driving subtasks which consist of roadway subtask and vehicle interaction subtask. Each of the subtasks can be further decomposed in a longitudinal component and a lateral component, i.e., longitudinal roadway subtask, lateral roadway subtask, longitudinal vehicle interaction subtask and lateral vehicle interaction subtask. This thesis only categorizes driver behaviors into longitudinal and lateral behavior.

The study on the longitudinal behavior includes research on the interaction between two successive vehicles on a road, free-flow speed choice, and the acceleration towards this free speed. This interaction is described as a relation between the vehicular speed and the spacing between these two vehicles. The spacing means the distance from the follower to its predecessor. There exists a threshold spacing, which is called the critical spacing. If the spacing is larger than the critical spacing, vehicles drive at free-flow speed. Otherwise, the vehicular speed decreases as the spacing decreases. Cassidy (1998) argues that in nearly stationary traffic, this relation is a bivariate one. However, in reality drivers do not always behave following this relation. This phenomenon is usually described in terms of the instability. The instability of car-following traffic flow is due to the physical delay in adapting the speed to the actual traffic conditions

(Treiber & Kesting, 2013). The physical delay can be interpreted as a consequence of human errors (Yeo & Skabardonis, 2009; Laval et al., 2014).

The lateral behavior denotes the lane changing behavior on a multi-lane road section. Surveys show that the origins of lane changing decisions lie in four motivations: (1) speed leading; (2) speed leading with overtaking; (3) lane leading; and (4) traffic leading; see Keyvan-Ekbatani et al. (2015). Drivers proceed the lane changing in two steps: first anticipation process on the original lane and then relaxation process on the targeted lane (Zheng et al., 2013), which means the process of lane changing can influence the traffic condition on two lanes. In most of European countries , drivers follow a "Keep Right Unless Overtaking" rule, while in UK drivers keep left unless overtaking.

Traffic jams has two fronts i.e., downstream and upstream front. The distance between these two fronts determines the queue length. When two fronts meet, technically, it means congestion is eased. Generally, we can observe two kinds of congestion: (a) standing queues whose heads are fixed at one location (e.g., bottlenecks); and (b) stop-and-go waves which propagate upstream with both fronts moving. There are more different categorizations of congestion patterns in literature. For more information, we refer to Kerner (2004) and Treiber and Kesting (2013).

In most of empirical observations, the reason for the congestion formation is that demand (e.g., traffic flow) exceeds supply (e.g., road capacity). In most of empirical observations, it is a bottleneck that triggers traffic jams. The bottleneck includes the fixed bottleneck due to geographical change of road (e.g., lane drop and lane block) and moving bottlenecks (e.g., trucks and merging vehicles), which explains a mismatch between a demand (e.g., traffic flow) and a supply (e.g., road capacity). Another important reason for the traffic jam is the instability of traffic flow. Because of the absence of bottlenecks, the jam due to the instability is also named as "phantom jam" in some works.

To understand congestion characteristics, several relevant traffic phenomena, including hysteresis, oscillation and capacity drop, need to be considered. First, hysteresis are loops consisting of acceleration phases and deceleration phases in the fundamental diagram (e.g., flow-density relation). The acceleration phase means the transient states from congestion to free flow while the deceleration phase means that from free flow to congestion. Newell (1962) conjectures that the existence of two different congested branches is the reason for the hysteresis and the acceleration branch is below the deceleration branch. This conjecture is further discussed in Daganzo et al. (1999). Yeo and Skabardonis (2009) validate this conjecture with NGSIM data and explain it as a result of human errors. Zhang (1999) points out that the acceleration can be above the deceleration branch and a multi-loop hysteresis is also possible. Laval (2011) observes the clockwise and counter-clockwise hysteresis loops with new measurement and claim the aggressive and timid driving can explain these different loops. Deng and Zhang (2015) give the hysteresis in car-following models by incorporating relaxation and anticipation effects.

Second, oscillations refer to the phenomenon that vehicle movements in congested traffic states frequently present a cyclical pattern between slow and fast movements (Li & Ouyang, 2011). Lane changing can be a reason for the formulation of oscillation (Ahn & Cassidy, 2007; Laval & Daganzo, 2006). In the absence of lane changing, string instability is generally claimed to the justification for the oscillation. Laval and Leclercq (2010) as well as Chen et al. (2014) establish car-following model to reproduce the oscillation and the models are validated with the NGSIM dataset in Chen et al. (2012a) and Chen et al. (2012b). Laval et al. (2014) propose a parsimonious car-following model, as an extension of the Newell's first-order car-following model (Newell, 2002), to explain the string instability and give the oscillation as a result of the stochasticity of desired accelerations.

Third, capacity drop means the queue discharge rate is generally lower than the capacity. A detailed literature review on the capacity drop is given in Section 1.3. Understanding the hysteresis, oscillation and capacity drop is a crucial issue for traffic simulations and operations. For decades, researchers investigate those phenomena through analytical and empirical analysis. However, until now more empirical and analytical analysis are still required for understanding those phenomena.

1.2.3 Traffic flow models

Traffic modeling is an essential tool to computer the complex traffic dynamics in mathematical way for traffic flow analysis and simulations. In simulations, traffic models should have sufficient power to give relevant traffic phenomena. According to the level of aggregations, traffic models are generally categorized into macroscopic and microscopic traffic flow models.

Macroscopic traffic flow models

Macroscopic traffic flow models aim to describe traffic by computing the dynamics of aggregative variables, i.e., density, average speed and flow. Macroscopic models consists of continuum and discretized models. Generally, continuum models describe the dynamics of traffic as a continuum with partial differential equations while the discretized ones are usually solutions to those continuum models. Generally, macroscopic traffic flow model are suitable for model-based traffic management, especially in a large-scale network, due to its analytical properties. That is, solutions to continuum macroscopic models are usually formulated in a closed-form expression which requires only one run (Hoogendoorn & Bovy, 2001).

Continuum macroscopic model includes two relevant elements: a conservation law of vehicles and a fundamental diagram. The conservation law of vehicles means vehicles can not be created or disappear without entrances or exits. As argued in Papageorgiou (1998), the vehicular conservation law is the only 100% correct physical

law in traffic flow theory. The fundamental diagram indicates the relation among flow, density and speed in stationary traffic states. Without considering velocity dynamics, i.e., an equation describing the acceleration, the continuum macroscopic model is a first-order model. A most widely known example of the first-order model is the Lighthill-Whitham-Richards (LWR) model proposed by Lighthill and Whitham (1955) as well as Richards (1956)). Alternatively, a higher-order model (e.g., Payne's model by Payne, 1979) can be used.

Compared to first-order models, Papageorgiou (1998) argues that the second-order model has a higher accuracy. However, the so-called improved accuracy of the second order model comes at a cost of complexity (more model parameters) and assumptions. Daganzo (1995) once criticized the second order model for resulting in vehicles driving backward in 1995, following which some efforts are undertaken to address the problem (e.g., models by Aw & Rascle, 2000; Zhang, 2002). Lebacque et al. (2007) further generalize an invariant attached to drivers in the models proposed in Aw and Rascle (2000) and Zhang (2002) into the one attached to vehicles and dependent on traffic states. The generalization gives a whole family of second-order traffic flow models, i.e., Generic Second-Order Model (GSOM) family.

When numerically solving the continuum macroscopic traffic flow models, most of works propose or extend discretized macroscopic models with numerical approximations. Frequently, roads and time are divided into road cells and time instants, respectively. To ensure the stability, satisfying Courant-Friedrichs-Lewy (CFL) condition is required. That is, the length of each cell should be longer than the free-flow speed times the time step. Physically, it means no vehicles can pass one cell within one time step. A popular first-order discretized macroscopic model example is the Cell Transmission Model, proposed by Daganzo (1993, 1994) with a triangular fundamental flow-density diagram.

The first-order traffic flow model can be solved with respect to three 2-dimensional coordinate systems, i.e., Eulerian coordinates, Lagrangian coordinates and T coordinates (Laval & Leclercq, 2013). The categorization of the solutions depends on two of which continuous parameters (i.e., time t, location x and cumulative number of vehicles N) are discretized. Eulerian coordinates discrete the x and t, Lagrangian coordinates discrete the N and t, and T coordinates discrete the N and x. In each coordinate, those two discretized parameters are independent. The differentiation between any two of the continuous parameters gives representations of traffic flow. For example, in Lagrangian coordinates, speed and spacing are given to represent traffic flow.

Leclercq (2007) proposes a hybrid solution to the LWR model, which approximates the LWR model in the Lagrangian coordinates. Compared to the Eulerian expressions, the advantage of the Lagrangian description is a faster calibration on homogeneous road sections, while the disadvantage is more complicated source terms (Treiber & Kesting, 2013). In Lagrangian coordinates, traffic characteristics only move in one direction (upwind) to followers independent of traffic conditions. Lagrangian clusters only react to their leading Lagrangian clusters. By the contrast, the minimum supply demand

method, which is applied to solve most of LWR models (e.g., the Cell Transmission Model), switches between the upwind and downwind method, that depends on traffic states in the road cells (Van Wageningen-Kessels, 2013). Characteristics propagate in a switch of downstream (upwind) and upstream (downwind). Hence, compared to the Eulerian formulation, the Lagrangian one can more efficiently describe traffic dynamics. The knowledge of the macroscopic model will be used in this thesis for incorporating the capacity drop into a first-order model (see Chapter 3).

Microscopic traffic flow models

Microscopic traffic models present traffic dynamics by describing detailed behavior of each individual driver-vehicle unit. As described in Section 1.2.2, the behavior includes longitudinal and lateral behavior.

Models describing the longitudinal behavior is the so-called car-following model. A simple microscopic traffic flow model is the Newell's first-order car-following model (Newell, 2002). The Newell's car-following model describes the location of each vehicle at each time. However, the simplicity comes at a cost of accuracy. The hysteresis loops, oscillations and capacity drop cannot be given by the Newell's car-following model. Hence, some efforts are devoted to reproducing these phenomena in the framework of the Newell's first order car-following model, e.g., the L-L car-following model (Laval & Leclercq, 2010) and the parsimonious car-following model (Laval et al., 2014). There are also some other popular car-following models including the Optimal Velocity Model (Bando et al., 1995), the Intelligent Driver Model model (Treiber et al., 2000) and the Cellular automaton models (Kai, 1996). A more detailed review on car-following models can be found in Hoogendoorn and Bovy (2001) and Van Wageningen-Kessels (2013).

Regarding to the lane changing model, there are two branches: (1) modeling lane changing decisions and (2) modeling lane changing impacts. The lane changing decision is usually modeled as either probabilistic discretionary lane-change models or as deterministic ones. The popular theories include gap-acceptance theory and utility based theory. Examples on the lane changing decision model are Laval and Daganzo (2006); Schakel et al. (2012) and Roncoli et al. (2015). The model for giving impacts of lane changing focus on the interaction between the lane changing and car-following behaviors. Some researchers devote efforts to simulation models designed to characterize the behavior of lane changing (e.g., Schakel et al., 2012; Zheng et al., 2013). These simulation models give the impact of lane changing after running traffic flow system simulations. On the contrast, some other researchers propose analytical models (e.g., Leclercq et al., 2011, 2014, 2015, 2016) which directly give quantitative impacts of lane changing, such as the queue discharge rate calculation. More descriptions on lane changing models can be found in Zheng (2014).

1.3 Current knowledge on the capacity drop

This section provides a review of current knowledge on the capacity drop phenomenon. Section 1.3.1 presents a review on existing empirical findings on the capacity drop, which can offer a general view on the feature of the capacity drop. Since the capacity drop is essential for dynamic traffic management (see Section 1.1), and macroscopic traffic flow models are very suitable for large-network simulation as well as model-based operations (see Section 1.2.3), it is necessary to incorporate the capacity drop into macroscopic traffic flow models for traffic operation evaluations and traffic simulations. Therefore, Section 1.3.2 reviews literature on macroscopic traffic flow models which integrate the capacity drop. Capacity drop is a result of driver behaviors (e.g., lateral and longitudinal behaviors) on roads. Before proposing efficient operation strategies, understanding the driver behavioral mechanism behind the capacity drop is essential. Earlier literature has already proposed several hypotheses on the behavioral mechanism and integrated the hypothesis into microscopic traffic flow models, including car-following models, lane-changing models and analytical models. Section 1.3.3 gives a review on these hypotheses and models. Following the literature review on the capacity drop from Section 1.3.1 to Section 1.3.3, a current boundary of existing knowledge can be seen. More studies are needed to broaden current knowledge. Hence, several current challenges for properly understanding the capacity drop are proposed in Section 1.3.4.

1.3.1 Empirical studies on the capacity drop

The capacity drop phenomenon is revealed in empirical observations. Previous empirical research shows that the capacity drop magnitude ranges between 3% and 18% (Hall & Agyemang-Duah, 1991; Banks, 1991a; Cassidy & Bertini, 1999; Srivastava & Geroliminis, 2013; Chung et al., 2007; Cassidy & Rudjanakanoknad, 2005; Bertini & Leal, 2005; Oh & Yeo, 2012). In those studies, the capacity drop is reported based on data collected at active bottlenecks (e.g., on-ramp bottlenecks and lane-drop bottlenecks) on freeways. The active bottleneck means a bottleneck with a queue upstream and a free-flow state downstream. Cassidy and Bertini (1999) believe that the capacity drop magnitude at active bottlenecks (including an on-ramp and a lane-drop bottleneck) exhibit only small day to day deviations (i.e., smaller than 2%). However, actually in their observations, a wide range from 4% to 10% has been observed. They believe the difference is due to seasonal effects. Chung et al. (2007) find the capacity drop occur once the density near the bottleneck exceeds a threshold. Therefore, they suggest that a traffic control scheme regulating density can benefit averting the capacity drop. Kerner (2002) distinguishes capacity according to traffic phases (i.e., free flow, synchronized flow and wide moving jam). He argues that the queue discharge rate of synchronized flow changes in a wide range while the outflow of a wide moving jam is relatively stable. Srivastava and Geroliminis (2013) show that the change of capacity

drop can be related to on-ramp flows in empirical analysis, that is, the capacity drop increases as the on-ramp flow increases. In Oh and Yeo (2012), a negative relation between capacity drop and number of lanes, as well as a mitigation effect of an off-ramp are reported. By empirically analyzing NGSIM trajectory datasets, Oh and Yeo (2015) show that the contribution of traffic passing stop-and-go waves to discharge rates outperforms the contribution of lane change related traffic to discharge rates, and that it is found that the queue discharge rate could be related to the severity of congestion in the absence of lane changing.

In recent years, a popular approach to observe the capacity drop is to construct slanted cumulative curves at nearby upstream and downstream locations in the vicinity of a bottleneck (Cassidy & Bertini, 1999; Chung et al., 2007; Cassidy & Rudjanakanoknad, 2005; Bertini & Leal, 2005). Queuing diagrams and accumulation between detectors are applied for identifying the bottleneck activation. The slanted cumulative curves at the downstream nearby station is used for capacity and queue discharge rate measurements. A larger slope of a slanted cumulative curve indicates a larger flow. With slanted cumulative curves, it is intuitive that the slope of the slanted cumulative curves decreases when breakdown occurs at a bottleneck, which indicates a drop from the capacity to the queue discharge rate. However, this approach is criticized for possibly overestimating the capacity because the driver relaxation is not considered when identifying the bottleneck activation (Kim & Coifman, 2013). Kim and Coifman (2013) argue that the relaxation can temporally result in a supersaturated flow which is much higher than the free-flow capacity. Simulations in Kim and Coifman (2013) show that it is possible that the previous observed drop from the free-flow capacity to queue discharge rate using the slated cumulative curve is actually a drop from the supersaturated flow to the free-flow capacity. However, an empirical proof is absent.

1.3.2 Macroscopic models incorporating the capacity drop

Papageorgiou (1998) argues that incorporating the capacity drop into traffic flow models is crucial for the evaluation of dynamic traffic operations. Capacity drop is strongly related to the traffic delay calculation. Being unable to give the capacity drop, the traffic flow model cannot show benefits obtained from some traffic management strategies (e.g., on-ramp metering strategy) in simulations. Hence, this section reviews macroscopic traffic flow models which can give capacity drop.

Some earlier validations (e.g., Michalopoulos et al., 1992) suggest that a second-order model can be used to reproduce traffic dynamics. Nowadays researchers frequently use second order models to simulate the traffic state evolution on freeways (e.g., in Kotsialos et al. (2002) and Hegyi et al. (2005)). In Parzani and Buisson (2012), a second order junction model proposed by Haut and Bastin (2007) with the second order model by Aw and Rascle (2000) is tested to confirm the ability of generating capacity drop. In Hegyi (2004), the METANET model by Messmer and Papageorgiou (1990) is extended for giving better reproduction of the capacity drop. However, drawbacks of

higher-order models, including the high complexity and the absence of clear physical meanings of parameters, encourage contributions to incorporating the capacity drop into first-order models.

Some literature has tried some approaches to capture the capacity drop in the kinematic wave model, e.g., a discontinuous fundamental diagram proposed by Lu et al. (2009). However, firstly the discontinuous fundamental diagram will result in infinite shock wave speed and characteristic wave speed, as argued in Jin et al. (2015). Secondly, the discontinuity obtained from empirical data could be a result of non-stationary traffic (Cassidy, 1998). Therefore, many works argue that it is best to use a continuous fundamental diagram rather than a discontinuous one.

Following this, some efforts have been made, with a continuous fundamental diagram, to capture capacity drop in the kinematic wave model. Generally, those efforts are to revise the demand or supply function of cells in the vicinity of congestion. Maria et al. (2014) and Jin et al. (2015) reduce the supply of the immediate downstream cell of congestion by introducing an auxiliary variable illustrating the maximum flow for each cell. Srivastava and Geroliminis (2013) assume two values of capacity. If the cell is congested during the last time step, the lower-value capacity is used to restrict the supply; or the other higher value is the maximum supply. Alvarez-Icaza and Islas (2013) apply a hysteresis cycle to decide the supply function which is related to the wave speed.

Some works take account of bounded acceleration effects for giving capacity drop. Lebacque (2002, 2003a) show that when the density exceeds the critical density, the bounded acceleration effects modify the constant demand to a negative function of density. Combining the LWR-BA model, a bounded acceleration Lighthill-Whitham-Richards (LWR) model by Lebacque (2002, 2003a), and a node model proposed by Lebacque (2003b), can give capacity drop as a result of bounded traffic acceleration in the LWR framework (Khoshyaran & Lebacque, 2015; Monamy et al., 2012).

Some other works consider lane changing to give capacity drop. Muralidharan and Horowitz (2015) decrease the supply function at merge cells with a weaving parameter. The weaving parameter illustrates that the merging behavior from the on-ramp occupies more spacing than vehicles from the upstream cell. This concept is similar to the empirical observation in Srivastava and Geroliminis (2013).

Moreover, some contributions combine the bounded acceleration and lane changing effects for reproducing the capacity drop (e.g., Roncoli et al., 2015; Srivastava & Jin, 2016; Kontorinaki et al., 2016). Those works let the demand decrease as the over-critical density increases. However, they deal with lane changing effects in different ways. Roncoli et al. (2015) present a multi-lane model, giving both longitudinal and lateral behavior and dividing a multi-lane freeway into different groups of segment-lanes. The flow into and out of each cell in each lane is computed. Srivastava and Jin (2016) use a concept of perceived density which is higher than the actual density, that originates from an assumption that the lane changing contribute to the density on both

original and target lanes. Kontorinaki et al. (2016) give more capacity to merge cells by modifying the supply function to allow a density at the merge cells higher than the critical density.

1.3.3 Hypothesis on and modeling of driver behavioral mechanism

Many studies have investigated the capacity drop from a driver behavioral perspective in the past decades. Generally, existing proposed hypotheses about the mechanism behind the capacity drop can be divided into four categories: bounded acceleration capability, driver-vehicle-combination (DVC) heterogeneity, inter-driver spread, and intra-driver variation. The hypotheses on the driver behavior mechanism behind the capacity drop are summarized in Table 1.1.

The bounded acceleration capability means vehicles cannot accelerate instantaneously. Consequently lane change maneuvers can create voids in the traffic stream (Laval & Daganzo, 2006; Yeo, 2008; Duret et al., 2010; Leclercq et al., 2011, 2015). The limited acceleration causes that the lane changing vehicle cannot catch up with its new predecessor. Coifman and Kim (2011) show that lane changing in the far downstream of the congestion can result in the capacity drop, too. Insertions result in shock waves in the new lane and the divergences in the old lane create voids which cannot be filled in duo to the bounded acceleration capability. So an aggregated flow detected in the downstream of queue could be lower than the capacity. However, Oh and Yeo (2015), using NGSIM data collected from US101, find that the reduction due to stop-and-go waves in the absence of lane changing is much larger than that due to lane changing. Without lane changing, the queue discharge rate reduction is of dependence on the congestion severity (Oh & Yeo, 2015), which is related to Chapter 2.

In the absence of lane changing, driver-vehicle-combination (DVC) heterogeneity possibly can be a cause of the queue discharge rate reduction. The DVC heterogeneity indicates the variance of fundamental (e.g., flow-density) relations in homogeneous traffic states. Wong and Wong (2002) state that the variance of drivers' desires for free-flow speeds are the reason for the capacity drop. Because an extended LWR model, incorporating a distribution of heterogeneous DVC characterized by their choice of speeds in a traffic stream, replicates a reverse-lambda shape of fundamental diagram in simulations. However, a behavioral explanation on how the heterogeneity reduces the queue discharge rate is lacking. They also mention that their study is not complete due to an absence of direct observations. Chen et al. (2016) study the effect of heavy vehicles on driver behaviors through empirical analyses. The heavy vehicles are those with heavier weight and larger physical dimensions than passenger ones. It is found that there is a possibility that the heavy vehicles can benefit reducing the capacity drop by stabilizing surrounding traffic streams. They also comment that more confirmations on their findings are needed in the future. Some works, e.g., Coifman (2015), show the impacts of the DVC heterogeneity on the fundamental diagram, but no conclusions on the capacity drop is drawn. Hence, to what extent the DVC heterogeneity contributes

to the capacity drop and a behavioral explanation are not clear yet, which requires more empirical studies.

The inter-diver spread means heterogeneity among driver-vehicle-combinations with a same fundamental diagram. More specifically, the inter-driver spread in this work means the desired acceleration heterogeneity among vehicles. The desired acceleration means the acceleration driver intends to impose to a vehicle when in an absence of a predecessor. Some drivers intend to speed up with a large acceleration while some others prefer a relatively low acceleration. Papageorgiou et al. (2008) state that the capacity drop is due to the acceleration difference between two successive vehicles. Voids can be created between the low-acceleration vehicle and its high-acceleration predecessor. In this thesis, the inter-driver spread is also referred to as acceleration spread.

The fourth popular explanation, i.e., the intra-driver variation, assumes driver behavior varies depending on traffic conditions. For example, one driver's reaction time, desired time headway or desired gap-time during accelerating progress differs from that in decelerating process. The intra-driver variation mechanism means that the capacity drop is a result of traffic condition dependent variable driver behaviors.

Minderhoud (1999) proposes a hypothesis that gap distances of drivers increase after having experienced congested conditions. Similarly, Treiber et al. (2006) assume drivers would choose a longer time headway in congestion than that in free flow. The preferred time headway in congestion increases as density increases. This assumption, which is also called variance-driven time headways, is based on an empirical observation in Nishinari et al. (2003). The observation shows an increasing time gap between one vehicle's front bumper and the rear bumper of the preceding vehicle after a considerable queuing time. Zhang and Kim (2005) propose a multi-phase car-following traffic flow theory to reproduce the capacity drop. They highlight that the capacity drop is a result of driver behavior spread across three phases, i.e., acceleration, deceleration and coasting. Wu and Liu (2013) validate the acceleration and deceleration curves and further develop the asymmetric microscopic traffic flow theory based on empirical data in an urban environment, suggesting to explain the capacity drop as a difference of the maximum flow between the acceleration and the deceleration curve in density-flow fundamental diagram. The distinction between the acceleration and deceleration curve has been earlier discussed in Newell (1962) and Daganzo et al. (1999). Tampère (2004) gives capacity drop in simulations as a result of intra-driver variability - modeled as a temporary, traffic condition dependent variable "activation level". The low activation level used to accounted for a loss of motivation. Two behavioral assumptions about the activation level are made in Tampère et al. (2005). That is, desired time gaps are inversely related to the activation level, or alternatively the loss of motivation in low speed condition is related to a reduction of a maximum acceleration. Chen et al. (2014) explain the capacity drop as a result of the change of aggressiveness, that is, the intra-driver behavior variation is described as four different reaction (i.e., concave, convex, non-decreasing and constant) patterns to disturbances. Those different reac-

Table 1.1: Possible mechanisms of the capacity drop

Basis mechanisms	Assumed mechanisms	References
(a) Bounded acceleration capability	Lane changing	Laval and Daganzo (2006) Duret et al. (2010) Leclercq et al. (2011) Leclercq et al. (2015) Coifman and Kim (2011)
(b) DVC heterogeneity	Multi-class vehicles	Wong and Wong (2002)
(c) Inter-driver/vehicle spread	Acceleration spread	Papageorgiou et al. (2008)
(d) Intra-driver variation	Variance-driven time headways	Treiber et al. (2006)
	Multiphase car-following theory	Zhang and Kim (2005)
	Asymmetric driving behavior theory	Yeo (2008)
	Activation level	Tampère (2004)
	Reaction pattern	Chen et al. (2014)

tion patterns, which are validated with NGSIM dataset in Chen et al. (2012a, 2012b), can give capacity drop ranging from 8% to 23% in simulations.

1.3.4 Challenges for understanding the capacity drop

While earlier works have already made important achievements in investigating the capacity drop, there are still several tough challenges for deeply understanding the capacity drop. Those challenges include (1) characterizing more empirical features of the capacity drop, (2) incorporating the capacity drop into macroscopic models, (3) revealing the mechanism related to driver behaviors behind the capacity drop and incorporating the mechanism into microscopic models.

The first challenge is to broaden the current knowledge on the capacity drop phenomenon with more empirical observations. The existing knowledge shows that the reduction from the capacity to the queue discharge rate can be in a wide range. Some works (e.g., Srivastava & Geroliminis, 2013; Oh & Yeo, 2012) suggest that the capacity drop magnitude can differ under influences of several factors, such as the characteristics of the study site (e.g., number of lanes, traffic flow composition) and different operations. So it is probable that control strategies can increase discharge rates to evacuate vehicles in queues quickly and finally reduce traffic delays. Hence, principles illustrating a rational reduction in a specific traffic condition are required. A solid principle should be established in empirical analyses. However, few empirical analysis has contributed to revealing the dependence of queue discharge rates.

Macroscopic models are efficient tools for evaluating network-wide operations, espe-

cially the model-based dynamic traffic management (Hoogendoorn & Bovy, 2001). It is crucial to incorporate the capacity drop into the macroscopic model (Papageorgiou, 1998). Due to drawbacks of high-order models as described in Section 1.3.2, the incorporation of the capacity drop into the first-order model is required. Without the study on the dependence of the capacity drop, it is difficult to precisely describe the capacity drop with macroscopic models. Some of current works even handle the discharge rate as a constant variable in models, that means after calibrations the capacity drop magnitude in simulations is always constant. It is notable that a constant capacity drop magnitude means completely avoiding congestion or dissolving congestion entirely as soon as possible is the only way to decrease the capacity drop magnitude, which differs from the empirical results in Srivastava and Geroliminis (2013) and Oh and Yeo (2012). Hence, correctly reproducing the capacity drop in a first-order traffic flow model is relevant, and this second challenge is addressed in this thesis (see Chapter 3).

The third challenge is to identify significant driver behaviors for the capacity drop. As presented in Section 1.3.3, while lane changing is believed to be a possible behavior accounting for the capacity drop (Laval & Daganzo, 2006; Yeo, 2008; Duret et al., 2010; Coifman & Kim, 2011; Leclercq et al., 2011, 2015, 2016), a recent observation reveals that the capacity also drops in the absence of lane changing (Oh & Yeo, 2015). After a comparison between effects of lateral and longitudinal behaviors, it is found the queue discharge rate reduction in the absence of lane changing greatly exceeds the reduction due to the lane changing (Oh & Yeo, 2015). Hence, it is important to explore the longitudinal behavior mechanism for the formation of voids as well as the dependence of the capacity drop. If the influence of vehicle class heterogeneity is excluded, there are two longitudinal mechanisms giving the capacity drop. That is, the inter-driver spread (Papageorgiou et al., 2008) and the intra-driver variation (Tampère, 2004; Zhang & Kim, 2005; Treiber et al., 2006; Chen et al., 2014). This thesis aims to identify which longitudinal mechanism is relevant for giving the capacity drop, as well as incorporate the mechanism into a car-following model.

1.4 Research objectives and questions

This thesis aims to provide new insights into the capacity drop, including its relevant features (Objective 1) and driver behavioral explanations of the capacity drop (Objective 2). Meanwhile, the work in this thesis aims to describe these insights mathematically (Objective 3). More specifically, the objective 1 is to unveil the dependence of the capacity drop magnitude, understanding the queue discharge rate reduction by looking at congestion characteristics. The objective 2 is figure out what driver behavior results in the queue discharge reduction. The objective 3 is to mathematically describe the capacity drop and its dependence (i.e., a macroscopic traffic flow model), as well as the hypothesized driver behavior mechanism behind the capacity drop (i.e., a microscopic traffic flow model).

It should be noticed that this research not only focus on bottleneck queue discharge rates, but also stop-and-go wave discharge rates. This book refer the free flow detected downstream of congestion to as queue discharge rate.

To reach the main research objectives, four research questions have to be answered:

> **Research question 1:** What is the relation between the queue discharge rate reduction and the congestion characteristics (e.g., speeds, densities)? (Objective 1)
>
> **Research question 2:** What is the relevant driver behavioral mechanism behind the capacity drop? (Objective 2)
>
> **Research question 3:** How to incorporate the capacity drop into a first-order macroscopic traffic flow model? (Objective 3)
>
> **Research question 4:** How to incorporate the relevant driver behavioral mechanism of the capacity drop into a microscopic simulation model? (Objective 3)

Objective 1 focuses on understanding the mechanism behind the wide range of capacity drop magnitudes. Answering research question 1 aims to investigate what is the relation between the congested state and the capacity drop, that can show quantitative understanding on the queue discharge rate reduction. Research question 2 is about an explanation of the capacity drop from a driver behavioral perspective, corresponding to the objective 2. After finding out relevant features of and mechanisms behind the capacity drop, this thesis tries to describe them mathematically (which is the objective 3). Research question 3 investigates how to incorporate the capacity drop and its features into a macroscopic model, and research question 4 wonders how to give a microscopic simulation model integrating the behavioral mechanism behind the capacity drop phenomenon. The macroscopic model can be used for traffic operation evaluations and model-based traffic management. The microscopic model is used to support the hypothesized behavioral explanation of the capacity drop, following the process presented in Figure 1.1.

1.5 Main research contributions

This section summarizes the main contributions in this thesis. The contributions are categorized into those are scientific nature in Section 1.5.1 and those are practical nature in Section 1.5.2.

1.5.1 Scientific contributions

The main scientific contributions are the following:

1. *Characterizing the difference on the queue discharge rate between the standing queue and stop-and-go waves* (Chapter 2). This thesis compares the queue discharge rate of standing queues and stop-and-go waves. It is found that the queue discharge rate of standing queue is much higher than that of the stop-and-go waves. The difference of the queue discharge rate indicates that even with less lane changing behaviors there could be lower queue discharge rate.

2. *Presenting and justifying traffic flow distribution across lanes around a lane-drop bottleneck* (Chapter 2). This thesis shows the flow distribution as a function of average density across lanes. The observation is conducted at a lane-drop bottleneck where the number of lanes decreases from four to three, see Figure 2.3. It is found that particularly near head of a standing queue more vehicles can merge into the lane adjacent to the ending lane, thereby locally increasing the capacity of that lane. The capacity of the shoulder lane is markedly wasted when in congestion. The reason for the low flow distribution in congestion on the shoulder lane is the large spacing between successive vehicles.

3. *Observing the queue discharge rate reduction on each individual lane* (Chapter 2). Observations show that the capacity drop exist on each individual lane. The difference on the queue discharge rate between the standing queue and the stop-and-go wave can also be observed on individual lanes. Please note that there are lane changing behaviors on multi-lane freeways. Please note that there are lane changing behaviors on multi-lane freeways. The merging and diverging flow have been subtracted from and added into the outflow on each lane, respectively.

4. *Presenting a new methodology of confirming the capacity drop qualitatively and observing the queue discharge rate variance quantitatively* (Chapter 2). This thesis provides an approach for observing a wide range of queue discharge rates at a same freeway bottleneck. Because the capacity is defined to be a fixed maximum free flow, and that the maximum observed queue discharge rate is no higher than the capacity, the other observed queue discharge rates, which are lower than the maximum queue discharge rate, will definitely be lower than the capacity. That is, this work qualitatively confirms the queue discharge rate reduction regardless of the capacity estimation.

5. *Identifying quantitatively a relation between the congestion characteristic and the queue discharge rate* (Chapter 3). For the first time, a relation between the speed in congestion and the queue discharge rate is given through empirical analysis. This relation explains the difference on the queue discharge rate between the standing queue and the stop-and-go wave.

6. *Proposing a new kinematic wave model incorporating the capacity drop* (Chapter 4). The new model incorporates the queue discharge rate reduction and the

dependence of the capacity drop on the congestion characteristic. A more accurate capacity drop magnitude can be given in the first-order kinematic wave model.

7. *Providing a proof showing the inter-driver spread slightly reduce the queue discharge rate* (Chapter 5). This thesis finds that the inter-driver spread can only slightly reduce the queue discharge rate. No matter the reaction time is considered or not, the reduction from the capacity due to the inter-driver spread is not comparable to that due to the intra-driver variation.

8. *Identifying a crucial longitudinal driver behavior for giving the capacity drop* (Chapter 6). Through empirical analysis and numerical simulations, this thesis finds that the stochastic feature, including mean and standard deviation, of the desired acceleration at different vehicular speeds is a relevant explanation of the capacity drop.

9. *Developing a new car-following model incorporating the stochastic features of the desired acceleration* (Chapter 6). The thesis extends the parsimonious car-following model (Laval et al., 2014) by incorporating the stochastic feature of desired accelerations. The extended parsimonious car-following model gives the capacity drop following previous observations, verifying that the stochastic desired acceleration is relevant for giving the capacity drop.

1.5.2 Practical contributions

The main practical contribution of this thesis consists of traffic operations and simulations.

1. *Indicating the possibility of resolving congestion by increasing the queue discharge rate* (Chapter 2 and Chapter 3). The empirical findings in this thesis show that the queue discharge rate is strongly related to the congested states. That is, the queue discharge rate can be operated. The work in this thesis indicates that it is possible to propose a dynamic traffic management strategy to maximize bottleneck throughput by increasing the queue discharge rate.

2. *Denoting a possible direction for increasing the bottleneck throughput with advanced technology* (Chapter 6). This thesis finds that the intra-driver desired acceleration deviation is a relevant factor for the queue discharge rate reduction. It denotes that a development in vehicular engines for automatically modifying acceleration (i.e, minimize human errors) could facility maximizing queue discharge rates.

3. *Providing efficient simulation models for evaluating traffic operation strategies* (Chapter 4 and Chapter 6). Both macroscopic and microscopic traffic flow models presented in this these can reproduce all features of the capacity drop. It

means they have sufficient power for evaluating past and promising future traffic operation strategies in a precise way.

1.6 Outline of thesis

The remainder of this thesis is structured as shown in Figure 1.2.

Chapter 2 and Chapter 3 contribute to exploring more relevant macroscopic features of the capacity drop. Chapter 2 investigates the main characteristics of the capacity drop, especially and the difference on the queue discharge rate between the standing queue and stop-and-go waves, based on loop detector data. It is found that the queue discharge rate of stop-and-go waves are greatly lower than that of standing queues. The observations show that the difference on the queue discharge rate between the standing queue and the stop-and-go wave can also be observed on individual lane. The different queue discharge rate distribution across lanes could be due to the different average spacing rather than the speed difference among lanes. Following the findings in the Chapter 2, Chapter 3 focuses on the dependence on the characteristic (e.g., speed) in congestion. Through empirical analysis, a relation between the speed in congestion and the queue discharge rate is concluded. The queue discharge rate will increase if the speed in congestion increases. Chapter 4 then gives a kinematic wave model incorporating the capacity drop and its dependence. For efficiency, the model is formulated in the Lagrangian coordinates. The new model can give a wide range of capacity drop magnitude, following the empirical relation revealed in Chapter 3. In Chapter 5, an analytical investigation on the longitudinal mechanism behind the capacity drop shows that intra-driver variance is crucial for giving the queue discharge rate reduction. Chapter 6 then models the intra-driver variance by integrating drivers' desired acceleration stochasticity into a car-following model in the Newell's first order car-following framework, after empirically observing the stochasticity through floating car data. It is concluded that the stochastic features of the desired acceleration is a quite relevant factor determining the queue discharge rate reduction. Finally, the conclusions and discussions of this thesis, as well as its prospects for future work are summarized in the final Chapter 7.

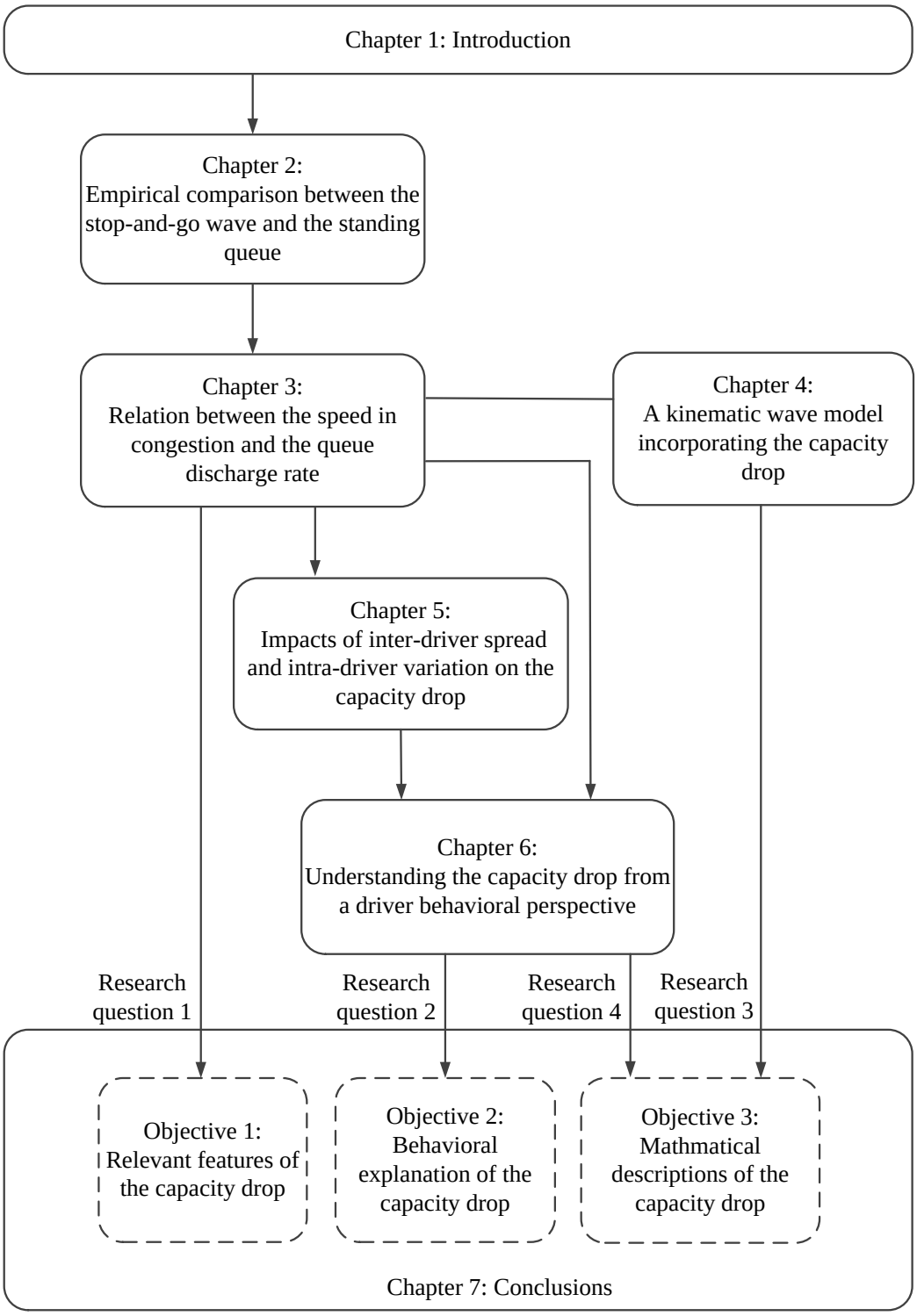


Figure 1.2: Outline of the thesis.

Chapter 2

Empirical comparison between stop-and-go wave and standing queue

In this chapter, a comparison on the queue discharge rate between the stop-and-go waves and standing queues is made. The difference in the queue discharge rate indicates whether the queue discharge rate is related to congestion category. It is found that the queue discharge rate of stop-and-go waves is considerably lower than that of standing queues, which can be observed on every individual lane. This thesis argues that the dependence of the queue discharge rate on the congestion category can be further understood as a dependence on the speed in congestion, which is discussed in Chapter 3.

This chapter is an edited version of the article: Yuan, K., Knoop, V.L., Leclercq, L., & Hoogendoorn, S.P. (2016). Capacity drop: a comparison between stop-and-go wave and standing queue at lane-drop bottleneck. *Transportmetrica B: Transport Dynamics*. DOI: 10.1080/21680566.2016.1245163

2.1 Introduction

Even though a large amount of research effort has been put into studying the capacity drop, some significant macroscopic features on the capacity drop are still unclear. For example, it is not clear to what extent the capacity reduces when different congestion occurs upstream. Moreover, it is not clear what is the amount of traffic on each lane (flow distribution over lanes), especially at the downstream of a bottleneck with compulsory merging behaviors upstream. Hence, this chapter tries to show more empirical observations to forward traffic research to reveal more empirical features. These findings can contribute to a better understanding of the traffic processes, possibly leading to control principles mitigating congestion. Moreover, it also gives an indication of the lane change behavior at the bottleneck locations.

The aim of this chapter is to figure out the difference between traffic states downstream of stop-and-go waves and that downstream of standing queues at the same site. To achieve this goal, following four subquestions need to be addressed.

First, to what extent is the queue discharge rate downstream of a stop-and-go wave reduced? Most of previous research observe capacity drop phenomenon at active bottlenecks. Few of those studies reveals features of capacity drop downstream of a stop-and-go wave. Kerner (2002) observes that the outflow of wide moving jam can be higher than minimum outflow of synchronized flow, and lower than the maximum outflow of synchronized flow. This thesis categorizes congestion into stop-and-go waves and standing queues, showing present empirical observations of capacity drop in stop-and-go waves.

Second, to what extent does the outflow of congestion vary at the same road section without other disturbances such as weather and road layouts? In short, this subquestion hence discusses the variation of the outflow of the queue. Previous research shows that discharging flows of standing queues at one bottleneck only exhibit small deviations (Cassidy & Bertini, 1999). But those research only target standing queue at an active bottleneck. In contrast to the standing queue, whose traffic states are limited in a narrow range because the road layout dictates the congested traffic state upstream, different stop-and-go waves can result in different congestion states. The study of stop-and-go waves can enlarge observation samples.

Third, what is the flow on each lane in queue discharging conditions? This might shed light to the capacity drop as well.

Four, what is the traffic flow distribution over lanes downstream of an bottleneck with compulsory merging behaviors upstream, especially locations near bottlenecks? The study of the flow distribution can show the utilization of lanes when the capacity drop is observed, which can benefit increasing queue discharge rates with multi-lane dynamic management.

To answer those questions, this chapter studies a traffic scenario where a standing queue forms immediately after a stop-and-go wave passes. It seems that the standing

queue is induced by the stop-and-go wave. In this scenario, there can be at least two congestion states and two outflow states observed at the same road section at the same day.

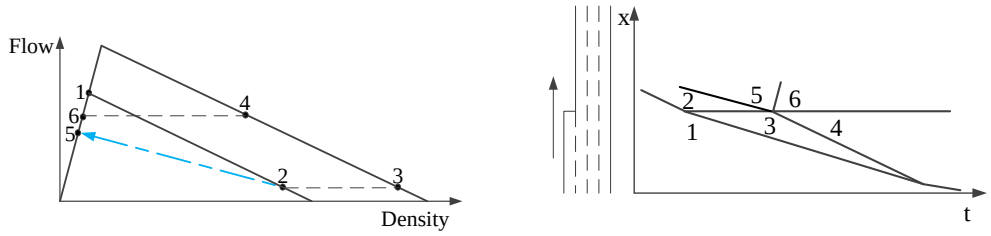
The remainder of the chapter is set up as follows. Section 2.1 introduces research motivations and research questions. The methodology used to address the questions are described in Section 2.2. Section 2.3 introduces the data analyzed in this chapter. In Section 2.4, empirical observations are presented, including various traffic states and flow distribution in each lane. Finally, this chapter ends with conclusions in Section 2.5.

2.2 Methodology

This chapter targets a homogeneous freeway section with a lane-drop bottleneck upstream. In the expected scenario, a standing queue forms immediately after the passing of a stop-and-go wave. It seems like the bottleneck is activated by the stop-and-go wave. In this way, we can compare the outflows of congestion at that location and possible location specific influences are excluded from the analysis.

Since the differences in the capacity drop (in standing queues) between any two days at the same bottleneck lies in a small range among days (Cassidy & Bertini, 1999), it is difficult to observe standing queues in distinctly different congestion states at the same bottleneck. However, the congestion level in stop-and-go wave is considerably different from the congestion in a standing queue. Congestion level is represented by vehicle speed in the congestion and density. Previous research (Laval & Daganzo, 2006; Chung et al., 2007) shows that the capacity drop is strongly related to the congestion level, hence it is expected that downstream of a stop-and-go wave traffic states differ from that downstream of a standing queue. In this way, several state points at the same road stretch can be observed empirically, including free flow and congestion states. Shock wave analysis is applied to identify those congestion states qualitatively.

By comparing the outflows downstream of congestion, this chapter shows the capacity drop corresponding to the two different congestion types, stop-and-go wave and standing queue. The key of the traffic state analysis is to identify those traffic states. To avoid unnecessary deviations, this chapter applies slanted cumulative counts to calculate flow. The slanted cumulative curve, also known as oblique cumulative curves, is drawn by subtracting a reference flow from the cumulative number of passing vehicles. The slanted cumulative curve can promote the visual identification of changing flows (Cassidy & Bertini, 1999). Both of these two outflows are flow detected downstream of the congestion. There are repetitive observations. For the duration of congestion until the congestion is dissolved, there are no other influences from downstream. The outflow of a stop-and-go wave can be detected at some location where the speed returns to the free flow speed after the break down phenomenon, and the discharging flow can be detected at each location downstream of an active bottleneck.



(a) Fundamental diagram (b) x-t plot

Figure 2.1: Shock wave analysis on one traffic scenario at a lane-drop bottleneck.

2.2.1 Shock wave analysis

The states which occur are determined using shock wave analysis. Figure 2.1 shows the resulting traffic states, including the regions in space-time where the outflows can be measured. For the sake of simplicity, we choose triangular fundamental diagrams, Figure 2.1(a) shows these fundamental diagrams, the smaller one for three-lane section and the larger one for the four-lane section. The outflow of a stop-and-go wave, shown as state 5, and discharging flow of standing queue, shown as state 6, both lie in the free flow branch, see Figure 2.1. The flows in both of these two states are lower than the capacity shown as state 1 to represent the capacity drop. A stop-and-go wave, state 2 in Figure 1, propagates upstream to the bottleneck and this triggers a standing queue, state 4. Figure 2.1(b) shows that once the bottleneck has been activated both of state 5 and 6 can be observed in the downstream of the bottleneck. The further away from the bottleneck, the longer time state 5 can be observed. Note that because state 5 and 6 are always located in the free flow branch, the shock wave between these two states are always a positive line parallel to the free flow branch. Therefore, in Figure 2.1(b) the shock wave between state 5 and 6 are always the same in x-t plot, no matter which state shows a higher flow. All those states are predicted theoretically by shock wave analysis, which should be observed in empirical observations.

Hence, for measuring the outflow, observations at locations far away from the bottleneck are preferred. In that case, the outflow of stop-and-go wave can be measured for a long enough time and compared clearly there.

With the same methodology, different outflow features in different lanes are analyzed. This shows the performance of each lane during the transition from outflow of stop-and-go wave to queue discharging flow. This chapter applies slanted cumulative counts to calculate the outflow in each lane. Note that in the Netherlands the rule is Keep Right Unless Overtaking. This asymmetric rule might lead to a different lane choice, for instance for slugs and rabbits (Daganzo, 2002), as well as leading to different traffic operations.

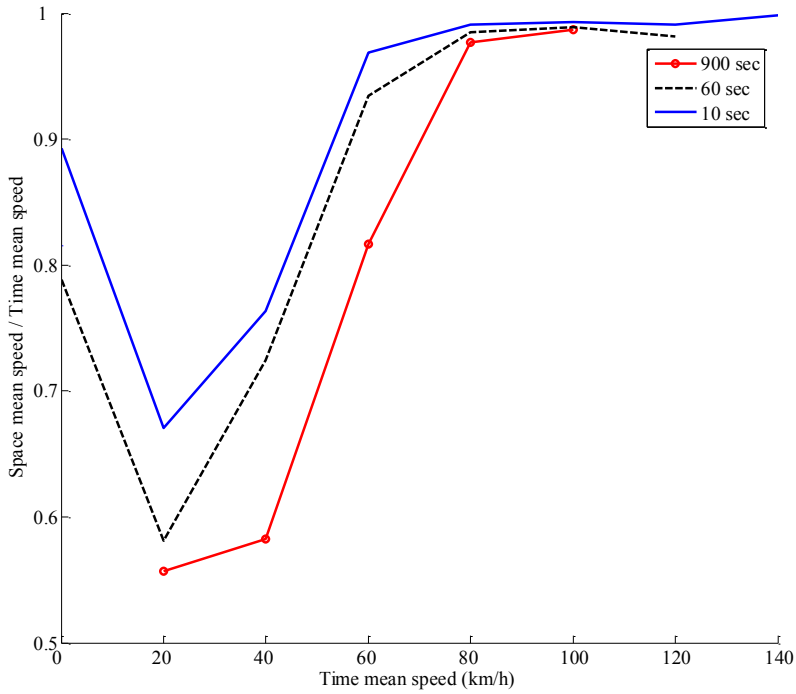


Figure 2.2: The impact of difference between time mean speed harmonic mean speed: 10 seconds aggregation (blue line), 60 seconds aggregation (black dashed line) and 900 seconds aggregation (red line with circles). (From Knoop et al. (2009)).

2.2.2 Data handling

This chapter reveals the flow distribution in each lane as a function of average density over lanes in Section 2.4.4. The density (ρ) which is estimated through dividing flow (q) by space mean speed (v_s) is necessary.

In the Netherlands, loop detector data is time mean speed (v_t) and flow (q). Knoop et al. (2009) point out the substantial difference between the time mean speed (v_t) and space mean speed (v_s), especially when the speed in congestion. Yuan et al. (2010) present a correction algorithm based on flow-density relations to calculate space mean speed. This method requires that traffic states should lie on the linear congested branch of the fundamental diagram. However, this chapter considers acceleration states downstream a bottleneck, so an another method is needed.

Knoop et al. (2009) show an empirical relation between space mean speed and time mean speed, see Figure 2.2. The space mean speed actually is estimated as harmonic speed. This relation is applied to space mean speed calculation in Ou (2011). This chapter also applies the relation to calculate the space mean speed and the density.

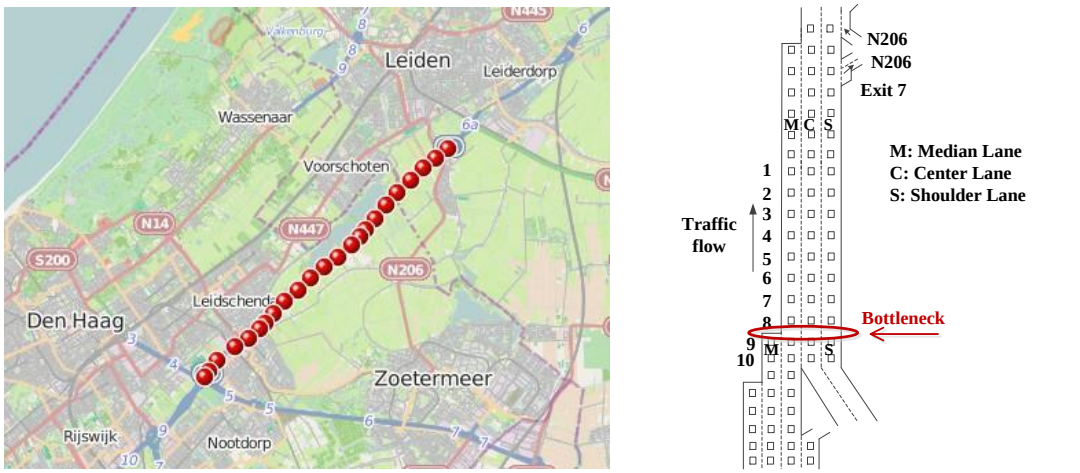


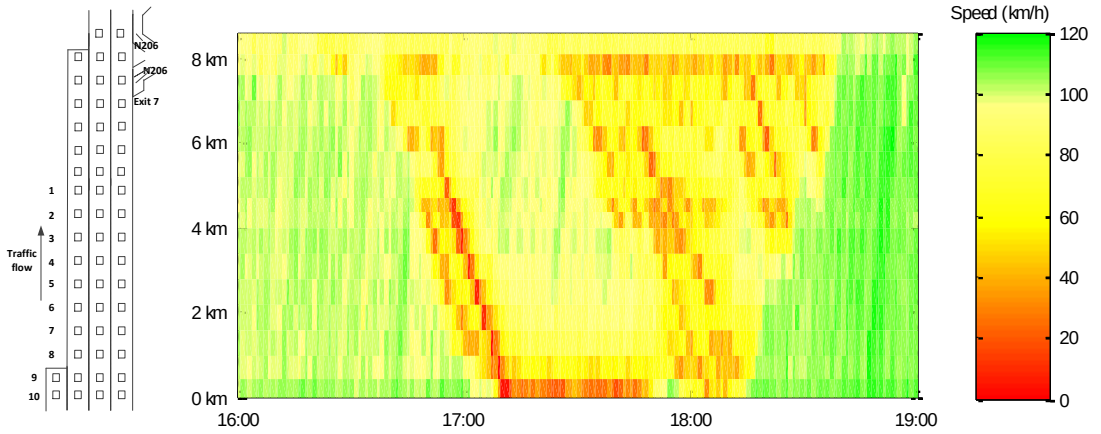
Figure 2.3: Open street figure of targeted section in freeway A4 in the Netherlands (left) shown in red dots and the layout of the study site (right). The bottleneck is a lane-drop bottleneck highlighted with a red circle. This chapter only targets 10 locations. The total distance from the location 1 to location 10 in the freeway is approximately 4.5 km. The bottleneck is around 6.5 km away from the downstream off-ramp.

2.3 Data

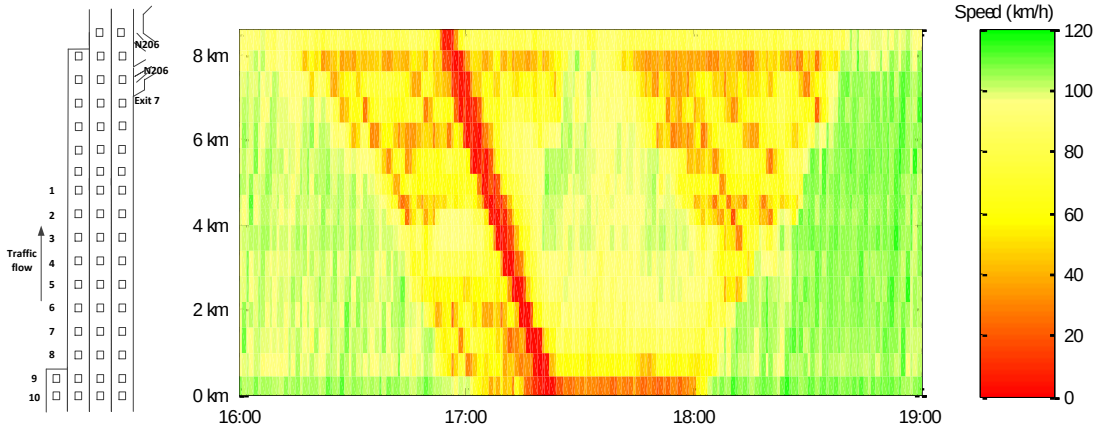
The data analyzed is one minute aggregated, collected around a lane-drop bottleneck on the freeway A4 in the Netherlands. This chapter considers the northbound direction just around Exit 8 (The Hague) in A4 shown in Figure 2.3. The layout of the study site is shown in the right part of Figure 2.3. The targeted bottleneck is a lane-drop bottleneck which is circled in Figure 2.3. Downstream of this bottleneck, there is another lane-drop bottleneck next to Exit 7. Drivers in the targeted road section are driving from a four-lane section to a three-lane section (the upward direction in Figure 2.3), so a lane-drop bottleneck occurs.

The data is collected from 10 locations with approximately 500m spacing between them, giving a total length of around 5 km. There are 2 detectors in the four-lane section, followed by 8 in the three-lane section. This chapter does not consider detectors further downstream because vehicles will change into shoulder lane to leave freeway through Exit 7, possibly leading to external disturbances, for instance lane changing near the off-ramp.

Data for analysis is collected on two days, Monday 18 May 2009 and Thursday 28 May 2009. Figure 2.4 shows the speed contour plots in the study section on two days. There are two similar traffic situations in both of days. The first event is a group of stop-and-go waves. The main interest in this chapter is in the last stop-and-go wave in the first event. On 18 May the stop-and-go wave originated from the lane-drop bottleneck near Exit 7 at about 16:45. On 28 May the stop-and-go wave enters the selected stretch from further downstream at around 16:55. At 17:40 and 17:50 (18 and 28 May re-



(a) Speed contour on 18 May 2009



(b) Speed contour on 28 May 2009

Figure 2.4: Layout of the study site and data on two days (18 May and 28 May 2009) for study. The lane-drop bottleneck located between Detector 8 and 9 is activated by a stop-and-go wave from downstream. The numbers show locations of detectors. This study restricts to 10 locations around the targeted lane-drop bottleneck.

spectively), the next stop-and-go wave reaches the lane-drop bottleneck which is close to Exit 7. Downstream of the second stop-and-go wave there is congestion. When calculating the outflows, this study analyzes the data before the entering of the second stop-and-go wave in order to avoid influences of this congestion. The bottleneck queue discharge rate is calculated after the bottleneck has been active. When analyzing the flow distribution, this chapter analyzes the data collected from 16:00 to 19:00.

2.4 Results

This section first presents the different states, then the capacity estimates, and then in section 4.3 and 4.4 the lane-specific features are discussed.

2.4.1 State identification

This section describes empirical observations. Figure 2.5 shows empirical slanted cumulative counts across three lanes at 8 locations downstream the bottleneck on two study days. The arrow in each figure shows the shock wave which propagates downstream from the bottleneck. This means the traffic is in a free flow state, and not influenced by the off-ramp downstream. The outflow of the stop-and-go wave and the discharging flow of the standing queue are clearly distinguishable with the shock wave between these two states, see the upward arrows in Figure 2.5. Generally, the empirical observations are in line with the expectations presented in in section 2.

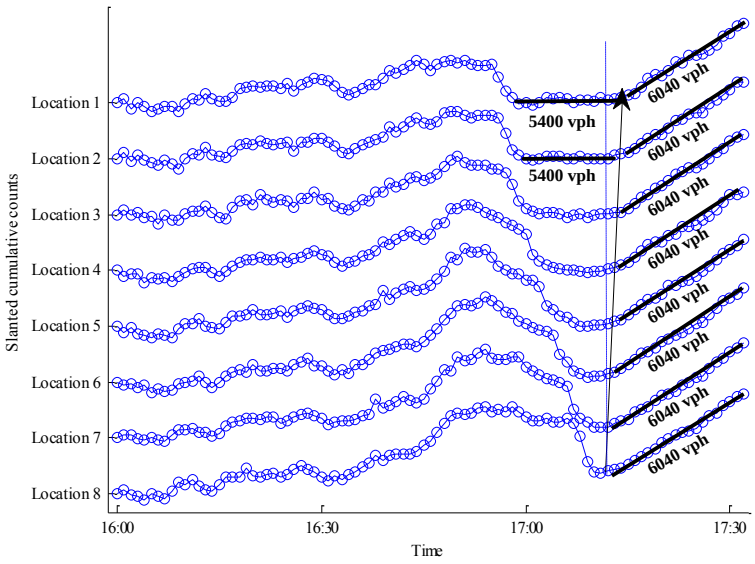
This shock wave separates the outflow of stop-and-go wave from the discharging flow of standing queue. This shock wave has been expected in section 2 (see Figure 2.1(b)). At one location, we first observe the outflow of the stop-and-go wave and then observe the discharging flow of the standing queue. First, we find the outflow of the stop-and-go wave only directly downstream of the stop-and go wave. The wave travels upstream, from location 1 to location 8. Once it reaches location 8, the traffic state will change, with a wave propagating downstream, which takes some time before it reaches location 8. During that whole time, at location 1 the outflow of the stop-and-go wave can be detected.

The discharging flows found for the two days are constant for each day, at 6040 veh/h (18 May) and 5700 veh/h (28 May), see Figure 2.5. Although they are different for both days, the flows are remarkably constant over time. There is also a difference between the flows downstream of the standing queues at 18 and 28 May. This holds for all locations downstream of the bottleneck, including the acceleration phase. The flow is the different but constant for both days. During the acceleration process, the density continuously decreases. Since the flows differ for the two days, the speeds must differ for the two days for situations with an equal density. This means that drivers leave a larger gap than necessary in the day with the lower flow (28 May), since apparently - given the speed-density relationship for the other day - they can drive with lower speeds given the spacing.

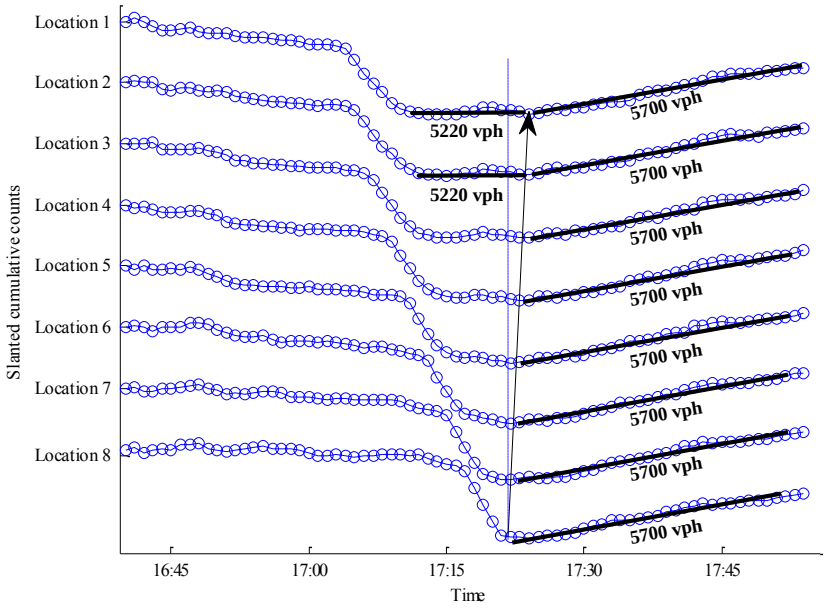
Moreover, the downstream direction of the shock wave implies that the off-ramp (Exit 7 in Figure 2.3) does not influence the discharging flow. Oh and Yeo (2012) imply that the off-ramp at the downstream location mitigates the capacity drop. In our study site, the off-ramp which is located far away has no effects. The shock waves propagating downstream imply no influence from downstream.

2.4.2 Capacity estimation

Figure 2.6 shows the capacities (with congestion upstream) which are the outflow of congestion at a homogeneous three-lane freeway section. In Figure 2.6, all red dashed lines show the slanted cumulative curves at the downstream locations and the blue bold

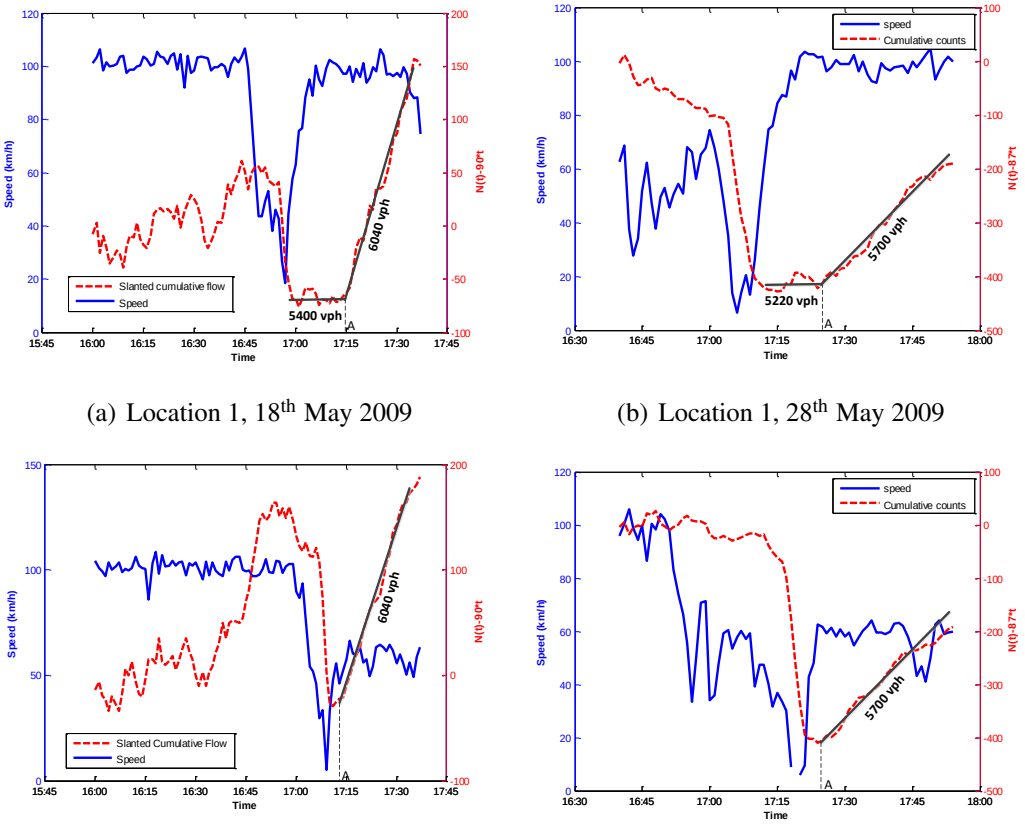


(a) A4, 18 May 2009



(b) A4, 28 May 2009

Figure 2.5: Slanted cumulative counts across three lanes at 8 locations downstream the bottleneck on two days, 18 May 2009 (left) and 28 May 2009 (right).



(a) Location 1, 18th May 2009

(b) Location 1, 28th May 2009

(c) Location 8, 18th May 2009

(d) Location 8, 28th May 2009

Figure 2.6: Average time mean speed (blue bold line) and slanted cumulative counts (red dash line) across three lanes at location 1 and location 8 on 18 May 2009 (a & c) and 28 May 2009 (b & d).

Table 2.1: Speed and flow in different traffic state points

	18 May 2009		28 May 2009	
	v_T (km/h)	q (veh/h)	v_T (km/h)	q (veh/h)
State 2	13.44	2182.50	6.34	885.00
State 5	98.73	5400.00	98.52	5220.00
State 4	30.80	6040.00	29.18	5700.00
State 6	98.32	6040.00	98.24	5700.00

lines represent speed evolution there. All figures in Figure 2.6 show firstly a decrease of flow (during the time the stop-and-go wave is present), indicated by a cumulative flow line with a negative slope. Afterwards, at location 1 the flow is constant for about 20 minutes, at approximately 5400 veh/h on 18 May and 5220 veh/h on 28 May. Figure 2.6(c) and Figure 2.6(d) show the slanted cumulative curves for the location 8, just downstream of the bottleneck. After the stop-and-go wave reaches location 8, the jam soon transforms into a standing queue and the outflow increases up to 6040 veh/h and 5700 veh/h respectively. These two discharging flows propagate downstream from the bottleneck and reaches location 1. In Figure 2.6, we label the moment when the higher discharge rate reached as "A". The higher outflow (6040 veh/h and 5700 veh/h) is not temporary and remains for at least 15 minutes at each location. The solid black line in each of the figures indicates a flow to which the slanted cumulative curve can be compared. In each figure, the increasing slope of black lines shows that the outflow of stop-and-go wave is lower than the discharging flow of the standing queue. Typically, we find that the outflow of the stop-and-go wave lies in the range of 5220 veh/h to 5400 veh/h and the outflow of the standing queue is in the range of 5700 veh/h to 6040 veh/h. All data points are shown in Table 2.1. The state number corresponds to Figure 2.1.

In Figure 2.6, there is a shift in the exact time of speed recovery and time of flow increase. It means flow stops growing before the speed recovers - a horizontal acceleration line in the flow-density plot, which can be a result of relaxation and anticipation effects (Deng & Zhang, 2015).

State 2, 4, 5 and 6 in Figure 2.1(a) are identified quantitatively. State 2 and 4 stand for congestion states. State 5 and 6 represent states of capacities. We thus find a correlation between the type of congestion and its outflow. In fact, the outflow of a stop-and-go wave is lower than the outflow of a standing queue at the same location.

It is believed that the flow estimation error in applications in this chapter is limited, smaller than 1%. In Figure 2.7, the slanted cumulative count is drawn with different reference flows. The horizontal line stands for the reference flow. It can be seen that the targeted flow which is circled by dashed line should be between 5340 vph and 5460

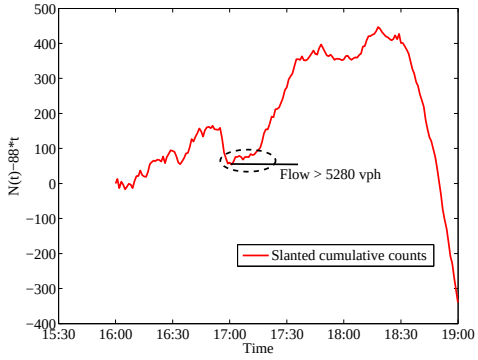
vph. That is, the error of the estimated flow 5400 vph in this chapter is limited up to 1%.

2.4.3 Outflows in each lane

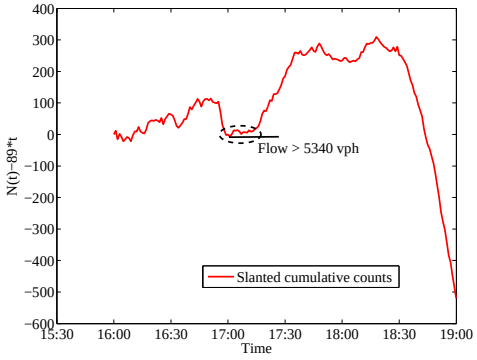
When congestion occurs, each lane presents different features regarding to outflows. It is worth to notice that the outflow on each lane here means the flow detected downstream of congestion on each lane, from which the merging and diverging flow have been subtracted from and added into, respectively. In Figure 2.8, slanted cumulative counts and speed in each lane are presented, shown as a red dashed line and a blue bold line respectively. Slow vehicles and trucks usually drive on the shoulder lane due to the Keep Right Unless Overtaking rule. Therefore, the flow and speed detected in each lane at the same location differ from each other. In both of Figure 2.6(a) and Figure 2.6(b), aggregated data over 3 lanes shows an increase of outflow at the moment the wave separating the outflow from the stop-and-go wave and the outflow from the standing queue reaches the detector. In Figure 2.8(a) and Figure 2.8(c), this increase of the outflow is observed on the median and center lane at location 1 on 18 May 2009, but not on the shoulder lane. At 28 May this increase is found on all lanes. The lack of change in flow on the shoulder lane is remarkable, but at the moment is it unclear what could be the reason.

2.4.4 Flow distribution over lanes

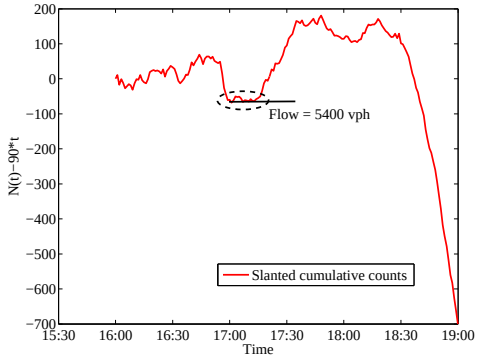
When the bottleneck has been active, there are several different traffic states in the downstream of the bottleneck. Along the distance, the density decreases. Therefore, in the targeted scenario, a large range of density can be detected, which can reveal the flow distribution as a function of density across lanes. The flow distributions are shown in Figure 2.9. Red lines show the fast lane (median lane), black lines show the center lane and the blue lines show the shoulder lane. Three bold lines (see Figure 2.9(a) & Figure 2.9(b)) represent average flow distribution at three lanes based on all data. Circles and triangles are the empirical data collected in each lane at location 1 (see Table 1 and Figure 2.8). Those circles and triangles stand for the state of the outflow in each lane at location 1, i.e., state 5 and state 6 (see Figure 2.1) respectively. Note that we at location 1 on 18 May 2009 there is no distinguish between the state 5 and state 6. Therefore, when calculating the flow distribution in these two states (state 5 and 6), we use the same flow, that is 1437 veh/h as shown in Figure 2.8(e). Note that, the lower flow in state 5 (compared to state 6) in the center lane (see Figure 2.8(d)) does not mean the flow distribution in state 5 should be lower than that in state 6. That explains why in the center lane the flow distribution in state 5 is higher than that in state 6 (see Figure 2.9(b)). The rest thin lines (in Figure 2.9(c) & Figure 2.9(d)) represent the flow distributions at each location. The lines with five-point stars stand for the distribution at location 8.



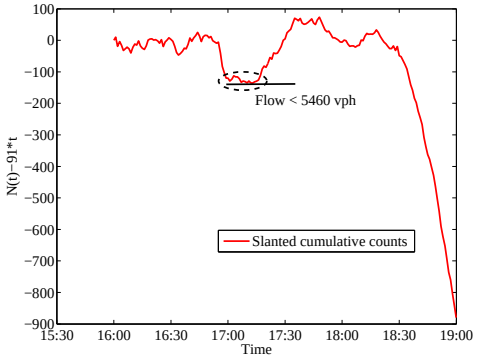
(a) Reference flow = 5280 vph



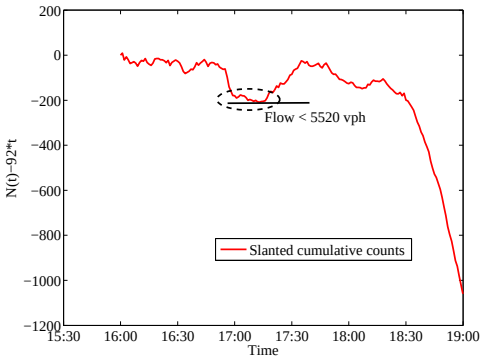
(b) Reference flow = 5340 vph



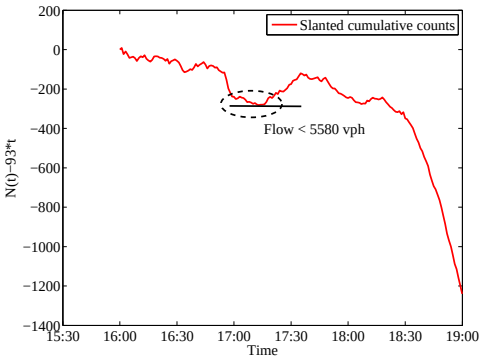
(c) Reference flow = 5400 vph



(d) Reference flow = 5460 vph

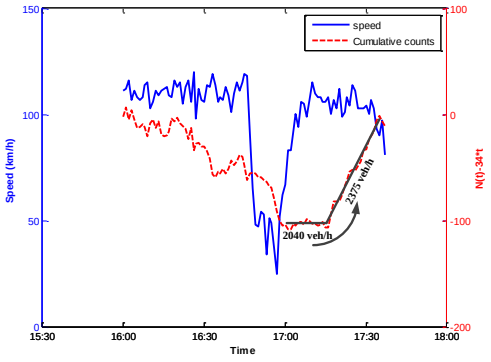


(e) Reference flow = 5520 vph

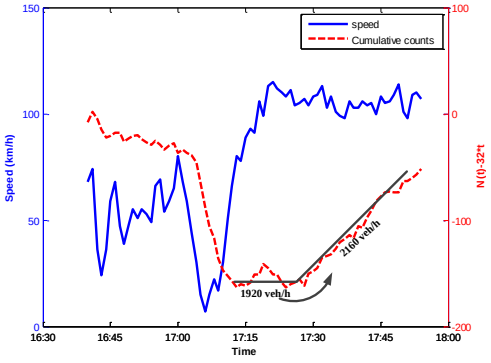


(f) Reference flow = 5580 vph

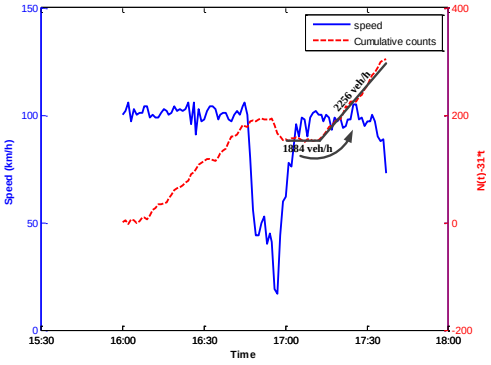
Figure 2.7: Slanted cumulative counts with different reference flow at location 1 on freeway A4. The horizontal line stands for a reference flow.



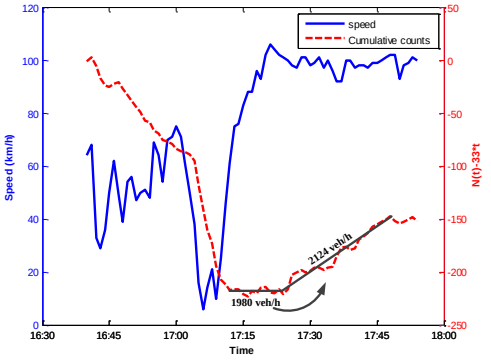
(a) Median lane, location 1, 18 May 2009



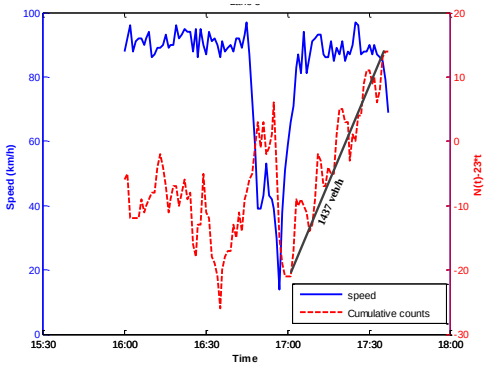
(b) Median lane, location 1, 28 May 2009



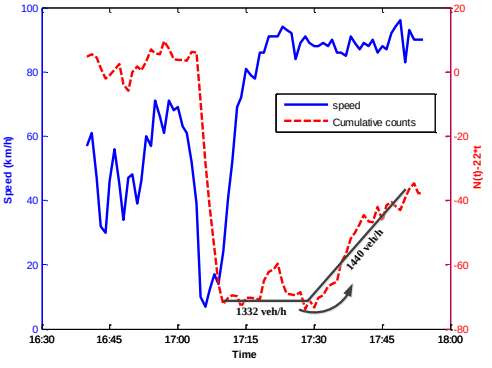
(c) Center lane, location 1, 18 May 2009



(d) Center lane, location 1, 28 May 2009

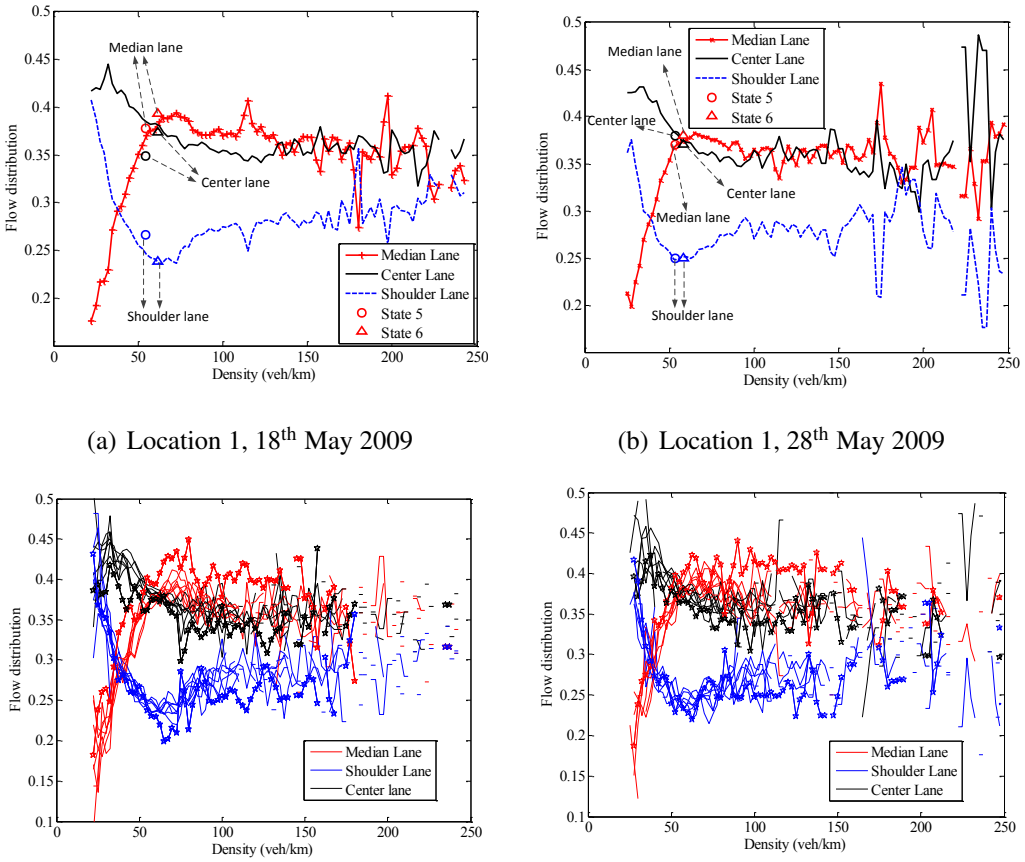


(e) Shoulder lane, location 1, 18 May 2009



(f) Shoulder lane, location 1, 28 May 2009

Figure 2.8: Speed and slanted cumulative count on each lane on 18 May 2009 (a, c & e) and 28 May 2009 (b, d & f) at location 1. Flows are shown next to the coinciding slanted cumulative counts (bold black lines).



(a) Location 1, 18th May 2009

(b) Location 1, 28th May 2009

(c) Eight locations, 18th May 2009

(d) Eight locations, 28th May 2009

Figure 2.9: Flow distributions at different densities at three-lane freeway section. a) and b) shows average flow distributions over 3 lanes, median lane (red), center lane (black), and shoulder lane (blue) on two days, 18 May (left) and 28 May (right). Circles and triangles show the performance of each lane in state 5 and state 6 respectively, corresponding to data in Figure 2.8. c) and d) shows flow distributions at each 8 locations. Each thin line shows a flow distribution at each location. Five-point stars represents the flow distribution at location 8.

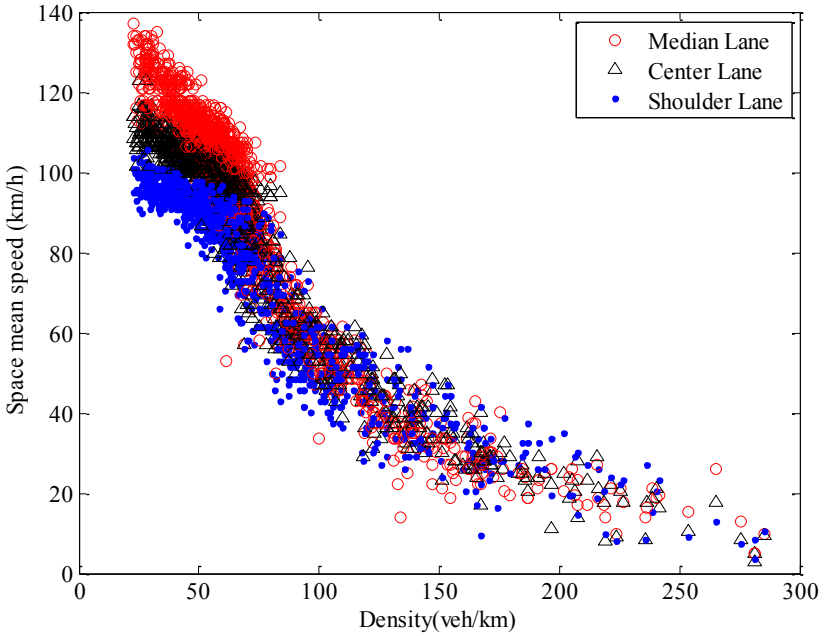
Figure 2.9(a) and Figure 2.9(b) shows flow distributions on two different days. Both figures show a common feature. When the density lies within the range 22 - 60 veh/km, the flow in the center lane is higher than that in both other lanes, although it keeps decreasing as density grows. When the density is around 60 veh/km, the fraction of the flow on the shoulder lane reaches the minimum at around 23%. For shoulder lane the decrease of the fraction of the flow was sharp, but afterwards the increase is only marginal. Meanwhile from 60 veh/km the fraction of the flow in median lane stops increasing with density and begins to stabilize at around 38%. Note that the density of 60 veh/km corresponds to a typical critical density, that is 20 veh/km/lane (Treiber & Kesting, 2013).

When the density exceeds 132 veh/km (18 May) and 95 veh/km (28 May), the fraction of the flow on the median lane is almost equal to the fraction of the flow on the center lane, at around 35% for each while the flow percentage on the shoulder lane is around 30%. So even in states with a very high density, flows on the shoulder lane are still lower than that on the other lanes. When density reaches up to 220 veh/km, the flow begins to be distributed evenly over three lanes on 18 May while the flow distribution is more unstable on 28 May. It is not surprising because in extremely high density situation standing vehicles can lead to some detection problems.

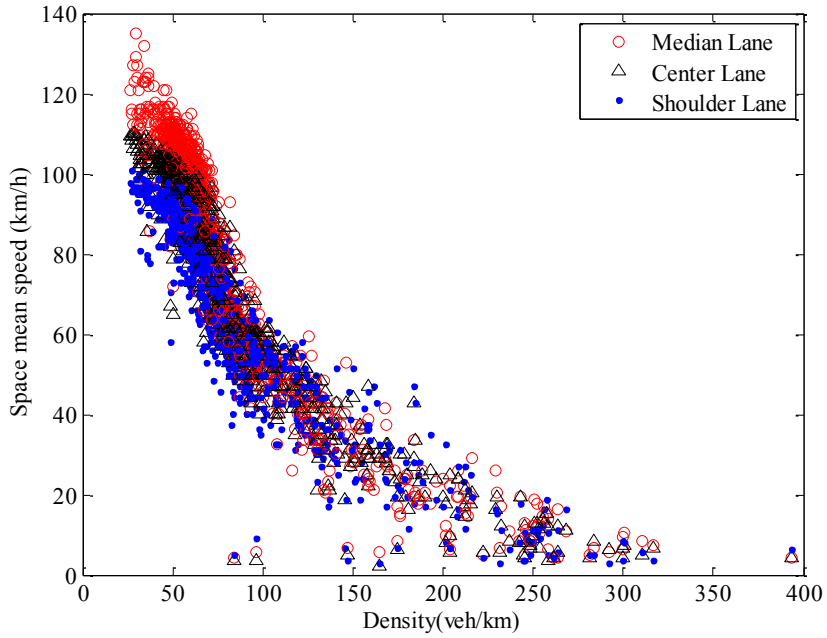
Figure 2.9(c) and Figure 2.9(d) show the flow distribution at 8 locations. The flow distribution in median lane (red line) at location 8 (marked as red five-point stars) is much higher than that at the other locations, see Figure 2.9(c) and Figure 2.9(d). In contrast, the flow distributions in the center and median lanes at location 8 are the lowest. That is because vehicles merge into median lane when passing through the lane-drop bottleneck. In the downstream of location 8, the flow distribution in median lane is lower than that at location 8. For the other locations, the distribution situations are similar to each other. This chapter explains this by the following. Vehicles force themselves into the traffic stream and it takes some time-and hence distance-before equilibrium distribution sets in again. Therefore, it is believed that a high percentage of vehicles choose to leave median lane by changing lane between location 8 and location 7. In the future research, more empirical data (especially trajectory data set) are needed for justifying the behavioral explanation on the different flow distributions at different locations.

Among three lanes, due to the Keep Right Unless Overtaking rule in the Netherlands, we can assume that the shoulder lane is first choice for drivers when the density is extremely low. As the density increases to around 20 veh/km, the fraction of traffic on center lane begins to be higher than that on the shoulder lane. The use of median lane (fast lane) is the least at that time. As the density increases, in contrast to the shoulder lane whose flow fraction reduces considerably, the use of median lane sharply grows. Finally, the median lane and center lane are highly made use of while the shoulder lane is being underutilized.

Figure 2.10 shows the speed in each lane at the same average density over three lanes. Circles, triangles and dots indicate the speed in the median lane, center lane and shoul-



(a) Location 1, 18th May 2009



(b) Location 1, 28th May 2009

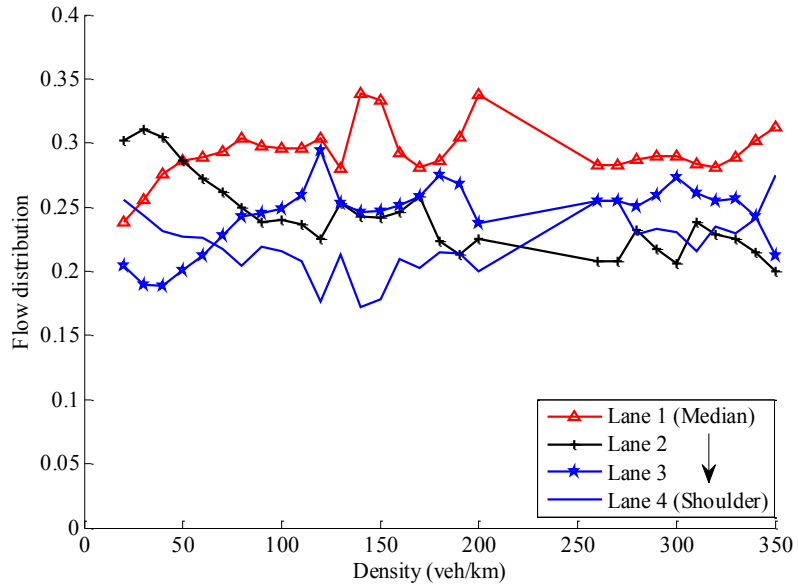
Figure 2.10: Speed-Density plot in each lane in the three-lane section on two study days, 18 May (left) and 28 May (right). The density is the average density over three lanes.

der lane, respectively. When the density is lower than around 70 veh/km, the speed decrease from the median lane towards the shoulder lane, that is due to the Keep Right Unless Overtaking rule. The median lane is the fastest lane. In Figure 2.10, when the average density is higher than 70 veh/km, circles, triangles and dots are greatly overlapped. That means the speed is becoming more equal among the lanes. Because in congestion the speeds are almost equal on all lanes (shown as the highly overlapped area among circles, triangles and dots), so the low flow on the shoulder lane must be due to a low density or large spacing. That means that microscopically in congestion the spacing between successive vehicles on the shoulder lane is the largest among three lanes.

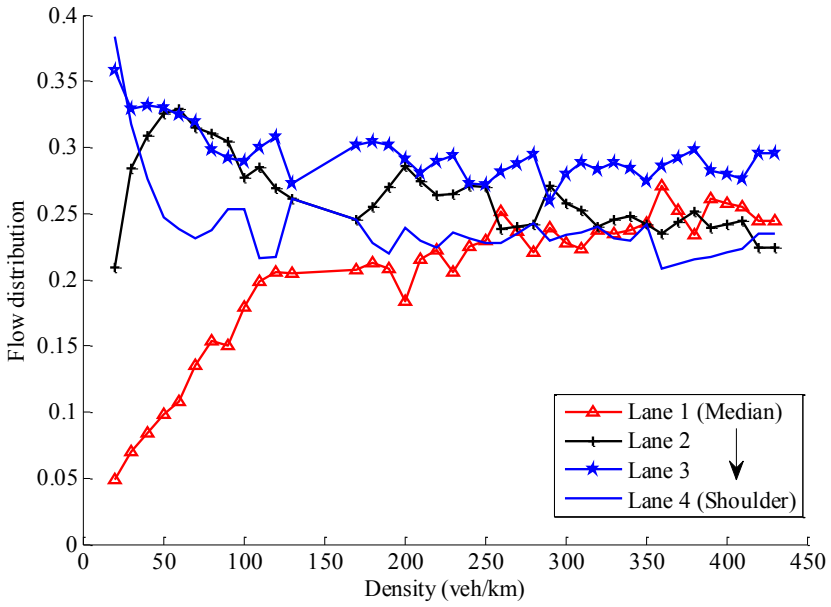
Figure 2.11 shows the flow distributions on the four-lane freeway section upstream the lane-drop bottleneck. Note that the outflow of the upstream four-lane freeway section is the inflow of the downstream three-lane freeway section. There are 2 locations for the data collection, location 9 and location 10 in Figure 2.3. Traffic flow moves from location 10 to location 9. The figure only shows the data for 18 May, the data for 28 May is similar.

In fact, we can distinguish two pairs of lanes. First, lane 1 and 2 are the median and shoulder lane of one of the upstream branches of the road. The flow distributions at lane 3 and 4 are similar to that of lane 1 and 2 respectively, also originating from a two lane road upstream. The flow distribution at two locations differs considerably. On one hand, in contrast to location 10 which is in the upstream of the location 9, location 9 shows a lower flow in the median lane, especially for low densities. On the other hand, at location 9 the flow on the shoulder lane is higher for low densities.

The non-compensated amount of lane changes can be estimated by the difference in flow per lane between the two detectors for a certain density (e.g., one can see how much lower the flow is). Compensation is possible be by other vehicles making opposite movements (e.g., vehicles moving into the lane). In lane 3, the right center lane, the flow is higher at location 9. Downstream of location 9, all vehicles in the median lane have to merge into lane 2. Drivers in lane 2 (the left center lane) might anticipate this and make space for the drivers merging from the median lane. These lane changes can be considered as an explanation for the changes in lane flow distribution we observe between location 10 and 9. The relative flow in lane 2 does not change as much, because there is a similar amount of lane changing from the median lane to lane 2; what is observed is a decrease of the utilization of the median lane. The number of lane changing decreases as the average density over lanes increases. The flow distribution at lane 2 and 4 is nearly stable for both locations and study days. At location 9 near the bottleneck, the flow in the lane 3 is always the highest for both study days. Note that the demand in the upstream two two-lane freeway sections could possibly greatly influence the flow distribution at location 10.



(a) Location 10, 18 May 2009



(b) Location 9, 18 May 2009

Figure 2.11: Flow distributions at different densities at four-lane freeway section on 18 May. The distribution on 28 May is the similar. The traffic flow is moving from location 10 to location 9.

2.5 Conclusion

This chapter compares the downstream states of a stop-and-go wave with that of a standing queue. The standing queue in this chapter is induced at a lane-drop bottleneck by a stop-and-go wave. Therefore, at one bottleneck there are two different congestion states observed. In the downstream of the congestion there are free flow states, that means the two outflows detected downstream of congestion are the capacities of the road section. This chapter applies shock wave analysis to find those two outflows at the same road section, which is well traceable in the real data. The most important finding is that the outflow of stop-and-go waves is be much lower than that of a standing queue. Therefore, the capacity with congestion upstream can vary in a rather wide range, e.g. from 5220 veh/h to 6040 veh/h at a three-lane road section. The various capacities could be related to congestion states, which means a promising traffic control strategies could increase the queue discharge rate and minimize traffic delays.

There are two other findings. First, different features of the outflow from congestion on different lanes can be found. Strong fluctuations occasionally can be observed on the shoulder lane, which might even trigger stop-and-go waves later on, for instance near a next bottleneck. Second, the flow distribution over three lanes is presented. This shows that particularly near head of a standing queue more vehicles can merge into the lane adjacent to the ending lane, thereby locally increasing the capacity of that lane. The capacity of the shoulder lane is markedly wasted when in congestion. The reason for the low flow distribution in congestion on the shoulder lane is the large spacing between successive vehicles.

Future research should show the mechanisms behind these features, from a behavioral perspective (whether people behave differently), from a vehicle perspective (what the influences of different acceleration profiles are) or from a flow perspective (what for instance the influence of voids is). In the future, a promising control strategy, based on our empirical research, should be proposed to maximize queue discharge rates and minimize traffic delays.

Chapter 3

Relation between the speed in congestion and the queue discharge rate

In Chapter 2, it is found that the queue discharge rate is related to the congestion category. Based on this finding, a hypothesis that the queue discharge rate is related to the congestion characteristic is made. The hypothesis requires empirical confirmations. This chapter fills the gap by empirically investigating the relation between the speed in congestion and the queue discharge rate. The research corresponds to research question 1. This chapter identifies a positive relation between the speed in congestion and the queue discharge rate. The investigation into the reason for the relation between the speed in congestion and the queue discharge rate is given in following chapters.

This chapter is an edited version of the article: Yuan, K., Knoop, V.L., & Hoogendoorn, S.P. (2015). Capacity drop: relation between the speed in congestion and the queue discharge rate. *Transportation Research Record: Journal of the Transportation Research Board*, 2 (2491), 72-80.

3.1 Introduction

Traffic congestion is a daily phenomenon in major urbanized areas. During peak hours, road capacity is insufficient for the traffic demand, so traffic jams occur. Due to the capacity drop, traffic delays increase once congestion sets in. There are control strategies trying to avoid the capacity drop by limiting the inflow. Another option would be to minimize the capacity drop magnitude after congestion sets in. It is however unclear what determines the size of the capacity drop.

This chapter studies the queue discharge rate, which is defined as the outflow of congestion without influences from downstream. Throughout the chapter, the word flow is used for the number of vehicles passing a location per unit of time; in other papers this is sometimes referred to as 'traffic volume' or 'flow rate'. Hence, the queue discharge rate is the maximum flow out of a queue. The term queue in this chapter refers to a general concept of congestion, including the standing queue with head fixed at the bottleneck and stop-and-go waves with congestion front moving upstream. The bottleneck means a fixed point upstream of which a queue forms.

From literature, the capacity drop itself, defined as the difference between the capacity and the queue discharge rate, is not a constant value; it differs under the influence of several factors, such as the characteristics of the study site (e.g. the number of lanes, the traffic flow composition, etc.) and also for different conditions for the same bottleneck. Literature on empirical data shows that the same location can produce different discharge rates (Chung et al., 2007) and that in the same link the discharge rate can vary in a wide range (see Chapter 2). These empirical observations reveal a high possibility that control strategies can promote discharge rate to evacuate vehicles in a queue quickly and finally reduce delays (Chung et al., 2007; Carlson et al., 2011; Cassidy & Rudjanakanoknad, 2005). To increase the discharge rate, it is important to know which factors influence the queue discharge rate.

Muñoz and Daganzo (2002) find a positive relation between the speeds of a moving bottleneck and the queue discharge rate for speeds of 50 km/h and lower. But the empirical data points are very limited and the speed range is very narrow. Moreover, the upper and lower bounds in their research are taken from other data sources in different traffic conditions. Laval and Daganzo (2006) extend this research by simulating the same experiment in a broader speed range. They show a positive relation between the capacity and moving bottleneck speed when speed is higher than 20 km/h and a negative one when speed in congestion is lower than 20 km/h. This result relies on the vehicular speed difference in the queue among lanes, which might be different for congestion without moving bottlenecks. As shown in Figure 2.10(b), vehicular speed difference is small in congested states in an absence of moving bottlenecks. Therefore, until now, there is still no thorough empirical analysis revealing a reliable relation between the outflow of congestion and the congested state, though this relation is relevant. This chapter fills in this gap.

This chapter expresses the congestion level as vehicular speed in congestion. The reason for the preference of speed in congestion is twofold. Firstly theoretically, previous models (e.g., Leclercq et al., 2011; Laval & Daganzo, 2006; Tampère, 2004) and empirical observations (e.g., Muñoz & Daganzo, 2002; Daganzo, 2002) hold a promising relation between the speed in congestion and the queue discharge rate. The hypothesis about the relation is also shown in Chapter 2. Secondly practically, some promising control strategies, e.g., mainstream metering (Carlson et al., 2011) and SPECIALIST dynamic speed limit strategy (Hegyi & Hoogendoorn, 2010), have a fundamental dependence on the relation between the speed in congestion and the queue discharge rate.

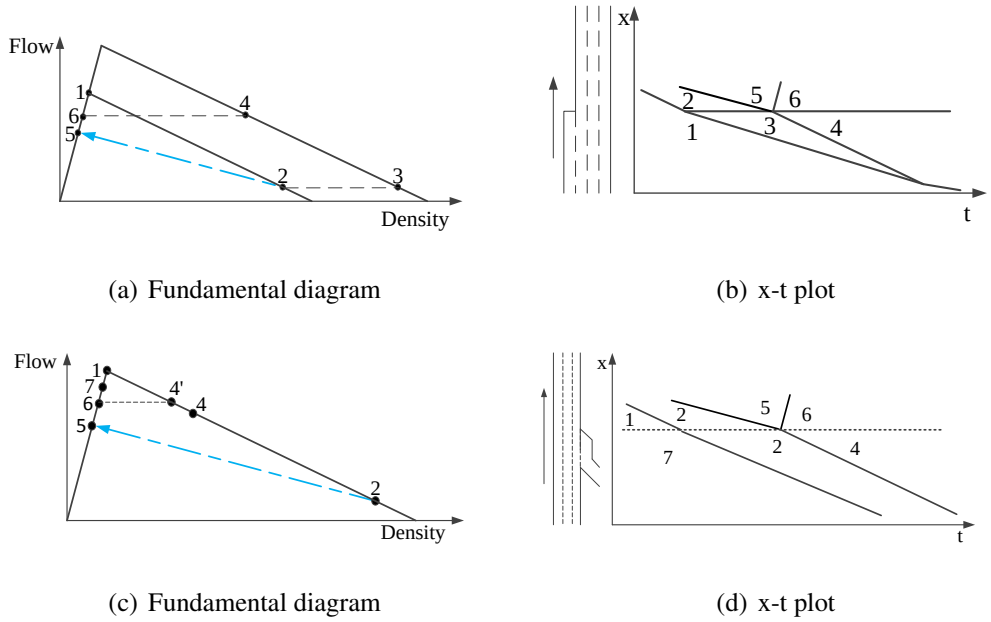
Speed has been proposed as a possible explanatory variable for the capacity drop. This relationship is tested and quantified in this chapter using empirical data. The influence of weather and site-specific calibration is also discussed.

3.2 Methodology

This chapter wants to analyze the queue discharge rate for speeds in the upstream congestion which vary strongly. The work considers a traffic situation with different types of congestion (standing queues and stop-and-go waves) and analyzes the queue discharge rate at the same location. Section 3.2.1 describes the traffic scenario this chapter targets. In this scenario, different traffic congestion states with various vehicle speed at the same location are observed. Section 3.2.2 presents some requirements about the data for the analysis. The requirements restrict the availability of data and the choice of study sites. Section 3.2.3 applies shock wave analysis to quantitatively and qualitatively identify the queue discharge rates and the speed in the corresponding congestion in the traffic scenario. Finally, a linear function is used to fit data to investigate the relation between the speed in congestion and the queue discharge rate, as described in Section 3.2.4.

3.2.1 Traffic scenario

To obtain a sufficiently wide range of speed in congestion, it is necessary to consider the capacity drop in stop-and-go waves because standing queues where vehicles' speed can not be as low as that in stop-and-go waves are not sufficient. First order traffic flow theory predicts that a bottleneck is activated immediately after a stop-and-go wave passing by. This traffic scenario is graphically presented in Figure 3.1. The occurrence of this traffic state has also been empirically confirmed by Chapter 2. In this scenario different congestion states, including standing queues and stop-and-go waves, and different outflows of congestion can be observed at the same location. This scenario can provide data of different congestion speeds at the same bottleneck. Hence, this chapter targets the data collected from this traffic scenario to collect data efficiently.



(a) Fundamental diagram

(b) x-t plot

(c) Fundamental diagram

(d) x-t plot

Figure 3.1: Shock wave analysis for distinguishing different outflows with different congestion upstream at a lane-drop (a & b) and an on-ramp (c & d) bottleneck.

At bottlenecks which are activated due to local break-down, this scenario can also be found because of boomerang effects. The so-called boomerang effect (Schönhof & Helbing, 2007) means that small perturbation in a free traffic flow first travels downstream. While doing so, it increases and traffic breaks down, downstream of the point where the disturbance has entered, close to a bottleneck. The congestion then propagates upstream. The boomerang effect usually can be observed around an on-ramp bottleneck (Cassidy & Bertini, 1999; Kim & Coifman, 2013). This effect can provide the stop-and-go wave we need if the standing queue forms spontaneously.

3.2.2 Data requirements

To reveal the relation between the speed in congestion and the discharge rate with thorough empirical analysis, there are several requirements about the data and study sites.

Firstly the data should present a wide range of speed in congestion, that can be solved with the traffic scenario presented in section 3.1. Secondly, to detect the discharge rate of the congestion, the state downstream of the congestion should be free flow. Thirdly, to ensure the detected discharge rate is stable, we need to observe the discharge rate for a certain time, for which we choose 10 mins. Therefore, if the stop-and-go wave originates downstream of a standing queue and propagates soon into the standing queue at the bottleneck, then the short-life discharge rate will not be considered as a stable discharge rate, and the speed data in that stop-and-go wave will be excluded. Meanwhile, as shown in Figure 3.1(b) & Figure 3.1(d), the long-time observation (10 min)

of queue discharge rate (for instance state 5) requires a long homogeneous road section in the downstream of the bottleneck. Last but not least, because capacity drop can be influenced by the number of lanes and the presence of off-ramp in the downstream (Oh & Yeo, 2012), so when choosing the appropriate data collection sites we have to ensure there are no such geometrical disturbances. So there should be a homogeneous freeway section downstream the bottleneck for instance at least 2.5 km to ensure vehicles have reached free-flow speed in the homogeneous section and state 5 in Figure 1 can be observed for a long time.

Due to the limited observation samples at one bottleneck, this study chose two different bottlenecks, a lane-drop bottleneck and an on-ramp bottleneck, to collect data. On the one hand, two different study sites impose two more restrictions. Firstly, we have to ensure both of bottlenecks meet the requirements of study sites. Secondly, the number of lanes downstream the bottleneck and the slope of the road section should be the same. On the other hand, two different bottlenecks can shed light on the discussion of site-specific calibration.

Moreover, to see the influence of weather, we also analyze data from a rainy day.

3.2.3 Analytical solution

The next step of the research, which is the key of the analysis, is to identify traffic states and their accompanying discharge rates. The analytical solution in this study for the identification of different traffic states is to apply shock wave analysis in the studied scenario. Figure 3.1 shows the shock wave analyses applied for identifying congestion states and their accompanying outflows. The fundamental diagram for the analysis is a triangular fundamental diagram. Two bottlenecks, lane-drop (Figure 3.1(a) & Figure 3.1(b)) and on-ramp (Figure 3.1(c) & Figure 3.1(d)) bottleneck, are analyzed. Chapter 2 presents that the outflow of a stop-and-go wave is lower than that of a standing queue. Note that in Chapter 2 the speed in the stop-and-go wave is lower than that in the standing queue. Therefore, this chapter expects the outflow of a standing queue (state 6) is higher than that (state 5) of a stop-and-go wave. When a stop-and-go wave passes one detector, we can observe a state transformation from state 2 to state 5 at one location. When a bottleneck is active, in the downstream of the bottleneck traffic state transformations from state 4 to state 6 can be observed in a sequence along the freeway.

Figure 3.1(b) & Figure 3.1(d) show the spatio-temporal plots of traffic situations. There is a forward moving shock wave between state 5 and state 6. Since these two free flow states, state 5 and 6, always lie in the free flow branch, the shock wave between them should always be positive no matter whose flow is higher. So the assumption that state 5 is below state 6 does not influence the analysis. Therefore, this chapter distinguishes these two capacities via this shock wave. Note that the targeted shock wave between state 5 and 6 is not influenced by the state 1. Therefore, the shock wave analysis

in Figure 3.1 can be applied to identify the outflow of stop-and-go wave originating downstream a standing queue.

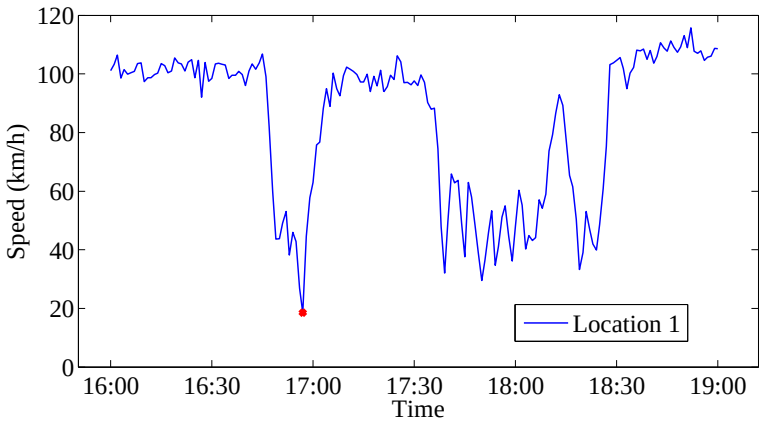
As shown in Figure 3.1, discharge rates of both stop-and-go waves and standing queue, i.e. state 5 and state 6 respectively, can be observed in the downstream of the bottleneck. However, the detection of discharge rate of these two different congestion differs slightly. In the downstream of a stop-and-go wave, the detected flow grows as speed increases while the discharge rate of a standing queue remains one value as speed increases. So in Figure 3.1 the state 5 close to the shock wave between state 2 and 5 actually should lie in the line connecting point 5 and point 2 in fundamental diagram, that is the flow in those states is lower than that in state 5. Only state 5 can show the discharge rate of the stop-and-go wave. Hence, the outflow of standing queue can be detected at any location downstream the bottleneck but that of stop-and-go wave should be detected far away from the bottleneck Overall, at downstream locations far away from the bottleneck both of outflows of stop-and-go waves and standing queues can be detected. Moreover, as shown in Figure 3.1, the location far away from the bottleneck can clearly show a long-period observation of two outflows, which benefits identifying the stable discharge rate.

3.2.4 Quantitative solution

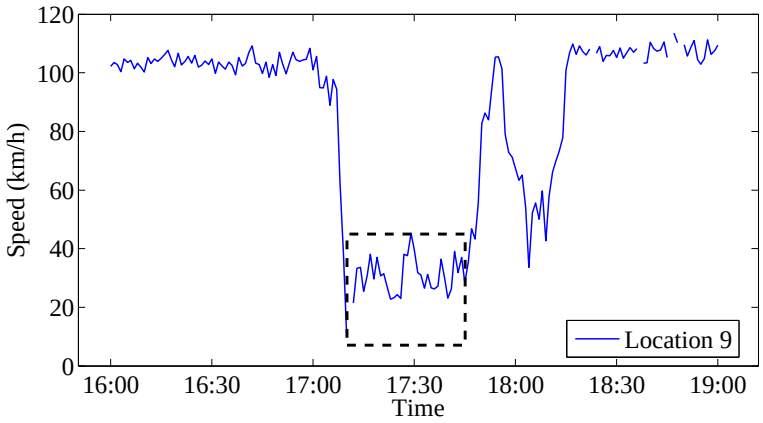
After the identification of traffic states, this chapter needs to investigate them quantitatively. Slanted cumulative curves are applied to calculate flow. The flow is the slope of each slanted cumulative count minus a reference flow. Because during the traffic state transition from state 5 into state 6, there is no remarkable speed in/decrease. Speed in both states is critical speed (maximum speed around critical density), so we cannot see the shock wave expected in Section 3.2.1. But we can observe the shock wave relying on the change of flow during the traffic state transition, that is we expect to observe the shock wave (between state 5 and state 6 in Figure 3.1) in the flow evolution plot presented as slanted cumulative curves.

The speed in stop-and-go wave is calculated as the average of all the lowest speed detected at each downstream locations when the studied stop-and-go wave passes (e.g., the red dot in Figure 3.2(a)) and the speed in standing queue is calculated as the average of speed detected at the location close to the downstream front of the standing queue (e.g., the data in the dashed box in Figure 3.2(b)). That means that for each observation, we have two, fairly accurate since averaged, data points. This chapter prefers this method over using all one-minute aggregated data points individually since in this way, each day has the same weight and each traffic condition has the same weight. Otherwise, congestion which lasts longer becomes more influential.

After obtaining the empirical data, this chapter fits the flow as function of speed in congestion. A linear function is used, and it is tested to what extent the linear function can represent the relation.



(a) Location 1



(b) Location 9

Figure 3.2: Speed detected at location 1 and location 9. In (a), the speed is highlighted by a red dot; while in (b), the averaged speed is shown in a dashed box.

3.3 Data collection

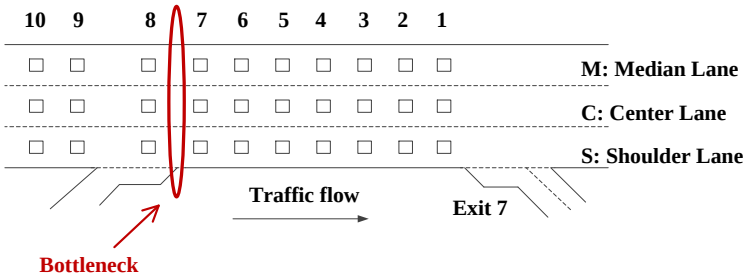
This chapter analyzes the queue discharge rate for speeds in the upstream congestion which vary strongly, considering a traffic situation with different types of congestion (standing queues and stop-and-go waves) and analyzing the queue discharge rate at the same location. Section 3.2.1 describes the traffic scenario this chapter targets. In this scenario, we can observe different traffic congestion states with various vehicle speed at the same location. Section 3.2.2 presents some requirements about the data for the analysis. The requirements restrict the availability of data and the choice of study sites. Section 3.2.3 applies shock wave analysis to quantitatively and qualitatively identify the discharge rates and the speed in the corresponding congestion in the traffic scenario. Finally, we choose to fit data with linear function to investigate the relation between the speed in congestion and the queue discharge rate, described in Section 3.3.

3.3.1 Data collection sites

On the freeway A4 (see Figure 3.3(a)), the data is collected around a lane-drop bottleneck in the northbound direction just downstream of Exit 8 (The Hague). Drivers in the targeted road section are driving from a four-lane section to a three-lane section. Thus, the outflow of congestion should be representative for the queue discharge rate of a three-lane freeway. In the downstream end of this bottleneck, there is another lane-drop bottleneck next to Exit 7, which is around 6.5 km further downstream. The data is collected from 10 locations around 5 km, of which 2 are located in the four-lane section and 8 are located in the three-lane section. In this chapter, we restrict our study to 10 locations because the speed at location 1 should have reached the critical speed which is vehicles' possible maximum speed after accelerating from congestion and there the state of outflow of congestion can last long enough for a clear observation. In the considered data set (May 2009 and September 2012), three days fulfill the requirement that a stop-and-go wave triggered a standing congestion. These days are (18 May, 28 May and 11 September). At all these dates, there was no precipitation.

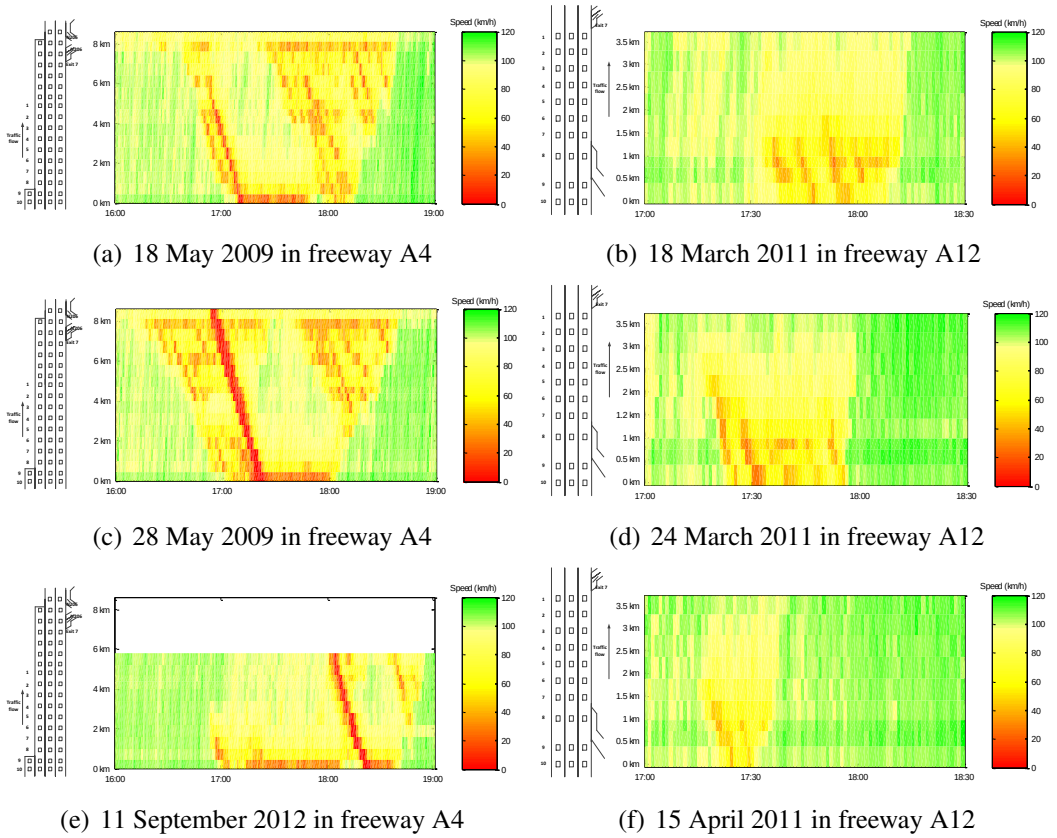
On the freeway A12 (see Figure 3.3(b)), this chapter considers an on-ramp bottleneck in the eastbound direction upstream the Exit 6 (Zoetermeer city center). The study sections are three-lane section upstream and downstream of the bottleneck. Hence, the outflow of congestion at this site should be representative for the discharge rate of a three-lane freeways, too. The data is collected from ten locations around 5 km, of which there are 2 upstream of the acceleration lane, 1 in the acceleration lane area as well as 7 in the downstream of the bottleneck. The on-ramp bottleneck is around 2.5 km away from the off-ramp in the downstream end. At location 1 the speed has reached the critical speed and the states of capacities can be identified clearly. The data for March and April 2011 have been checked and 3 days have been found to fulfill the requirements of a stop-and-go wave (included by the boomerang effect) leading to a

(a) Data collection site on freeway A4



(b) Data collection site on freeway A12

Figure 3.3: Data collection site of freeway (a) A4 and (b) A12.



(a) 18 May 2009 in freeway A4

(b) 18 March 2011 in freeway A12

(c) 28 May 2009 in freeway A4

(d) 24 March 2011 in freeway A12

(e) 11 September 2012 in freeway A4

(f) 15 April 2011 in freeway A12

Figure 3.4: Speed contour plots of study traffic situations on freeway A4 (a, c, e) and freeway A12 (b, d, f).

standing queue, being 18 March, 24 March and 15 April. Note that on 18 March 2011 there is 8.8 mm precipitation. The other two observations are made on sunny day.

3.3.2 Traffic conditions

To observe various congestion states at the same location, this chapter targets both of standing queues and stop-and-go waves. Figure 3.4 shows the speed contour plots of the traffic operations on freeway A4 (Figure 3.4(a), Figure 3.4(c) and Figure 3.4(e)) and A12 (Figure 3.4(b),Figure 3.4(d),Figure 3.4(f)).

On freeway A4, the targeted bottleneck is the lane-drop bottleneck between the 4-lane section and the 3-lane section. The observations on freeway A4 show the scenario that the lane-drop bottleneck is activated when a stop-and-go wave passes. After the activation of the lane-drop bottleneck, there comes a second stop-and-go wave which is not taken into consideration in this work. On 11 September 2012, the lane-drop bottleneck was activated at around 17:10 before the stop-and-go wave arrived the bottleneck. Therefore, these three days' data provides 7 congestion states and accompanying discharge rates in total.

In freeway A12, the study bottleneck is an on-ramp bottleneck. The bottleneck is the

original location where break down occurs. On 24 March and 15 April 2011, before the break down at the bottleneck, a stop-and-go wave originates in the downstream of bottleneck. This observation could due to boomerang effects (Schönhof & Helbing, 2007) or the drivers relaxation impacts (Kim & Coifman, 2013). On 18 March, the stop-and-go wave originated very close to the downstream front of the following standing queue, so we believe there is only standing queue counting for the discharge rates. Therefore, there are 5 congestion states observed on freeway A12.

These congestion states correspond to a broad range of speed, from 5km/h to 60 km/h, that means the data can provide a reliable empirical relation between the speed in congestion and the outflow of the congestion. According to section 3, all the outflows of congestion are identified at location 1 on both of freeways, A4 and A12.

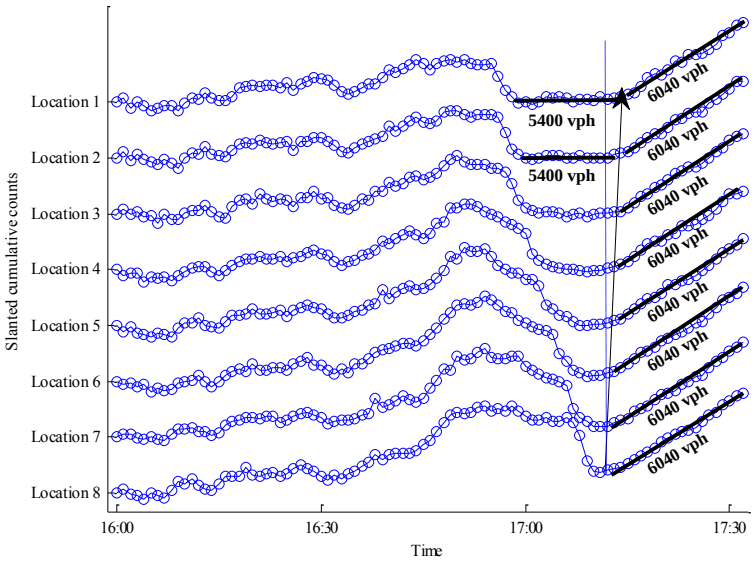
3.4 Results

3.4.1 Empirical observations

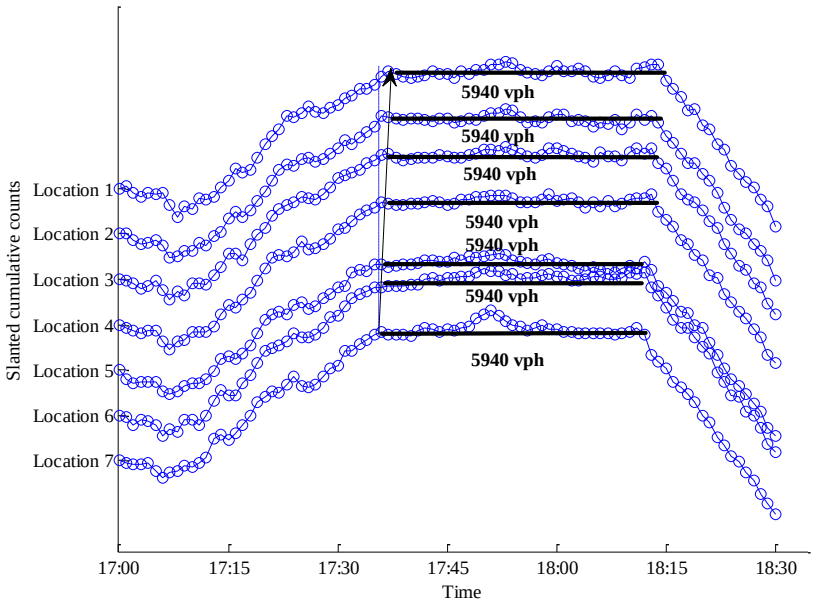
Figure 3.5 presents slanted cumulative counts over three lanes at 8 locations downstream the lane-drop bottleneck in A4 on 18 May 2009 (Figure 3.5(a)) and 7 locations downstream the on-ramp bottleneck in A12 on 18 March 2012 (Figure 3.5(b)). The arrow in each figure shows a clear shock wave which propagates downstream from the bottleneck. Secondly, because the speed before the off-ramp have been over 100km/h (see Figure 3.6), the off-ramp (Exit 7 in A4 and Exit 6 in A12) is believed to have negligible or even no influence on the queue discharge rate.

The empirical observations match the shock analysis in section 3. At the upstream end of the shock wave, we can see the corresponding congestion. Then the speed in corresponding congestion is extracted.

Figure 3.6 shows all the stable discharge rates and the average speed detected on at location 1 on both of three-lane freeways. In Figure 3.6, blue lines stand for speed at location 1 and red dashed lines are the slanted cumulative counts. The black bold lines highlight the stable discharge rates. The value of the discharge rate is attached next to the corresponding black bold line. Note that Figure 3.4(b)) shows on 18 March 2011 several clear stop-and-go waves during the activation period of the on-ramp bottleneck, but all those stop-and-go waves originates near location 7 which is only around 0.5 km away from the bottleneck, that means the discharge rate of those stop-and-go waves only persist for a quite short time and hardly influence the standing queue discharge rate detected at location 1. Therefore, in contrast to the observations in other days, there is only one discharge rate indicated on 18 March 2011 (see Figure 3.6(b)). Figure 3.6 shows there are 12 discharge rates extracted in total, including 7 discharge rates on freeway A4 and 5 discharge rates on freeway A12. The twelve discharge rates and the speed in the corresponding congestion are listed in Table 3.1.

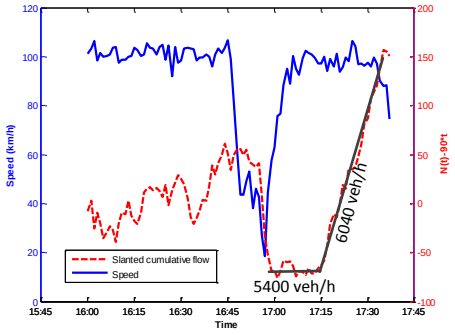


(a) A4, 18 May 2009

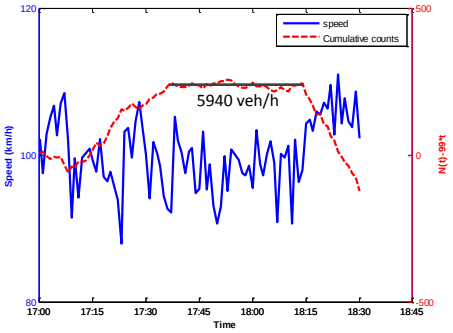


(b) A12, 18 March 2011

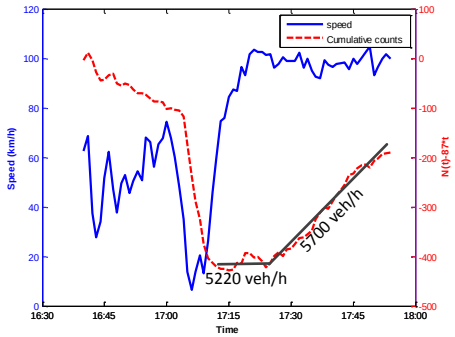
Figure 3.5: Slanted cumulative counts over three lanes at locations downstream the lane-drop (a) and on-ramp (b) bottlenecks on two study days.



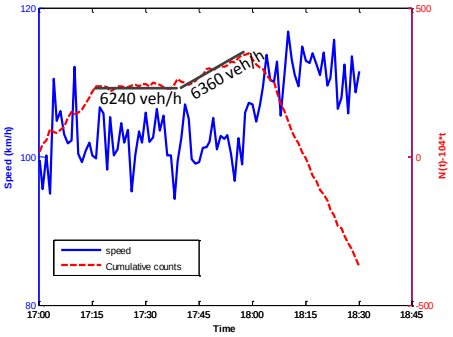
(a) A4, 18 May 2009



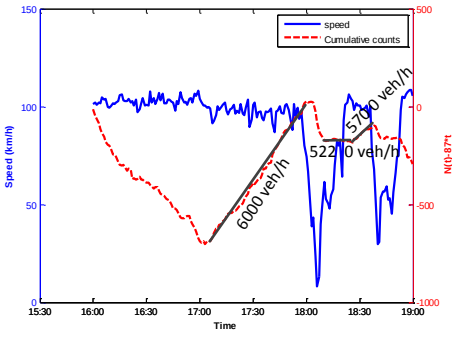
(b) A12, 18 March 2011



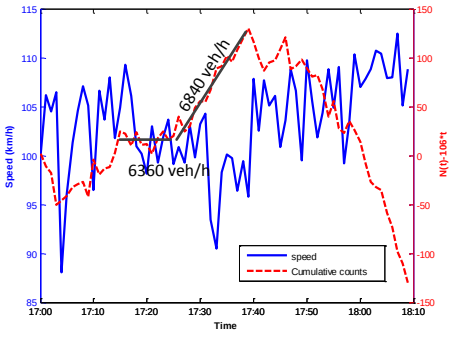
(c) A4, 28 May 2009



(d) A12, 24 March 2011



(e) A4, 11 September 2009



(f) A12, 15 April 2011

Figure 3.6: Discharge rates and the average time mean speed detected at location 1 on different study days at freeway A4 (a, c, e) and A12 (b, d, f).

Table 3.1: Empirical speed in congestion and the outflow of congestion

Freeway	Date	Speed in congestion (km/h)	Queue discharge rate (veh/h)
A4	18 May 2009	13.4	5400
		30.8	6000
	28 May 2009	6.3	5220
		29.2	5700
	11 September 2012	34.0	6000
		7.0	5220
		30.1	5700
A12	18 March 2011	45.0	5940
	24 March 2011	37.6	6240
		48.7	6360
	15 April 2011	48.7	6360
		61.2	6840

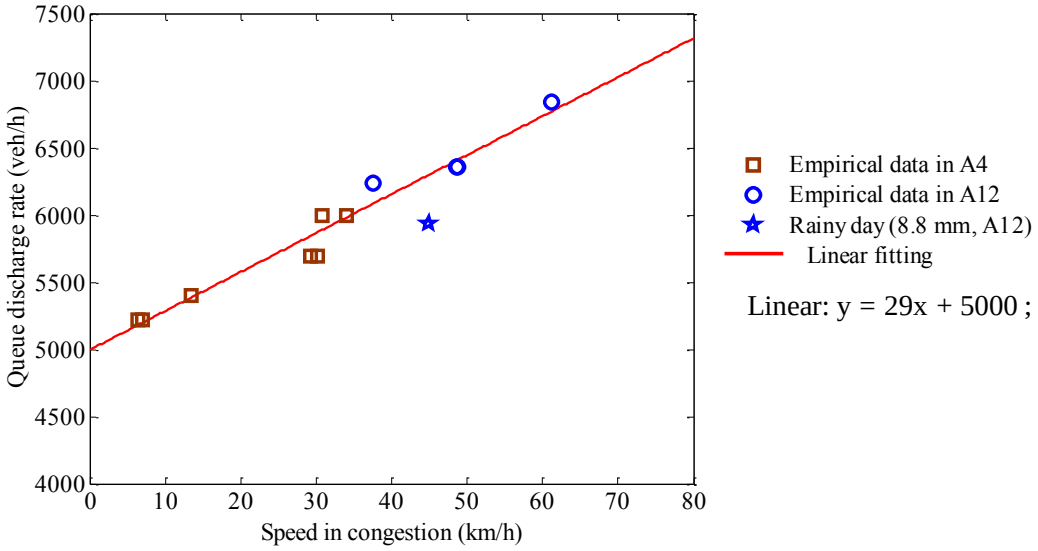


Figure 3.7: Relation between queue discharge rate and the speed in congestion.

3.4.2 Relation between speed in congestion and capacities

Empirical observations provide 12 data points (listed in Table 3.1) to show the relation between the speed in congestion and the corresponding discharge rate or outflow of congestion. Their relation is graphically presented in Figure 3.7. The data collected on sunny days is shown as circles (collected on A4) and squares (collected on A12) while the data presenting the discharge rate on rainy day is shown as a five-point star. The reason there is only 11 circles in Figure 3.7 is that two data points corresponding to a same discharge rate 6360 veh/h overlap (see Table 3.1), which is a remarkable coincidence.

In contrast to previous observations and simulations, this observation shows a broad speed range, from 6 km/h to 60 km/h. In Figure 3.7 within the wide range of speed, the outflow of congestion also range broadly from 5220 veh/h to 6840 veh/h. Note that the observations of the outflow is much higher than that in Muñoz and Daganzo (2002), which might be due to the different traffic flow compositions, different driving regulations, different set up of observations and even different drivers' characteristics in different countries. Meanwhile, the discharge rate in this chapter, for instance 6840 veh/h, can be even substantially higher than the handbook-value for free-flow capacity on a three-lane freeway (with 15% proposition of trucks) 6300 veh/h (Goemans et al., 2011) in the Netherlands. The capacity is estimated through Product Limit Method (Brilon et al., 2005). Though there is no data showing the traffic flow composition in A4 and A12, personal experience shows that the proposition of trucks in A4 and A12 is not as high as 15%. So it is believed that the discharge rate can be influenced considerably by the proposition of trucks. It is even possible that the discharge rate might increase as the proposition of trucks decreases.

At first, the size of the capacity drop is remarkable. The flows go almost as low as 5000 veh/h, which equals almost a 25% capacity drop. Moreover, the measurements from both locations seem to match quite well. There is a clear influence of speed, but apart there is not much noise.

To quantify the influence of speed, this chapter fits a linear function to the empirical data (excluding the one collected in rainy day). The linear function fits the data very well. The correlation coefficient r is 0.9819. The functions are listed in the right of Figure 3.7. Clearly, that the queue discharge rate increases as the speed in congestion increases. Even when the speed in congestion is lower than 20km/h the discharge rate still decreases as the speed in congestion decreases, that is different from the simulation results in Laval and Daganzo (2006).

Because the data in this study is collected from road sections downstream two different bottlenecks, the qualitative trend that the outflow of congestion increases as the speed in congestion increases might can be applied to lane-drop and on-ramp bottlenecks. But the quantitative function could be greatly influenced by the site characters, such as traffic flow compositions and weather, so it is necessary to calibrate the relation in different set up of traffic conditions.

Moreover, the observation in rainy day, shown as the five-point star in Figure 3.7, shows a lower discharge rates than that in days without precipitation. More empirical data analyses are required in the future.

3.5 Conclusion

This chapter reveals a relation between the speed in congestion and the outflow of the queue. This relation shows that as the speed in congestion decreases, the outflow decreases substantially. This research targets empirical data on three-lane freeway. The range of speed in congestion is broad enough, from 6 km/h to 60 km/h. The flow at three-lane section ranges from 5220 veh/h to 6840 veh/h. Compared to previous research on the relation between the congestion levels and the queue discharge rate, this chapter presents sufficiently large empirical observation samples with a broad speed range.

The most important finding is the large influence of the vehicular speed in the upstream congestion on the queue discharge rate. Depending on the speed, the queue discharge rate might drop up to 25%. The qualitative trend of the relation between the speed in congestion and discharge rate could be applied to lane-drop bottleneck and on-ramp bottleneck. In fact, the relation is shown for data collected from these two different bottlenecks. However, the quantitative relation requires calibration because this study finds that the discharge rate is greatly influenced by local traffic situations, such as weather and proposition of trucks. The rainy day in this study shows an exception with a lower queue discharge rate than the other observations. The queue discharge

rate here is also considerably different from other research results in different traffic situations.

Regarding to the causality between the speed in congestion and the queue discharge rate, it is argued here that the speed in congestion determines the queue discharge rate. Because the time headway between two successive vehicles after accelerating cannot change the vehicular speed in a queue. however, the speed in the queue can influence the time headway after speeding up. More investigation on the behavioral understanding of the capacity drop are shown in Chapter 5 and Chapter 6.

In the future, the study of the influence of the relation on the fundamental diagram is relevant, which can lead to a better capacity drop prediction. Meanwhile, it is necessary to see how the other conditions such as number of lanes, slope of freeway and weather influence the relation.

Chapter 4

A kinematic wave model incorporating the capacity drop

A relation between the speed in congestion and the queue discharge rate has been revealed in Chapter 3. This finding can be relevant for traffic simulations and traffic operation evaluations. However, an efficient model to reproduce this finding is still lacking. This chapter proposes an approach of incorporating the capacity drop and its dependence on the speed in congestion into the kinematic wave model, addressing research question 3. The solution to the kinematic wave model is formulated in the Lagrangian coordinates. The new kinematic wave model is believed to be able to give more precise capacity drop magnitude than before.

This chapter is an edited version of the following paper: Yuan, K., Knoop, V. L., & Hoogendoorn, S. P. (2016). A kinematic wave model in Lagrangian coordinates incorporating capacity drop: application to homogeneous road stretches and discontinuities. *Physica A Statistical Mechanics & Its Applications*, 465, 472-485.

4.1 Introduction

Researchers have devoted much effort to finding solutions to congestion problems. The dynamic traffic operation is possibly one of the most cost-effective strategies. Empirical research provides theoretical bases for promising freeway operations. One of those most crucial empirical findings is the capacity drop phenomenon.

Incorporating the capacity drop into traffic flow models is crucial for the evaluation of dynamic traffic operations, as argued in Papageorgiou (1998). Traffic delays strongly depend on the capacity drop. A model which does not include the capacity drop cannot show benefits obtained from some traffic management strategies (e.g., on-ramp metering strategy). Following earlier works on the empirical observations on the capacity drop, we argue that a model that well incorporates the capacity drop should (i) give a queue discharge rate reduction from the pre-queue capacity and (ii) be able to produce a relation between the queue discharge rate and the speed in congestion, see Figure 3.7. The relation is presented in Chapter 3 with loop detector data collected on freeway A4 and A12 in the Netherlands. Giving those two phenomena are relevant for traffic flow models because the accuracy of the reproduced capacity drop influences the values of control variables and even the strategy performances.

However, until now no existing kinematic wave model can capture the two crucial features, and at the same time keep the simplicity of the kinematic wave model. This chapter fills in this research gap, aiming to find out how to incorporate the capacity drop into the kinematic wave model. In this contribution, the kinematic wave model incorporating the capacity drop is solved in Lagrangian coordinates.

The approach used in this chapter is to apply a hysteresis loop for the incorporation of capacity drop. For the introduction of the hysteresis loop, this chapter refers to Treiterer and Myers (1974). Some behavioral interpretations of the hysteresis loop are given in Deng and Zhang (2015) and Laval (2011). The congested branch in the fundamental diagram is divided into acceleration and deceleration branches. A clockwise hysteresis circle in the flow-density fundamental diagram (i.e., the acceleration branch is below the deceleration one) is utilized to describe the process of passing through traffic congestion. The deceleration branch is fixed, overlapping with the original triangular fundamental diagram. So the traffic state evolution at the upstream front of jams is the same as the original kinematic wave model. But the shock wave speed of the acceleration branch is a function of the congestion states. The acceleration and deceleration branch intersect each other at the congested state. The model is expressed in Lagrangian coordinates because the Lagrangian formulation can show more efficiency than the Eulerian coordinates in applications (e.g. traffic state estimation Yuan et al., 2012). In Lagrangian coordinates, traffic characteristics only move in one direction (upwind) to followers independent of traffic conditions. The simplified solution to the Lagrangian formulation of kinematic wave model is called upwind method. The work in this thesis also proposes solutions at nodes to extend the ability of the model to model traffic in a network, which is relevant and necessary for traffic simulations and

control evaluations.

This chapter begins with an introduction of incorporating the capacity drop into macroscopic traffic flow models. The new kinematic wave model formulation and its solutions in Lagrangian coordinates are presented in Section 4.2 and Section 4.3, respectively. Section 4.4 presents the simulation setting and simulation results, followed by Section 4.5 which draws conclusions.

4.2 Model formulation

This section firstly describes the principles of the extended kinematic wave model (Section 4.2.1), which is presented in Eulerian coordinates for an easy access. The new kinematic wave model formulation is presented in Lagrangian coordinates in Section 4.2.2, since this gives more accurate results with less numerical errors.

4.2.1 Principles

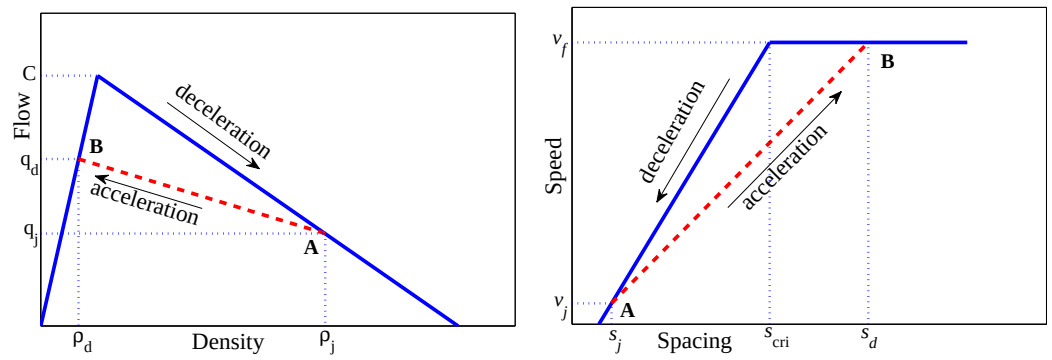
On a homogeneous freeway without entering or exiting traffic, there is a conservation law of vehicles. The conservation law is expressed in terms of an equation relating the traffic flow q, speed v and density ρ on a road section, shown as (4.1) which is also called the conservation equation.

$$\frac{\partial \rho}{\partial t} + \frac{\partial q}{\partial x} = 0 \tag{4.1}$$

t is the time and x is the location. In equilibrium states, traffic flow, density and speed follow:

$$v(x,t) = \frac{q(x,t)}{\rho(x,t)} \tag{4.2}$$

Except the conservation law, the other part of the kinematic wave model is the fundamental diagram. In the Cell Transmission model, which does not give capacity drop, the fundamental diagram is a flow-density relation with a free-flow branch and a congestion branch. To integrate the capacity drop, acceleration branches are added to the triangular flow-density fundamental diagram, which lie below the congestion branch. The congestion branch and the acceleration branch form a clockwise hysteresis loop in the flow-density diagram, as shown in Figure 4.1(a). When vehicles accelerate from congestion, the fundamental flow-density relation follows the acceleration branch which is shown as a dashed line in Figure 4.1(a). In other words, when the speed remains the same or decreases, the fundamental relation will follow the congestion branch in the triangular fundamental diagram, see the solid line in Figure 4.1(a).



(a) Eulerian coordinates (b) Lagrangian coordinates

Figure 4.1: Fundamental diagram with capacity drop in (a) Eulerian and (b) Lagrangian coordinates

Throughout this chapter, the congestion branch is referred to as deceleration branch, which has a slope $-w$. Q denotes the flow-density relation in the deceleration branch, $q = Q(\rho)$, while the acceleration branch is subscripted by a, $q = Q_a(\rho)$.

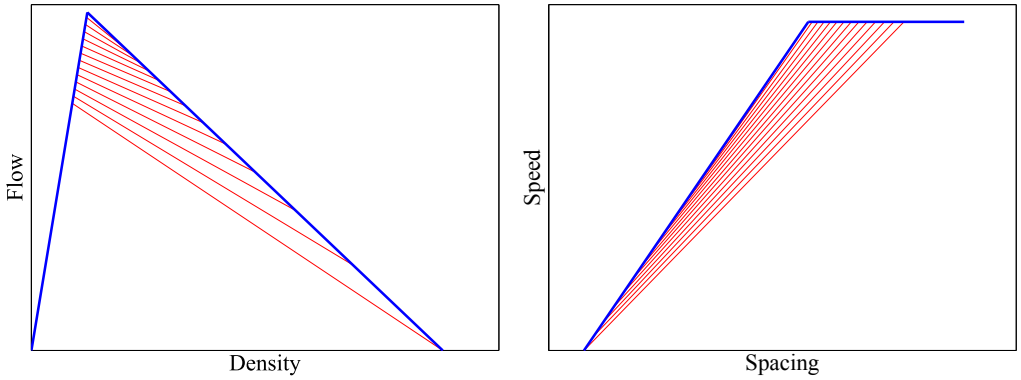
The acceleration branch intersects with both of the free-flow branch and the congestion branch in the triangular flow-density fundamental diagram at point B and A (see Figure 4.1(a)), respectively. State B at the free-flow branch indicates a free-flow state in the downstream of a vehicular queue. The congested state in the queue is shown as state A. The flow detected at traffic state B is the outflow of congested state A, i.e., the queue discharge rate q_d. The clockwise hysteresis loop indicates that the queue discharge rate q_d is lower than the capacity C. That satisfies the definition of the capacity drop.

Traffic state A can be any point in the deceleration branch, indicating different congestion levels. For each state A, there should be an unique acceleration branch which intersects with the deceleration branch at point A. We can have this acceleration branch by calculating the queue discharge rate (corresponding to state B) which is a function of the speed in congested state A, following:

$$q_d(\rho_j) = \min\left(C, \alpha \cdot v_j(\rho_j) + q_0\right) \tag{4.3}$$

which is the relation between the speed in congestion v_j and the queue discharge rate, as suggested by Chapter 3. C is the free-flow capacity, which bounds the queue discharge rate in case of over-high queue discharge rates when v_j approaches v_f. In the linear relation in Equation (4.3), two parameters, α and q_0, are applied. α indicates by how much the traffic jam discharge rate will increase when the speed in the jam grows by 1 unit. q_0 is the queue discharge rate of wide moving jam in which vehicular speed equals zero. The slope of the acceleration branch $-w_a$, which is a variable, is:

$$w_a(\rho_j) = \frac{q_d(\rho_j) - q_j(\rho_j)}{\rho_d(\rho_j) - \rho_j} \tag{4.4}$$



(a) Eulerian coordinates (b) Lagrangian coordinates

Figure 4.2: Correlations between acceleration branches and congestion in (a) Eulerian and (b) Lagrangian coordinates

ρ_j and $q_j(\rho_j)$ is the density and flow in the congestion state A, respectively. $\rho_d(\rho_j)$ and $q_d(\rho_j)$ are the density and flow in the free-flow state B, respectively. $\rho_d(\rho_j) = q_d(\rho_j)/v_f$ and v_f is the free-flow speed.

Equation (4.3) gives different queue discharge rates (or capacity drop) for different levels of congestion. The corresponding acceleration branch can be found with Equation (4.4). Once the value for the α and q_0 are calibrated, all acceleration branches have been predefined. Figure 4.2(a) displays more acceleration branches (shown as light lines), one of which was drawn for each congestion state. In short, this newly proposed kinematic wave model can give both of the capacity drop and the relation between the speed in congestion and the queue discharge rate.

Apart from normal parameters of triangular fundamental diagram, the new kinematic wave model only need to calibrate two parameters, α and q_0. These two parameters have been found in Chapter 3. Note that the assumption of the triangular density-flow fundamental diagram can be relaxed in principle because Equation (4.3) is independent of the shape of the continuous density-flow fundamental diagram.

4.2.2 Formulations in Lagrangian coordinates

This section reformulates the new kinematic wave model in the system of Lagrangian coordinates. The conservation equation which is expressed as Equation (4.1) in Eulerian coordinates is reformulated as Equation (4.5):

$$\frac{\partial s(x,t)}{\partial t} + \frac{\partial v(x,t)}{\partial N} = 0 \tag{4.5}$$

where N is the cumulative number of vehicles, decreasing in space. v is the vehicular speed, which follows a speed-spacing fundamental diagram:

$$v = \begin{cases} V_a(s), & \text{during acceleration} \quad (4.6) \\ V(s), & \text{otherwise} \quad (4.7) \end{cases}$$

with $V_a(s) \neq V(s)$, s is the spacing. The subscript indicates the acceleration branch. The speed-spacing fundamental diagram is shown in Figure 4.1(b). The bold line is the deceleration branch while the dashed line indicates the acceleration branch. The clockwise hysteresis circle in the flow-density diagram is translated into a counter-clockwise circle in the Lagrangian speed-spacing diagram. The deceleration branch is fixed with a slope:

$$\max \left| \frac{\partial v}{\partial s} \right| = w\rho_{\max} \tag{4.8}$$

where $\rho_{\max}$ is the maximum jam density, i.e., speed in the queue is zero km/h. The acceleration branch varies depending on the congested state A. Figure 4.2(b) presents more acceleration branches which are shown as light lines. Corresponding to different congestion states, different acceleration lines are predefined with α and q_0. The maximum speed v_f is the free-flow speed. The minimum spacing is the inverse of the maximum jam density $1/\rho_{\max}$.

The queue discharge rate is formulated into the vehicular spacing downstream of the congestion. The lower the queue discharge rate, the larger the spacing. The coordinates of state B in the Lagrangian coordinate are (s_d, v_f) and

$$s_d = \frac{1}{\rho_d(\rho_j)} \tag{4.9}$$

s_d is the vehicular spacing in the downstream free flow. With (4.3), we can rewrite (4.9):

$$s_d(v_j) = \frac{v_f}{\min(C, \alpha \cdot v_j + q_0)} \tag{4.10}$$

4.3 Solutions to the Lagrangian kinematic wave model

To solve the kinematic wave model, this chapter applies an upwind method in Lagrangian coordinates. In Lagrangian coordinates, Lagrangian clusters only react to their leading Lagrangian clusters. Vehicles are divided into vehicular groups, which are referred to as Lagrangian clusters in this manuscript. Characteristics only move upwind to their followers. Only an upwind method is needed which results in more efficient simulations (Van Wageningen-Kessels, 2013). By the contrast, the minimum supply demand method, which is applied in most of numerical implementations of the LWR model (e.g., the Cell Transmission Model), switches between the upwind and downwind method, that depends on traffic states in the road cells (Van Wageningen-Kessels, 2013). Characteristics propagate in a switch of downstream (upwind) and upstream (downwind). More details about the upwind method in the Lagrangian coordinates can be found in Leclercq et al. (2007), Leclercq (2007) and Van Wageningen-Kessels (2013). For more information about the upwind method in Eulerian coordinates, this chapter refers to Helbing and Treiber (1999).

This section describes the solutions to the Lagrangian kinematic wave model on links in Section 4.3.1, followed by Section 4.3.2 presenting the solutions at nodes. Section 4.3.3 briefly discuss the numerical diffusion.

4.3.1 Links

This section firstly describes the upwind method in Lagrangian coordinates on links briefly. Then the speed updating process is adjusted for solving the Lagrangian kinematic wave model.

Basic upwind method

Time are divided into time steps. Each Lagrangian cluster, indicated by a subscript i, is characterized by three quantities: number of vehicles ΔN, spacing $s(k)$ and speed $v(k)$. k is the time instant, $t = k \cdot \Delta t$, and Δt is the time step.

Lagrangian cluster i follows Lagrangian cluster $i-1$. The upwind method is expressed in time discretization as:

$$s_i(k+1) = s_i(k) + \frac{\Delta t}{\Delta N} \cdot (v_{i-1}(k) - v_i(k)) \tag{4.11}$$

The numerical scheme (4.11) is equivalent to the following expression:

$$x_i(k+1) = x_i(k) + v_i(k) \cdot \frac{\Delta t}{\Delta N} \tag{4.12}$$

which updates the new position x of the Lagrangian cluster i.

For stability and convergence, the time step Δt should satisfy the Courant-Friedrichs-Lewy (CFL) condition as shown in (4.13):

$$\frac{\Delta N}{\Delta t} \geqslant \max \left| \frac{\partial v}{\partial s} \right| \tag{4.13}$$

For the interpretation of the CFL condition, this chapter refers to Van Wageningen-Kessels (2013). If the CFL condition is satisfied as an equality, the upwind method can be free of numerical errors (Leclercq et al., 2007).

Speed updating process

The Lagrangian cluster speed updating process is presented in Figure 4.3. V_i^* is the default fundamental speed-spacing relation, following which Lagrangian cluster i updates its speed at present. V_i^* is updated every time step, too. This process firstly updates the Lagrangian cluster's position and spacing. Then, it checks which speed-spacing branch the Lagrangian cluster follows. This checking process is introduced

with a dashed rectangle in Figure 4.3. If $v_i(t) = v_f$, the speed updating process should follow the deceleration branch $V_i^* = V$ (Process 4). If not, firstly the process is going to update the speed with the default speed-spacing relation which has been updated in the last time step. Secondly, this work checks whether an acceleration branch should be or have been activated. If the acceleration branch has been activated, then the default relation V_i^* remains the same as that in the last time step (Process 1). If the acceleration branch should be but has not been activated yet, then it means $v_i(t)$ is the speed in congestion. The default relation will be updated into an acceleration branch (Process 2). If the acceleration should not be activated, the default relation V_i^* equals V (Process 3). Except process 1, the rest processes will update the Lagrangian cluster speed following the newly updated speed-spacing relation V_i^*. Process 1 and process 4 only update speed for one time, while process 2 and process 3 need to update speed twice. For process 2 and 3, the first speed updating checks whether the cluster is speeding up, and the second speed updating determines by how many the speed should increase.

4.3.2 Nodes

A freeway network consists of links and nodes. Section 4.3.1 only considers links, or homogeneous road sections. This section discusses how to solve the Lagrangian kinematic wave model at nodes (a lane-drop (Section 4.3.2) and on-ramp (Section 4.3.2) nodes). Van Wageningen-Kessels (2013) presents a methodology of modeling nodes using the basic kinematic wave model in Lagrangian coordinates. Based on Van Wageningen-Kessels (2013) and Van Wageningen-Kessels et al. (2013b), this section presents how to model nodes in Lagrangian coordinates when incorporating the capacity drop.

Lane-drop nodes

When a Lagrangian cluster reaches a lane-drop node, a change of the fundamental diagram occurs. Usually this change occurs when the Lagrangian cluster passed the node rather than the beginning or end of a time step. This work applies the new fundamental diagram to the Lagrangian cluster at the end of the time step during which the Lagrangian clusters the node. With different fundamental diagrams, the parameters α and q_0 change, accordingly.

Meanwhile, the speed in congestion $v_j(\rho_j)$ does not change. If one Lagrangian cluster passes the lane-drop node during acceleration process (or the Lagrangian cluster keeps accelerating at the end of the time step during which the Lagrangian clusters the node), when drawing the acceleration branch with (4.10) for the downstream road section, we should still apply the speed in the queue which is in the upstream of the node.

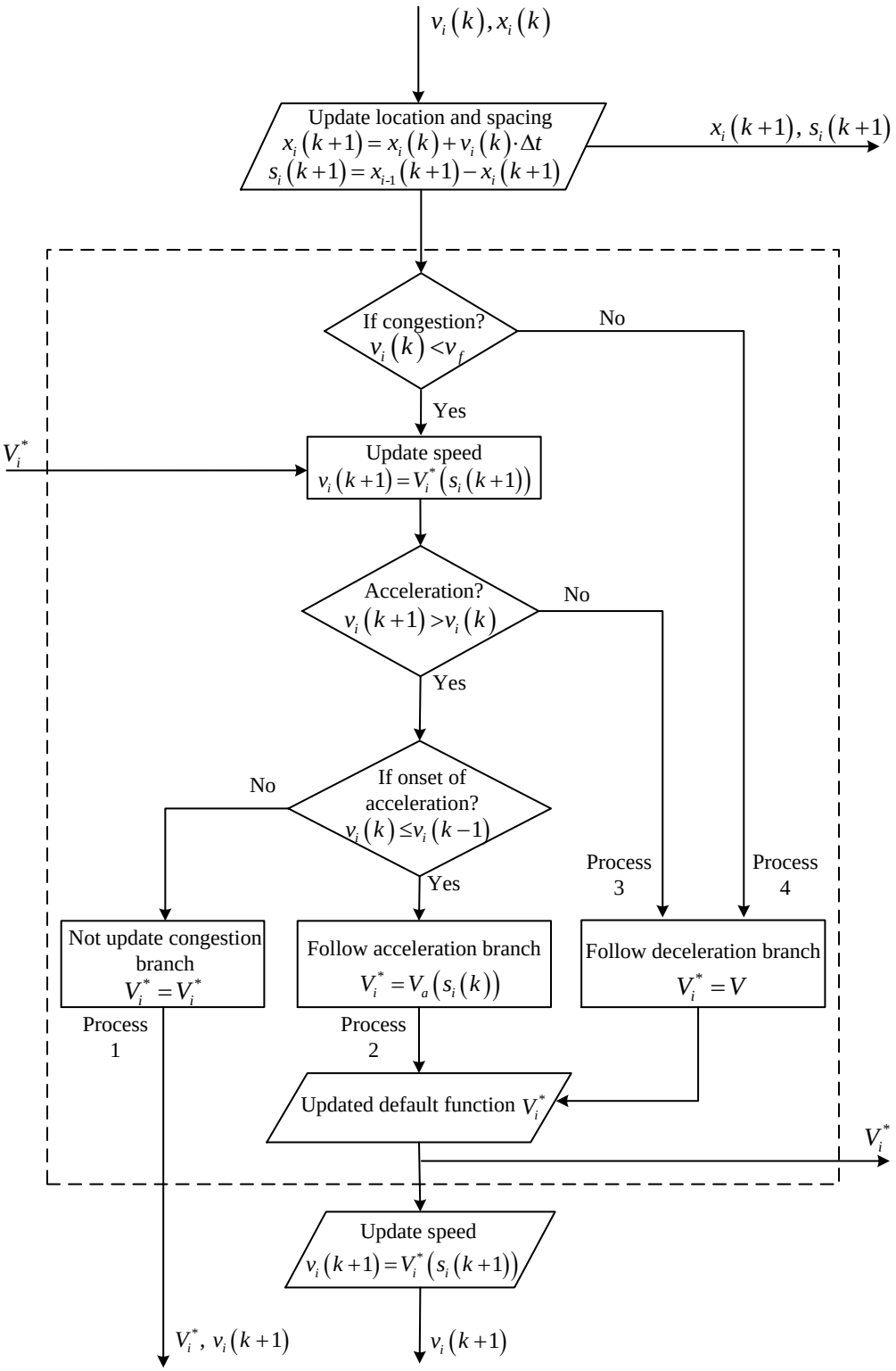


Figure 4.3: The process of updating Lagrangian cluster speed

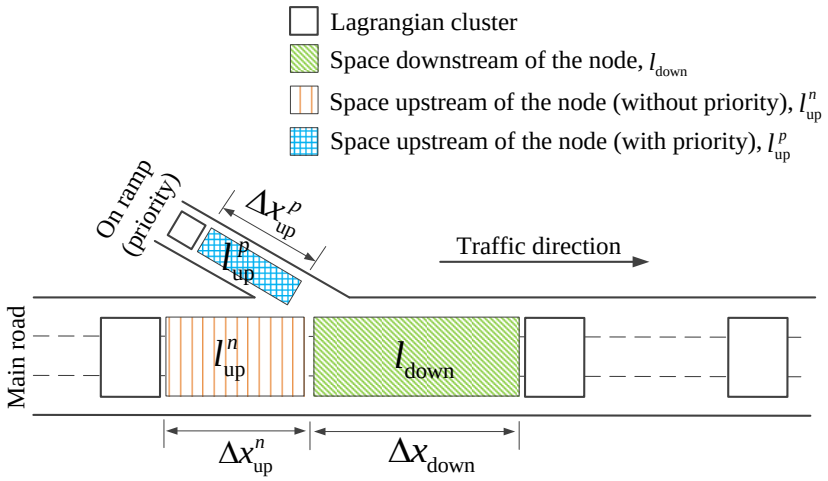


Figure 4.4: An illustration of available space

On-ramp nodes

The on-ramp node is a merging node. Without incorporating capacity drops, two issues occur. Firstly, on-ramp nodes add vehicles into the flow on the main road. Secondly, when there is a conflict between the merging Lagrangian cluster and the Lagrangian cluster on the main road, a merging ratio determines which Lagrangian cluster (i.e., the one from the main road or the one from the on ramp) can merge into the main road more easily. Next, both issues are discussed.

For the first issue, Van Wageningen-Kessels (2013) addresses it by modeling that the trajectory of every Lagrangian cluster is continuous. When the number of vehicles on the on-ramp reaches ΔN, the Lagrangian cluster is added to the flow on the freeway. For the second issue, Van Wageningen-Kessels et al. (2013b) distribute available space between incoming Lagrangian clusters on the main road and the on ramp. For more elaborations, this chapter refers to Van Wageningen-Kessels (2013) and Van Wageningen-Kessels et al. (2013b). The approach applied here to solve the Lagrangian kinematic wave model at on-ramp nodes is comparable to the contributions in Van Wageningen-Kessels (2013) and Van Wageningen-Kessels et al. (2013b). This section briefly presents the distributing approach with equations.

In this work, the "space" denotes the room within a certain spacing or the spacing in each lane, that is spacing $\times$ number of lanes on one road section. It is the spacing that is used for calculating the speed of the Lagrangian cluster. As illustrated in Figure 4.4, the available space at the on-ramp node consists of three space: the space on the outgoing link l_{down}, the space upstream of the node on the link with priority l^p_{up} and the space upstream of the node on the other link l^n_{up}. The superscript p indicates the parameter for the Lagrangian cluster with priority while the n denotes the parameter for the Lagrangian cluster without priority. The space l^p_{up} and l^n_{up} are only available to the Lagrangian cluster with and without priority, respectively. The space l_{down} is

distributed according to the priority. For the Lagrangian cluster with priority, the solution distributes as much as critical space $\Delta N s_c^p m^p$ to the Lagrangian cluster with priority, including l_{up}^p. s_c^p is the critical spacing on the incoming link with priority. If $l_{\text{down}} + l_{\text{up}}^p < \Delta N s_c^p m^p$, all space $l_{\text{down}} + l_{\text{up}}^p$ is distributed to the Lagrangian cluster with priority. Then, the distributed space is divided by $\Delta N m^p$ to get the spacing s_i^p, that is:

$$s_i^p = \frac{\max\left(l_{\text{up}}^p, \min\left(\Delta N s_c^p m^p, l_{\text{up}}^p + l_{\text{down}}\right)\right)}{\Delta N m^p} \tag{4.14}$$

m^p is the number of lanes on the link with priority. m_{down} is the number of lanes on the outgoing link. $l_{\text{up}}^p = \Delta x_{\text{up}}^p m^p$ and $l_{\text{down}} = \Delta x_{\text{down}} m_{\text{down}}$. Δx_{up}^p is the distance between the most downstream Lagrangian cluster on the link with priority and the node. Δx_{down} is the distance between the node and the most upstream Lagrangian cluster on the outgoing link, see Figure 4.4.

The remaining space is distributed to the Lagrangian cluster without the priority. Hence, the Lagrangian cluster without priority gets all the remaining space $l_{\text{up}}^p + l_{\text{up}}^n + l_{\text{down}} - \Delta N m^p s_i^p$. The speed of the Lagrangian cluster without the priority is calculated according to the distributed spacing. The spacing for the Lagrangian cluster without priority equals to the distributed space divided by $\Delta N m^n$. That is, the spacing of the Lagrangian cluster without priority is calculated as follows:

$$s_i^n = \frac{l_{\text{up}}^p + l_{\text{up}}^n + l_{\text{down}} - \Delta N m^p s_i^p}{\Delta N m^n} \tag{4.15}$$

m^n is the number of lanes on the incoming link without priority. $l_{\text{up}}^n = \Delta x_{\text{up}}^n m^n$ is the available space on the incoming link without priority. Δx_{up}^n is the distance between the most downstream Lagrangian cluster on the link without priority and the node.

The distribution of the priority depends on the merging ratio δ. Priority is given to the on ramp when $N_n/N < \delta$. Otherwise the main road gains the priority. N_n is the number of Lagrangian clusters from the on ramp within the last N Lagrangian clusters recently passing the node.

After calculating the spacing of each Lagrangian cluster from different upstream links, the Lagrangian kinematic wave model applies the process diagram in Figure 4.3 to update Lagrangian cluster speed on the main road. Since the fundamental diagram does not change, which is different from the lane-drop node, the queue discharge rate in response to the standing queue can be given accordingly.

4.3.3 Numerical errors

In the Lagrangian kinematic wave model, a temporary change in the fundamental diagram is caused by a change of traffic situations. The change of the fundamental diagram indicates that the CFL condition cannot be satisfied as an equality all the time.

When applying the upwind discretization method in the Lagrangian coordinate to solve the kinematic wave model, numerical errors arise inevitably. Without capacity drop,

the upwind method can be free of numerical errors by setting an equal sign in CFL condition (Leclercq et al., 2007), see (4.16).

$$\Delta t = \Delta N \cdot \max \left| \frac{\mathrm{d}S(v)}{\mathrm{d}v} \right| \tag{4.16}$$

$S(v)$ is the inverse of $V(s)$, i.e., $S(v) = V^{-1}(s)$.

However, in the extension the slope of the acceleration branch $V_a(s)$ is smaller than that of the deceleration branch $V(s)$, see Figure 4.1(b), so:

$$\Delta t < \Delta N \cdot \max \left| \frac{\mathrm{d}S(v)}{\mathrm{d}v} \right| \tag{4.17}$$

The fact that there is no equal sign in (4.17) means that when vehicles accelerate from congestion, numerical errors will be inevitable.

4.4 Simulation

This chapter simulates the new kinematic wave model with adaptive capacity drop. The simulation wants to highlight the capability of reproducing the two important phenomena which have been described in Section 4.1. Simulation set-up and results are presented in Section 4.4.1 and Section 4.4.2, respectively.

4.4.1 Simulations set-up

In simulations, three scenarios are designed, see Table 4.1. In different scenarios, different road structures are simulated. In scenario 1, all cases are studied on a three-lane homogeneous road section. The lane-drop node and the on-ramp node are studied in scenario 2 and 3, respectively. The case descriptions in Table 4.1 are also shown in Figure 4.5, which shows traffic states we expect to find.

In the first scenario, a three-lane homogeneous freeway section and two stop-and-go waves are considered: heavy congestion J_{slow} (density $\rho_{\text{slow}} = 400$ veh/km) and light congestion J_{fast} (density $\rho_{\text{fast}} = 200$ veh/km). In case 1, the heavy congestion is in the upstream of light congestion, while in case 2, the heavy congestion is in the downstream of light congestion. Figure 4.5(a) and Figure 4.5(b) show shock wave analyses in case 1 and case 2, respectively. State 1 and 2 indicate the stop-and-go wave J_{slow} and J_{fast}, respectively. The darker shading in J_{slow} indicates a higher density than that in J_{fast}. State 3 and 4 are the downstream free-flow states of state 1 and 2, respectively. Flow in state 3 is lower than that in state 4. State 5 indicates the capacity. The shock wave analysis predicts the difference on the traffic state evolution between those two case simulations, which should be expected in our simulations. This difference indicates the importance of simulating capacity drop accurately.

Table 4.1: Simulation scenarios

Scenarios	Road structures	Descriptions
1	Homogeneous road	Case 1: J_{slow} upstream of J_{fast} Case 2: J_{fast} upstream of J_{slow}
2	Lane-drop node	Case 3: After J_{slow} passing through the node, a standing queue is emerged.
3	On-ramp node	Case 4: A standing queue is emerged due to the high total demand from both of the main road and the on ramp.

To generate stop-and-go waves, a spacing profile over time of the first Lagrangian cluster is predefined. For the first Lagrangian cluster, the spacing is set to be critical spacing s_{cri} (which is the critical spacing for a three-lane freeway) in free-flow state while the spacings in congestion are 2.5 m (in J_{slow}) and 5 m (in J_{fast}) respectively. Note that, since there are 3 lanes, 2.5 m spacing means the space is 2.5×3=7.5m.

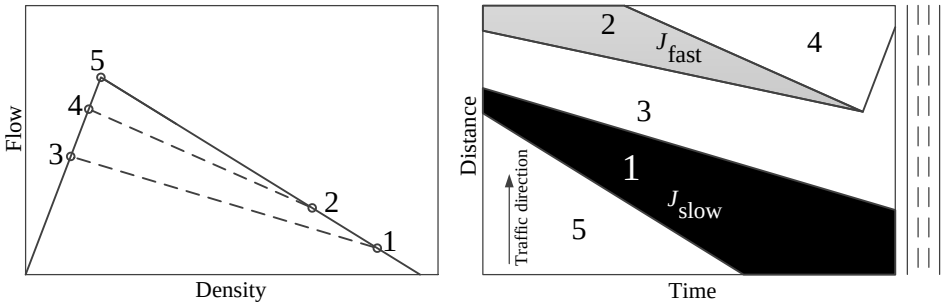
The remaining two scenarios simulate traffic situations with nodes. In the second scenario, a lane-drop node, which connects an upstream four-lane freeway section and a three-lane freeway section, is considered. In the third scenario, the traffic model is simulated around an on-ramp node on a three-lane freeway.

In case 3, a standing queue forms at the lane-drop node after a stop-and-go wave passes through. A shock wave analysis for the case 3 is presented in Figure 4.5(c). The stop-and-go wave J_{slow} is shown as state 1. The downstream state of J_{slow} is state 3. A standing queue is state 8, whose queue discharge rate is shown as state 6. A shock wave between state 3 and state 6 is expected to distinguish different queue discharge rates of different queues. In case 4, a standing queue (shown as state 9 in Figure 4.5(d)) forms when the on-ramp flow enter into the main road where the density has already reached a critical density. The shock wave analyses for case 4 is shown in Figure 4.5(d). The studied cases around nodes in this chapter have been empirically observed in Chapter 2 and Chapter 3. In Case 4, the on ramp is one-lane and the merging ratio $\delta = 0.35$. In both of case 3 and case 4, the nodes are located at $x = 0$ m.

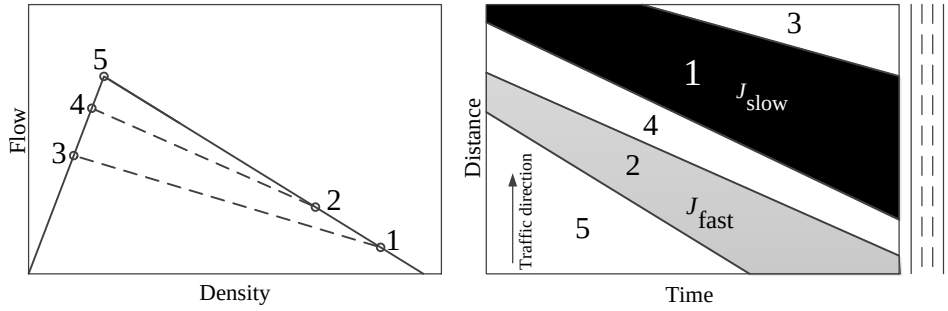
In all cases, at the beginning of simulations ($t = 0$ s), the spacing of all Lagrangian clusters on the main road are set to be s_{cri}.

The fundamental diagrams in our simulations are summarized in Table 4.2. In all cases, $\Delta N = 1$. To ensure the stability, this chapter sets Δt following (4.13):

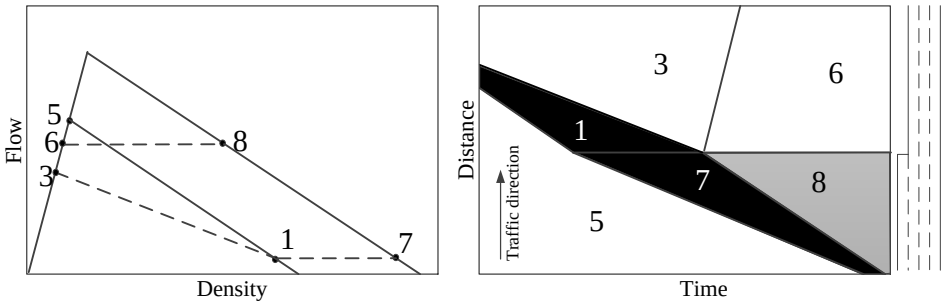
$$\Delta t \approx \begin{cases} 0.45s, & \text{in scenario 1} \quad (4.18) \\ 0.34s, & \text{in scenario 2} \quad (4.19) \\ 0.45s, & \text{in scenario 3} \quad (4.20) \end{cases}$$



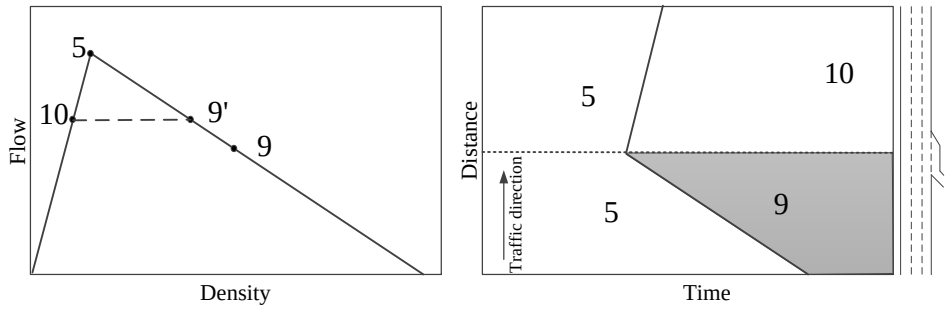
(a) Case 1: Heavy congestion J_{slow} upstream



(b) Case 2:Heavy congestion J_{slow} downstream



(c) Case 3: Lane-drop node scenario



(d) Case 4: on-ramp node scenario

Figure 4.5: Shock wave analysis in (a) case 1, (b) case 2, (c) case 3 and (d) case 4. In each case, the left figure shows the fundamental diagram and the right one the speed/density contour

Table 4.2: Fundamental diagrams

Road section	Fundamental Diagram	α	q_0
3-lane section	$v_f = 114$km/h, $C = 6840$veh/h, $\rho_{cri} = 60$veh/km, $w = 18$km/h	29veh/km	5000veh/h
4-lane section	$v_f = 114$km/h, $C = 9120$veh/h, $\rho_{cri} = 80$veh/km, $w = 18$km/h	39veh/km	6667veh/h
On ramp	$v_f = 114$km/h, $q_c = 2280$veh/h, $\rho_{cri} = 20$veh/km, $w = 18$km/h	N/A	N/A

Δt calculated by (4.13) has over 2 decimal digits, so only rounding is shown here.

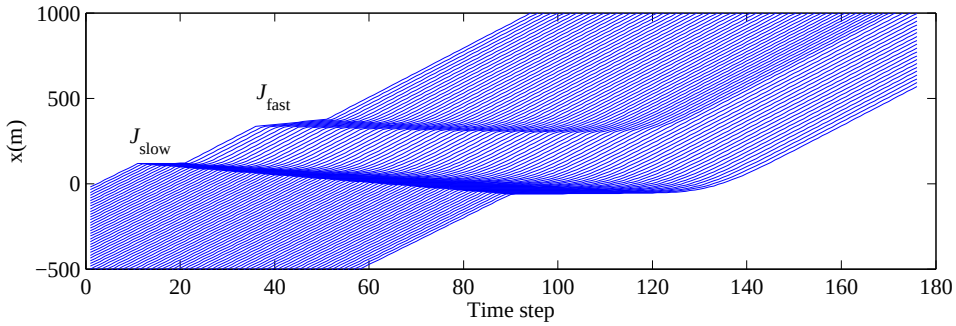
4.4.2 Simulations results

When simulating in Lagrangian coordinates, we can draw the trajectories of each Lagrangian cluster, as presented in Figure 4.6. A space contour plot (see Figure 4.7 and Figure 4.8(a)) shows different traffic states in space-time. The simulation results all fit the expectations shown in Figure 4.5.

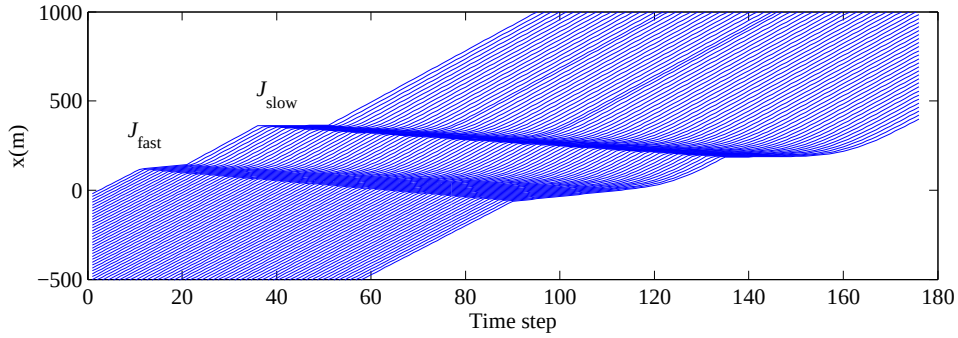
In Figure 4.6, the blue lines are the Lagrangian clusters. The queue discharge rate of J_{slow} is 5052 veh/h (capacity drop = 26%) and that of J_{fast} is 5626 veh/h (capacity drop = 18%). In Case 1, the queue length of J_{slow} increases while the jam J_{fast} diminishes. That is because for J_{slow} the inflow (equals to capacity) is higher than the outflow (i.e., the definition of capacity drop), while the queue discharge rate of J_{fast} is lower than the inflow which is the outflow of J_{slow}. In case 2 both of queues grow because the queue discharge rates of both jams are lower than their inflows. Moreover, simulations indicate that the lower the speed limit, the more difficult to dissolve the triggered congestion, i.e., J_{slow} requires lower inflow to diminish than J_{fast}. The simulation results in Figure 4.6 is the same as expected in Figure 4.5(a) and Figure 4.5(b). The numerical errors predicted in Section 4.3.3 are shown as sluggish accelerations in Figure 4.6.

The simulation results in case 3 are shown in the spacing contour plot in Figure 4.7. The spacing decreases as the color goes darker, see the legend in Figure 4.7. It is shown that as a stop-and-go wave passes the lane-drop node, a standing queue forms. In Figure 4.7, The congestion is shown as dark colour. In the downstream of congestion, a dashed line distinguishes the state downstream of the stop-and-go wave and the one downstream of the standing queue. The shock wave shown as dashed line has been expected in Figure 4.5(c). Note that, this shock wave has been frequently observed on freeways, see Chapter 2 and Chapter 3.

Finally, Figure 4.8 shows the simulation results on an on-ramp node. Figure 4.8(a) shows the spacing contour. To better understand the model, the trajectories of clusters



(a) Case 1: Low-speed jam J_{slow} upstream



(b) Case 2: High-speed jam J_{fast} upstream

Figure 4.6: Simulation results in (a) case 1 and (b) case 2 in Lagrangian coordinates

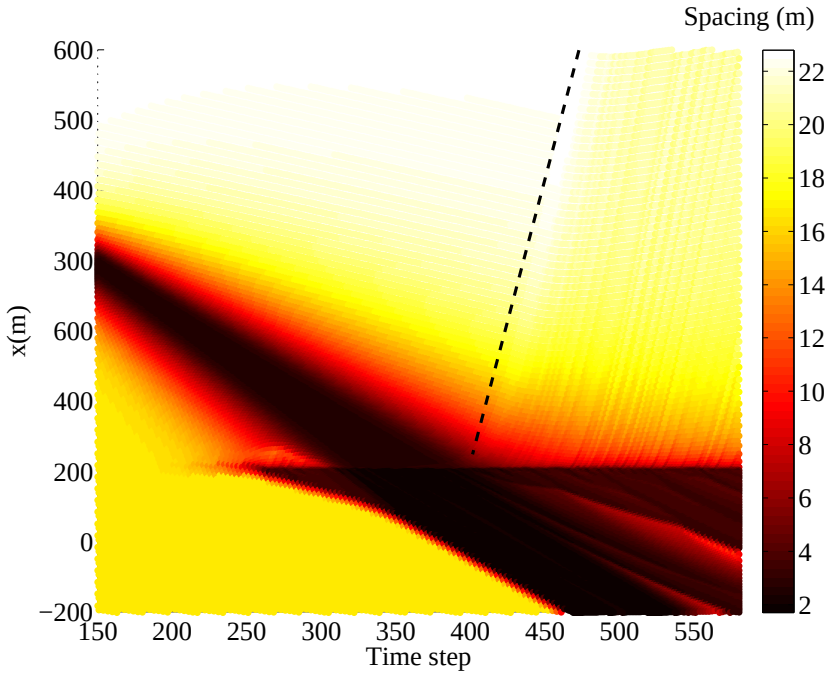


Figure 4.7: Spacing contour in case 3: simulations around a lane-drop node

from both of main road and the on ramp in a rectangle area (highlighted by a blue rectangle) in Figure 4.8(a) are shown in Figure 4.8(b), where the dashed lines are the trajectories of on-ramp clusters while the solid ones are the clusters from the main road. Around time step 60, the first Lagrangian cluster merges into the main road. A standing queue begins forming from that time. In Figure 4.8(a), before the merging of the first Lagrangian cluster from the on ramp, the spacing between clusters are critical spacing shown as orange. Since the onset of a standing queue, in the downstream of the standing queue, the spacing between Lagrangian clusters are larger than the critical spacing.

It is noticed that every time when a Lagrangian cluster passes the bottleneck, the spacing distributed to those clusters can be different, that results in different congested states in the upstream of bottlenecks. As shown in Figure 4.6, the sequence of J_{fast} and J_{slow} on links can influence the traffic state transition. So it is shown in Figure 4.7 and Figure 4.8 that different traffic state transitions can appear.

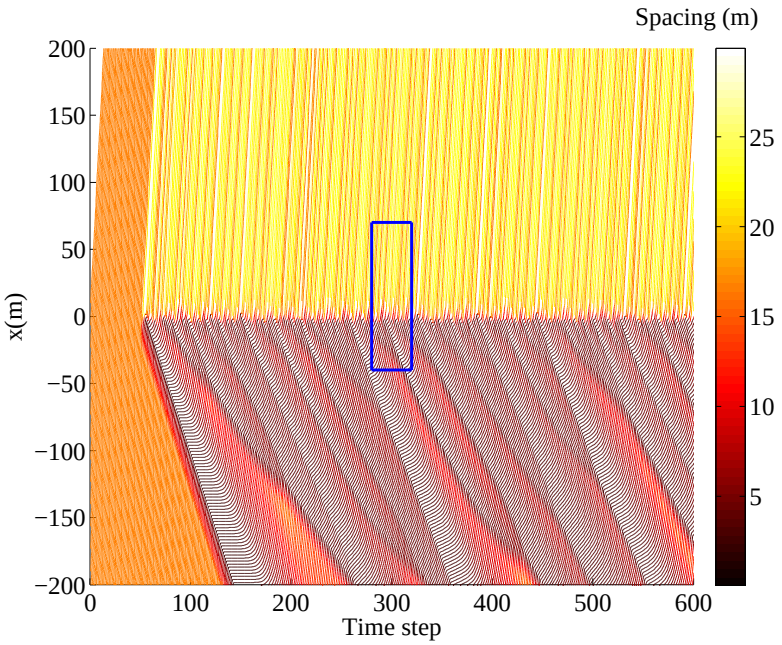
In summary, the new model works when reproducing the capacity drop and the relation between the congestion and its queue discharge rate.

4.5 Conclusion

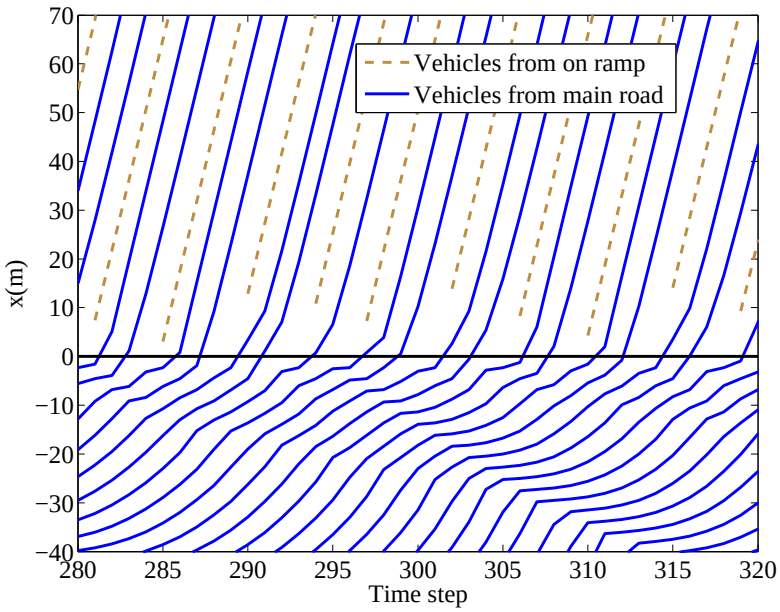
This chapter presents a new kinematic wave model incorporating the capacity drop phenomenon to improve the validity and applicability of the classical kinematic wave model. It is solved in Lagrangian coordinates with an upwind scheme for its simplicity and high accuracy, compared to the Eulerian coordinates. The simplicity due to that the traffic character propagates only in one-direction to followers in Lagrangian coordinates. The high accuracy is shown as smaller numerical errors, compared to the Eulerian coordinates.

This new model uses predefined hysteresis loops to capture the capacity drop – clockwise in Eulerian coordinates and counter clockwise in Lagrangian coordinates. The loop is determined by the an empirical relation between the speed in congestion and the queue discharge rate. In contrast to other first order models, it can reproduce the capacity drop and its important features, such as the relation between the speed in congestion and the queue discharge rate. The model is applied to jams on both of a homogeneous road section and nodes (including lane-drop and on-ramp nodes) in this work, which means it is possible to apply the model in a traffic network. More efforts will be made in the future to give a diverging node model.

The Lagrangian kinematic wave model is suitable for traffic applications, such as traffic state estimations, traffic delay calculations and model based dynamic traffic management. The Lagrangian kinematic wave model can contribute to accurate knowledge of accumulated vehicles by providing accurate capacity drop magnitude, which benefits traffic state estimations and traffic delays calculations greatly. The increasing accurate



(a) Spacing contour in case 4: simulations around an on-ramp node



(b) Trajectories of Lagrangian clusters in case 4: simulations around an on-ramp node

Figure 4.8: Simulation results in case 4 is shown as (a) spacing contour and (b) trajectories of Lagrangian clusters

knowledge of the traffic system benefits active traffic management. For locations with speed limits and ramp meter installations, the dynamic speed limit and ramp metering rate can be more precise and situation-oriented. For modeling speed limits, we can revise the fundamental spacing-speed relation by reducing the free-flow speed in the new model. Note that, in the upstream of the speed limited road section, a stop-and-go wave might form if the upstream demand is high enough. This model can give accurate capacity drop for this stop-and-go wave, which is relevant for evaluating the speed limit consequences and determining the dynamic speed limit values.

In the future, to what extent the numerical errors, revealed in this chapter, will influence the model performance should be investigated.

Chapter 5

Impacts of inter-driver spread and intra-driver variation on capacity drop

Previous chapters show that the queue discharge rate is strongly related to the congested state. This finding indicates a possibility of maximizing the queue discharge rate through traffic operations. The research about proposing an efficient strategy to increase the queue discharge rate, needs an investigation into the longitudinal behavioral mechanism behind the capacity drop and its dependence on the congestion. In this chapter, efforts are devoted to understanding the mechanism. Firstly, analytical investigations into the impact of bounded acceleration spread and reaction extension are conducted, respectively. Secondly, numerical simulations are carried out to confirm the analytical results and investigate the combined effects of the acceleration spread and the reaction time extension.

This chapter is an edited version of the following paper: Yuan, K., Knoop, V.L., & Hoogendoorn, S.P. (in print). A microscopic investigation into the capacity drop: impacts of longitudinal behavior on the queue discharge rate. *Transportation Science*.

5.1 Introduction

The capacity drop is usually observed at fixed bottlenecks, especially merging bottlenecks (e.g., on-ramp and lane-drop bottlenecks). Before the standing queue forms (i.e., before a break down) at a fixed merging bottleneck, loop detectors can measure flow approaching the capacity. So empirical data at fixed merging bottlenecks can show the reduction from the capacity to the queue discharge rate, i.e., the capacity drop phenomenon. However, a reduction of capacity does not only occur at bottlenecks. In fact, some research (Kerner, 1998; Oh & Yeo, 2015) and Chapter 2 report on queue discharge rates out of stop-and-go waves on homogeneous freeway sections, which are also lower than the capacity. It is worthwhile to know that the mechanism behind the capacity drop accompanying stop-and-go waves could differ from that behind the bottleneck capacity drop. This chapter argues that the difference is the void created on the targeted lane due to the merging behaviors (e.g., from on ramp). That is, the bottleneck capacity drop (take on-ramp bottleneck as an example) can be seen as a combination of two types of reductions: (1) reduction due to voids created by the merging behavior from the on ramp and (2) capacity drop accompanying stop-and-go waves. Since there could be lane changing in stop-and-go waves, this chapter categorizes the mechanism of the capacity drop into lateral mechanisms (i.e., the capacity drop due to lane changing behavior) and longitudinal mechanisms (i.e., the capacity drop due to longitudinal behaviors).

Oh and Yeo (2015) show that the contribution of traffic passing stop-and-go waves to discharge rates outperforms the contribution of lane change related traffic to discharge rates. This chapter can hence be seen in the perspective shaped by Oh and Yeo (2015), and therefore focus on the study on the longitudinal behavior mechanism behind the capacity drop.

As shown in the Chapter 1, there are two categories of longitudinal behavior mechanism which can reduce the queue discharge rate: (1) inter-driver spread and (2) intra-driver variation. The inter-driver spread mechanism shows that the driver character composition, such as timid or aggressive behaviors, might be important for a model intending to give the capacity drop. While the intra-driver variation mechanism indicates that it is crucial to understand and model drivers' different behavior patterns in different traffic conditions in attempts of giving the capacity drop phenomenon. It is worthwhile to know that model-based simulations are frequently used to understand the traffic dynamic. So different approaches of modeling the longitudinal mechanism behind the capacity drop might influence the understanding of traffic dynamics and operations. Therefore, it is relevant to distinguish these two longitudinal mechanisms when understanding the capacity drop phenomenon from a driver behavioral perspective.

This chapter investigates to what extent each of these two mechanisms contributes to the queue discharge rate reduction. As described in Section 1.3.3, the inter-driver spread means the inter-driver acceleration spread among vehicles, while the intra-

driver variation can be understood in several driver behavioral perspectives (e.g., driver time headways, reaction pattern). This chapter understands the intra-driver variation from a perspective of reaction time. The reaction time indicates how long a following vehicle needs to take to react to the change of its leader's driving behavior. Voids can also be created if the follower's reaction time is longer than Newell's reaction time. In this chapter, the long reaction time is called the extended reaction time. The difference between the extended reaction time and the Newell's reaction time is called the reaction time extension. At present, to what extent the inter-driver acceleration spread and the extended reaction time contribute to the capacity drop are unknown. In this chapter, the impacts of the acceleration spread and the extended reaction time on the queue discharge rate are studied.

This chapter develops analytical models to investigate the independent impact of inter-driver acceleration spread and reaction time extension in relation to longitudinal driving behavior. Furthermore, numerical experiments are designed. First, the experiment is used to validate the analytical model to verify the accuracy of the approximation in the analytical model. Second, the experiment is used to see the combined effects of the inter-driver acceleration spread and reaction time extension on the queue discharge rate. The empirical relation revealed in Chapter 3 is the reference used in our analyses, see Figure 3.7.

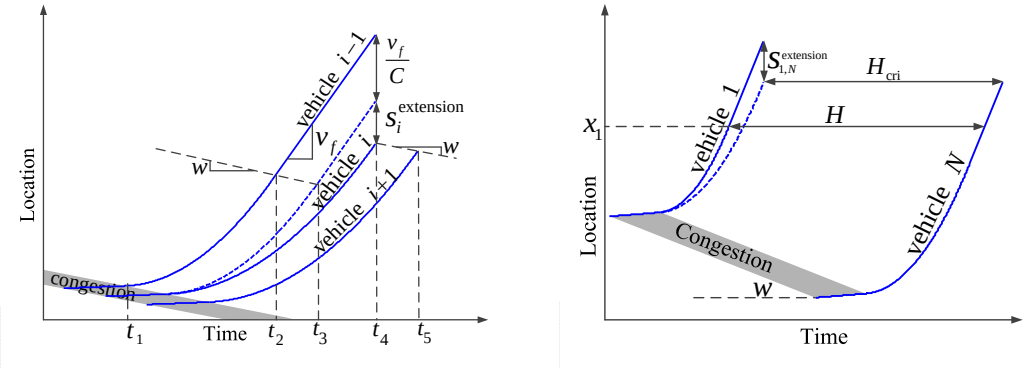
In this study, the driver-vehicle-combination (DVC) heterogeneity is not considered. It is argued in this thesis that the driver-vehicle-combination (DVC) heterogeneous is not a necessary condition for reducing the queue discharge rate. Behaviors of all studied vehicles can be described with a same fundamental diagram.

The outline of this chapter is as follows: Section 5.2 presents analytical investigations on the independent impact of the inter-driver acceleration spread and the reaction time extension on the capacity drop. In Section 5.3, numerical experiments are conducted for validating the analytical model accuracy and investigating the combination of the inter-driver acceleration spread and reaction time extension. Finally, this chapter ends with discussions and conclusions in Section 5.4.

5.2 Analytical investigations

This section analytically investigates to what extent the inter-driver acceleration spread (Section 5.2.1) and reaction time extension (Section 5.2.2) can independently reduce the queue discharge rate. In each of Section 5.2.1 and Section 5.2.2, a numerical expression of the queue discharge rate is presented, followed by an analysis of the model property.

The studies in this chapter show that the queue discharge rate reduction as a consequence of the inter-driver acceleration spread is smaller than that found in empirical studies for reasonable assumptions on the driver behavior parameters. An intra-driver reaction time extension mechanism can model similar queue discharge rates as reality.



(a) Spacing extensions due to acceleration variability.

(b) Queue discharge rates measurements with acceleration spread.

Figure 5.1: Measurements of the capacity drop due to the acceleration spread.

5.2.1 Capacity drop due to inter-driver acceleration spread

This section derivates analytical formulas for the queue discharge rate reduction due to the inter-driver spread, and finds the inter-driver spread does not give sufficient queue discharge rate reduction compared to empirical observations.

Analytical expressions of queue discharge rates

Let us consider a stop-and-go wave moving upstream on a homogeneous road section, see Figure 5.1. The grey block is the stop-and-go wave and bold lines are vehicular trajectories. The traffic in the scenario is described by a triangular fundamental diagram with a wave speed $-w$, free-flow speed v_f and capacity C. The critical density and maximum jam density are given by ρ_{cri} and ρ_{max}, respectively. There are N vehicles in total in the queue on a single lane, obeying the first-in-first-out (FIFO) rule. Each vehicle is numbered $i(i = 1,2,...N)$, increasing from the head of the queue $(i = 1)$ to the tail $(i = N)$. The speed and density in the queue are v_j and ρ_j , respectively. When all vehicles reach the free-flow speed after leaving the queue, the free-flow spacing and time headway between vehicle i and $i-1$ is given by s_i and h_i, respectively. The minimum free-flow spacing s_{cri} for all vehicles should be $\frac{1}{\rho_{\text{cri}}}$ (or the minimum time headway $h_{\text{min}} = \frac{1}{C}$), indicating no capacity drop at all. Each vehicle i is described by two constants, its desired acceleration a_i^{desire} and acceleration a_i. In principle, every vehicle accelerates with its desired acceleration. However, s_i is at the low end bounded by s_{cri}. Therefore, if $a_i = a_i^{\text{desire}}$ will result in $s_i < s_{\text{cri}}$, we set $a_i < a_i^{\text{desire}}$ to ensure $s_i = s_{\text{cri}}$. Note that $a_1 = a_1^{\text{desire}}$. Desired accelerations fall within the interval $[a_{\text{min}}, a_{\text{max}}]$. The reaction time of vehicle i is denoted as t_r. All vehicles have reached free-flow speed at x_1 where the sum of free-flow time headways from the second vehicle to the last vehicle is denoted as H, see Figure 5.1(b). H_{cri} denotes the sum of free-flow time headways from vehicle 2 to vehicle N in no-capacity-drop condition, i.e., $H_{\text{cri}} = \frac{N-1}{C}$.

If all vehicles follow the continuous Newell's car-following model (Newell, 2002), constructed by shifting its predecessor's trajectory by spacing $\Delta s = \frac{1}{\rho_{\max}}$ and time $\tau = \frac{\Delta s}{w} = \frac{1}{w\rho_{\max}}$, there is no capacity drop.

It is impossible that all vehicles have the same desired acceleration. It is assumed that the desired acceleration follows an uniform distribution bounded by $a_{\min}$ and $a_{\max}$, i.e., $a_{\text{desire}} \sim U[a_{\min}, a_{\max}]$. The uniform distribution is denoted by U. This section excludes the impact of reaction time extensions by setting $t_r = \tau$. When the desired acceleration of vehicle i is higher than its leader's acceleration, $a_i = a_{i-1}$ can ensure the follower can neither overtake nor be too close ($s_i < s_{\min}$) to its leader. Otherwise, $a_i = a_i^{\text{desire}}$.

A void is created between two successive vehicles if the follower's desired acceleration is lower than the predecessor's acceleration. In Figure 5.1(a), a dashed line is the Newell trajectory of vehicle i. Note the void between the Newell trajectory and the real trajectory of vehicle i. The void indicates that the free-flow spacing is extended by $s_i^{\text{extension}}$:

$$s_i^{\text{extension}} = \frac{\left(v_j - v_f\right)^2 (a_{i-1} - a_i)}{2a_{i-1}a_i} = \frac{1}{2}\left(v_j - v_f\right)^2 \left(\frac{1}{a_i} - \frac{1}{a_{i-1}}\right) \tag{5.1}$$

Now let us consider N vehicles within a stop-and-go wave as in Figure 5.1(b). The average queue discharge rate $\mathrm{E}(q_d)$ is expressed as:

$$\mathrm{E}(q_d) = \frac{N-1}{\mathrm{E}(H)} \tag{5.2}$$

$\mathrm{E}(H)$ is the mean value of H. Now let us calculate $\mathrm{E}(H)$ in order to calculate $\mathrm{E}(q_d)$.

Theorem 1

$$\begin{aligned}\mathrm{E}(H) =& \frac{N-1}{C} + \frac{\left(v_j - v_f\right)^2 (1+N)}{2v_f\left(a_{\max} + a_{\min}\cdot N\right)} - \frac{\left(v_j - v_f\right)^2 \ln\left(\frac{a_{\max}}{a_{\min}}\right)}{2v_f\left(a_{\max} - a_{\min}\right)} + \\ & \frac{(1+N)^3\left(v_j - v_f\right)^2}{2v_f(a_{\max} + a_{\min}N)^3}\left[\frac{N\left(a_{\max} - a_{\min}\right)^2}{2+N} + \right. \\ & \left. \frac{a_{\max}^2(-N+1) + 2Na_{\max}a_{\min}}{1+N} - \left(\frac{a_{\max} + Na_{\min}}{1+N}\right)^2\right]\end{aligned} \tag{5.3}$$

Proof. In Figure 5.1(b), the dashed line is a shifted trajectory from the trajectory of vehicle N by spacing $(N-1)\Delta s$ and time $(N-1)\tau$. According to Equation (5.1), the void between the trajectory of vehicle 1 and the dashed trajectory is $s_{1,N}^{\text{extension}} = \frac{1}{2}\left(v_j - v_f\right)^2\left(\frac{1}{a_N} - \frac{1}{a_1}\right)$. At location x_1, all vehicles have reached the free-flow speed.

So vehicle N reaches x_1 after the dashed trajectory has passed x_1 for $H_{cri} = \frac{N-1}{C}$. From Figure 5.1(b), we derive

$$\begin{aligned} \mathrm{E}(H) &= \mathrm{E}\left(H_{cri} + \frac{s_{1,N}^{\text{extension}}}{v_f}\right) \\ &= \frac{N-1}{C} + \frac{1}{2v_f}\left(v_j - v_f\right)^2 \left[\mathrm{E}\left(\frac{1}{a_N}\right) - \mathrm{E}\left(\frac{1}{a_1}\right)\right] \end{aligned} \tag{5.4}$$

which indicates that the calculation of $\mathrm{E}\left(\frac{1}{a_N}\right)$ and $\mathrm{E}\left(\frac{1}{a_1}\right)$ can give $\mathrm{E}(q_d)$. For $\mathrm{E}\left(\frac{1}{a_1}\right)$, since $a_1 = a_1^{\text{desire}} \sim U[a_{\min}, a_{\max}]$, we can have:

$$\mathrm{E}\left(\frac{1}{a_1}\right) = \frac{\ln\left(\frac{a_{\max}}{a_{\min}}\right)}{a_{\max} - a_{\min}} \tag{5.5}$$

Now the next step is to calculate $\mathrm{E}\left(\frac{1}{a_N}\right)$.

This chapter uses a Taylor series expansion to estimate the second-order approximation of $\mathrm{E}\left(\frac{1}{a_N}\right)$, giving:

$$\mathrm{E}\left(\frac{1}{a_N}\right) \approx g\left(\mathrm{E}\left(a_N\right)\right) + \frac{1}{2} g''\left(\mathrm{E}\left(a_N\right)\right) \mathrm{Var}\left(\mathrm{E}\left(a_N\right)\right) \tag{5.6}$$

where $g(x) = \frac{1}{x}$, and $\mathrm{E}(a_N)$, $\mathrm{Var}(a_N)$ are the mathematical expectation and the standard deviation of a_N, respectively. Now $\mathrm{E}(a_N)$ and $\mathrm{Var}(a_N)$ need to be calculated before the estimation of $\mathrm{E}\left(\frac{1}{a_N}\right)$, which requires the probability distribution function f_N of a_N.

Let $\left(a_{(1)}^{\text{desire}}, ..., a_{(N)}^{\text{desire}}\right)$ denote the corresponding order statistics of the random sample $\left(a_1^{\text{desire}}, ..., a_N^{\text{desire}}\right)$ so that $a_{(1)}^{\text{desire}} \leq a_{(2)}^{\text{desire}} \leq ... \leq a_{(N)}^{\text{desire}}$. Because $a_N = a_{(1)}^{\text{desire}}$, the probability density function of a_N equals to the probability density function of the smallest order statistic $a_{(1)}^{\text{desire}}$. Order statics (Reiss, 1989) gives the probability distribution function f_A of $a_{(1)}^{\text{desire}}$:

$$f_A\left(a_{(1)}^{\text{desire}}\right) = N\left[1 - F\left(a_{(1)}^{\text{desire}}\right)\right]^{N-1} f\left(a_{(1)}^{\text{desire}}\right) \tag{5.7}$$

where $F\left(a_{(1)}^{\text{desire}}\right)$ is the cumulative distribution function of the smallest desired acceleration and $f\left(a_{(1)}^{\text{desire}}\right)$ is the probability distribution function of the smallest desired acceleration. For $a_{(1)}^{\text{desire}} \sim U[a_{\min}, a_{\max}]$, we get $F\left(a_{(1)}^{\text{desire}}\right) = \frac{a^{\text{desire}} - a_{\min}}{a_{\max} - a_{\min}}$ and $f\left(a_{(1)}^{\text{desire}}\right) = \frac{1}{a_{\max} - a_{\min}}$. So the probability distribution function f_N of a_N can be given from (5.7) to:

$$\begin{aligned} f_N\left(a_N\right) &= f_A\left(a_{(1)}^{\text{desire}}\right) \\ &= \left(\frac{a_{\max} - a_N}{a_{\max} - a_{\min}}\right)^{N-1} \frac{N}{a_{\max} - a_{\min}}, \text{ for } \mathrm{a}_{(1)}^{\text{desire}} = \mathrm{a_N} \in [\mathrm{a_{min}}, \mathrm{a_{max}}] \end{aligned} \tag{5.8}$$

Equation (5.8) gives $\mathrm{E}(a_N) = \frac{a_{\max}+Na_{\min}}{1+N}$ and $\mathrm{Var}(a_N) = \frac{N(a_{\max}-a_{\min})^2}{(2+N)} + \frac{a_{\max}^2(-N+1)+2Na_{\min}a_{\max}}{(1+N)} - \frac{(a_{\max}+Na_{\min})^2}{(1+N)^2}$. So $g''(\mathrm{E}(a_N)) = 2\left(\frac{1+N}{a_{\max}+a_{\min}\cdot N}\right)^3$. Now the estimation of $\mathrm{E}\left(\frac{1}{a_N}\right)$ is:

$$\mathrm{E}\left(\frac{1}{a_N}\right) \approx \frac{1+N}{a_{\max}+Na_{\min}} + \left(\frac{1+N}{a_{\max}+Na_{\min}}\right)^3 \left[\frac{N(a_{\max}-a_{\min})^2}{2+N} + \frac{a_{\max}^2(-N+1)+2Na_{\max}a_{\min}}{1+N} - \left(\frac{a_{\max}+Na_{\min}}{1+N}\right)^2\right] \tag{5.9}$$

Incoporating Equation (5.5) and Equation (5.9) into Equation (5.4), we get $\mathrm{E}(H) = \frac{N-1}{C} + \frac{(v_j-v_f)^2(1+N)}{2v_f(a_{\max}+a_{\min}\cdot N)} - \frac{(v_j-v_f)^2\ln\left(\frac{a_{\max}}{a_{\min}}\right)}{2v_f(a_{\max}-a_{\min})} + \frac{(1+N)^3(v_j-v_f)^2}{2v_f(a_{\max}+a_{\min}N)^3}\left[\frac{N(a_{\max}-a_{\min})^2}{2+N} + \frac{a_{\max}^2(-N+1)+2Na_{\max}a_{\min}}{1+N} - \left(\frac{a_{\max}+Na_{\min}}{1+N}\right)^2\right]$ ■

Incorporating Equation (5.3) into Equation (5.2) can give the estimated average queue discharge rate $\mathrm{E}(q_d)$.

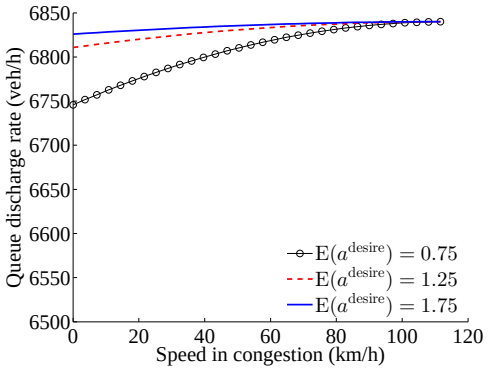
Model property analysis

This chapter gives a triangular fundamental diagram with $w = 18$ km/h, $v_f = 114$ km/h, $C = 6840$ veh/h, $\rho_{\mathrm{cri}} = 60$ veh/km and $\rho_{\max} = 440$km/h. This fundamental diagram indicates a similar traffic situation as that in Chapter 3. Different bounds for accelerations have been reported: for instance 0.5m/s^2 - 3m/s^2 (Leclercq et al., 2011), or 1.5m/s^2 - 2m/s^2 (Lebacque, 2003a). Here the limits for desired accelerations is assumed to range from 0.5m/s^2 to 2m/s^2.The range will be limited further according to different purposes.

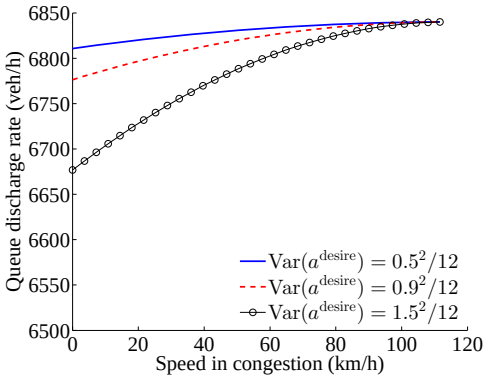
Consider a stop-and-go wave that propagates at speed w for $\Omega = 10$ minutes. Variational theory (Daganzo, 2005) gives the number of vehicles in the stop-and-go wave as $N = \left\lfloor \frac{w\rho_{\max}\Omega}{60} \right\rfloor = 1320$veh. This section analyses the queue discharge rate for this queue.

As shown in Equation (5.3), $\mathrm{E}(H)$ is a function of $a_{\min}$, $a_{\max}$ and N. The sensitivity of the queue discharge rate to the average desired accelerations, standard spread of desired accelerations and number of vehicles are evaluated with Equation (5.2) and (5.3), presented in Figure 5.2.

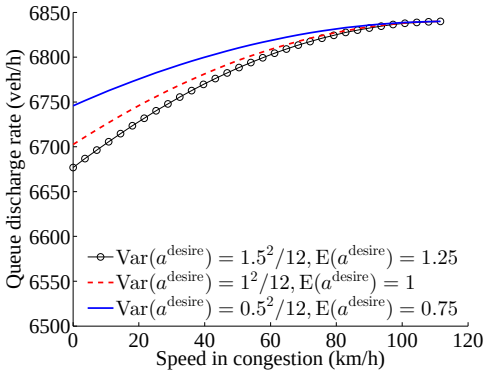
Figure 5.2(a) presents a relation between the speed in congestion v_j and the queue discharge rate q_d when setting $\mathrm{E}\left(a^{\mathrm{desire}}\right)$ as 0.75m/s^2, 1.25m/s^2, 1.75m/s^2 respectively and $\mathrm{Var}\left(a^{\mathrm{desire}}\right) = \frac{0.5^2}{12}$m/s^2. The $\mathrm{E}\left(a^{\mathrm{desire}}\right)$ is obtained by setting the pair $(a_{\min}, a_{\max})$ to $\left(0.5\mathrm{m/s}^2, 1\mathrm{m/s}^2\right)$, $\left(1\mathrm{m/s}^2, 1.5\mathrm{m/s}^2\right)$ and $\left(1.5\mathrm{m/s}^2, 2\mathrm{m/s}^2\right)$. It is seen that that the faster the average desired acceleration, the higher the queue discharge rate.



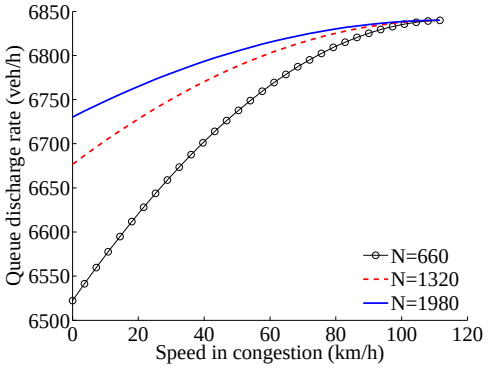
(a) Sensitivity of the analytical model to the mathematic expectation of desired accelerations.



(b) Sensitivity of the analytical model to the standard deviations of desired accelerations.



(c) Comparisons of impacts on queue discharge rates between the mean and deviation of desired accelerations.



(d) Sensitivity of the analytical model to vehicle numbers.

Figure 5.2: Sensitivity of queue discharge rates when capacity drop is due to the acceleration spread.

Figure 5.2(b) presents the relation between v_j and q_d when $\mathrm{Var}\left(a^{\mathrm{desire}}\right)$ equals to $\frac{0.5^2}{12}\mathrm{m/s^2}$, $\frac{0.9^2}{12}\mathrm{m/s^2}$ and $\frac{1.5^2}{12}\mathrm{m/s^2}$, setting $\mathrm{E}\left(a^{\mathrm{desire}}\right) = 1.25\mathrm{m/s^2}$. That is, the pair $(a_{\min}, a_{\max})$ are chosen to be ($1\mathrm{m/s^2}$,$1.5\mathrm{m/s^2}$), ($0.8\mathrm{m/s^2}$,$1.7\mathrm{m/s^2}$) and ($0.5\mathrm{m/s^2}$,$2\mathrm{m/s^2}$) respectively. It indicates that the larger the spread, the lower the queue discharge rate.

If $a_{\min} = 0.5\mathrm{m/s^2}$ and $a_{\max}$ decreases from $2\mathrm{m/s^2}$ to $1\mathrm{m/s^2}$, then both of $\mathrm{E}\left(a^{\mathrm{desire}}\right)$ and $\mathrm{Var}\left(a^{\mathrm{desire}}\right)$ decreases. Figure 5.2(c) shows that the decrease of $a_{\max}$ increases the queue discharge rates. Since the decrease of $\mathrm{E}\left(a^{\mathrm{desire}}\right)$ and $\mathrm{Var}\left(a^{\mathrm{desire}}\right)$ will decrease and increase the queue discharge rate respectively, the increase of queue discharge rates in Figure 5.2(c) indicates that $\mathrm{Var}\left(a^{\mathrm{desire}}\right)$ has more influences on the queue discharge rate than $\mathrm{E}\left(a^{\mathrm{desire}}\right)$.

Figure 5.2(d) shows the sensitivity to N with $a_{\min} = 0.5\mathrm{m/s^2}$ and $a_{\max} = 2\mathrm{m/s^2}$. The more vehicles, the higher queue discharge rates. It is not a surprise because the follower's acceleration is always limited by its leader's acceleration, that makes the acceleration spread decrease as the vehicle number increases. Since $N = 1320$veh means the congestion only propagates for 10min, the queue discharge rate can be even higher when setting a longer time of congestion propagation.

Setting $a_{\min} = 0.5\mathrm{m/s^2}$, $a_{\max} = 2\mathrm{m/s^2}$ and $N = 660$veh gives a considerable influence of the acceleration spread on queue discharge rates, shown as the line with circles in Figure 5.2(d). However, the contribution of acceleration spread to the queue discharge rate reduction is still small relatively. In Figure 3.5 when $v_j = 0$km/h, the minimum queue discharge rate (6522veh/h) is still much higher than the empirical value (5000veh/h) shown in Figure 3.7.

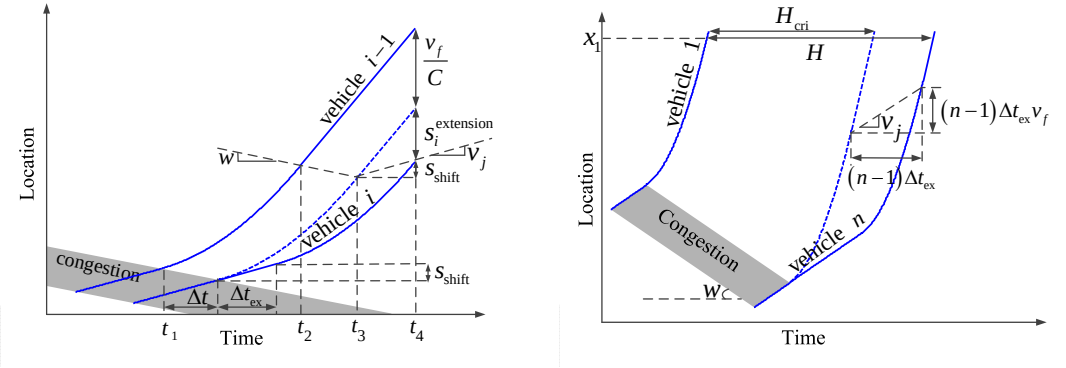
Note that the hypothesis about the uniform desired acceleration distribution has already maximized the $\mathrm{Var}\left(a^{\mathrm{desire}}\right)$. In reality, the desired acceleration could follow some distribution with peaks (such as shown in Koutsopoulos and Farah (2012)) which will have smaller $\mathrm{Var}\left(a^{\mathrm{desire}}\right)$. Therefore, it is concluded that the inter-driver desired acceleration spread is not a determinant of capacity drop size.

5.2.2 Capacity drop due to reaction time extension

This section shows that the reaction time extension can considerably influence the queue discharge rate. A negative relation between the reaction time extension and the speed in congestion could result in a similar queue discharge rates as empirical findings. Analytical expressions of queue discharge rates and the sensitivity analyses are given in this section.

Analytical expressions of queue discharge rates

A reaction time longer than τ will lead to a larger spacing than $\frac{1}{\rho_{\mathrm{cri}}}$, so the queue discharge rate will be lower than the capacity which is the capacity drop. Therefore, this



(a) Spacing extensions due to reaction time extensions.

(b) Queue discharge rates measurements with reaction time extensions.

Figure 5.3: Measurements of the capacity drop due to the reaction time extension.

chapter only considers the cases when the reaction time is longer, and it is defined that $t_r = \tau + \Delta t_{\text{ex}}$ where τ is considered as a fixed reaction time (related to the fundamental diagram) and Δt_{ex} as a reaction time extension. As shown in Figure 5.3(a), two bold solid lines are trajectories of two successive vehicles accelerating from speed v_j up to free speed v_f. The follower's reaction time is extended by Δt_{ex} from τ. The dashed line is the follower's trajectory when $\Delta t_{\text{ex}} = 0$s. The follower's trajectory can be considered as a shifted trajectory from the dashed line in time (by Δt_{ex}) and space (by s_{shift}). Geometry (see Figure 5.3(a)) gives us $s_i^{\text{extension}} = \left(v_f - v_j\right)\Delta t_{\text{ex}}$. So the free-flow spacing is extended to:

$$s_i = \frac{1}{\rho_{\text{cri}}} + s_i^{\text{extension}} = \frac{v_f}{C} + \left(v_f - v_j\right)\Delta t_{\text{ex}} \tag{5.10}$$

Considering N vehicles accelerating from a queue with the same acceleration (see Figure 5.3(b)), we can calculate the spacing between the first and the last vehicle as follows:

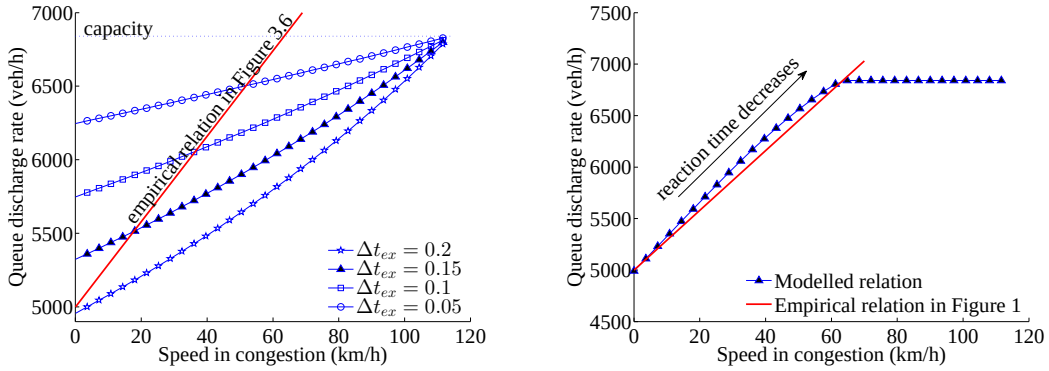
$$s_{1,N} = H_{\text{cri}} v_f + \sum_{i=1}^{N-1} s_i^{\text{extension}} = \frac{N-1}{\rho_{\text{cri}}} + (N-1)\left(v_f - v_j\right)\Delta t_{\text{ex}} \tag{5.11}$$

which means the queue discharge rate q_d is:

$$q_d = \frac{n-1}{H} = \frac{(N-1)v_f}{s_{1,N}} = \frac{v_f \rho_{\text{cri}}}{1 + \rho_{\text{cri}}\left(v_j - v_f\right)\Delta t_{\text{ex}}} \tag{5.12}$$

Analysis of model properties

The independent impact of the reaction time extension is evaluated with Equation (5.12), see Figure 5.4(a). The relation between the speed in congestion and the queue discharge rate is examined by setting reaction time extension Δt_{ex} to 0.05s, 0.1s, 0.15s and 0.2s. Figure 5.4(a) firstly indicates that reaction time extension Δt_{ex} can give a positive relation between the speed in congestion and the queue discharge rate. As the reaction time extension increases even slightly, the queue discharge rate will decrease



(a) Sensitivity of the analytical model to the reaction time extension.

(b) An intra-driver reaction time extension mechanism for giving empirical observations.

Figure 5.4: Sensitivity of queue discharge rates to reaction time extensions.

considerably. When $\Delta t_{\text{ex}} = 0$s, the queue discharge rate equals to the capacity. Secondly, a dynamic reaction time extension can model the empirical observation. The bold line in Figure 5.4(a) is the empirical relation revealed in Chapter 3 (see Figure 3.7). The intersections between the bold line and the other lines indicates that to give empirical observations it is needed to decrease the reaction time extension as the speed in congestion increases. Note that in Chapter 3 (see Figure 3.7) when the vehicular speed in congestion reached around 63km/h, the queue discharge rate has reached 6840 veh/h which is much higher than a three-lane dutch freeway capacity 6300veh/h (with 15% proposition of trucks) claimed in Goemans et al. (2011). Be more specific, that 6840 veh/h is the highest point measured in the data. This chapter believes 6840veh/h is close to a three-lane freeway capacity. So it is believed that when the vehicular speed in congestion exceeds 63km/h, there is no capacity drop. $v_j^{\max}$ indicates the lowest speed in congestion leading to no capacity drop. When $v_j \geqslant v_j^{\max}$, the queue discharge rate equals to the capacity. No capacity drop means the reaction time extension might be zero. Hence, we set

$$\Delta t_{\text{ex}} = \max\left(0, \gamma - \frac{\gamma v_j}{v_j^{\max}}\right) \tag{5.13}$$

γ is a parameter indicating the reaction time extension when the speed in congestion is 0 km/h. If $\gamma = 0.195$ s, a relation matching the empirical finding can be established. The modelled relation with Equation (5.13) is shown as dark triangles in Figure 5.4(b). The bold line is the empirical relation as in Figure 3.7. The modelled relation can fit the empirical relation quite well, see Figure 5.4(b).

5.3 Numerical experiments

In this section, numerical experiments are conducted to firstly validate the analytical model presented in Section 5.2.1. The estimation of queue discharge rate in Sec-

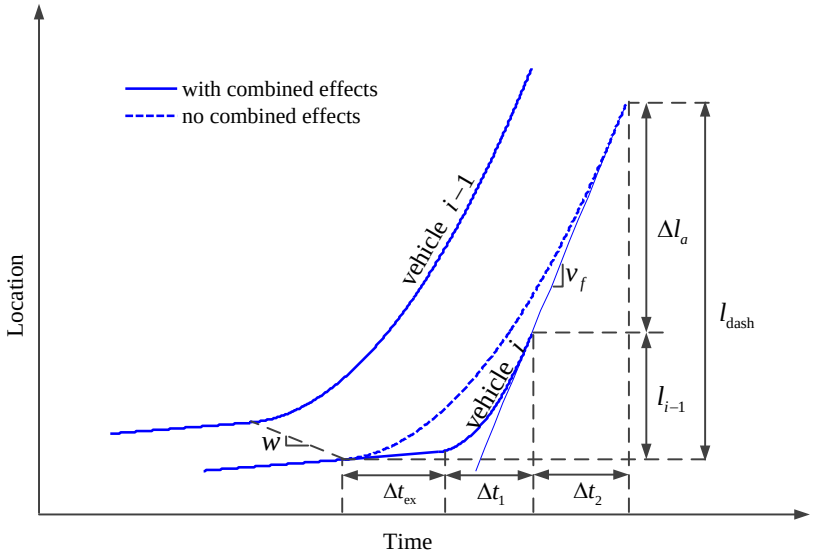


Figure 5.5: Measurement of accelerations when reaction time is extended.

tion 5.2.1 is an approximation with the Delta method. So it is necessary to check the accuracy of the approximation.

Secondly, the combined effects of inter-driver acceleration spread and the reaction time extensions are presented. A positive reaction time extension can allow a following vehicle to have a faster-than-predecessor acceleration. So the acceleration of the last vehicle in the queue a_N does not have to be the slowest acceleration among all vehicles in the queue. The distribution of the last vehicle's acceleration is difficult to deduce, so this chapter decides to use numerical experiments to see the combined effects of the inter-driver acceleration spread and the reaction time extension.

Thirdly, the experiment investigates how to give a same relation between the speed in congestion and the queue discharge rate as empirical observations, considering combined effects of the acceleration spread and the reaction time extension.

The above three key questions are answered in three experiments denoted by Ex1, Ex2 and Ex3, respectively. The simulation results in this section correlate quite well with the analytical findings in Section 5.2. No matter whether the reaction time is included or not, the inter-driver acceleration spread is not a determinant of the capacity drop. No matter whether the inter-driver acceleration spread is considered or not, a negative relation between the reaction time and the speed in congestion can give similar queue discharge rates as empirical observations.

5.3.1 Simulation model

Figure 5.5 shows trajectories of two vehicles accelerating from congestion, $i-1$ and i. Let us set an acceleration difference Δa. The free-flow spacing between vehicle i and

$i-1$ will be $\frac{1}{\rho_{\text{cri}}}$ if $a_i = a_{i-1} + \Delta a$. In a special case that reaction time of all vehicles are not extended (i.e., $t_r = \tau$), $\Delta a = 0\text{m/s}^2$. So if $a_i > a_{i-1} + \Delta a$, the free-flow spacing between two vehicles will be smaller than the critical spacing $\frac{1}{\rho_{\text{cri}}}$. In Figure 5.5 we use a dashed line to present a trajectory of vehicle i according to Newell's model. Finally the trajectory of vehicle i will overlap with the dashed line. Vehicle i reached the free-flow speed earlier than the dashed trajectory by Δt_2. The whole acceleration process of vehicle i lasts Δt_1. So we can describe v_f as follows:

$$v_f = v_j + (a_{i-1} + \Delta a)\,\Delta t_1 \tag{5.14}$$

Meanwhile, vehicle $i-1$ spends $\Delta t_{\text{ex}} + \Delta t_1 + \Delta t_2$ accelerating up to v_f, i.e.,

$$v_f = v_j + a_{i-1}\,(\Delta t_{\text{ex}} + \Delta t_1 + \Delta t_2) \tag{5.15}$$

For the whole acceleration process, vehicle i travels $l_{i-1} = v_j \Delta t_{\text{ex}} + \frac{v_f^2 - v_j^2}{2(a_{i-1}+\Delta a)}$ which is less than the distance ($l_{\text{dash}} = \frac{v_f^2 - v_j^2}{2a_{i-1}}$) the dashed trajectory travel by $\Delta l_a = l_{\text{dash}} - l_{i-1}$. Vehicle i spends Δt_2 traveling Δl_a at free-flow speed v_f:

$$v_f = \frac{\frac{v_f^2 - v_j^2}{2a_{i-1}} - \left(v_j \Delta t_{\text{ex}} + \frac{v_f^2 - v_j^2}{2(a_{i-1}+\Delta a)} \right)}{\Delta t_2} \tag{5.16}$$

Combination of Equation (5.14),(5.15) and (5.16) can give:

$$\Delta a = \frac{2a_{i-1}^2 \Delta t_{\text{ex}}}{v_f - v_j - 2a_{i-1}\Delta t_{\text{ex}}}, \quad \text{for} \quad v_f - v_j > 2a_{i-1}\Delta t_{\text{ex}} \tag{5.17}$$

Equation (5.17) shows that the following vehicle can catch up with its predecessor with $a_i = a_{i-1} + \Delta a$ when $v_f - v_j > 2a_{i-1}\Delta t_{\text{ex}}$. If $a_{i-1} + \Delta a \leqslant a_i^{\text{desire}}$, then $a_i = a_{i-1} + \Delta a$, the free-flow spacing between vehicle i and $i-1$ will be critical spacing. If $a_{i-1} + \Delta a > a_i^{\text{desire}}$, then $a_i = a_i^{\text{desire}} < a_{i-1} + \Delta a$, the free-flow spacing between two successive vehicles is $s_i = \frac{1}{\rho_{\text{cri}}} + \frac{\left(\frac{1}{a_i} - \frac{1}{a_{i-1}}\right)\left(v_j - v_f\right)^2}{2} + \left(v_f - v_j\right)\Delta t_{\text{ex}}$.

When $v_f - v_j \leqslant 2a_{i-1}\Delta t_{\text{ex}}$, i.e., the reaction time is too long, and it is impossible for the follower to catch up with the leader. In this case, the follower's acceleration will not be limited by its predecessor, i.e., $a_i = a_i^{\text{desire}}$. The free-flow spacing between two vehicles will be larger than the critical spacing, calculated as $s_i = \frac{1}{\rho_{\text{cri}}} + \frac{\left(\frac{1}{a_i} - \frac{1}{a_{i-1}}\right)\left(v_j - v_f\right)^2}{2} + \left(v_f - v_j\right)\Delta t_{\text{ex}}$. In summary, each vehicle accelerates with acceleration:

$$a_i = \min\left(a_{i-1} + \Delta a, a_i^{\text{desire}}\right) \tag{5.18}$$

and the free-flow spacing between any vehicle and its predecessor is:

$$s_i = \begin{cases} \frac{1}{\rho_{\text{cri}}}, & \text{for } v_f - v_j > 2a_{i-1}\Delta t_{\text{ex}} \text{ and } a_{i-1} + \Delta a \leqslant a_i^{\text{desire}} \\ \frac{1}{\rho_{\text{cri}}} + \frac{\left(\frac{1}{a_i} - \frac{1}{a_{i-1}}\right)\left(v_j - v_f\right)^2}{2} + \left(v_f - v_j\right)\Delta t_{\text{ex}}, & \text{for others} \end{cases} \tag{5.19}$$

In the numerical experiment we calculate the queue discharge rate as:

$$q_d = \frac{v_f}{E(s_i)}, \quad i = 2, ..., n \tag{5.20}$$

Equation (5.19) and (5.20) are general expressions for estimating queue discharge rates in two experiments, that is for the validation of analytical models (in Ex1) and the examination of combined effects (in Ex2).

Since in Section 5.2, it is found that the independent impact of acceleration on the queue discharge rate is marginal. This chapter hypothesizes that when considering reaction time extensions, the inter-driver acceleration spread cannot contribute to queue discharge rate reduction greatly, either. The consequence of the hypothesis is that to obtain the empirically observed queue discharge rate (Figure 3.7), it is more important to model the impact of the intra-driver variation than that of the inter-driver spread. Equation (5.13) is used to give the queue discharge rate in the third experiment Ex3.

5.3.2 Simulation set-up

This section describes the set-ups of the three experiments. For validations of the analytical model in Section 5.2.1, we let $\Delta t_{\text{ex}} = 0\text{s}$ in Ex1. For examining combined effects of the acceleration spread and the reaction time extension (Ex2), we set two scenarios, i.e., $\Delta t_{\text{ex}} = 0.1\text{s}$ and $\Delta t_{\text{ex}} = 0.2\text{s}$.

At the beginning of each experiment, we set all vehicleś desired acceleration and reaction time extension. With Equation (5.18), (5.19) and (5.20), we can directly have the final queue discharge rate. The notations of models and the set-up of fundamental diagram are the same as those in Section 5.2. To draw the relation between the speed in congestion and the queue discharge rate, in each scenario set-up we run one simulation with newly distributed desired accelerations for each speed in congestion. We run the simulation for 1000 times to get the expected value and standard spread of queue discharge rates. We use $N = 660\text{veh}$, $a_{\min} = 0.5\text{m/s}^2$ and $a_{\max} = 2\text{m/s}^2$.

5.3.3 Validations of analytical models

We approximate the mean queue discharge rate by approximating the expected value of the time-headway in Section 5.2. So we need to check whether the approximations are accurate enough to draw conclusions on the independent impacts of accelerations, which is finished in the first experiment Ex1. The comparison between the numerical experiment result and the analytical result is shown in Figure 5.6. In Figure 5.6, we use error bars and plus signs to indicate the standard spread and the expected value of queue discharge rates respectively for experiment results. Circles show the analytical approximations of queue discharge rates from Section 5.2.1.

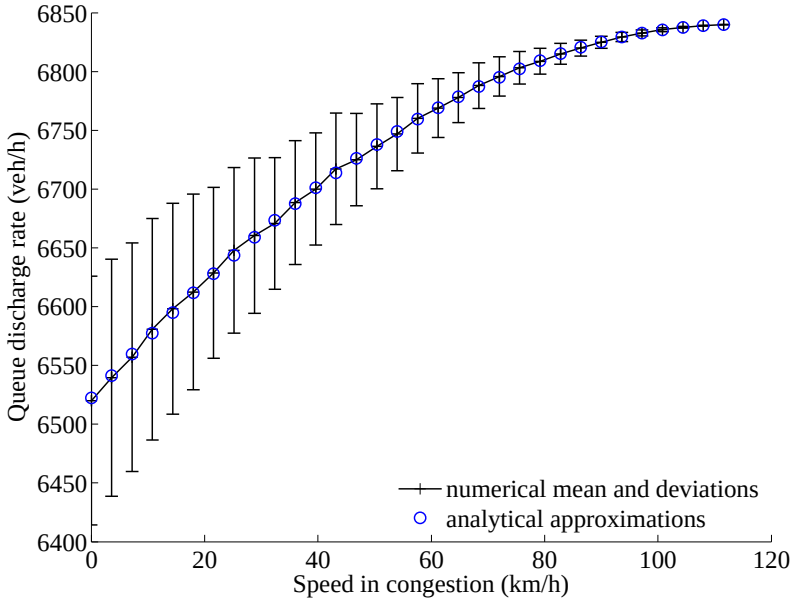


Figure 5.6: Validation of the analytical model for the inter-driver acceleration spread. Mean and standard spreads of queue discharge rates are shown as plus signs and error bars, respectively

The analytical approximations of queue discharge rates fit the numerical experiment results very well. Secondly, the queue discharge rate spread increases as the speed in congestion decreases. The fluctuation of queue discharge rate might be a related to the order of desired accelerations. More importantly, the analysis of the independent impacts of accelerations on the queue discharge rate in Section 5.2 is correct, that is, the acceleration spread cannot contribute sufficiently to the capacity drop independently.

5.3.4 Combined effects of the inter-driver acceleration spread and the reaction time extension

The combined effects of the inter-driver acceleration spread and the reaction time extension are examined in numerical experiments (Ex2), shown in Figure 5.7. Circles are used to indicate the mean queue discharge rate with the independent impact of reaction time, which is the same as shown in Figure 5.4(a).

Figure 5.7 shows that the inter-driver acceleration spread hardly contributes to the queue discharge rate reduction. When setting $\Delta t_{\text{ex}} = 0.1$s and $v_j = 0$km/h, the inter-driver acceleration spread can only reduce the queue discharge rate by 180 veh/h (around 3% reduction) at most. That observation is shown as the small difference between the plus signs and the circles in Figure 5.7. Meanwhile, increasing Δt_{ex} from 0.1s to 0.2s decreases the queue discharge rate considerably by 13% (with inter-driver acceleration spread) and 14% (without the inter-driver acceleration spread) in a case of

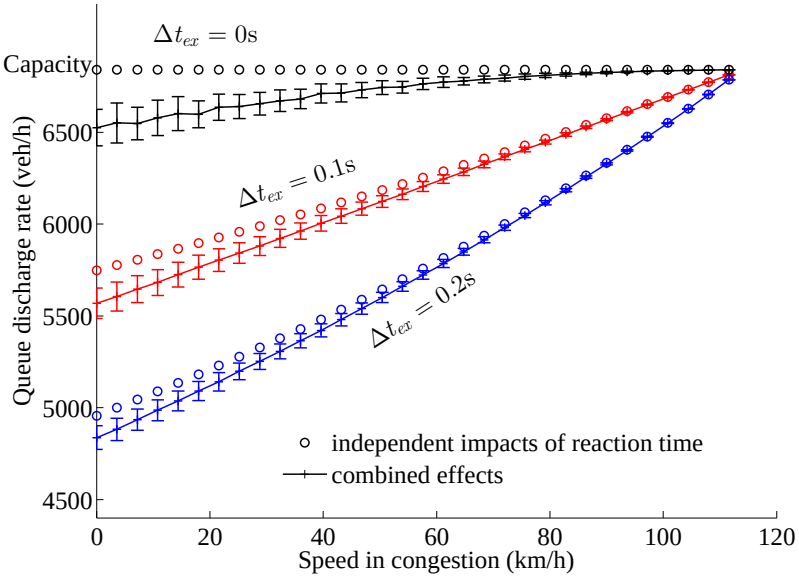


Figure 5.7: Combined effects of inter-driver acceleration spread and reaction time extension on queue discharge rates.

$v_j = 0$km/h. It also means a slight decrease of reaction time can contribute a considerable increase of the queue discharge rate.

Because the inter-driver acceleration spread can only reduce the queue discharge rate slightly, a speed-dependent reaction time extension is given (see Equation (5.13)) to model mechanism behind the capacity drop (Ex3). Figure 5.8 shows the experiment results. If reaction time decreases when congestion gets lighter, queue discharge rates can fit empirical observations. Hence, it is argued that when understanding the capacity drop, studying the intra-driver variety is more important than modeling the inter-driver spread.

5.4 Conclusion

This chapter reveals the impacts of inter-driver spread and intra-driver variation on the queue discharge rate. Firstly, it is found that the impact of inter-driver acceleration spread on the queue discharge rate is rather small (around 3% at most). No matter whether the reaction time extension is considered or not, the inter-driver acceleration spread does not decrease the queue discharge rate as much as determined empirically. Secondly, a speed-dependent reaction time extension mechanism, that is the reaction time decreases as the speed in congestion increases, yields a similar relation between the speed in congestion and the queue discharge rate as found in empirical observations.

This research assumes the free-flow branch in the flow-density fundamental diagram

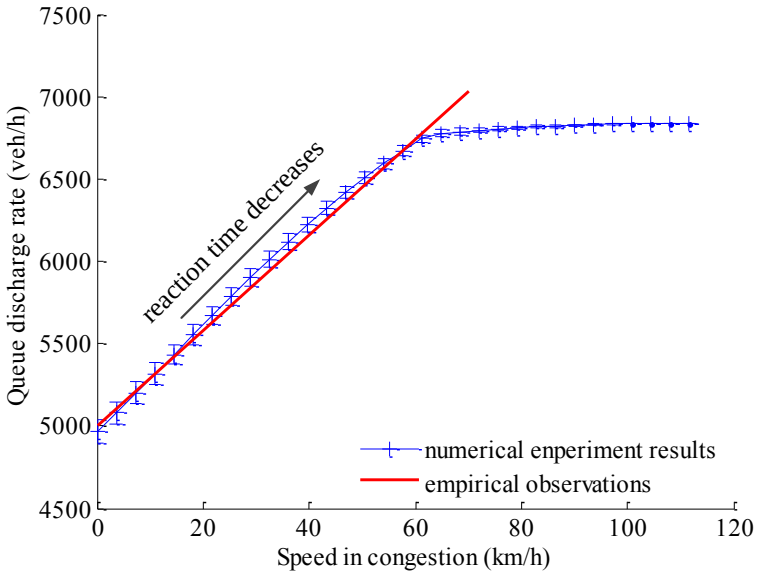


Figure 5.8: Combined effects on queue discharge rates with intra-driver reaction time extension mechanism.

as a straight line. Some researchers might argue that the free-flow speed in the vicinity of capacity is lower than the maximum free-flow speed, according to empirical observations. Because in empirical observations, the average free-flow speed in the vicinity of capacity is not very sensitive to the average density across lanes (see Chapter 2), relaxing the linear free-flow branch assumption by letting v_f as the free-flow speed in the vicinity of capacity and the maximum free-flow speed as $v_{\max}$ ($v_{\max} > v_f$), does not influence the research results at all.

Therefore, it is concluded that including the inter-driver acceleration spread when modeling the capacity drop within car-following models is not essential, but including intra-driver variation is. This chapter cannot be seen as a complete work on understanding the capacity drop from a driver behavioral perspective. A further understanding on the intra-driver the reaction time extension variation from a driver behavioral perspective is discussed in Chapter 6.

Chapter 6

Understanding capacity drop from a driver behavioral perspective

The final step in this research is to understand the behavioral reason for the capacity drop and incorporate it into a microscopic model for verification. Before that, the relation between the speed in congestion and the queue discharge rate has been proposed in Chapter 3 and it is found that the intra-driver variation is the most plausible relevant mechanism behind the capacity drop in Chapter 5, which provide fundamentals for the research in this chapter. To finally unveil the behavioral explanation of the capacity drop, this work analyses floating-car data to see how the desired acceleration stochasticity reduces the queue discharge rate. To verify the hypothesis that the stochasticity of desired accelerations is crucial for the queue discharge rate reduction, A Geometric Brownian motion car-following model is presented.

This chapter is an edited version of the following paper: Yuan, K., Laval, J., Knoop, V.L., Jiang, R., & Hoogendoorn, S.P. (under review). A Geometric Brownian motion car-following model: towards a better understanding of capacity drop. *Transportation Research Part B: Methodological*.

6.1 Introduction

While lane changing is believed to be a possible behavior accounting for the capacity drop (Laval & Daganzo, 2006; Yeo, 2008; Coifman & Kim, 2011; Leclercq et al., 2011, 2015), recent observations reveal that the capacity also drops in an absence of lane changing (Oh & Yeo, 2015). After analyzing time headways in the downstream free-flow state of congestion when lane changing is absent, Oh and Yeo (2015) also find that the queue discharge rate decreases as the congestion severity increases. When a vehicle accelerates from a queue, the size of the void in its front is strongly related to its speed in the queue. Hence, it is plausible that the relation between the speed in congestion and the queue discharge rate in Chapter 3 is a result of car-following behaviors and not of lane changing. It is important to explore the longitudinal behavior mechanism for the formation of voids as well as the relation between the void and the vehicular speed in the queue, and incorporate the mechanism into a car-following model.

Except lane changing, driver-vehicle-combination (DVC) heterogeneity possibly can be a cause of the queue discharge rate reduction. The DVC heterogeneity indicates the variations in fundamental (e.g., flow-density) relations in homogeneous traffic states. While acknowledging the possible effects of DVC heterogeneity, this thesis argues that the DVC heterogeneity is not a necessary condition for the capacity drop. First, it is worthwhile to note that Chapter 2 observes the capacity drop and its dependence on congestion states in individual lane with loop detector data collected on a three-lane freeway section in the Netherlands. Drivers generally follow a "Keep right unless overtaking" rule in there. Because the percentage of trucks in there is rather small, so the variation in desires for the free-flow speed during the speeding up process, especially on the shoulder lane, should be small. Second, Chapter 3 analyzes more loop detector data on freeways with little heterogeneity and see considerable queue discharge rate reductions. Hence, this chapter believes the key justification of the capacity drop phenomenon is hidden in DVC homogeneous behaviors, assuming all drivers follow a same fundamental relation.

As shown in Section 1.3.3, excluding the influence of vehicle class heterogeneity, two longitudinal mechanisms in the car-following framework for giving the capacity drop have been proposed. That is, the inter-driver spread (Papageorgiou et al., 2008) and the intra-driver variation (see Chapter 5, Tampère, 2004; Zhang & Kim, 2005; Treiber et al., 2006; Chen et al., 2014). Chapter 5 has shown that the queue discharge rate reduction due to the inter-driver acceleration spread is rather small provided bounded accelerations through analytical analysis and numerical simulations. The independent effect of the inter-driver acceleration spread can only reduce the discharge rate slightly, and hardly influence the reduction that is due to the intra-driver reaction time extension. Following this contribution, this research focus on studying incorporating the intra-driver variation into a car-following framework for the interpretation of the capacity drop.

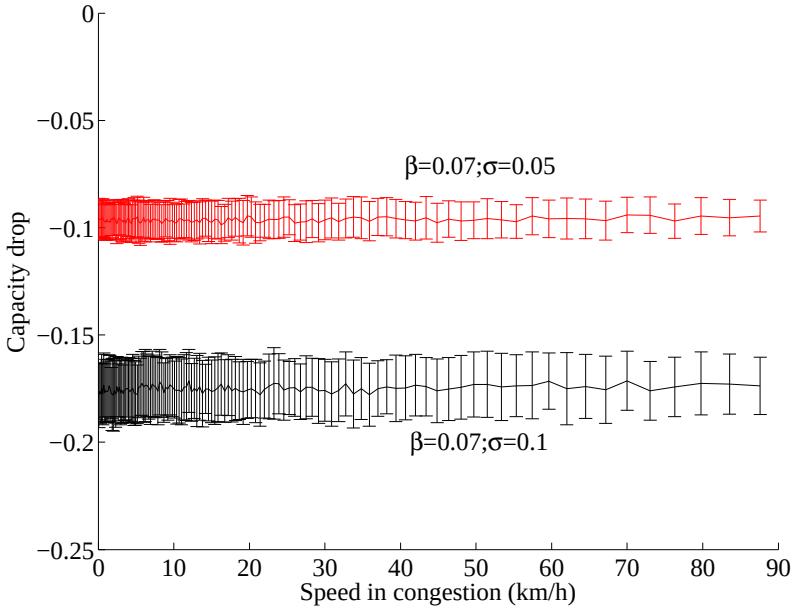


Figure 6.1: Relation between the speed in congestion and the capacity drop given by the parsimonious car-following model. The error bar indicates the standard deviation of the capacity drop. β and σ are two constant parameters in the model. A higher σ gives a larger capacity drop.

To achieve the research goal, there are two research questions to be addressed. First, (1) what specific behavioral characteristics should we describe? Second, (2) how to incorporate it into a car-following model? The first question is related to an overarching behavioral explanation of the queue discharge rate reduction. After taking a closer look at the trajectory dataset analyzed in Laval et al. (2014), it is noticed that both of the mean value and the standard deviation of the driver desired acceleration decrease as the vehicular speed increases. The desired acceleration means the acceleration drivers want to impose to cars if there is no predecessor in front. This chapter believes this stochastic feature during accelerating is relevant for reducing queue discharge rates. Laval et al. (2014) propose a parsimonious car-following model in the Newell's first-order car-following framework, giving the formation and propagation of oscillations. But that model cannot give the capacity drop which depends on the speed in congestion. Figure 6.1 shows the relation between the capacity drop and the speed in congestion given by the parsimonious car-following model. β and σ are the inverse relaxation time and diffusion coefficient with units of $[\text{time}]^{-1}$ and $[\text{distance}]^2[\text{time}]^{-1}$, respectively. The reason is that the original model only consider the relation between the mean acceleration and the vehicular speed, excluding the negative relation between the standard deviation and the vehicular speed. Hence, this chapter extends the parsimonious car-following model, considering both of the mean and standard deviation of the acceleration. The approach is to use a geometric Brownian motion to describe the stochastic speeding up process, which answers the second research question.

Two additional independent parameters are added to the Newell's first order car-following model, resulting in a model with five parameters in total. A dimensional analysis shows that four dimensionless parameters in total drive the extended parsimonious model. In simulations, this chapter firstly examines the model ability of showing features of platoons. Then it is showed that the extended parsimonious model is able to give the reduction of queue discharge rate and its dependence on the vehicular speed. This work can also be seen as a contribution to justifying the capacity drop from a driver perspective. The outline of this chapter is as follows: the extended parsimonious car-following model is presented in Section 6.2. The power of the car-following model over showing longitudinal drivers' behaviors is investigated in Section 6.3, with comparisons to earlier empirical findings. Section 6.4 simulates the capacity drop and its relation with congested states using the extended model. Finally, this chapter ends with conclusions and discussions in Section 6.5.

6.2 The car-following model

This section illustrates how to model the intra-driver variable longitudinal behavior variation for giving the capacity drop. The methodology is to incorporate the stochasticity into drivers' desired acceleration model (see Section 6.2.1), which will be incorporated into Newell's first-order car-following framework in Section 6.2.2. The extended parsimonious car-following is analyzed in terms of dimensionless parameters in Section 6.2.3.

6.2.1 The desired acceleration model

One vehicle speeds up when leaving a vehicular queue with a certain acceleration. Due to physical engine ability, the acceleration is bounded. For safety, the acceleration is always constrained by its predecessor to avoid collisions. Without the predecessor, the acceleration which the driver imposes to the vehicle is called as the desired acceleration. This section shows a desired acceleration model incorporating stochasticity.

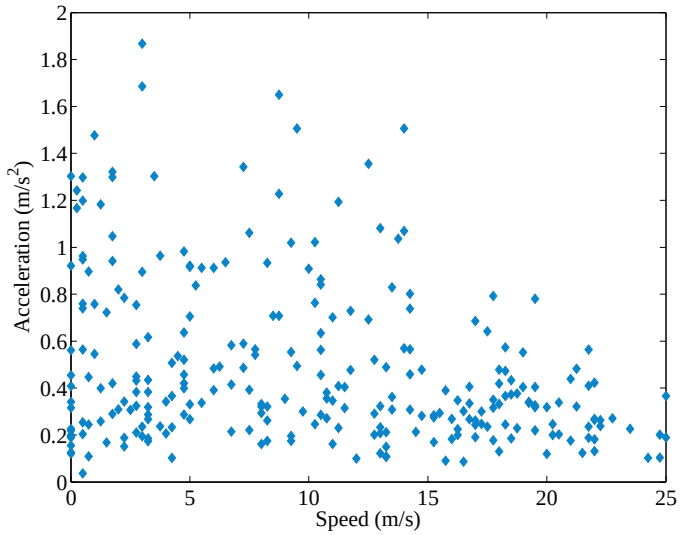
From empirical investigations, Laval et al. (2014) show that when a vehicle is traveling at speed $v(t)$ at time t, its mean desired acceleration $\mathrm{E}[a(v)]$ is a function of $v(t)$, that is:

$$\mathrm{E}[a(v(t))] = (v_f - v(t))\beta \tag{6.1}$$

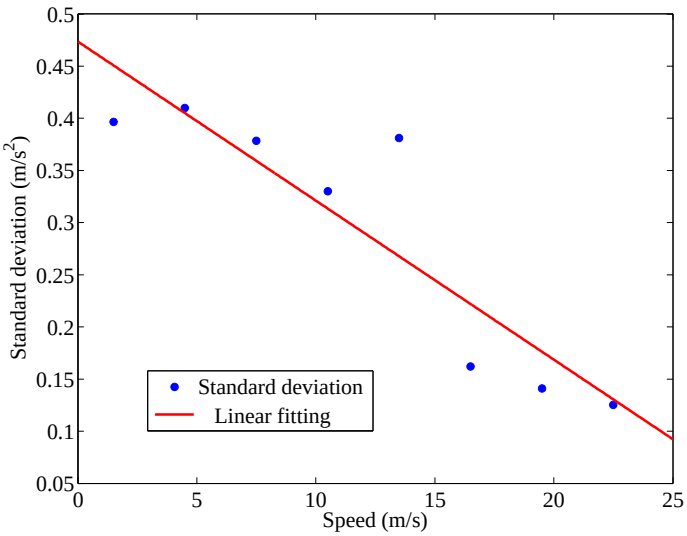
where v_f is the free-flow speed drivers want to speed up to and β is a positive parameter with a unit of $[\text{time}]^{-1}$.

Meanwhile, empirical data indicate that the desired acceleration standard deviation $\mathrm{Std}[a(v(t))]$ decrease as the speed $v(t)$ increases, see Figure 6.2(a).

The data analyzed in Figure 6.2 are collected GPS, accelerometer and barometer sensor information from Android smartphones by an Android application. The application is



(a) Drivers' desired acceleration when traveling at different speed.



(b) Standard deviation of drivers' desired accelerations at different vehicular speeds v.

Figure 6.2: Data collections for justifying the function between the vehicular speed and the desired accelerations. In Figure 6.2(a), all data are collected during one vehicle's accelerating process, including 15 acceleration processes. Note that the datasets analyzed here is collected when the vehicle is the first one in a platoon stopped in front of red signal at a junction, to ensure the acceleration used here is the desired acceleration without disturbances from further downstream. The data is the same as analyzed in Laval et al. (2014). Figure 6.2(b) shows the standard deviation of the desired accelerations collected around vehicular speed v (i.e., $[v-2.5\ \mathrm{m/s}, \mathrm{v}+2.5\ \mathrm{m/s}]$), shown as dot points. A linear function $\mathrm{Std}(a(v(t))) = -0.015v(t)+0.47$ is applied to fit the data in Figure 6.2(b).

programmed by authors of Laval et al. (2014), which can get position, speed, acceleration and altitude information. The data was collected during the acceleration process where the vehicle was the leader of the platoon at signalized intersections.

The methodology this chapter applies for incorporating the stochasticity into the desired acceleration model is to add a random term, that is, a standard Brownian motion $W(t)$, to (6.1). The standard Brownian motion $W(t)$ is also called the Wiener process, with covariance $\mathrm{Cov}(W(0),W(t)) = 0$ and correlation $\mathrm{Corr}(W(0),W(t)) = 0$. Hence, the desired acceleration is expressed as follows:

$$\mathrm{d}v(t) = (v_f - v(t))\beta \mathrm{d}t + (v_f - v(t))\sigma \mathrm{d}W(t) \tag{6.2}$$

where σ is a constant. Note that σ^2 has a unit of $[\text{time}]^{-1}$. Using $S(t) = (v_f - v(t))\beta$, we reformulate (6.2) as a stochastic differential equation in (6.3) :

$$\mathrm{d}S(t) = -\beta S(t)\mathrm{d}t + (-\sigma)S(t)\mathrm{d}W(t) \tag{6.3}$$

where $S(0) = S_0$ and $v(0) = v_0$. (6.3) indicates that the stochastic process $S(t)$ follows a Geometric Brownian motion. According to Øksendal (2010), the solution $S(t)$ is a log-normal distributed random variable with expected value $\mathrm{E}[S(t)] = S_0 e^{-\beta t}$ and variance $\mathrm{Var}[S(t)] = S_0^2 e^{-2\beta t}(e^{\sigma^2 t} - 1)$.

Hence, $v(t)$ follows log-normal distribution with $\mathrm{E}[\cdot]$ and $\mathrm{Var}[\cdot]$ given by:

$$\mathrm{E}[v(t)] = v_f - (v_f - v_0)e^{-\beta t} \tag{6.4a}$$

$$\mathrm{Var}[v(t)] = (v_f - v_0)^2 e^{-2\beta t}(e^{\sigma^2 t} - 1) \tag{6.4b}$$

The log-normal distribution of $v(t)$ means the speed of vehicle is always non-negative. A higher β gives higher E[v(t)] and lower Var[v(t)]. The Var[v(t)] is a positive function of σ^2. The E[v(t)] is independent of σ^2.

One behavioral assumption, that drivers cannot precisely control the pressure they give on the accelerator, can be made to help understand the acceleration error.

6.2.2 Incorporation into car-following framework

This section shows the incorporation of the desired acceleration model into the Newell's first order car-following framework. The Newell's first order car-following model can be expressed as:

$$x_{i+1}(t) = \min\left\{x_{i+1}(t-\tau) + \tau v_f, x_i(t-\tau) - \frac{1}{\rho_j}\right\} \tag{6.5}$$

where $\tau = 1/(w\rho_j)$ is the shortest time for congestion characteristic propagating upstream between two successive vehicles, $x_i(t)$ is the location of vehicle i at time t. The jam density ρ_j, shock wave speed $-w$ and free-flow speed v_f are three parameters of a triangular flow-density fundamental diagram.

To incorporate the $v(t)$ in to (6.5), the term τv_f is reformulated into $\tau \cdot \min(v_f, v_{i+1}(t))$. That is:

$$x_{i+1}(t) = \min\left\{x_{i+1}(t-\tau) + \tau \cdot \min(v_f, v_{i+1}(t)), x_i(t-\tau) - \frac{1}{\rho_j}\right\} \tag{6.6}$$

To implement the stochasticity into (6.6), the $v(t)$ in (6.6) is obtained by generating a random number which follows log-normal distribution with mean value and variance given by (6.4). The log-normal distribution ensures $v(t)$ is always non-negative.

6.2.3 Dimensional analysis

This section studies the extended parsimonious car-following model using dimensional analyses. The dimension means minimum number of coordinates required for mathematical descriptions. The analysis can facilitate understanding how independent parameters drive the model to reduce the queue discharge rate.

The trajectories of vehicles can be formulated as $x = \Theta\left(\beta, \sigma^2, t, v_f, \rho_j, w\right)$, according to (6.4) and (6.6). In this chapter, the capital Greek symbols (e.g., Θ) denote functions. To calculate the flow q, one more independent parameter N denoting the number of vehicles, is needed. That is, $q = \Phi\left(\beta, \sigma^2, t, v_f, \rho_j, w, N\right)$ is a function of seven independent parameters.

A dimensional analysis is applied for the reduction of parameters. In this work there are three fundamental dimensions, i.e., time, length and number. The units of the seven independent parameters are combinations of [time], [distance] and [number of vehicle]. Units can be rescaled and parameters are reformulated as $\tilde{x} = x\beta/v_f$, $\tilde{q} = q/\left(v_f\rho_j\right)$, $\tilde{\beta} = 1$, $\tilde{\sigma}^2 = \sigma^2/\beta$, $\tilde{t} = \beta t$, $\tilde{v}_f = 1$, $\tilde{\rho}_j = 1$, $\tilde{w} = w/v_f$ and $\tilde{N} = N\beta/\left(v_f\rho_j\right)$. Finally, we can have $\tilde{x} = \tilde{\Theta}\left(\tilde{\sigma}^2, \tilde{t}, \tilde{w}\right)$ and $\tilde{q} = \tilde{\Phi}\left(\tilde{\sigma}^2, \tilde{t}, \tilde{w}, \tilde{N}\right)$, which are dimensionless representations of the car-following model and the flow. Hence,

$$x = \frac{v_f}{\beta} \cdot \tilde{\Theta}\left(\frac{\sigma^2}{\beta}, \beta t, \frac{w}{v_f}\right) \tag{6.7a}$$

$$q = v_f\rho_j\tilde{\Phi}\left(\frac{\sigma^2}{\beta}, \beta t, \frac{w}{v_f}, \frac{N\beta}{v_f\rho_j}\right) \tag{6.7b}$$

In a specific scenario reproducing queue discharging, the estimated queue discharge rate q_{dis} should be independent of time t. With a predefined fundamental diagram, we can give the queue discharge rate as a function in (6.8):

$$q_{\text{dis}} \approx v_f\rho_j\tilde{\Phi}\left(\frac{\sigma^2}{\beta}, \frac{N\beta}{v_f\rho_j}\right) \tag{6.8}$$

which indicates when calculating the queue discharge rate, the dimensionless variable $\tilde{\sigma}^2$ and $\tilde{N}$ are the two variables determining queue discharge rate quantitatively. Different combinations of σ, β and N does not affect the queue discharge rate unless $\tilde{\sigma}^2$ or $\tilde{N}$ is changed. Because both of $\tilde{\sigma}^2$ and $\tilde{N}$ are strongly related to β, the sensitive analyses of $\tilde{\sigma}^2$ and $\tilde{N}$ can be transformed into that of σ and N, with a constant β.

6.3 Car-following platoon simulations

This section aims to examine to what extent the extended parsimonious car-following model can reproduce the car-following behavior in simulations. The methodology is to set up car-following platoon simulations. The reference is the oscillation propagation characterizations revealed in literature, that is the concavity of traffic oscillations (Jiang et al., 2014, 2015; Tian et al., 2016) and the growing-and-flattening pattern (Li et al., 2010; Li & Ouyang, 2011). Both of concavity and the oscillatory series, show that the oscillation magnitude, simulated by the extended parsimonious car-following model, grows to a bounded value and then stabilize there. This simulated oscillation propagation pattern is consistent with empirical observations in Li et al. (2010).

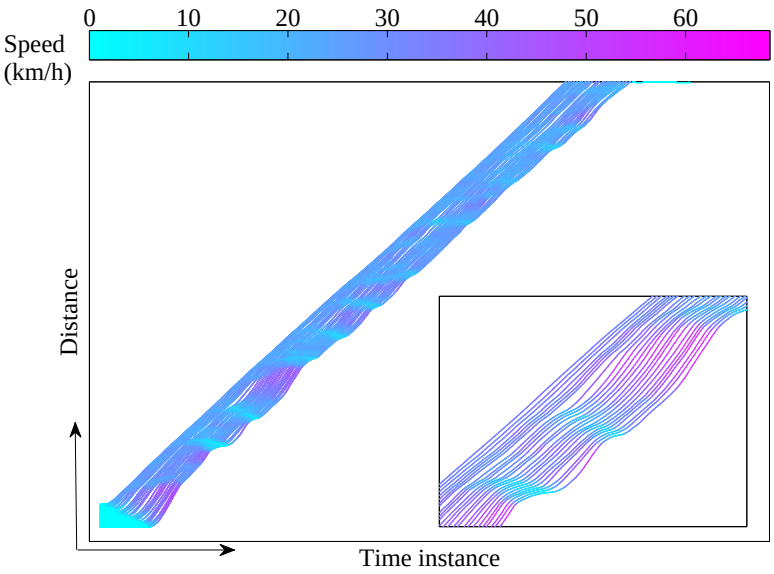
The concavity means the standard deviation of vehicular speed in a platoon with a constant-speed leader increases along the platoon in a concave way (Jiang et al., 2014, 2015). The concavity is revealed in a 25-vehicle car-following platoon experiment, which is reported for the first time in Jiang et al. (2014). Figure 6.3 shows two samples of the car-following experiments. Without lane changing, oscillations form and develop, see the subfigures in the right corner of each figure in Figure 6.3. From now on, this chapter calls the data collected in this experiment 'experiment data' for short. Tian et al. (2016) find that when the first vehicle's speed is in a range between 30km/h and 55km/h, the concavity revealed in both of car-following experiment datasets and NGSIM trajectory datasets can be fitted by a single concave curve. The extended parsimonious car-following model is expected to give the concavity of traffic oscillations and its independence on traffic speed if leading vehicle of the platoon is between 30 km/h and 55 km/h.

Li and Ouyang (2011) assume the vehicle trajectory is a superposition of a nominal series and oscillatory series. In Newell's first order car-following framework, the nominal series can be seen as the Newell's trajectory. That is, the following trajectory is a shift from its predecessor's trajectory by τ (in time) and $w\tau$ (in distance). The oscillatory series is obtained by subtracting the nominal series from the original trajectory. As characterized within a mathematical framework proposed in Li and Ouyang (2011), the oscillation amplitude which is shown as the standard deviation of the oscillatory series, grows to a certain bound value and then flattens out (i.e., the growing-and-flattening pattern). Following this finding, this chapter examines the propagation of the oscillatory series along the car number in simulations, too.

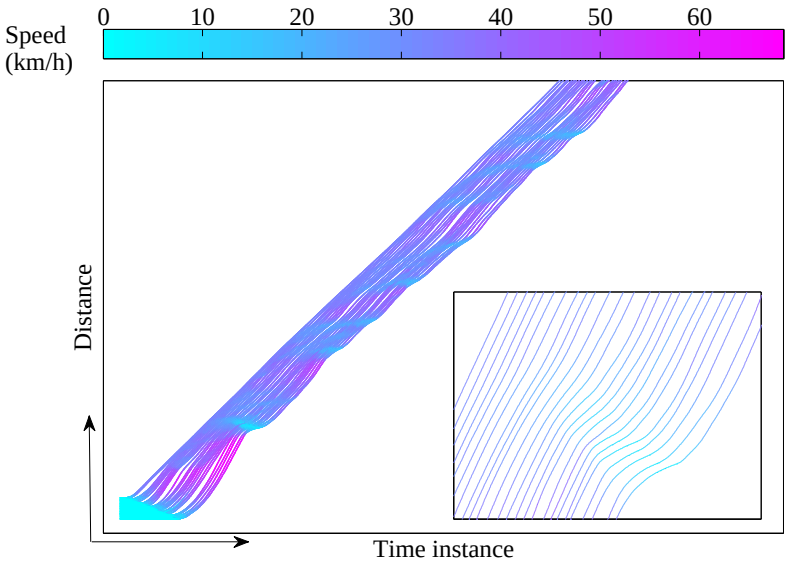
The simulation set-ups are presented in Section 6.3.1, followed by Section 6.3.2 which illustrates simulation results.

6.3.1 Simulation set-up

Following the experiment set-up in Jiang et al. (2014, 2015), this chapter simulates a 250-vehicle platoon on a one-lane road with the same fundamental diagram as proposed in Section 6.4. One more virtual leading vehicle is designed which can drive at



(a) $v_{\text{lead}} = 25$ km/h



(b) $v_{\text{lead}} = 35$ km/h

Figure 6.3: Samples of vehicular trajectories in car-following experiments. Without lane changing, oscillations can be observed. The subfigure at the right corner of each figure shows the formation and the development of oscillations.

a constant speed v_{lead} without fluctuations. In the following simulations, the v_{lead} is set to be 30 km/h, 40 km/h and 50 km/h in different simulations. With each v_{lead}, simulations run for 500 times. Each simulation runs for 300τ long. For different illustrations, this chapter sets different β and σ.

Finally, the experiment data are loaded into simulations. That is, the first leading vehicle drives along the trajectory of the first vehicle in the experiment. The rest 250 vehicles' trajectories are simulated according to the extended parsimonious car-following model.

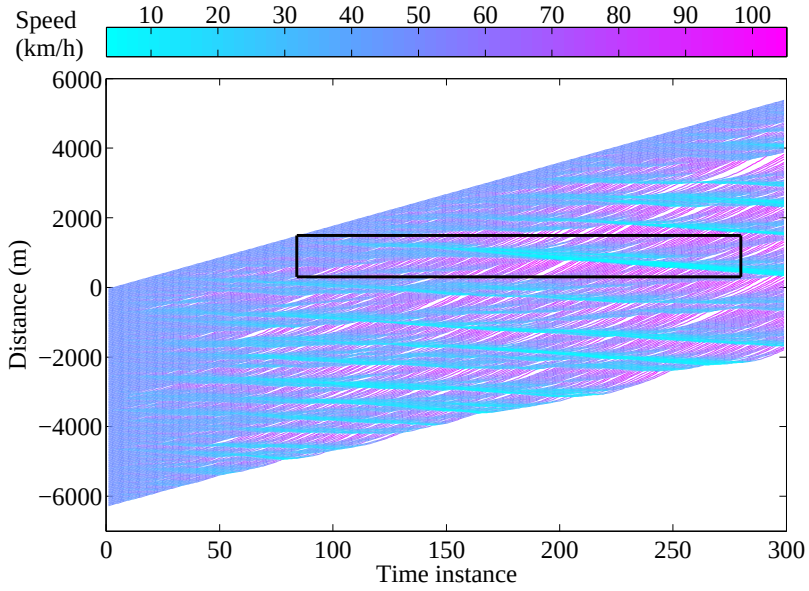
6.3.2 Simulation results

This section shows the simulation results in the car-following platoon simulation, studying the properties of the extended parsimonious model.

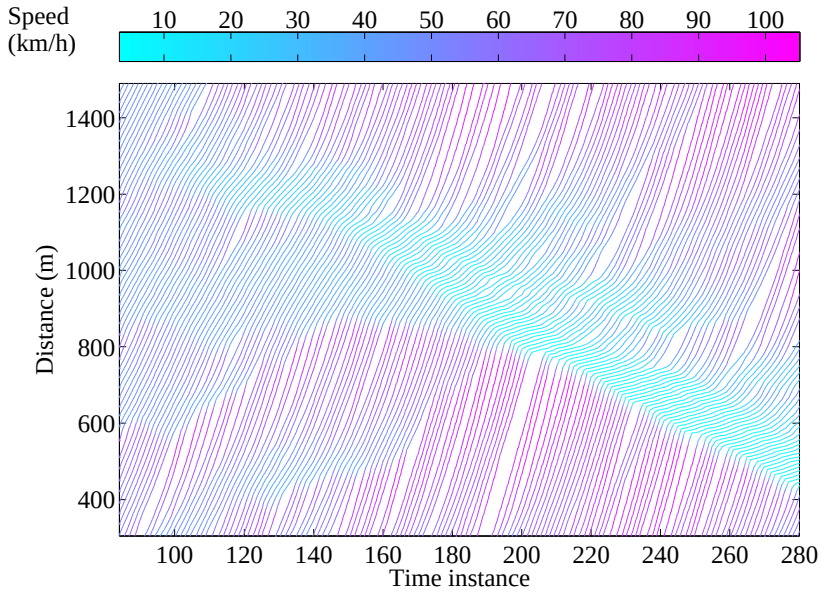
Figure 6.4 shows a sample of trajectories in simulations. The color bar indicates the vehicular speed. Even though there is no external disturbance (such as lane changing) simulated to trigger oscillations, oscillations can be observed, which is consistent with this empirical finding (see Figure 6.3). A rectangle is drawn in Figure 6.4(a). The trajectories in the rectangle is shown in Figure 6.4(b).

Figure 6.5 shows the relation between the speed standard deviation and the car number. The shaded areas illustrate the 95%-probability (interval between 97.5^{th} and 2.5^{th} percentile) and 66%-probability (interval between 83^{th} and 17^{th} percentile) of simulation results after 500-time simulations. The bold curve is the mean value of the simulation results. The v_{lead} increases from Figure 6.5(a) to Figure 6.5(c). It is shown that the standard deviation of vehicular speed increases in a concave curve along the platoon when the car number is small, such as smaller than 100 in simulations. As the vehicle number keeps growing, the standard deviation of speed flattens out around a bounded value. The bounded value increases as the v_{lead} increases. The dots in Figure 6.5 are the speed standard deviation obtained from the car-following experiments. It is found that the experiment data match the simulation results highly. Matching empirical data also means the difference of concavity due to the change of v_{lead} is trivial, as unveiled in Tian et al. (2016). Figure 6.6 shows how the oscillatory series propagate along the platoon. A growing-and-flattening pattern, as predicted in Li and Ouyang (2011), is given. In the model, the standard deviation of the oscillatory series stabilize at a bounded value. The bounded value increases as the v_{lead} increases. The growing-and-flattening pattern indicates that the extended parsimonious car-following model can show the nonlinearity property, according to Li and Ouyang (2011).

In both of the concavity and the oscillatory series, the characterization of traffic oscillation propagation given in simulations is the same as revealed or predicted in literature. Being consistent with earlier findings shows that the extend parsimonious car-following model can give properties of longitudinal behaviors.

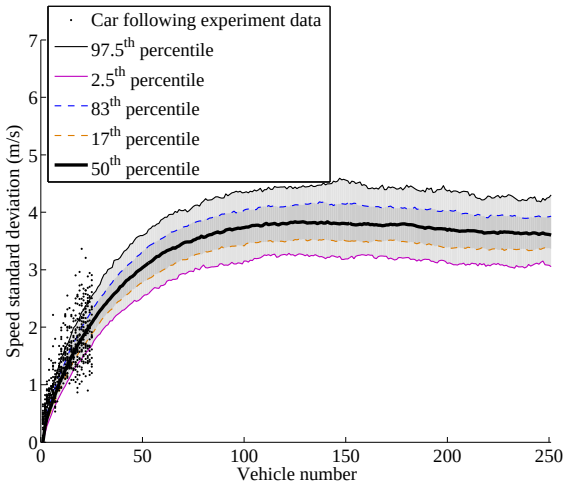


(a) $v_{\text{lead}} = 50$ km/h

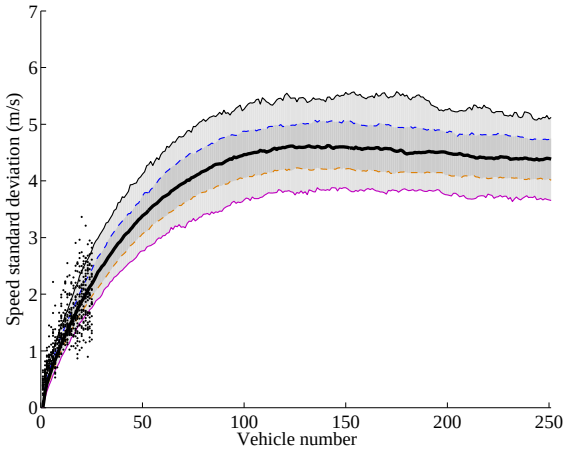


(b) A zoomed-in version of trajectories in the rectangle.

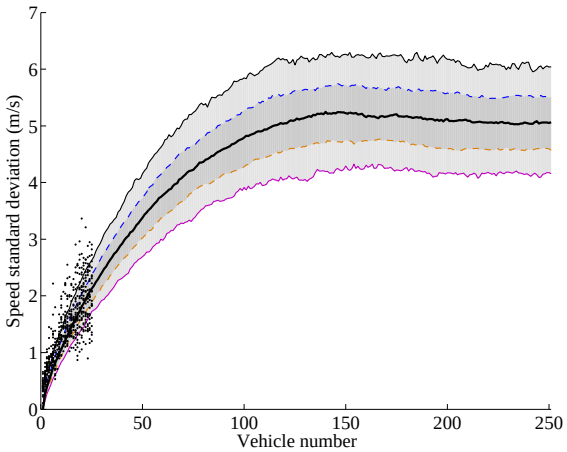
Figure 6.4: Samples of vehicular trajectories in simulations. The color map indicates the vehicular speed. $v_{\text{lead}} = 50$ km/h in this sample. An overview of the simulation results are given in (a). The trajectories in the rectangle in (a) is shown in (b) for a better visualizationn. In simulations, $\beta = 0.06$ and $\sigma = 0.06$.



(a) Concavity with $v_{\text{lead}} = 30$ km/h.

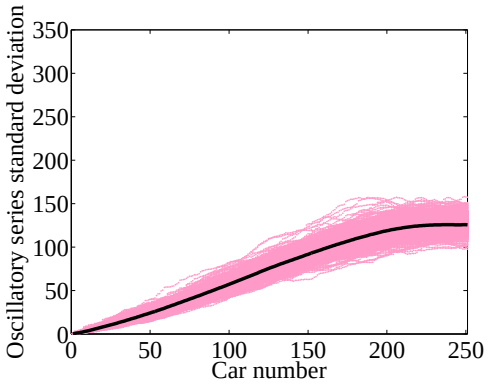


(b) Concavity with $v_{\text{lead}} = 40$ km/h.

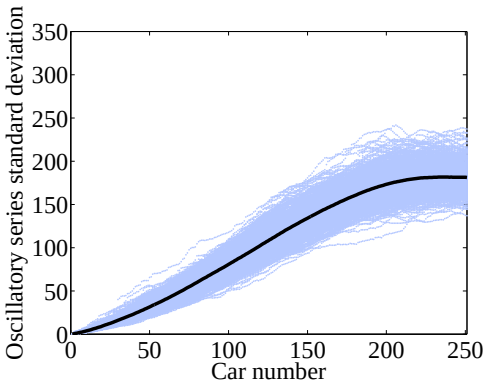


(c) Concavity with $v_{\text{lead}} = 50$ km/h.

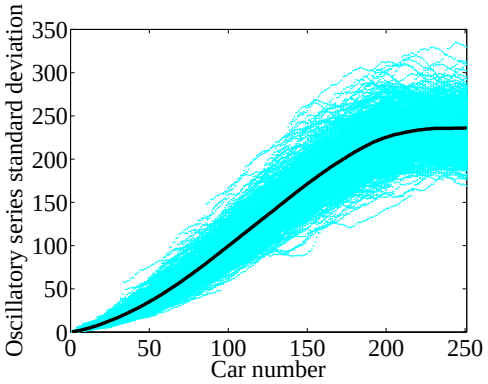
Figure 6.5: Concavity revealed in car-following platoon simulations. $\beta = 0.06$ and $\sigma = 0.055$.



(a) Standard deviation of oscillatory series with $v_{\text{lead}} = 30$ km/h.



(b) Standard deviation of oscillatory series with $v_{\text{lead}} = 40$ km/h.



(c) Standard deviation of oscillatory series with $v_{\text{lead}} = 50$ km/h.

Figure 6.6: Standard deviation of oscillatory series revealed in car-following platoon simulations when the leading vehicle in the platoon drives at (a) $v_{\text{lead}} = 30$ km/h, (b) $v_{\text{lead}} = 40$ km/h and (c) $v_{\text{lead}} = 50$ km/h. In simulations, $\beta = 0.06$ and $\sigma = 0.055$.

6.4 Queue discharge rate reductions

This section studies to what extent the extended car-following model can reproduce the capacity drop in simulations. As observed in Oh and Yeo (2015), the queue discharge rate reduction in absence of lane changing, is a large proportion of the reduced capacity at bottlenecks, outperforming the reduction due to lane changing. The capacity drop in this research is referred to the reduction of the outflow of stop-and-go waves, excluding the impacts of lane changing around bottlenecks (e.g., on-ramp and lane-drop bottlenecks). At the onset of each simulation, all simulated vehicles are driving in a queue with a speed $v_i(0)$. A virtual vehicle is assumed to drive in the front of the simulated queue, which can speed up with infinite acceleration. This virtual vehicle can be simulated with Newell's first order car-following model. After a while, the virtual vehicle accelerates to the free-flow speed immediately. All the rest vehicles are simulated with the extended parsimonious car-following model. In short, the virtual vehicle helps to simulate a scenario where a group of vehicles speed up to leave congestion.

In simulations, three features of the capacity drop are needed to be given:

- Feature 1: the queue discharge rate is lower than the free-flow capacity, according to the definition of the capacity drop phenomenon
- Feature 2: the queue discharge rate increases as the vehicular speed in congestion increases, see Figure 3.7
- Feature 3: The magnitude of the capacity drop at should be plausible, e.g., able to range up to 25% with a severe wide moving jam.

The simulation set-up is described in Section 6.4.1, followed by simulation results in Section 6.4.2.

6.4.1 Simulation set-up

This section describes the simulation set-up for examining the ability of giving the capacity drop in the extended parsimonious car-following model.

The fundamental diagram on simulated segment is a triangular fundamental diagram with $w = 18$ km/h, $v_f = 114$ km/h, and capacity $C = 2280$ veh/h. Accordingly, $\rho_j \approx 146.7$ veh/km and critical density $\rho_{\text{cri}} = 20$ veh/km. At time $t = 0$, all vehicles are driving in a congested state with an equilibrium spacing $1/\rho_0$. ρ_0 is the density in the congestion. That is, $v_i(0) = \frac{w(\rho_j - \rho_0)}{\rho_0}$, $i = 1,2,...N$. At $t = 10\tau$ (where $\tau = 1/w\rho_j$ is the time step), the first vehicle begins speeding up.

In simulations, ρ_0 varies from ρ_{cri} to ρ_j. With each ρ_0, the situation is simulated for 100 times. Every time, the queue discharge flow q_{dis} is calculated as:

$$q_{\text{dis}} = \frac{v_f}{\sum_{i=2}^{N} s_i(t)} \tag{6.9}$$

where $s_i(t) = x_{i-1}(t) - x_i(t)$ is the spacing in front of vehicle i at time t. $x_i(t)$ is the location of vehicle i at time t. The spacing between the first vehicle and the virtual vehicle is excluded.

6.4.2 Simulation results

One simulation for giving the capacity drop is illustrated in Section 6.4.1. The simulation results are presented in this section.

Figure 6.7 shows two samples of trajectories in simulations. At time $t = 0$, the density of the vehicular queues are $\rho_0 = 39$ veh/km in Figure 6.7(a) and $\rho_0 = \rho_j \approx 146.7$ veh/km in Figure 6.7(b).

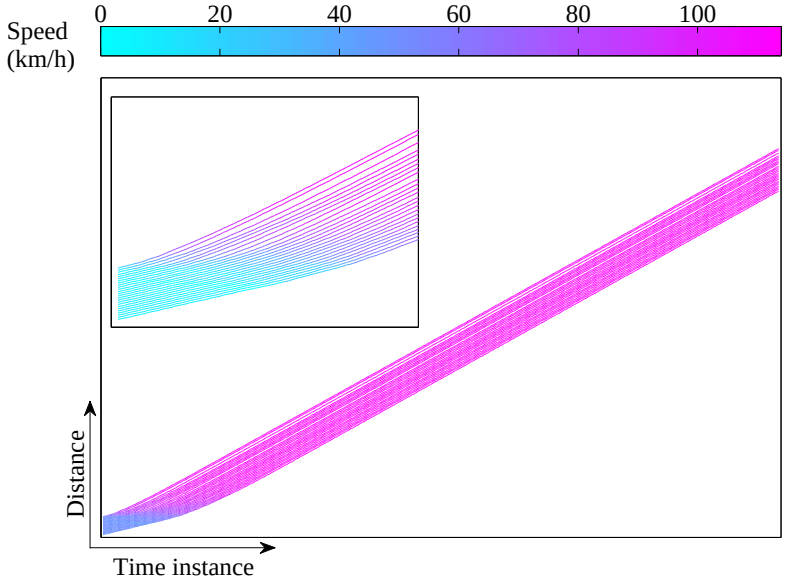
Figure 6.8 presents the simulation results on the sensitive analysis of the parameters N and σ on the queue discharge rate. In Figure 6.8(a), $\beta = 0.07\text{s}^{-1}$ and $\sigma = 0.05\text{s}^{-1/2}$. $\tilde{N}$ is a function of one variable N. It is shown that when $N \geqslant 441$ veh (or $\tilde{N} \geqslant 0.018$), the queue discharge rates are almost independent of N, flattening out as N keeps increasing. In the subplot inside, when $v_j = 0$ km/h, the standard deviations of the queue discharge rates are shown as error bars. The deviations decrease as the N increases, which is the same as in simulations with different v_j. The finding in Figure 6.8(a) indicates that, with a predefined fundamental diagram, the queue discharge rate is a function of one parameter $\tilde{\sigma}^2$ if the number of simulated vehicles N is large enough. Figure 6.8(b) shows the effects of σ on the queue discharge rate. It is found that the queue discharge rate decreases as the σ increases.

As shown in Section 6.2.3, it is $\tilde{\sigma}^2$ and $\tilde{N}$ that determine the queue discharge rate. σ and N can affect the queue discharge rate, as shown in Figure 6.8, because they can change $\tilde{\sigma}^2$ and $\tilde{N}$. That is, the influence of a multiplied β by δ on the queue discharge rate equals to that of the multiplications of σ by $1/\sqrt{\delta}$ and N by δ. The increase of β can result in the increase of the queue discharge rate.

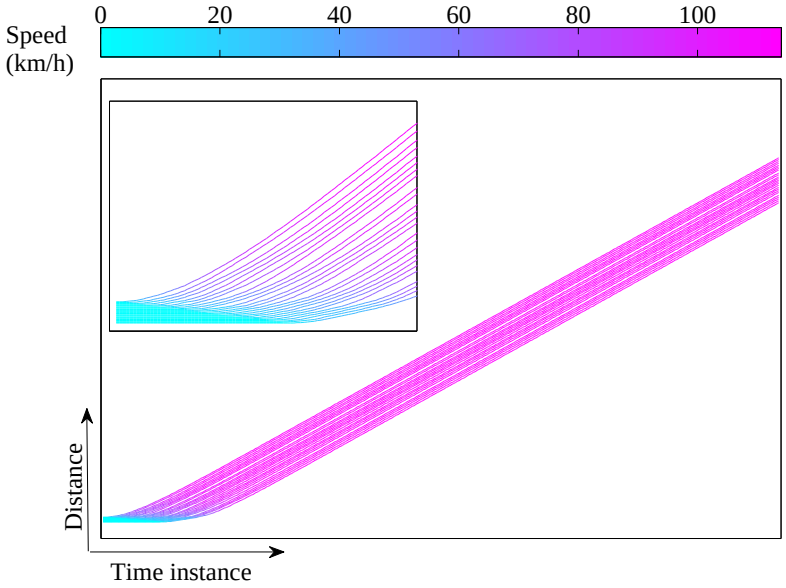
Figure 6.9 presents the ability of the extended parsimonious car-following model to give the capacity drop, which shows the positive relation between the speed in congestion and the queue discharge rate. When setting $\tilde{\sigma}^2 = 0.06$, the capacity drop can reach around 22% with a wide moving jam (i.e., $v_j = 0$ km/h), which is similar to the capacity drop magnitude shown in Figure 3.7. Hence, by simulating the stochasticity of the acceleration, the observed relation between the speed in congestion and the queue discharge rate can be given. The intra-driver variation mechanism can explained as the intra-driver desired acceleration variation.

6.5 Conclusion

This paper extended the parsimonious car-following model in (Laval et al., 2014) to reproduce empirical speed-capacity relationships. This was achieved by modeling driver

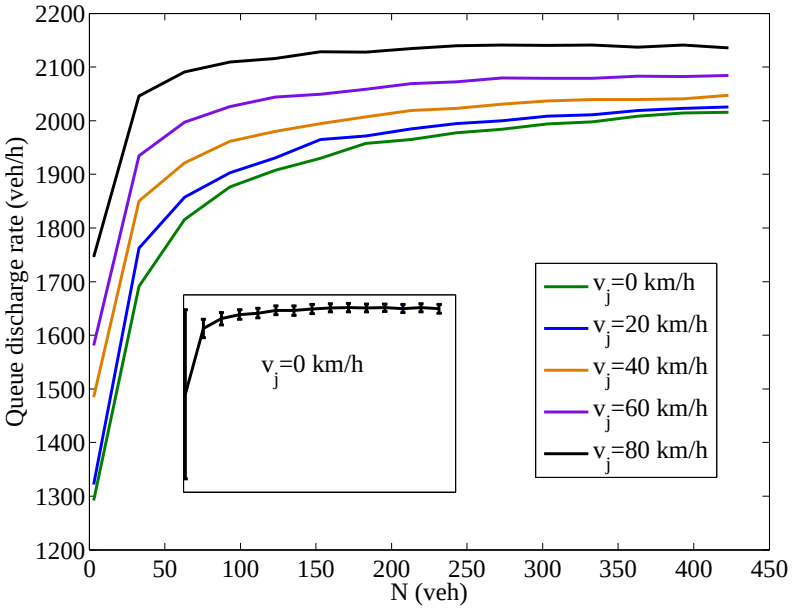


(a) $\rho_0 = 39$ veh/km

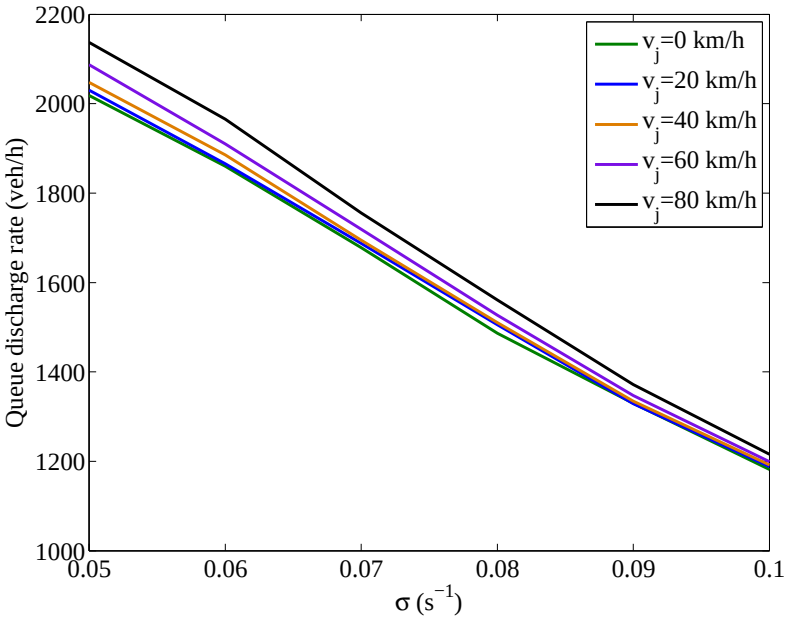


(b) $\rho_0 = \rho_j \approx 146.7$ veh/km.

Figure 6.7: Samples of vehicular trajectories in simulations. The color map indicates the vehicular speed. The density at $t = 0$ is (a) $\rho_0 = 39$ veh/km and (b) $\rho_0 = \rho_j \approx 146.7$ veh/km. $N = 50$ vehicles, excluding the virtual vehicle, are simulated on a one-lane road segment.



(a) Sensitive analyses of parameter N with constant $\beta = 0.07\text{s}^{-1}$ and $\sigma = 0.05\text{s}^{-1/2}$.



(b) Sensitive analyses of parameter σ with $\beta = 0.07\text{s}^{-1}$ and $N = 441$ veh.

Figure 6.8: Simulation results by means of observed queue discharge rates as a function of parameter N and σ.

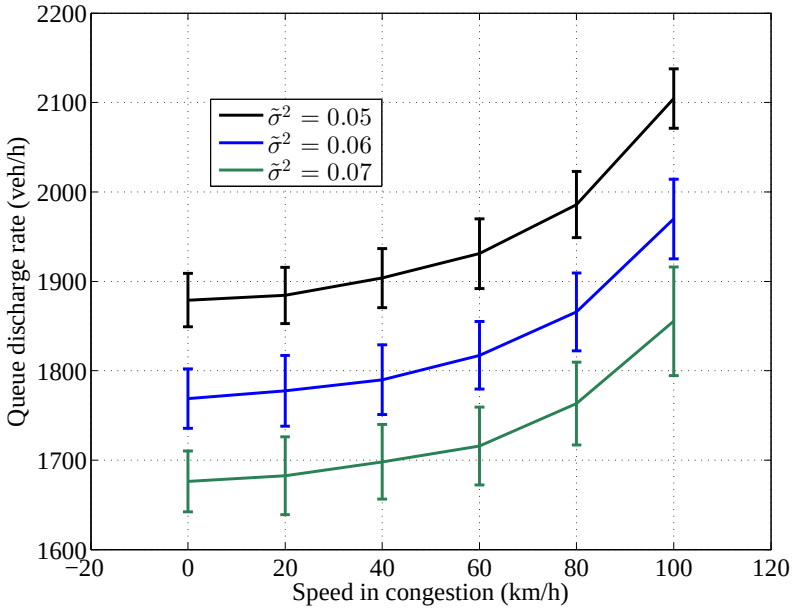


Figure 6.9: Relation between the speed in congestion and the capacity drop given by the extended parsimonious car-following model. The error bar indicates the standard deviation of the queue discharge rate.

acceleration errors as a Geometric Brownian motion, which is consistent with empirical data showing that the standard deviation of the desired acceleration is a decreasing function of the speed.

Notably, the extended model remains a single-parameter model if one is interested in flow or delay calculations; in particular, the speed-capacity relationship depends on a single parameter, $\tilde{\sigma}^2 = \sigma^2/\beta$. Similarly, the queue discharge rate was found to be a decreasing linear function of $\tilde{\sigma}$. From the driver behavior perspective, this parameter encapsulates human error, σ^2 and how fast acceleration changes with speed, β. Therefore, queue discharge rates can be increased by (i) decreasing human error as much as possible, especially at the beginning of the acceleration process, and/or (ii) increasing the acceleration at low speeds, by increasing engine power and/or educating drivers to be more aggressive when accelerating from a full stop.

Based on these findings, we can argue that newer vehicle technologies should be beneficial for reducing or maybe even eliminating capacity drop. Autonomous vehicles should be able to substantially reduce the variability of acceleration processes, which is consistent with (i) above; electric vehicles offer much greater accelerating power than conventional combustion engine vehicles, especially at lower speeds, which accords well with (ii).

Chapter 7

Conclusions

This thesis aims to properly understand the capacity drop phenomenon, motivated by the challenges in understanding the capacity drop phenomenon described in Chapter 1. This chapter firstly summarizes all findings and conclusions which have been presented through Chapter 2-6. This research in this thesis also sees challenges and opportunities. Following the summary, implications and recommendations for future research are discussed in this chapter.

7.1 Findings and conclusions

This thesis has three main objectives, that is, (1) unveiling features of the capacity drop, (2) investigating mechanisms behind the capacity drop from driver behavioral perspective, and (3) describing the feature and the mechanism mathematically. Four research questions arise in these three objectives. This section summarizes answers to those four research questions.

7.1.1 An empirical relation between the speed in congestion and the queue discharge rate

An empirical relation the speed in congestion and the queue discharge rate is unveiled to answer a question of what is the relation between the queue discharge rate reduction and the congestion characteristics (e.g., speeds, densities). It is found that there is a positive relation between the speed in congestion and the queue discharge rate. In Chapter 3, a linear relation is used to fit this relation, with a correlation coefficient $r = 0.9819$. This relation not only benefits understanding the observed wide range of capacity drop magnitude in literature, but also indicates the capacity drop size can be as high as around 25% (see Chapter 3).

The relation between the speed in congestion and the queue discharge rate also explains the correlation between the queue discharge rate and the congestion categories (i.e., stop-and-go waves and standing queues). Generally, the congestion in the stop-and-go wave is frequently much more sever than that in the standing queue, e.g., due to instability. Chapter 2 shows that without influences from the downstream, the outflow of a stop-and-go wave is generally lower than that of a standing queue. It is possible that the difference on the queue discharge rate between these two categorizes of congestion can be explained as a result of the different traffic speed in two types of congestion.

7.1.2 Justification for the queue discharge rate reduction from a longitudinal driver behavioral perspective

An explanation of queue discharge rate reduction from a longitudinal driver behavioral perspective is proposed, addressing the question of what is the relevant driver behavioral mechanism behind the capacity drop. It is believed that the queue discharge rate is reduced due to the stochasticity of desired accelerations. The desired acceleration is the acceleration which the driver imposes to the vehicle without a predecessor. The stochasticity of the desired acceleration consists two relations: (1) the mean value of a driver's desired acceleration decreases as the vehicular speed increases, and (2) the standard deviation of a driver's desired acceleration decreases as the vehicular speed increases. The stochasticity of the desired acceleration is verified from empirical

floating-car data, see Chapter 6. The stochasticity of the desired acceleration specifies how the intra-driver variation reduces the queue discharge rate.

On the contrary, it is identified that the other longitudinal explanation, i.e., inter-vehicle spread or maximum acceleration spread among vehicles, can only slightly decrease the queue discharge rate. It is worthwhile to know that no matter whether the effect of intra-driver variation is considered or not, the impact of the inter-vehicle spread on the capacity drop is marginal. Both of analytical analysis and numerical simulations indicate that modeling intra-driver variance, rather than the inter-driver spread, in car-following is relevant for giving the capacity drop (see Chapter 5).

7.1.3 A kinematic wave model in Lagrangian coordinates incorporating capacity drop

A kinematic wave model is presented and solved in the Lagrangian coordinates in thesis, which answers the question of how to incorporate the capacity drop into a first-order macroscopic traffic flow model. The new model is motivated by the need for a first-order model which can accurately present traffic state evolutions and give traffic delays. Such accurate descriptions are crucial for proposing and evaluating dynamic traffic management strategies. The main idea in this new kinematic wave model is that the fundamental flow-density relation differs in accelerating and non-accelerating (i.e., cruise driving or decelerating) conditions. A concept of hysteresis is used for giving the relation between the speed in congestion and the queue discharge rate. In the flow-density fundamental diagram, the acceleration branch is assumed to be lower than the deceleration and those two branches intersect at a stable congested state. There is only one deceleration branch and its relationship is fixed. There are many acceleration branches and each of them corresponds to a stable congested state. Those acceleration branches are predefined according to a function describing the relation between the speed in congestion and the queue discharge rate. With a linear relation as suggested in Chapter 2, only two additional parameters are required next to those describing a fundamental diagram.

7.1.4 An extended parsimonious car-following model

A car-following model is presented in this thesis, which is an extended version of a parsimonious car-following model proposed in Laval et al. (2014). Presenting the extended parsimonious car-following answers the question of how to incorporate the relevant driver behavioral mechanism of the capacity drop into a microscopic simulation model. As identified in Chapter 6, the stochasticity of the desired acceleration is a relevant longitudinal behavior for justifying the capacity drop. Thus, the extended parsimonious car-following model aims to incorporating the stochasticity of the desired acceleration into Newell's first-order car-following model. This new model adds two

additional parameters to Newell's first-order car-following model. In simulations with a predefined fundamental diagram, if the simulation time is long enough and the sample size of the simulated vehicles is sufficiently large, the queue discharge rate is function of a single dimensionless parameter. The extended parsimonious car-following model can capture other car-following features revealed in literature, including the concavity of traffic oscillations and the growing-and-flattening pattern of oscillations.

7.2 Implications for practice

The work in this thesis have some important practical implications and contributions, opening a door to developing effective dynamic on-ramp metering strategies to diminish traffic jams, providing directions for using intelligent vehicles to increase road performance and improving traffic simulation tools.

First, the dependence of the queue discharge rate on the congested state provides a direction for designing on-ramp metering strategies. The relation between the speed in congestion and the queue discharge rate indicates that the future on-ramp metering should focus on controlling traffic states by managing merging behaviors at on-ramp bottlenecks. It is advised to assess the impacts of merging (e.g., location, speed and number) on congested states before proposing an on-ramp metering strategy.

Second, the unveiled behavioral explanation of the capacity drop indicates possible solutions to increasing queue discharge rate. Newer vehicle technologies should be beneficial for reducing or maybe even eliminating capacity drop. For example, autonomous vehicles should be able to substantially reduce the variability of acceleration processes; or electric vehicles offer much greater accelerating power than conventional combustion engine vehicles, especially at lower speeds, which can also contribute to increasing the queue discharge rate. What is more, a speed limit immediately downstream of a fixed bottleneck could help increase the queue discharge rate, by speeding up acceleration changes with speed.

Third, the kinematic wave model and the car-following model presented in this thesis show new approaches to efficiently and effectively describe traffic state transitions and traffic behaviors. Applying these new model or extending them are expected to improve the predictive and reproductive power of traffic simulation tool and promote the accuracy in evaluating the effects of traffic operations.

7.3 Recommendations for future research

The findings and conclusions following from the analyses presented in this thesis lead to the several recommendations for future research.

The first recommendation is to calibrate the extended parsimonious car-following model. The calibration is relevant for the practical implication of the car-following model. Second, the calibration can provide more solid evidences on to what extent the stochasticity of desired accelerations contribute to the capacity drop. Currently, the data applied is collected in urban environment, which can be quite different from that on freeways. It is important to remark that there is no traffic light on freeways, which might be difficult to detect the desired acceleration on freeways directly.

Second, it is recommended to study the impact of the Driver-vehicle-combination (DVC) heterogeneity on the capacity drop. This thesis does not consider the heterogeneity because it is believed that the heterogeneity is not a necessary condition for giving the capacity drop. However, the DVC heterogeneity could be a sufficient condition for reducing the queue discharge rate. Current knowledge cannot explain explicitly how the heterogeneity reduces the queue discharge rate. Filling the gap can facilitate promoting road performances, e.g., by proposing vehicle-class specific control strategies or lane-specific management.

Third, the development of strategies for maintaining higher throughputs on freeway operations are required. Both of the macroscopic features of the capacity drop and the relevant driver behavior for the queue discharge rate reduction have already been identified in this thesis, that have shown some fundamental directions for minimizing the capacity drop through traffic operations. Especially as the development of the ITS, it is possible to try managing individual vehicle behavior for maximizing traffic throughputs.

The last but not least, the impacts of mixed traffic including automatic driving, connected driving and legacy vehicles on the queue discharge rate need investigations. The investigation includes coming up with theories, modeling and simulations for understanding the impact, as well as proposing control strategies for increasing the queue discharge rate.

References

Ahn, S., & Cassidy, M. J. (2007). Freeway traffic oscillation and vehicle lane-change maneuvers. In *Proceeding of the 17th international symposium on transportation and traffic theory (isttt17).* London, UK.

Alvarez-Icaza, L., & Islas, G. (2013, Oct). Hysteretic cell transmission model. In *Intelligent transportation systems - (itsc), 2013 16th international ieee conference on* (p. 578-583). doi: 10.1109/ITSC.2013.6728293

Aw, A., & Rascle, M. (2000). Resurrection of "second order" models of traffic flow. *SIAM Journal on Applied Mathematics*, *60*(3), 916-938. doi: 10.1137/S0036139997332099

Bando, M., Hasebe, K., Nakayama, A., Shibata, A., & Sugiyama, Y. (1995, Feb). Dynamical model of traffic congestion and numerical simulation. *Phys. Rev. E*, *51*, 1035–1042. doi: 10.1103/PhysRevE.51.1035

Banks, H., James. (1991a). The two-capacity phenomenon: Some theoretical issues. *Transportation Research Record*(1320), p. 234-241.

Bertini, R. L., & Leal, M. T. (2005). Empirical study of traffic features at a freeway lane drop. *Journal of Transportation Engineering*, *131*(6), 397-407. doi: 10.1061/(ASCE)0733-947X(2005)131:6(397)

Brilon, W., Geistefeldt, J., & Regler, M. (2005). Reliability of freeway traffic flow: a stochastic concept of capacity. In *Proceedings of the 16th international symposium on transportation and traffic theory* (Vol. 125143).

Carlson, R. C., Papamichail, I., & Papageorgiou, M. (2011, Dec). Local feedback-based mainstream traffic flow control on motorways using variable speed limits. *IEEE Transactions on Intelligent Transportation Systems*, *12*(4), 1261-1276. doi: 10.1109/TITS.2011.2156792

Cassidy, M. J. (1998). Bivariate relations in nearly stationary highway traffic. *Transportation Research Part B: Methodological*, *32*(1), 49 - 59. doi: http://dx.doi.org/10.1016/S0191-2615(97)00012-X

Cassidy, M. J., & Bertini, R. L. (1999). Some traffic features at freeway bottlenecks. *Transportation Research Part B: Methodological*, *33*(1), 25 - 42.

Cassidy, M. J., & Rudjanakanoknad, J. (2005). Increasing the capacity of an isolated merge by metering its on-ramp. *Transportation Research Part B: Methodological*, *39*(10), 896 - 913.

Chen, D., Ahn, S., Bang, S., & Noyce, D. (2016). Car-following and lane-changing behavior involving heavy vehicles. In *inproceedings of transportation research board 95th annual meeting.* Washington DC, United States.

Chen, D., Ahn, S., Laval, J., & Zheng, Z. (2014). On the periodicity of traffic oscillations and capacity drop: The role of driver characteristics. *Transportation Research Part B: Methodological*, *59*, 117 - 136.

Chen, D., Laval, J., Zheng, Z., & Ahn, S. (2012a). A behavioral car-following model that captures traffic oscillations. *Transportation Research Part B: Methodological*, *46*(6), 744 - 761. doi: http://dx.doi.org/10.1016/j.trb.2012.01.009

Chen, D., Laval, J. A., Ahn, S., & Zheng, Z. (2012b). Microscopic traffic hysteresis in traffic oscillations: A behavioral perspective. *Transportation Research Part B: Methodological*, *46*(10), 1440 - 1453. doi: http://dx.doi.org/10.1016/j.trb.2012.07.002

Chung, K., Rudjanakanoknad, J., & Cassidy, M. J. (2007). Relation between traffic density and capacity drop at three freeway bottlenecks. *Transportation Research Part B: Methodological*, *41*(1), 82 - 95.

Coifman, B. (2015). Empirical flow-density and speed-spacing relationships: Evidence of vehicle length dependency. *Transportation Research Part B: Methodological*, *78*, 54 - 65. doi: http://dx.doi.org/10.1016/j.trb.2015.04.006

Coifman, B., & Kim, S. (2011). Extended bottlenecks, the fundamental relationship, and capacity drop on freeways. *Procedia - Social and Behavioral Sciences*, *17*(0), 44–57.

Daganzo, C. F. (1993). *The cell-transmission model. part I: A simple dynamic representation of highway traffic* (Tech. Rep.). Institute of Transportation Studies, University of California, Berkeley.

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

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. doi: http://dx.doi.org/10.1016/S0191-2615(00)00042-4

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

Daganzo, C. F., Cassidy, M., & Bertini, R. (1999). Possible explanations of phase transitions in highway traffic. *Transportation Research Part A: Policy and Practice*, *33*(5), 365 - 379. doi: http://dx.doi.org/10.1016/S0965-8564(98)00034-2

Deng, H., & Zhang, H. M. (2015). On traffic relaxation, anticipation, and hysteresis. *Transportation Research Record: Journal of the Transportation Research Board*, *2491*, 90-97. doi: 10.3141/2491-10

Duret, A., Bouffier, J., & Buisson, C. (2010). Onset of congestion from low-speed merging maneuvers within free-flow traffic stream. *Transportation Research Record: Journal of the Transportation Research Board*, *2188*, 96-107.

Goemans, J., Daamen, W., & Heikoop, H. (2011). Dutch highway capacity manual ("handboek capaciteitswaarden infrastructuur autosnelwegen (cia) volledig vernieuwd") [Computer software manual].

Hall, F. L., & Agyemang-Duah, K. (1991). Freeway capacity drop and the definition of capacity. *Transportation Research Record*(1320), p. 91-98.

Haut, B., & Bastin, G. (2007). A second order model of road junctions in fluid models of traffic networks. *Networks and Heterogeneous Media*, *2*, 227.

Hegyi, A. (2004). *Model predictive control for integrating traffic control measures* (PhD thesis). Delft University of Technology.

Hegyi, A., De Schutter, B., & Hellendoorn, H. (2005). Model predictive control for optimal coordination of ramp metering and variable speed limits. *Transportation Research Part C*, *13*(3), 185-209.

Hegyi, A., & Hoogendoorn, S. P. (2010, Sept). Dynamic speed limit control to resolve shock waves on freeways - field test results of the specialist algorithm. In *Intelligent transportation systems (itsc), 2010 13th international ieee conference on* (p. 519-524). doi: 10.1109/ITSC.2010.5624974

Helbing, D., & Treiber, M. (1999). Numerical simulation of macroscopic traffic equations. *Computing in Science and Engineering*, *1*(5), 89-99. doi: http://doi.ieeecomputersociety.org/10.1109/5992.790593

Hoogendoorn, S. P. (2007). *Traffic flow theory and simulation*. Delft University of Technology.

Hoogendoorn, S. P., & Bovy, P. H. (2001, 6). State-of-the-art of vehicular traffic flow modelling. *Journal of Systems and Control Engineering*, *215*(4), 283-303.

Jiang, R., Hu, M.-B., Zhang, H., Gao, Z.-Y., Jia, B., & Wu, Q.-S. (2015). On some experimental features of car-following behavior and how to model them. *Transportation Research Part B: Methodological*, *80*, 338 - 354. doi: http://dx.doi.org/10.1016/j.trb.2015.08.003

Jiang, R., Hu, M.-B., Zhang, H. M., Gao, Z.-Y., Jia, B., Wu, Q.-S., ... Yang, M. (2014, 04). Traffic experiment reveals the nature of car-following. *PLoS ONE*, *9*(4), e94351. doi: 10.1371/journal.pone.0094351

Jin, W.-L., Gan, Q.-J., & Lebacque, J.-P. (2015). A kinematic wave theory of capacity drop. *Transportation Research Part B: Methodological*, *81, Part 1*, 316 - 329. doi: http://dx.doi.org/10.1016/j.trb.2015.07.020

Kai, N. (1996). Particle hopping models and traffic flow theory. *Physical Review E: Statistical, Nonlinear, and Soft Matter Physics*, *53*, 4655-4672.

Kerner, B. S. (1998, Oct). Experimental features of self-organization in traffic flow. *Phys. Rev. Lett.*, *81*, 3797–3800.

Kerner, B. S. (2002, Apr). Empirical macroscopic features of spatial-temporal traffic patterns at highway bottlenecks. *Phys. Rev. E*, *65*, 046138. doi: 10.1103/PhysRevE.65.046138

Kerner, B. S. (2004). *The physics of traffic: Empirical freeway pattern features, engineering applications, and theory*. Springer.

Keyvan-Ekbatani, M., Knoop, V. L., & Daamen, W. (2015). Categorization of the lane change decision process on freeways. *Transportation Research Part C: Emerging Technologies*, -. doi: http://dx.doi.org/10.1016/j.trc.2015.11.012

Khoshyaran, M., & Lebacque, J. (2015). Capacity drop and traffic hysteresis as a consequence of bounded acceleration. *IFAC-PapersOnLine*, *48*(1), 766 - 771. (8th Vienna International Conferenceon Mathematical ModellingMATHMOD 2015) doi: http://dx.doi.org/10.1016/j.ifacol.2015.05.160

Kim, S., & Coifman, B. (2013). Driver relaxation impacts on bottleneck activation, capacity, and the fundamental relationship. *Transportation Research Part C: Emerging Technologies*, *36*, 564 - 580. doi: http://dx.doi.org/10.1016/j.trc.2013.06.016

Knoop, V., Hoogendoorn, S. P., & Zuylen, H. (2009). Traffic and granular flow '07. In C. Appert-Rolland, F. Chevoir, P. Gondret, S. Lassarre, J.-P. Lebacque, & M. Schreckenberg (Eds.), (pp. 351–356). Berlin, Heidelberg: Springer Berlin Heidelberg.

Kontorinaki, M., Spiliopoulou, A., Roncoli, C., & Papageorgiou, M. (2016, January). Capacity drop in first-order traffic flow models: Overview and real-data validation. In *Proceedings of the 95th annual meeting of the transportation research board (TRB).* (paper no. 16-3541)

Kotsialos, A., Papageorgiou, M., Diakaki, C., Pavlis, Y., & Middelham, F. (2002, Dec). Traffic flow modeling of large-scale motorway networks using the macroscopic modeling tool METANET. *Intelligent Transportation Systems, IEEE Transactions on*, *3*(4), 282-292. doi: 10.1109/TITS.2002.806804

Koutsopoulos, H. N., & Farah, H. (2012, June). Latent class model for car following behavior. *Transportation Research Part B: Methodological*, *46*(5), 563–578.

Laval, J. A. (2011). Hysteresis in traffic flow revisited: An improved measurement method. *Transportation Research Part B: Methodological*, *45*(2), 385 - 391.

Laval, J. A., & Daganzo, C. F. (2006). Lane-changing in traffic streams. *Transportation Research Part B: Methodological*, *40*(3), 251 - 264.

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 of London A: Mathematical, Physical and Engineering Sciences*, *368*(1928), 4519–4541. doi: 10.1098/rsta.2010.0138

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. doi: http://dx.doi.org/10.1016/j.trb.2013.02.008

Laval, J. A., Toth, C. S., & Zhou, Y. (2014). A parsimonious model for the formation of oscillations in car-following models. *Transportation Research Part B: Methodological*, *70*, 228 - 238. doi: http://dx.doi.org/10.1016/j.trb.2014.09.004

Lebacque, J. (2002). A two phase extension of the lwr model based on the boun-bounded of traffic acceleration. In *Proceedings of the 15th international symposium on transportation and traffic theory (ISTTT).* South Australia.

Lebacque, J. (2003a). Two-phase bounded-acceleration traffic flow model: Analytical solutions and applications. *Transportation Research Record: Journal of the Transportation Research Board*, *1852*, 220–230.

Lebacque, J. (2003b). Traffic and granular flow. In S. P. Hoogendoorn, S. Luding,

P. H. L. Bovy, M. Schreckenberg, & D. E. Wolf (Eds.), (pp. 261–278). Berlin, Heidelberg: Springer Berlin Heidelberg.

Lebacque, J., Mammar, S., & Salem, H. H. (2007). Generic second order traffic flow modelling. In R. Allsop, M. Bell, & B. Heydecker (Eds.), *Proceeding of the 17th international symposium on transportation and traffic theory (isttt17).*

Leclercq, L. (2007, August). Hybrid approaches to the solutions of the "Lighthill-Whitham-Richards" model. *Transportation Research Part B: Methodological*, *41*(7), 701–709.

Leclercq, L., Knoop, V., Marczak, F., & Hoogendoorn, S. (2014). Capacity drops at merges: New analytical investigations. In *Intelligent transportation systems (ITSC), 2014 ieee 17th international conference on* (pp. 1129–1134).

Leclercq, L., Knoop, V. L., Marczak, F., & Hoogendoorn, S. P. (2015). Capacity drops at merges: New analytical investigations. *Transportation Research Part C: Emerging Technologies*.

Leclercq, L., Laval, J. A., & Chevallier, E. (2007). The lagrangian coordinates and what it means for first order traffic flow models. In B. Heydecker, M. Bell, & R. Allsop (Eds.), *17th international symposium on transportation and traffic theory (isttt).* New York: Elsevier.

Leclercq, L., Laval, J. A., & Chiabaut, N. (2011). Capacity drops at merges: An endogenous model. *Transportation Research Part B: Methodological*, *45*(9), 1302 - 1313.

Leclercq, L., Marczak, F., Knoop, V. L., & Hoogendoorn, S. (2016). Capacity drops at merges: Analytical expressions for multilane freeways. In *Proceedings of the 95th annual meeting of the transportation research board (TRB).* Washington D.C..

Li, X., & Ouyang, Y. (2011). Characterization of traffic oscillation propagation under nonlinear car-following laws. *Transportation Research Part B: Methodological*, *45*(9), 1346 - 1361. (Select Papers from the 19th {ISTTT}) doi: http://dx.doi.org/10.1016/j.trb.2011.05.010

Li, X., Peng, F., & Ouyang, Y. (2010). Measurement and estimation of traffic oscillation properties. *Transportation Research Part B: Methodological*, *44*(1), 1 - 14. doi: http://dx.doi.org/10.1016/j.trb.2009.05.003

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*(1178), 317–345. doi: 10.1098/rspa.1955.0089

Lu, Y., Wong, S. C., Zhang, M., & Shu, C.-W. (2009). The entropy solutions for the lighthill-whitham-richards traffic flow model with a discontinuous flow-density relationship. *Transportation Science*, *43*(4), 511-530.

Maria, T. J., Soriguera, F., & Geroliminis, N. (2014, January). Coordinated active traffic management freeway strategies using capacity-lagged cell transmission model. In *Proceedings of the 93rd annual meeting of the transportation research board (TRB).* Washington DC.

May, A. D. (1990). *Traffic flow fundamentals*. Prentice Hall.

Messmer, A., & Papageorgiou, M. (1990). Metanet: A macroscopic simulation program for motorway networks. *Traffic Engineering and Control*.

Michalopoulos, P. G., Yi, P., & Lyrintzis, A. S. (1992). Development of an improved high-order continuum traffic flow model. *Transportation Research Record*.

Minderhoud, M. (1999). *Supported driving: Impacts on moterway traffic flow* (PhD thesis). Delft University of Technology.

Monamy, T., Haj-Salem, H., & Lebacque, J.-P. (2012). A macroscopic node model related to capacity drop. *Procedia - Social and Behavioral Sciences*, *54*, 1388 - 1396. (Proceedings of EWGT2012 - 15th Meeting of the EURO Working Group on Transportation, September 2012, Paris) doi: http://dx.doi.org/10.1016/j.sbspro.2012.09.853

Muñoz, J., & Daganzo, C. (2002). Moving bottlenecks: A theory grounded on experimental observation. In *Proceedings of the 15th international symposium on transportation and traffic theory* (p. 441-461).

Muralidharan, A., & Horowitz, R. (2015). Computationally efficient model predictive control of freeway networks. *Transportation Research Part C: Emerging Technologies*, *58, Part C*, 532 - 553. (Special Issue: Advanced Road Traffic Control) doi: http://dx.doi.org/10.1016/j.trc.2015.03.029

Newell, G. (1962). Theories of instability in dense highway traffic. *The Operations Research Society of Japan*.

Newell, G. (2002, March). A simplified car-following theory: a lower order model. *Transportation Research Part B: Methodological*, *36*(3), 195–205.

Nishinari, K., Treiber, M., & Helbing, D. (2003, December). Interpreting the wide scattering of synchronized traffic data by time gap statistics. *Phys. Rev. E*, *68*(6), 067101–.

Oh, S., & Yeo, H. (2012). Estimation of capacity drop in highway merging sections.

Transportation Research Record: Journal of the Transportation Research Board, *2286*, 111-121. doi: 10.3141/2286-13

Oh, S., & Yeo, H. (2015). Impact of stop-and-go waves and lane changes on discharge rate in recovery flow. *Transportation Research Part B: Methodological*, *77*, 88 - 102.

Øksendal, B. (2010). *Stochastic differential equations: An introduction with applications*. Springer.

Ou, Q. (2011). *Fusing heterogeneous traffic data: Parsimonious approaches using data-data consistency* (PhD Thesis). Delft University of Technology.

Papageorgiou, M. (1998). Some remarks on macroscopic traffic flow modelling. *Transportation Research Part A: Policy and Practice*, *32*(5), 323 - 329.

Papageorgiou, M., Papamichail, I., Spiliopoulou, A., & Lentzakis, A. (2008, October). Real-time merging traffic control with applications to toll plaza and work zone management. *Transportation Research Part C: Emerging Technologies*, *16*(5), 535–553.

Parzani, C., & Buisson, C. (2012). Second-order model and capacity drop at merge. *Transportation Research Record: Journal of the Transportation Research Board*, *2315*, 25-34. doi: 10.3141/2315-03

Payne, H. J. (1979). Freflo: A macroscopic simulation model of freeway traffic. *Transportation Research Record: Journal of the Transportation Research Board*.

Reiss, R.-D. (1989). *Approximate distributions of order statistics: with applications to nonparametric statistics*. Springer-Verlag.

Richards, P. I. (1956). Shock waves on the highway. *Operations Research*, *4*(1), 42-51. doi: 10.1287/opre.4.1.42

Roncoli, C., Papageorgiou, M., & Papamichail, I. (2015). Traffic flow optimisation in presence of vehicle automation and communication systems - part i A first-order multi-lane model for motorway traffic. *Transportation Research Part C: Emerging Technologies*, *57*, 241 - 259. doi: http://dx.doi.org/10.1016/j.trc.2015.06.014

Schakel, W., Knoop, V., & van Arem, B. (2012). Integrated lane change model with relaxation and synchronization. *Transportation Research Record: Journal of the Transportation Research Board*, *2316*, 47-57. doi: 10.3141/2316-06

Schönhof, M., & Helbing, D. (2007). Empirical features of congested traffic states and their implications for traffic modeling. *Transportation Science*, 135-166.

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.

Srivastava, A., & Jin, W. (2016, January). A lane changing cell transmission model for modeling capacity drop at lane drop bottlenecks. In *Proceedings of the 95th annual meeting of the transportation research board (TRB).* Washington DC.

Tampère, C. (2004). *Human-kinetic multiclass traffic flow theory and modelling with application to advanced driver assistance systems in congestion* (PhD thesis). Delft University of Technology.

Tampère, C., Hoogendoorn, S. P., & Arem, B. V. (2005). A behavioural approach to instability, stop and go waves, wide jams and capacity drop. In (p. 205-228).

Tian, J., Jiang, R., Jia, B., Gao, Z., & Ma, S. (2016). Empirical analysis and simulation of the concave growth pattern of traffic oscillations. *Transportation Research Part B: Methodological*, *93, Part A*, 338 - 354. doi: http://dx.doi.org/10.1016/j.trb.2016.08.001

Treiber, M., Hennecke, A., & Helbing, D. (2000, Aug). Congested traffic states in empirical observations and microscopic simulations. *Phys. Rev. E*, *62*, 1805–1824. doi: 10.1103/PhysRevE.62.1805

Treiber, M., & Kesting, A. (2013). *Traffic flow dynamics: Data, model and simulation*. Springer.

Treiber, M., Kesting, A., & Helbing, D. (2006, July). Understanding widely scattered traffic flows, the capacity drop, and platoons as effects of variance-driven time gaps. *Phys. Rev. E*, *74*(1), 016123–.

Treiterer, J., & Myers, J. (1974). The hysteresis phenomenon in traffic flow. *Transportation and traffic theory*, *6*, 13-38.

Van Wageningen-Kessels, F. (2013). *Multi-class continuum traffic flow model: Analysis and simulation methods* (PhD thesis). Delft University of Technology.

Van Wageningen-Kessels, F., Yuan, Y., Hoogendoorn, S. P., Van Lint, H., & Vuik, K. (2013b). Discontinuities in the lagrangian formulation of the kinematic wave model. *Transportation Research Part C: Emerging Technologies*, *34*, 148 - 161.

Wong, G., & Wong, S. (2002). A multi-class traffic flow model: an extension of lwr model with heterogeneous drivers. *Transportation Research Part A: Policy and Practice*, *36*(9), 827–841.

Wu, X., & Liu, H. X. (2013, June). The uncertainty of drivers' gap selection and

its impact on the fundamental diagram. In *In proceeding of 20th international symposium on transportation and traffic theory (ISTTT 2013)* (Vol. 80, pp. 901–921).

Yeo, H. (2008). *Asymmetric microscopic driving behavior theory* (PhD thesis). University of California Transportation Center. UC Berkeley: University of California Transportation Center.

Yeo, H., & Skabardonis, A. (2009). Understanding stop-and-go traffic in view of asymmetric traffic theory. In W. H. K. Lam, S. C. Wong, & H. K. Lo (Eds.), *Transportation and traffic theory 2009: Golden jubilee* (pp. 99-115–). Springer US.

Yuan, Y., van Lint, J., Wilson, R., van Wageningen-Kessels, F., & Hoogendoorn, S. (2012, March). Real-time lagrangian traffic state estimator for freeways. *Intelligent Transportation Systems, IEEE Transactions on*, *13*(1), 59-70. doi: 10.1109/TITS.2011.2178837

Yuan, Y., van Lint, J. W. C., Schreiter, T., Hoogendoorn, S. P., & Vrancken, J. L. M. (2010, April). Automatic speed-bias correction with flow-density relationships. In *Networking, sensing and control (icnsc), 2010 international conference on* (p. 1-7). doi: 10.1109/ICNSC.2010.5461559

Zhang, H. (1999). A mathematical theory of traffic hysteresis. *Transportation Research Part B: Methodological*, *33*(1), 1 - 23. doi: http://dx.doi.org/10.1016/S0191-2615(98)00022-8

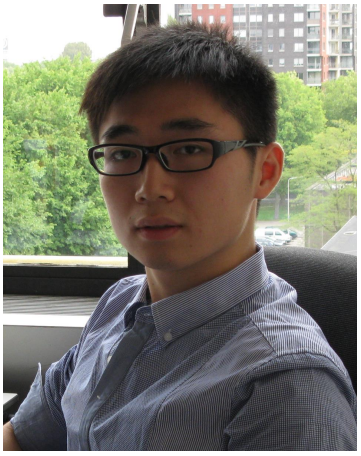
Zhang, H. (2002). A non-equilibrium traffic model devoid of gas-like behavior. *Transportation Research Part B: Methodological*, *36*(3), 275 - 290. doi: http://dx.doi.org/10.1016/S0191-2615(00)00050-3

Zhang, H., & Kim, T. (2005, June). A car-following theory for multiphase vehicular traffic flow. *Transportation Research Part B: Methodological*, *39*(5), 385–399.

Zheng, Z. (2014). Recent developments and research needs in modeling lane changing. *Transportation Research Part B: Methodological*, *60*, 16 - 32. doi: http://dx.doi.org/10.1016/j.trb.2013.11.009

Zheng, Z., Ahn, S., Chen, D., & Laval, J. (2013). The effects of lane-changing on the immediate follower: Anticipation, relaxation, and change in driver characteristics. *Transportation Research Part C: Emerging Technologies*, *26*, 367 - 379. doi: http://dx.doi.org/10.1016/j.trc.2012.10.007

About the author



Kai Yuan was born in 1988 in Jiangxi Province, China. He received his Bachelor of Engineering degree from the School of Traffic and Transportation, Beijing Jiaotong University in Beijing, China in 2010. At the same year, he began his study for a master of Engineering degree, with specialization in Transport Planning and Management there.

From the September 2012, he began performing his PhD research at the Department of Transport and planning in Delft University of Technology in Delft, the Netherlands. During his first PhD year, he worked on the development of back-pressure algorithms to improve the urban traffic network performance with signal control. From the middle of 2013, he began working on the capacity drop phenomenon on freeways, broadening the existing knowledge on the capacity drop phenomenon and describing it mathematically. His PhD research is under the supervision of (Promotor) Prof. Serge P. Hoogendoorn and (Co-promotor) Dr. Victor L. Knoop. Part of his PhD was performed at the Georgia Institute of Technology, where he worked with Dr. Jorge Laval. Apart from research, he also assisted in master course and served as a referee for international conferences.

Kai's main research interests include empirically inspired traffic dynamics, traffic modeling and dynamic traffic operations.

Summary

Earlier studies on the traffic flow on freeways reveal that queue discharge rate cannot reach as high as the free-flow capacity. This important phenomenon is called the "Capacity drop", which indicates that the potential freeway capacity cannot be fully utilized when discharging traffic jams. Even though a majority of empirical and analytical research has already been conducted to understand the capacity drop, several relevant challenges still need to be addressed. Those challenges include (1) characterizing more empirical features of the capacity drop, (2) incorporating the capacity drop into macroscopic models, (3) revealing mechanism related to driver behaviors behind the capacity drop and incorporating the mechanism into microscopic models. This thesis fills the research gap.

Four research questions are addressed:

(1) What is the relation between the queue discharge rate reduction and the congestion characteristics (e.g., speeds, densities)?

This thesis contributes to the understanding of the dependence of the capacity drop by analyzing empirical loop detector data on freeways. Firstly, a comparison on the queue discharge rate between the stop-and-go waves and the standing queues shows that the queue discharge rate of stop-and-go waves are much lower than that of standing queues. Secondly, this thesis shows the queue discharge rate increases as the speed in congestion increases. These two conclusions extend current knowledge of the capacity drop magnitude, and indicate a high probability that the queue discharge rate can be operated.

(2) What is the relevant driver behavioral mechanism behind the capacity drop?

After applying analytical analysis and numerical simulations, this thesis finds intra-driver variation mechanism is a significant longitudinal behavioral mechanism behind the capacity drop and its dependence. By contrast, the inter-driver variation can only slightly decrease the queue discharge rate. Floating car data analysis shows that both the mean and the standard deviation of the desired acceleration decrease as the vehicular speed increases. In simulations, it is found that such stochasticity of driver desired accelerations is an quite relevant determinant on the queue discharge rate reduction. Hence, this thesis concludes that the intra-driver desired acceleration variation considerably determines the capacity drop magnitude.

(3) How to incorporate the capacity drop into a first-order macroscopic traffic flow model?

A concept of hysteresis is used to reproduce the relation between the speed in congestion and the queue discharge rate. Both of link model and node (on-ramp and lane-drop node) models are presented in this work. The main idea is that the fundamental flow-density relation differs in accelerating and non-accelerating (i.e., cruise driving or decelerating) conditions. In the flow-density fundamental diagram, the acceleration branch is assumed to be lower than the deceleration. Those two branches intersect at a stable congested state. There is only one deceleration branch and its function is fixed. There are many acceleration branches and each of them corresponds to a stable congested state. Those acceleration branches are predefined according to a function describing the relation between the speed in congestion and the queue discharge rate. With a linear relation, only two additional parameters are required next to those describing a fundamental diagram. The kinematic wave model is presented and solved in the Lagrangian coordinates.

(4) How to incorporate the relevant driver behavioral mechanism of the capacity drop into a microscopic simulation model?

One car-following model is extended from a parsimonious car-following model, for incorporating the intra-driver desired acceleration variation. This car-following model is used to certificate the hypothesis about the relevant longitudinal behavioral explanation of the capacity drop phenomenon. The extended parsimonious car-following model is shown in the Newell's first order car-following framework. This new model adds two additional parameters to Newell's first-order car-following model. In simulations with a predefined fundamental diagram, if the simulation time is long enough and the sample size of the simulated vehicles is sufficiently large, the rescaled queue discharge rate is function of a single dimensionless parameter. The extended parsimonious car-following model can capture other car-following features revealed in earlier literature, including the concavity and the growing-and-flattening pattern of the traffic oscillation.

Conclusions

The relation between the speed in congestion and the queue discharge rate indicates a possibility of increasing queue discharge rate through traffic operations. The proposed macroscopic first-order model in this thesis is a useful simulation tool for evaluating strategies maximizing queue discharge rates. The understanding of the queue discharge rate reduction from a driver behavioral perspective denotes the primary behavior determining the queue discharge rate, which is fundamental to operation strategies aiming to maximize the queue discharge rate.

Samenvatting

Eerdere studies naar verkeersstromen op snelwegen tonen aan dat de afrijcapaciteit de vrije capaciteit niet kan evenaren. Dit belangrijke verschijnsel staat bekend als de "capaciteitsval", met als gevolg dat de potentiële capaciteit van een snelweg niet volledig kan worden benut wanneer automobilisten uit files wegrijden. Hoewel een meerderheid van het empirische en analytische onderzoek om de capaciteitsval te begrijpen reeds is uitgevoerd, is een aantal relevante vragen vooralsnog onbeantwoord gebleven. Deze uitdagingen omvatten (1) het karakteriseren van meer empirisch eigenschappen van de capaciteitsval, (2) het inbedden van de capaciteitsval in macroscopische modellen, (3) het onthullen van het mechanisme ten grondslag aan de capaciteitsval met betrekking tot de gedragingen van automobilisten en het inbedden van dit mechanisme in microscopische modellen. Deze scriptie vult deze kennisleemten.

Vier onderzoeksvragen worden behandeld:

(1) Wat is het verband tussen de daling van de afrijcapaciteit en de karakteristieken van de congestie (bijv. snelheden of dichtheden)?

Deze scriptie draagt bij aan het begrip van de afhankelijkheid van de capaciteitsval door empirische gegevens van inductielussen in snelwegen te analyseren. Ten eerste toont een vergelijking van de afrijcapaciteit tussen filegolven en staande files aan dat de afrijcapaciteit van filegolven veel lager ligt. Ten tweede laat deze scriptie zien dat dat de afrijcapaciteit toeneemt naarmate de snelheid in congestie toeneemt. Deze twee bijdragen breiden de huidige kennis over de grootte van de capaciteitsval uit en wijzen op een grote waarschijnlijkheid dat de afrijcapaciteit kan worden bestuurd.

(2) Wat is het relevante rijgedragsmatige mechanisme achter de capaciteitsval?

Toepassing van analytische analyses en numerieke simulaties in deze scriptie wijzen op variatie bínnen automobilisten als een belangrijk longitudinaal mechanisme achter de capaciteitsval en zijn afhankelijkheid. Variatie tússen automobilisten kan de afrijcapaciteit daarentegen slechts in kleine mate laten afnemen. Uit analyse van gegevens van individuele voertuigen blijkt dat zowel het gemiddelde als de standaarddeviatie van de gewenste acceleratie afneemt naarmate de voertuigsnelheid toeneemt. Uit simulaties blijkt zulke stochasticiteit in de wensacceleratie een tamelijk relevante determinant van de afname van de afrijcapaciteit te zijn. Hieruit concludeert deze scriptie dat de variatie in wensacceleratie bínnen bestuurders de omvang van de capaciteitsval voor een aanzienlijk deel bepaalt.

(3) Hoe dient de capaciteitsval in een eersteordemacroscopisch verkeersstroommodel te worden opgenomen?

Een notie van hysterese wordt gebruikt om het verband tussen de snelheid in congestie en de afrijcapaciteit te reproduceren. Dit werk verschaft zowel een schakelmodel als knoopmodellen (oprit en einde rijstrook). Het globale idee is dat de fundamentele intensiteit-dichtheidrelatie verschilt tussen accelereer- en niet-accelereeromstandigheden (d.w.z. constante snelheid of deceleratie). In het fundamentele intensiteit-dichtheiddiagram wordt aangenomen dat de acceleratietak lager ligt dan de deceleratietak. Deze twee takken snijden elkaar in een stabiele congestietoestand. Er is slechts één deceleratietak; de bijbehorende functie ligt vast. Er zijn meerdere acceleratietakken, van welke elke correspondeert met een stabiele congestietoestand. Deze acceleratietakken zijn vooraf gedefinieerd door een functie die de relatie tussen de snelheid in congestie en de afrijcapaciteit beschrijft. Indien deze relatie lineair is, zijn slechts twee extra parameters bovenop de parameters van het fundamentele diagram nodig. Het kinematischegolfmodel wordt geformuleerd en opgelost in Lagrangiaanse coördinaten.

(4) Hoe dient het relevante rijgedragsmatige mechanisme achter de capaciteitsval in een microscopisch model te worden opgenomen?

Een parsimoon voertuigvolgmodel wordt uitgebreid om de variatie in wensacceleratie bínnen bestuurders erin op te nemen. Dat voertuigvolgmodel wordt gebruikt om de hypothese te onderbouwen betreffende de relevante longitudinaal-rijgedragsmatige verklaring voor het verschijnsel van de capaciteitsval. Het uitgebreide parsimone voertuigvolgmodel is inpasbaar binnen Newell's eersteordevoertuigvolgraamwerk. Dit nieuwe model voegt twee extra parameters aan Newell's eersteordevoertuigvolgmodel toe. In simulaties met een vooraf gedefinieerd fundamenteel diagram, van voldoende duur en met een voldoende grote steekproef van gesimuleerde voertuigen, is de herschaalde afrijcapaciteit een functie van één deceleratietak; de bijbehorende functie ligt vast. Er zijn meerdere acceleratietakken, van welke elken enkele dimensieloze parameter. Het uitgebreide parsimone voertuigvolgmodel kan overige, in eerdere literatuur genoemde eigenschappen van voertuigvolgen reproduceren, met inbegrip van de concaviteit en het groei-en-vervlakpatroon van de verkeersoscillatie.

Conclusies

De relatie tussen de snelheid in congestie en de afrijcapaciteit wijst op een mogelijkheid om met operationele ingrepen de afrijcapaciteit te verhogen. Het in deze scriptie voorgestelde macroscopische eersteordemodel is een nuttig simulatiehulpmiddel voor de evaluatie van strategieën om de afrijcapaciteit te maximaliseren. Het begrip van de afrijcapaciteitsreductie vanuit rijgedragsmatig perspectief wijst het primaire gedrag aan dat de afrijcapaciteit bepaalt; dit is essentieel voor operationele strategieën die de afrijcapaciteit trachten te maximaliseren.

List of Publications

Journal articles

- Yuan, K., Knoop, V.L., Leclercq, L., and Hoogendoorn, S.P. (2016) Capacity drop: a comparison between stop-and-go wave and standing queue at lane-drop bottleneck. *Transportmetrica-B*. DOI:10.1080/21680566.2016.1245163
- Yuan, K., Knoop, V. L., and Hoogendoorn, S.P. (2016). A kinematic wave model in Lagrangian coordinates incorporating capacity drop: application to homogeneous road stretches and discontinuities. *Physica A: Statistical Mechanics & Its Applications*, 465, 472-485.
- Yuan, K., Knoop, V.L., and Hoogendoorn, S.P. (2015). Capacity drop: relation between the speed in congestion and the queue discharge rate. *Transportation Research Record: Journal of the Transportation Research Board*, 2 (2491), 72-80.
- Yuan, K., Knoop, V.L., and Hoogendoorn, S.P. (in print) A microscopic investigation into the capacity drop: impacts of longitudinal behavior on the queue discharge rate. *Transportation Science*

The following article is currently under review:

- Yuan, K., Laval, J., Knoop, V.L., Jiang, R., and Hoogendoorn, S.P. (under review). A Geometric Brownian motion car-following model: towards a better understanding of capacity drop. *Transportation Research Part B: Methodological*.

Peer-reviewed conference contributions

- Yuan, K., Laval, J., Knoop, V.L., Jiang, R., and Hoogendoorn, S.P. (2017). A better understanding of the capacity drop: from a driver behavioral perspective. *Presented at the 96th Annual Meeting of the Transportation Research Board*, 8-12 January 2017, Washington D.C.

- Yuan, K., Knoop, V.L., and Hoogendoorn, S.P. (2016). Optimal dynamic green time for distributed signal control. *In: Proceedings of the IEEE 19th International Conference on Intelligent Transportation Systems (ITSC)*, 1-4 November 2016, Rio de Janeiro, Brazil.

- Yuan, K., Knoop, V.L., and Hoogendoorn, S.P. (2016). A microscopic investigation into the capacity drop: impacts of the bounded acceleration and reaction time. *In: Proceedings of the 95th Annual Meeting of the Transportation Reseach Board*, 10-14 January 2016, Washington D.C.

- Yuan, K., Knoop, V.L., Schreiter, T., and Hoogendoorn, S.P. (2015). A hybrid kinematic wave model incorporating capacity drops. *In: Proceedings of the IEEE 18th International Conference on Intelligent Transportation Systems (ITSC)*, 15-18 September 2015, Las Palmas de Gran Canaria, Spain.

- Yuan, K., Knoop, V.L., and Hoogendoorn, S.P. (2015). Capacity drop: a relation between the speed in congestion and the queue discharge rate. *In: Proceedings of the 94th Annual Meeting of the Transportation Research Board*, 11-15 January 2015, Washington D.C.

- Yuan, K., Knoop, V.L., Leclercq, L., and Hoogendoorn, S.P. (2015). Capacity drop: a comparison between stop-and-go wave and standing queue at lane-drop bottleneck. *In: Transportation Reseach Circular, Symposium of Celebrating 50 Years of Traffic Flow Theory*, Number E-C197, Portland.

- Lin, L., Yuan, K., and Ren, S.Y. (2014). Analysis of urban freeway traffic flow characteristics based on frequent pattern tree. *In: Proceedings of the IEEE 17th International Conference on Intelligent Transportation Systems (ITSC)*, 8-11 October 2014, Qingdao, China.

TRAIL Thesis Series

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.

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

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

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

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

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

Campanella, M., *Microscopic modelling of walking behavior*, T2016/20, November 2016, TRAIL Thesis Series, the Netherlands

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

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

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

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

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

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

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

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

Psarra, I., *A Bounded Rationality Model of Short and Long-Term Dynamics of Activity-Travel Behavior*, T2016/11, June 2016, TRAIL Thesis Series, the Netherlands

Ma, Y., *The Use of Advanced Transportation Monitoring Data for Official Statistics*, T2016/10, June 2016, TRAIL Thesis Series, the Netherlands

Li, L., *Coordinated Model Predictive Control of Synchromodal Freight Transport Systems*, T2016/9, June 2016, TRAIL Thesis Series, the Netherlands

Vonk Noordegraaf, D.M., *Road Pricing Policy Implementation*, T2016/8, June 2016, TRAIL Thesis Series, the Netherlands

Liu, S., *Modeling, Robust and Distributed Model Predictive Control for Freeway Networks*, T2016/7, May 2016, TRAIL Thesis Series, the Netherlands

Calvert, S.C., *Stochastic Macroscopic Analysis and Modelling for Traffic Management*, T2016/6, May 2016, TRAIL Thesis Series, the Netherlands

Sparing, D., *Reliable Timetable Design for Railways and Connecting Public Transport Services*, T2016/5, May 2016, TRAIL Thesis Series, the Netherlands

Rasouli, S, *Uncertainty in Modeling Activity-Travel Demand in Complex Urban Systems*, T2016/4, March 2016, TRAIL Thesis Series, the Netherlands

Vries, J. de, *Behavioral Operations in Logistics*, T2016/3, February 2016, TRAIL Thesis Series, the Netherlands

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

Ardiç, Ö., *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