PARAMETER ESTIMATION AND STATISTICAL INFERENCE OF A TWO-REGIME STOCHASTIC CAR-FOLLOWING MODEL

A Thesis
Presented to
The Academic Faculty

By

Tu Xu

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy in the
School of Civil and Environmental Engineering

Georgia Institute of Technology

August 2020

PARAMETER ESTIMATION AND STATISTICAL INFERENCE OF A TWO-REGIME STOCHASTIC CAR-FOLLOWING MODEL

Approved by:

Dr. Jorge A. Laval, Advisor
School of Civil and Environmental Engineering
Georgia Institute of Technology

Dr. Patricia L. Mokhtarian
School of Civil and Environmental Engineering
Georgia Institute of Technology

Dr. Michael P. Hunter
School of Civil and Environmental Engineering
Georgia Institute of Technology

Dr. Haobing Liu
School of Civil and Environmental Engineering
Georgia Institute of Technology

Dr. Siva Theja Maguluri
School of Industrial and Systems Engineering
Georgia Institute of Technology

Date Approved: June 15, 2020

To my parents with all my love.

ACKNOWLEDGEMENTS

I feel I cannot express my gratitude to Dr. Jorge A. Laval in words. He really guided me into the world of independent research. He gives me freedom to explore the traffic flow theory world while he is always there to give me advice, comments and support. He is my boss while he is also a cool person to work with. When I was stuck in Switzerland in 2020 because of the pandemic, he not only keep supporting me financially but really showed empathy with me. Even if we cannot meet with each other in person for the last months of my Ph.D. program, our relationship gets more solid. I really wish to visit the U.S. soon and meet with him.

I also had the luck to meet a lot of great teachers and classmates at Georgia Tech. It was a special experience that I did my thesis defense online in China because of the pandemic. However, I still saw a lot of colleagues and friends joined my thesis defense from different time zones, for which I really appreciated. Particularly, I spent a lot of precious time with Dr. Cho, Rafegh, Hao, Haobing and Zhaocheng, and I would like to say thank you to them.

I am very grateful to Dr. Dianhai Wang and Dr. Chuankun Rao at Zhejiang University, who led me into the world of traffic. I remembered the lectures I expected most each week at Zhejiang University was their traffic network class. I am also very grateful to Dr. Xiaoyuan Hou at Fudan University, we had a lot of fun conversations not only about physics in his office.

I am very grateful to my supervisors and colleagues at Pond & Company. They always give me trust and we worked on a lot of interesting projects together. I am also very lucky to meet a group of students in Hangzhou. Under the supervise of Dr. Zhenhui Li, they are working on deep reinforcement learning for traffic signal control, which motivates me and led me to make a decision to join an AI Research Lab for my future career.

This thesis work has luckily received the support from Dr. Jiang at Beijing Jiaotong University. I thank you for their willingness to share their data.

Also, I want to say thank you to my friends. Xiaoma, Renjing and Jin Xu, we always discussed interesting coding topics and state-of-the-art together. Rundong, when I first learned Mathematica, you were always there to help me patiently. Finally, my love goes to my parents, you are wonderful and always give me love and support. This is for you.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

SUMMARY

This thesis presents the formulation of a family of two-regime car-following models where both free-flow and congestion regimes obey random processes. This formulation generalizes previous efforts based on Brownian and geometric Brownian acceleration processes, each reproducing a different feature of traffic instabilities.

We show that the unified model is able to capture virtually all types of traffic instabilities consistently with empirical data, including oscillations, speed-capacity relationship, and the concave growth pattern of vehicle speeds along a platoon.

The probability density of vehicle positions turns out to be analytical in our model, and therefore parameters can be estimated using maximum likelihood. This allows us to test a wide variety of hypotheses using statistical inference methods, such as the homogeneity of the driver/vehicle population and the statistical significance of the impacts of roadway geometry.

Using data from two controlled car-following experiments and one uncontrolled car-following dataset, we find that (i) model parameters are similar across repeated experiments within the same dataset but different across datasets, (ii) the acceleration error process is closer to a Brownian motion, and (iii) drivers press the gas pedal harder than usual when they come to an upgrade segment. The model is flexible so that newer vehicle technologies can be incorporated to test such hypotheses as differences in the car-following parameters of automated and regular vehicles, when data becomes available.

CHAPTER 1
INTRODUCTION

Traffic congestion is a big concern for motorists. Due to traffic congestion, financial costs, emission, environmental pollution and the likelihood of car crashes increase (Bilbao-Ubillos, 2008). In most cases, it is not desirable to increase roadway capacity by expanding the road from a resource allocation perspective. Therefore, traffic operations become a crucial method to mitigate traffic congestion around the world (Ahmed, 1999). However, the mechanisms of traffics are very complex, especially in congestion. Studies have shown that the performance of the transportation system is subject to weather, roadway geometry, driver behavior and many other factors (Kilpeläinen and Summala, 2007; Li et al., 2007; Chen et al., 2012b). Meanwhile, several traffic phenomena are not well explained by existing models. For example, a model to explain traffic oscillation and capacity drop simultaneously is still lacking (Laval and Leclercq, 2010; Yuan et al., 2018). This challenge is described in detail in Chapter 2. Therefore, in order to perform better traffic operations and management, a comprehensive understanding of the underlying reasons behind traffic congestion is needed. This motivates us to take a detailed analysis of car-following models at the microscopic level.

Toward this end, a two-regime microscopic-level car-following model is developed in Chapter 3. The model combines previous modeling work based on Brownian motion (Laval, Toth, and Zhou, 2014) and geometric Brownian motion (Yuan et al., 2018) vehicle acceleration process. The model takes advantage of both types of acceleration processes and can reproduce different features of traffic instabilities consistent with empirical data, including traffic oscillations, capacity drop and the concave growth pattern of vehicle speeds along a platoon, as reported in Jiang et al. (2014). A detailed analysis of the model is also provided in Chapter 3. Not only we are able to simulate traffic instabilities, consis-

tently with field data but we prove the ability of the model to explain these phenomena analytically.

In Chapter 4, we show that another highlight of the model is that the free-flow and congestion regimes follow stochastic process and the probability density function of the vehicle positions can be solved analytically for the first time. This allows us to use maximum likelihood estimation to estimate model parameters and opens the door for statistical inference, also for the first time in our discipline.

In particular, we can use statistical inference methods to test a wide variety of hypotheses, such as the homogeneity of the driver/vehicle population and the statistical significance of the impacts of roadway geometry or weather. In Chapter 5, by taking advantage of the model and the NGSIM datasets (James Colyar, 2007), we found evidence that the current assumption in existing acceleration models does not apply to car drivers, who tend to overcome half of the gravitational effects by using more engine power. Truck drivers only compensate for 5% of the loss, possibly because of limited engine power. This finding is important because the results show that the current acceleration models are underestimating the emission and energy consumption by about 10 % in an acceleration process at uphill grades, as shown in Chapter 5.

In Chapter 6, we provide the main conclusions and future directions of this thesis work.

CHAPTER 2
BACKGROUND

2.1 Car-following Models

Car-following behavior is a basic microscopic-level driver behavior. Car-following models describe the driving behavior of the following vehicle in reaction to the leading vehicle without lane changing or traffic signals (Xu et al., 2013).

To the best of our knowledge, the concept of car-following is first mentioned in Pipes (1953), who study the dynamics of traffic based on ”law of separation”, where the desired ”following distance” is a linear function of velocity of the following vehicle. In the late fifties, Chandler, Herman, and Montroll (1958) and Herman et al. (1959) study the Gazis-Herman-Rothery (GHR) model. The model is calibrated and validated with empirical data. Also, this microscopic model is able to connect with macroscopic characteristics in traffic by integration (Gazis, Herman, and Potts, 1959; Gazis, Herman, and Rothery, 1961). On the other hand, Michaels (1963) notes that human factors could have a great impact on the driver behavior. This study builds a connection between car-following models and psychology. Later, many car-following models are developed to explain the nonlinear characteristics in congested traffic (Bando et al., 1995; Treiber, Hennecke, and Helbing, 2000; Newell, 2002). With the development of artificial intelligent, neural networks are used to model car-following behavior Dougherty (1995).

Most car-following models can be classified into the following categories: (i) stimulus-response models as in Newell (1961) and Gazis, Herman, and Rothery (1961), (ii) psycho-physical models, such as Widemann Model (Wiedemann and Reiter, 1992) and Fritzsche Model (Fritzsche and Ag, 1994), (iii) crash-avoidance models, such as Gipps Model (Gipps, 1981) and IDM Model (Kesting, Treiber, and Helbing, 2010; Treiber, Hennecke, and Hel-

bing, 2000), (iv) optimal-velocity models such as Optimal Velocity Model in Bando et al. (1995), Generalized Force Model in Helbing and Tilch (1998) and Full Velocity Difference Model in Jiang, Wu, and Zhu (2001). However, most of these models do not consider the variation among the behavior of different drivers while studies (Laval and Leclercq, 2010; Chen et al., 2012b) show that timid and aggressive drivers are responsible for traffic oscillations. The model proposed in this thesis is not classified into the four traditional categories listed above but extends Newell's model with bounded accelerations (Newell, 2002) such that stochastic vehicle acceleration process can be incorporated to model the variation among drivers.

2.2 Modeling Traffic Instabilities

2.2.1 Traffic oscillations

Stop-and-go driving conditions in traffic oscillations disturb motorists. Mauch and Cassidy (2002), Laval (2005), Laval and Daganzo (2006), Ahn and Cassidy (2007), and Zheng et al. (2011) have demonstrated that lane changing is a major factor causing traffic oscillations. However, a convincing explanation for this phenomenon on a one-lane road is still lacking. Stop-and-go traffic on a one-lane ring road without lane changing movements has been empirically found by Sugiyama et al. (2008).

To fill this gap, several analytical studies have proposed that a small driving instability may grow into mature traffic oscillations in the absence of lane changing (Treiber, Hennecke, and Helbing, 1999; Wilson, 2008; Wilson and Ward, 2011; Treiber and Kesting, 2012; Treiber and Kesting, 2017). However, these models require either further model validation with empirical data or a careful selection of parameter values.

On the other hand, Laval and Leclercq (2010), Chen et al. (2012a), and Chen et al. (2012b) argue that timid or aggressive car-following behaviors caused the formation and propagation of traffic oscillations. These models are validated with empirical data, but they require at least four parameters that are not directly observable. Laval, Toth, and Zhou

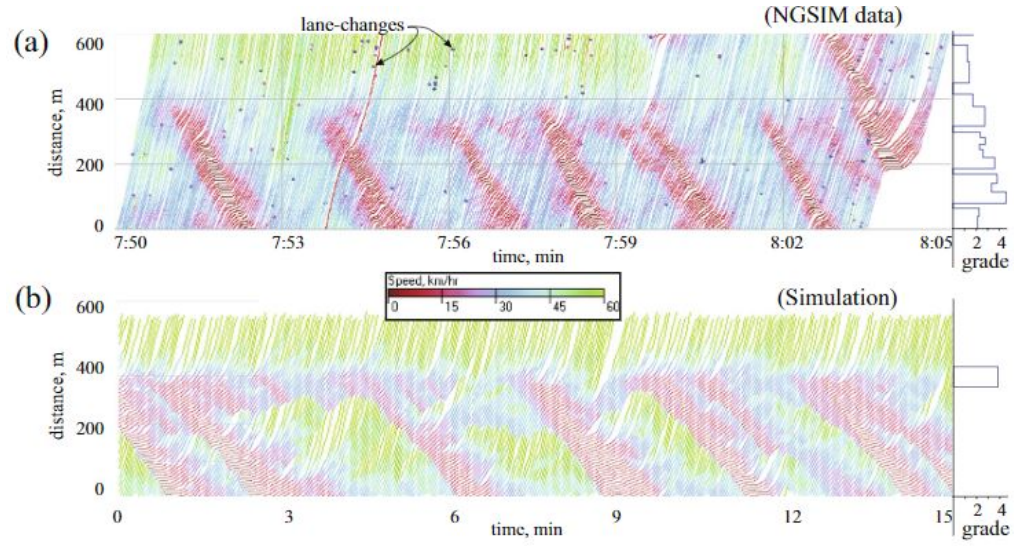

Figure 2.1: Time-space diagram on a 600m-long segment during 15 minutes. Grade profile is also shown in the figure. (a) NGSIM field data. (b) Simulation results based on the model with Brownian motion acceleration process. This figure is from Laval, Toth, and Zhou (2014).

(2014) assume a constant standard deviation of driver acceleration, resulting in a Brownian motion (BM) process only one unobservable parameter. There simulation results compared with field data can be seen in Figure 2.1. Delpiano et al. (2015) proposed a social force car-following model implementation of this theory that includes finite decelerations. These two models successfully replicate realistic traffic oscillations at uphill segments but are unable to explain capacity drop phenomenon introduced in the next section.

2.2.2 Capacity drop

It has been empirically observed for years that the queue discharge rate at bottlenecks is lower after the queue forms. This phenomenon is called capacity drop (Hall and Agyemang-Duah, 1991) has long been correlated with lane changing activity (Laval and Daganzo, 2006; Coifman and Kim, 2011; Leclercq, Laval, and Chiabaut, 2011; Yuan, Knoop, and Hoogendoorn, 2015; Chen and Ahn, 2018; Jin, 2017; Kontorinaki et al., 2017). But researchers have observed that capacity drop can occur in the absence of lane changing (Oh

and Yeo, 2015). To explain this, Wong and Wong (2002) takes the distribution of drivers' desired free-flow speeds in a traffic into account. However, empirically validation of this assumption is lacking. Yuan et al. (2018) argue that empirical data reveals that the standard deviation of the driver desired acceleration is not constant, as in the BM model (Laval, Toth, and Zhou, 2014), but decreases linearly with speed, leading to a geometric Brownian acceleration process. This geometric Brownian acceleration process successfully explains capacity drop but does not produce periodic traffic oscillations (Xu and Laval, 2018).

2.2.3 Concave growth of platoon oscillation

The concave growth of the vehicle platoon oscillation was first observed by Jiang et al. (2014) shown in Figure 2.2, where the standard deviation of vehicle speed grows concavely along the platoon. Tian et al. (2016) find that the growth pattern in different oscillations collapses into a single concave curve, which indicates the law of the oscillation growth is universal. Treiber and Kesting (2017) shows that the Brownian motion model in Laval, Toth, and Zhou (2014), and the IDM model with noise and indifference regions are all able to reproduce the concave growth pattern, indicating strongly that a stochastic component is vital for replicating this important phenomena.

For a theoretical proof of the concave growth of the speed standard deviation along a vehicle platoon, consider a vehicle platoon with vehicle id 0, 1, 2, ...,n from downstream to upstream. Let v_i denote the speed of the i-th vehicle. The lead vehicle (indexed by 0) travels at a constant speed v_0. We assume the following relationships:

$$v_1 = v_0 + \epsilon_1$$

$$v_2 = v_1 + \epsilon_2$$

.....

$$v_n = v_{n-1} + \epsilon_n$$

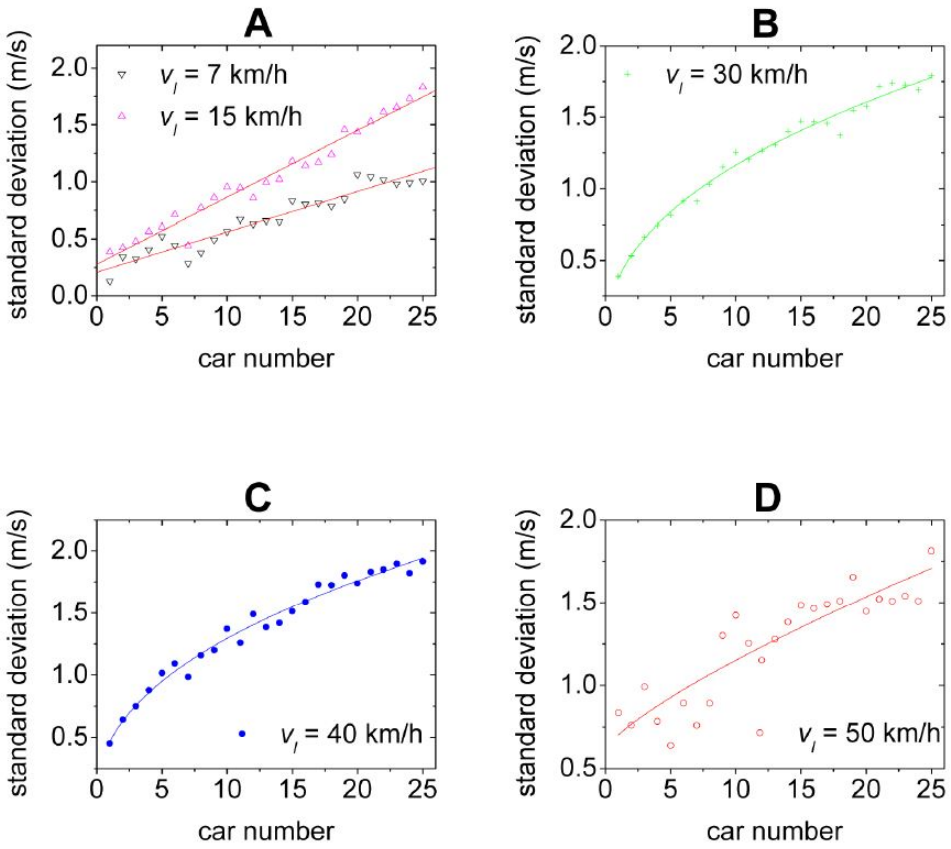

Figure 2.2: Figure is from Jiang et al. (2014). The relationship between the standard deviation of vehicle speeds against car number suggests there is a concave growth of the vehicle platoon oscillation.

where ϵ_is follow the same normal distribution $\epsilon_i \sim N(0, \text{Var}[\epsilon_i])$. We have $v_n = v_0 + \sum \epsilon_i$ and $\text{Var}[v_n] = \sum \text{Var}[\epsilon_i] = n\text{Var}[\epsilon_i]$. According to the Central Limit Theorem, v_n follows a normal distribution with a mean of v_0 and the following standard deviation:

$$\text{SD}[v_n] = \sqrt{n}\sqrt{\text{Var}[\epsilon_i]} \propto \sqrt{n},$$

which proves that the standard deviation of vehicle speed grows concavely along the platoon.

2.3 Traffic stream properties

There are three main macroscopic variables to describe a traffic stream: (i) flow q is the number of vehicles passing a position per unit of time, (ii) density k is the number of vehicles per unit length and (iii) speed v (average of all vehicles) is the distance moved per unit of time. There is a fundamental relationship between the three variables in traffic: $q = kv$.

In the microscopic level, there are also three variables to describe traffic: (i) headway h is the time between two consecutive vehicles at a given location, (ii) spacing s is the distance between 2 consecutive vehicles at a given time, and (iii) speed v of a single vehicle at a given time. For a homogeneous vehicle platoon, we have the following relationship: $h = \frac{1}{q}$, $s = \frac{1}{k}$ and $s = hv$.

2.4 Time-space diagrams

A time-space diagram is commonly used to show vehicle trajectories. The X-axis of the diagram represents time and the Y-axis represents distance. An example of time-space diagram is given in Figure 2.3, which shows the acceleration process of a vehicle. In the time-space diagram, the slope of the tangent of the trajectory represents the speed of the vehicle.

The time-space diagram is also used to show the trajectories of a vehicle platoon. An example is given in Figure 2.4, which shows the trajectory of a 3-vehicle platoon. In the figure, h represents the headway of vehicle 1 and 2 at location x_1, s is the spacing of vehicle 1 and 2 at time t_1, and $-w_1$, $-w_2$ are the wave speeds. The wave speed will change based on the field traffic conditions but a typical value of the wave speed is -14 mi/hr or -20 km/hr. The propagation of the traffic waves could be reproduced by the kinematic wave model to be introduced hereafter.

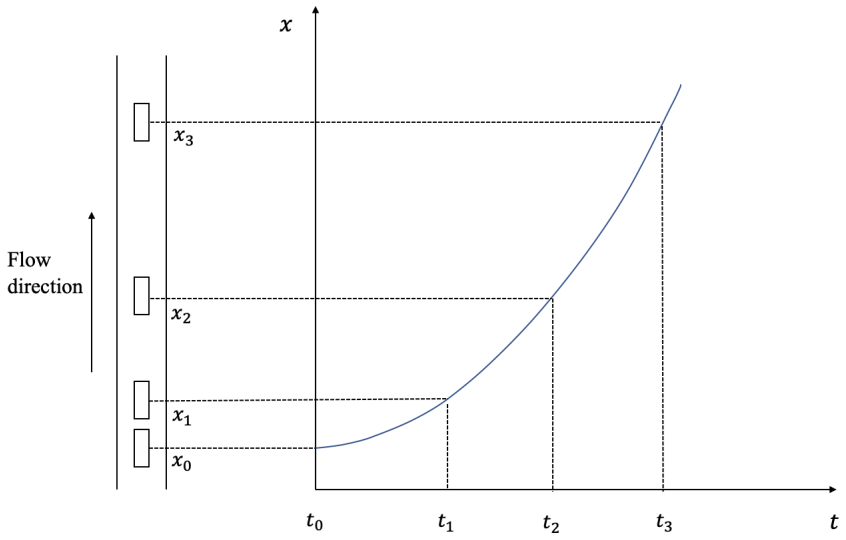

Figure 2.3: A time-space diagram showing the trajectory of an accelerating vehicle.

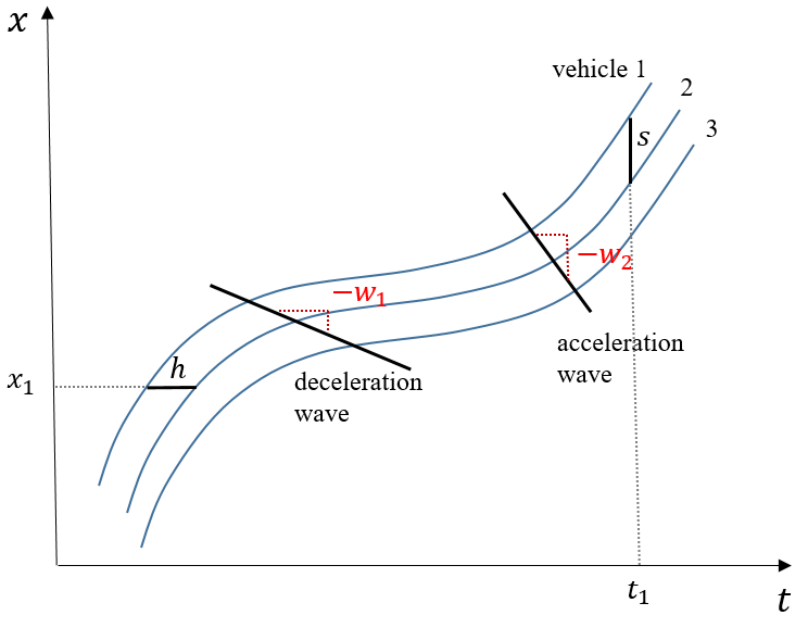

Figure 2.4: A time-space diagram showing the trajectory of a vehicle platoon consisting of 3 vehicles.

2.5 The Kinematic wave model

The Kinematic wave (KW) model is first developed by Lighthill and Whitham (1955). This model can predict travel time, queue length and many other travel behaviors at various of

bottlenecks. The KW model is the combination of the conservation law and the fundamental diagram $q = F(k)$.

Conservation law

For a road segment without entrances or exists the conservation law is a first-order PDE:

$$\frac{\partial k}{\partial t} + \frac{\partial q}{\partial x} = 0 \tag{2.1}$$

where k is the density and $q = F(k)$ is the flow as a function of density.

2.5.1 The Fundamental Diagram

The Fundamental Diagram in traffic flow is a diagram that illustrates the relationship between the flow q and density k. It contains two branches: (i) the free-flow branch and (ii) the congestion branch. A triangular fundamental diagram is widely used as an approximation. One example is given in Figure 2.5, where Q is the capacity, k_c is the critical density, k_j is the jam density, u is the free-flow speed and $-w$ is the wave speed.

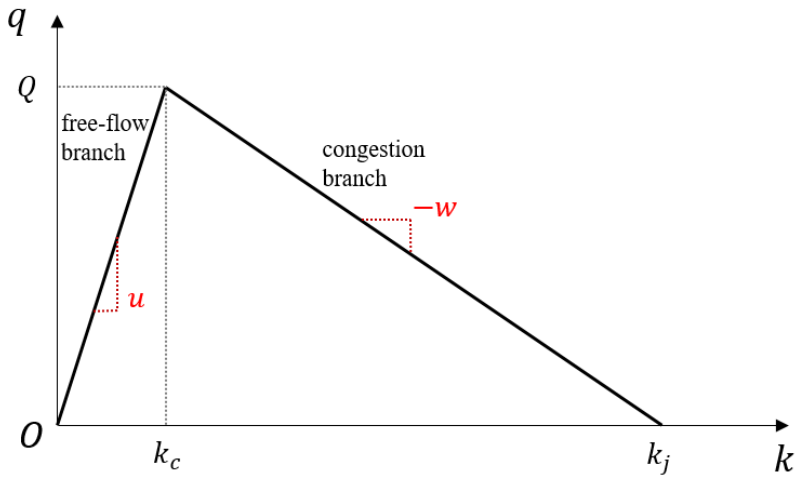

Figure 2.5: The triangular fundamental diagram for a one-lane road with a capacity of Q.

2.6 From the KW model to Newell's car-following model

In this section, we will introduce the deravation of Newell's car-following model from the KW model. We start with introducing the definition of cumulative count curves and maximum passing rate.

Cumulative count curves, $N(t, x)$, give the cumulative number of vehicles that have passed location x by time t. Cumulative count curves have the following properties:

$$q = \frac{\partial N}{\partial t}, k = -\frac{\partial N}{\partial x}. \tag{2.2}$$

The maximum passing rate is the maximum passing rate in terms of an observer:

$$R(v) = \sup_k \{F(k) - vk\}, \tag{2.3}$$

where $v \in [-w, u]$ is the speed of the observer. If we assume a triangular fundamental diagram, we have:

$$R(v) = Q - k_c v, \tag{2.4}$$

where k_c is the critical density in the fundamental diagram.

For a road segment, the fundamental diagram $q = F(k)$ gives the following PDE:

$$\begin{cases} \dfrac{\partial N}{\partial t} - F(-\dfrac{\partial N}{\partial x}) = 0, & \text{(fundamental diagram)} \quad (2.5a) \\ N(B) = G(B), \quad B \in \mathcal{B}, & \text{(boundary conditions)} \quad (2.5b) \end{cases}$$

which is of the Hamilton-Jacobi type. The solution is:

$$N(P) = \min_{B \in \mathcal{B}_P} \left\{ G(B) + (t - t_B) R\left(\frac{x - x_B}{t - t_B}\right) \right\}, \tag{2.6}$$

where $P \equiv (t_P, x_P)$ is a generic point and $B \equiv (t_B, x_B)$ is a point at the boundary $\mathcal{B}_P$. The

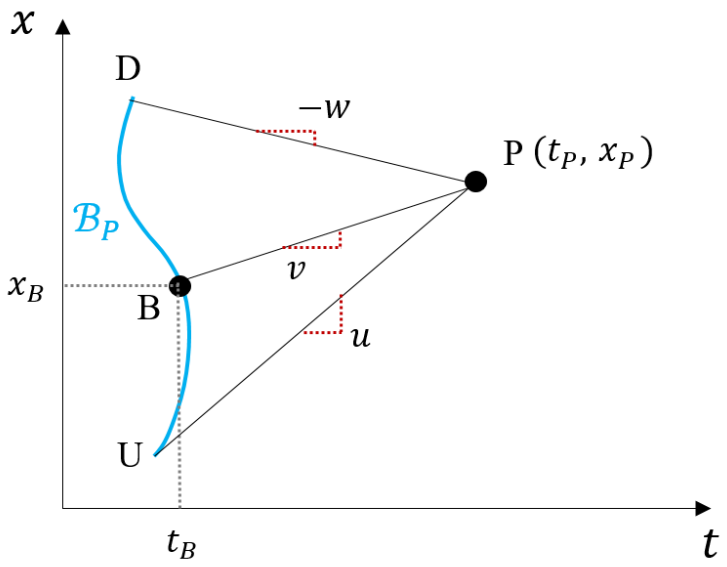

Figure 2.6: The solution of the Hamilton-Jacobi PDE. U means upstream of the boundary and D is the downstream. The path between P and B is minimized based on the solution.

point B needs to be selected along the boundary such that the cost from P to B is minimized. If the triangular fundamental diagram is assumed, one can simplify (2.6) as:

$$N(P) = \min_{B \in \mathcal{B}_P} \{G(B) + (t - t_B)Q - (x - x_B)k_c\}. \tag{2.7}$$

In traffic flow theory, the following three representations are equivalent:

- $N(t, x)$: number of vehicles that passed location x by time t,
- $X(t, n)$: position of vehicle n at time t.
- $T(n, x)$: time that the n th vehicle passes location x.

We showed how to solve $N(t, x)$ above and hereafter we follow similar procedure to solve $X(t, n)$. Notice the representations in traffic flow are equivalent, from (2.6), one can easily obtain:

$$X(t, n) = \min_{B \in \mathcal{B}_P} \left\{ G(t_B, n_B) + (t - t_B)R\left(\frac{n - n_B}{t - t_B}\right) \right\}, \tag{2.8}$$

if we assume a triangular fundamental diagram, we have the following relationships:

$$V(s) = \min\{u, s/\tau - w\} \quad \text{and} \quad R(\tilde{q}) = u - s_c\tilde{q}, \tag{2.9}$$

where $V(s)$ is speed v in a function of spacing s similar to $q = F(k)$, $s_c = \frac{1}{k_c}$ is the critical spacing, and $\tilde{q}$ is the wave flow similar to the speed of the observer $\tilde{v}$. Combining (2.8) and (2.9), one can have:

$$X(t,n) = \min\left\{\min_{y\in\mathcal{B}_P}\{X(0,y) + ut - (n-y)\}, X(t-n\tau, 0) - n\delta\right\}, \tag{2.10}$$

where jam spacing $\delta = u\tau - s_c$.

If the spacing of the vehicles are constant at $t = 0$, i.e. the initial condition is linear, one can rewrite (2.10) by interpreting the upstream of the boundary (0,0) as $(t - \Delta t, n - \Delta n)$:

$$X(t,n) = \min\left\{X(t-\Delta t, n) + u\Delta t, X(t - \Delta n\tau, n - \Delta n) - \Delta n\delta\right\}. \tag{2.11}$$

By setting $\Delta n = 1$ and choosing τ as Δt, one can rewrite (2.11) as:

$$x_j(t) = \min\{\underbrace{x_j(t-\tau) + u\tau}_{\text{free-flow term}}, \underbrace{x_{j-1}(t-\tau) - \delta}_{\text{congestion term}}\}, \tag{2.12}$$

Notice that we rewrite $X(t,n)$ as $x_j(t)$. The notation $x_j(t)$ represents the location of j-th vehicle at time t for the rest of the thesis.

Figure 2.7 is a graphical interpretation of Newell's car-following model.

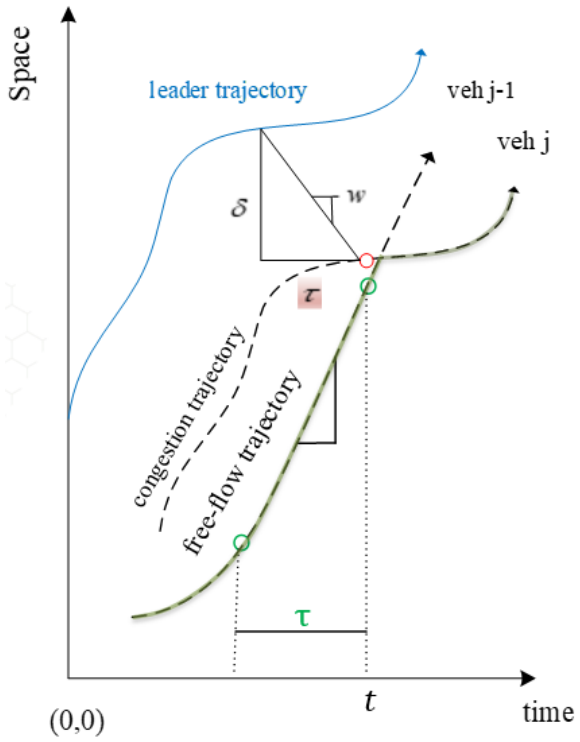

Figure 2.7: Newell's car following model.

CHAPTER 3
THE PROPOSED MODEL

This chapter is modified from the following published paper:
Xu, T., & Laval, J. A. (2019). Analysis of a Two-Regime Stochastic Car-Following Model: Explaining Capacity Drop and Oscillation Instabilities. Transportation Research Record. Vol. 2673, pp. 610-619

3.1 Introduction

Stop-and-go driving conditions (traffic oscillations) in traffic congestion increase not only financial costs and environmental pollution (Bilbao-Ubillos, 2008) but also the likelihood of car crashes (Zheng, Ahn, and Monsere, 2010). Intensive research has been done on its contributing factors and properties. However, we still have limited knowledge about traffic oscillations. Mauch and Cassidy (2002), Laval (2005), Laval and Daganzo (2006), Ahn and Cassidy (2007), and Zheng et al. (2011) have demonstrated that lane changing is a major factor causing traffic oscillations and Sugiyama et al. (2008) showed that stop-and-go traffic can occur on a one-lane ring road without lane changing movements. A convincing explanation for this phenomenon is still lacking.

Several analytical studies have proposed that a small driving instability may grow into mature traffic oscillations in the absence of lane changing (Treiber, Hennecke, and Helbing, 1999; Wilson, 2008; Wilson and Ward, 2011; Treiber and Kesting, 2012; Treiber and Kesting, 2017). However, these models require either further model validation with empirical data or a careful selection of parameter values. On the other hand, Laval and Leclercq (2010) and Chen et al. (2012a) argue that timid or aggressive car-following behaviors caused the formation and propagation of traffic oscillations. These models are

validated with empirical data, but they require at least four parameters that are not directly observable.

A more parsimonious family of car-following models, able to reproduce most traffic instabilities, is based on stochastic processes to describe drivers' desired accelerations. Laval, Toth, and Zhou (2014) assume a constant standard deviation of driver acceleration, resulting in a Brownian motion (BM) process that explained realistic traffic oscillations with only one unobservable parameter. However, this model is unable to explain the decreasing relationship between queue discharge rate and the speed in the queue, as observed empirically in (Yuan, Knoop, and Hoogendoorn, 2015), and called "speed-capacity relationship" hereafter.

Yuan et al. (2018) argue that empirical data reveals that the standard deviation of the driver desired acceleration is not constant, as in the BM model, but decreases linearly with speed, leading to a geometric Brownian acceleration process. This geometric Brownian acceleration process successfully explains empirical speed-capacity relationships but does not produce periodic traffic oscillations. This thesis aims to regulate the type of driver error on a scale from Brownian motion (BM) to geometric Brownian motion (g-BM) acceleration processes, and thus we can capture all said features of traffic instabilities.

3.2 Stochastic desired acceleration car-following framework

Car-following models based on stochastic desired acceleration models started with Laval, Toth, and Zhou (2014). They generalize Newell's car-following framework (Newell, 2002) with bounded accelerations (Leclercq et al., 2007; Leclercq and Laval, 2007; Laval and Leclercq, 2008; Laval and Leclercq, 2010; Leclercq, Laval, and Chiabaut, 2011) by including a stochastic process to account for driver error during acceleration. It was found that this component is crucial for capturing realistic formation and propagation of traffic

oscillations. The model can be expressed as:

$$x_j(t) = \min\{\underbrace{x_j(t-\tau) + \xi_j(\tau)}_{\text{free-flow}}, \underbrace{x_{j-1}(t-\tau) - \delta}_{\text{congestion}}\}, \tag{3.1}$$

where $x_j(t)$ is the position of j-th vehicle at time t, τ is the wave trip time between two consecutive vehicle trajectories and δ is the jam spacing. In this *two-regime* modeling framework the congestion term is deterministic while the free-flow term is stochastic. The $\xi_j(\tau)$ is a realization of the stochastic process describing the desired displacement of vehicle j during $t-\tau$ and t. This displacement is as a result of a stochastic desired acceleration model, which gives desired acceleration:

$$a = a(v), \tag{3.2}$$

that the driver imposes to the vehicle when traveling at speed $v(t)$ at time t under free-flow conditions, i.e., when unobstructed by the leading vehicle. The desired acceleration model captures both driver behavior and the physical limitations imposed by roadway geometry on the engine.

Based on empirical measurements, Laval, Toth, and Zhou (2014) suggests the following linear model for the mean desired acceleration $\mathrm{E}\left[a(v)\right]$ under free-flow conditions:

$$\mathrm{E}\left[a(v)\right] = \left(v_c - v(t)\right)\beta, \tag{3.3}$$

where $v(t)$ is the vehicle speed, v_c is the desired speed and β is the inverse relaxation time. (3.3) can be time-integrated twice to give

$$\mathrm{E}\left[\xi(t)\right] = v_c t - (1 - e^{-\beta t})(v_c - v_0)/\beta, \tag{3.4}$$

The key to this modeling framework is the specification of the acceleration error around

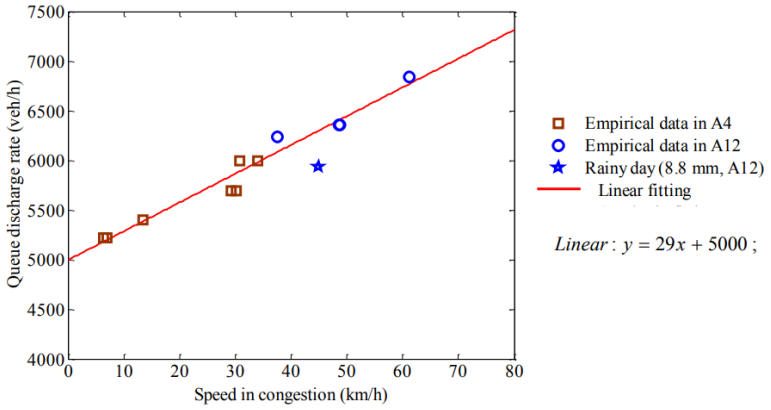

Figure 3.1: Relation between queue discharge rate and the speed at the bottleneck. The data is collected on three-lane sections on two freeways in the Netherlands by Yuan, Knoop, and Hoogendoorn (2015).

the mean desired vehicle acceleration E $[a(v)]$, which has profound implications in the ability of the model to replicate traffic instabilities. Laval, Toth, and Zhou (2014) assume that vehicle accelerations are normally distributed with a constant standard deviation, SD $[a(v)]$, which gives the following BM stochastic differential equation:

$$\textbf{BM:} \begin{cases} d\xi(t) = v(t)dt, \quad \xi(0) = 0, & \text{(3.5a)} \\ dv(t) = (v_c - v(t))\beta dt + \sigma dW(t), \quad v(0) = v_0, & \text{(3.5b)} \end{cases}$$

where $W(t)$ is a standard BM and σ is its diffusion coefficient. This is probably the simplest model that captures (i) bounded accelerations due to gravity and (ii) produce realistic traffic oscillations (due to driver random errors while accelerating).

However, Yuan et al. (2018) observe that model (3.1)-(3.5) is unable to reproduce the speed-capacity relationship, as observed empirically in (Yuan, Knoop, and Hoogendoorn, 2015), shown in Fig. 3.1. Notice that the speed-capacity relationship is intimately related to the hysteresis phenomenon in traffic flow (Treiterer and Myers, 1974; Ahn, Vadlamani, and Laval, 2013; Chen et al., 2012a). Hysteresis refers to the discrepancy in the traffic states before and after a platoon of vehicles passes through a congested wave, were most

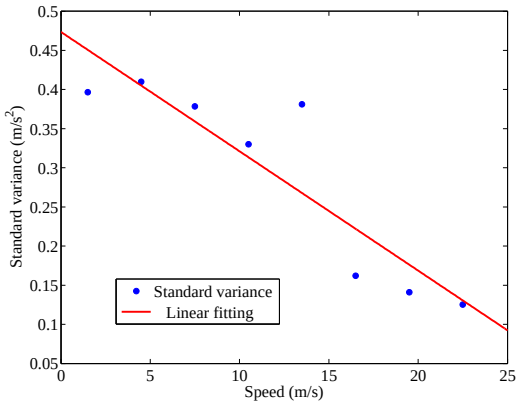

Figure 3.2: Standard deviation of the desired accelerations in 5-m/s speed bins. A linear function SD $[a(v(t))] = 0.015v(t) + 0.47$ is applied to fit the data. The data are from car-following experiments in Laval, Toth, and Zhou (2014), collected when the test vehicle is the platoon leader stopped in front of a red signal.

of the time the deceleration phase exhibits higher flow than the acceleration phase (Laval, 2011). The speed-capacity relationship can be considered as a special case where the flow in the acceleration phase is related to the speed inside the congested wave. Based on the empirical data shown in Fig. 3.2, they argued that the standard deviation of the driver desired acceleration, SD $[a(v(t))]$, is not constant but decreases linearly with the speed. Assuming that SD $[a(v(t))]$ vanishes at v_c, the problem can be modeled as a g-BM, i.e.:

$$\textbf{g-BM:} \begin{cases} d\xi(t) = v(t)dt, \quad \xi(0) = 0, & \text{(3.6a)} \\ dv(t) = (v_c - v(t))\beta dt + (v_c - v(t))\ \sigma\ dW(t), \quad v(0) = v_0, & \text{(3.6b)} \end{cases}$$

Although model (3.1)-(3.4) and (3.6) was shown to reproduce the bottleneck discharge rate-speed in queue relationship, it loses the ability to produce realistic traffic oscillations as shown in Fig. 3.3. Here we argue that this is due to the error term vanishing at the desired speed, per (3.6b), which implies that at high speeds oscillations are not likely to form.

To implement these models and the one presented momentarily, one has to solve (3.5)

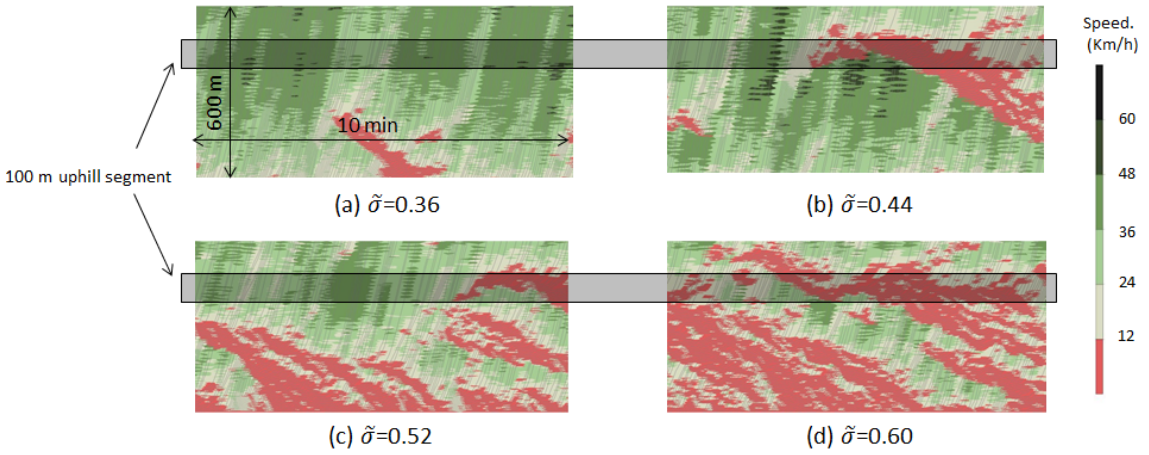

Figure 3.3: Traffic patterns generated using g-BM model, showing the model's inability to generate realistic traffic oscillations despite the different model parameter values. The simulations were conducted on a 1-lane, 600m road segment with a 100m, 5 percent upgrade.

and (3.6) to obtain the distribution $v(t)$ and $\xi(t)$. Then, we generate a realization of $\xi(t)$ and use it in the car-following model (3.1). Notice that $\xi(t) = \int_0^t v(s)ds$ tends to be normally distributed (by the Central Limit Theorem) all we need are the mean and variance of $\xi(t)$, which can be obtained analytically.

3.3 The proposed model

A general two-regime stochastic car-following model may be formulated as follows. Let the random process $X_j(t)$ represent the position of vehicle j at time t, and let $f(x;\Theta)$ be the probability density function of $X_j(t)$ given the set of model parameters $\Theta = (\theta_1, \theta_2, \ldots)$ and the data **x**=$\{x_j(t_i)\}$, which represents the observed trajectories of all vehicles $j = 1, 2, \ldots$ at times $t_i, i = 1, 2, \ldots$ up to time t. According to (3.1) we write

$$X_j(t) = \min\{Y(\mathbf{x};\Theta), Z(\mathbf{x};\Theta)\} \tag{3.7}$$

If random processes Y and Z are Bivariate normally distributed (BVM) with a correla-

tion ρ_0:

$$(Y, Z) \sim BVN(\mu_Y, \sigma_Y, \mu_Z, \sigma_Z, \rho_0),$$

where μ_- and σ_- are the mean and standard deviation of the variable in subscript, one can show that (Nadarajah and Kotz, 2008):

$$f(x;\Theta) = \frac{e^{-\frac{(x-\mu_Y)^2}{2\sigma_Y^2}}\,\text{erfc}\left(\frac{\frac{x-\mu_Z}{\sqrt{1-\rho_0^2}\sigma_z} - \frac{\rho_0(x-\mu_Y)}{\sqrt{1-\rho_0^2}\sigma_Y}}{\sqrt{2}}\right)}{2\sqrt{2\pi}\sigma_Y} + \frac{e^{-\frac{(x-\mu_Z)^2}{2\sigma_z^2}}\,\text{erfc}\left(\frac{\frac{x-\mu_Y}{\sqrt{1-\rho_0^2}\sigma_Y} - \frac{\rho_0(x-\mu_Z)}{\sqrt{1-\rho_0^2}\sigma_z}}{\sqrt{2}}\right)}{2\sqrt{2\pi}\sigma_z} \tag{3.8}$$

where $\text{erfc}(x) = \frac{2}{\sqrt{\pi}}\int_x^\infty e^{-t^2}\,dt$ is the complementary error function. This result is notable because the model is analytical and lends itself nicely to be estimated using MLE. In addition, it should have a much better explanatory power compared to existing models that are additive such as (4.1).

3.3.1 The congestion term

Based on our previous discussion, it is advantageous to consider both the free-flow (Y) and the congestion (Z) terms normally distributed, as opposed to the original model (3.1) where Z is deterministic. As demonstrated empirically in Ahn, Cassidy, and Laval (2003) parameters τ and δ can be assumed to follow the bivariate normal (BVN) distribution (Ahn, Cassidy, and Laval, 2003), i.e.

$$(\tau, \delta) \sim BVN(\mu_\tau, \mu_\delta, \sigma_\tau, \sigma_\delta, \rho).$$

Instead of taking the parameters directly from Ahn, Cassidy, and Laval (2003), here we estimate them in conjunction with the free-flow regime parameters, which are formulated next.

3.3.2 The free-flow term

We start by rewriting the proposed car-following model (3.1)-(3.13) as:

$$X_j(t) = \min\{\underbrace{x_j(t-\tau') + \xi_j(\tau')}_{\text{free-flow (Y)}}, \underbrace{x_{j-1}(t-\tau) - \delta}_{\text{congestion (Z)}}\}, \tag{3.9}$$

to emphasize that the time lag in free-flow, τ', and in congestion, τ, do not have to be the same, as customary in the literature, and should be statistically independent. Furthermore, we argue that τ' should not be interpreted as a parameter because it simply defines the initial conditions for the free-flowing component. For simplicity, we fix the value of τ' in the parameter estimation process to

$$\tau' = 1.2s,$$

which is a typical value for the time step in Newell-type car-following models, but any comparable value could be used instead. The main ingredients of the free-flow regime formulation is the displacement ξ_j, as explained next.

The extended desired acceleration model

Section 2 showed that there is a trade-off between the specification of the acceleration error (i.e, BM or g-BM) and the traffic feature that can be replicated (i.e., realistic traffic oscillations or speed-capacity relationship). In this thesis, we introduce a single extra dimensionless parameter $m \geq 1$, that produces models in a scale from the g-BM model ($m = 1$) to the BM model ($m >> 1$), and therefore generalizes the models in the previous section; see Fig. 3.4. The hope is that a suitable value of m will produce a model with the best features of the BM and g-BM models, namely:

1. realistic traffic oscillations: the BM model produces traffic oscillations even at high speeds because the acceleration error SD $[a(v(t))]$ is independent of the speed. In

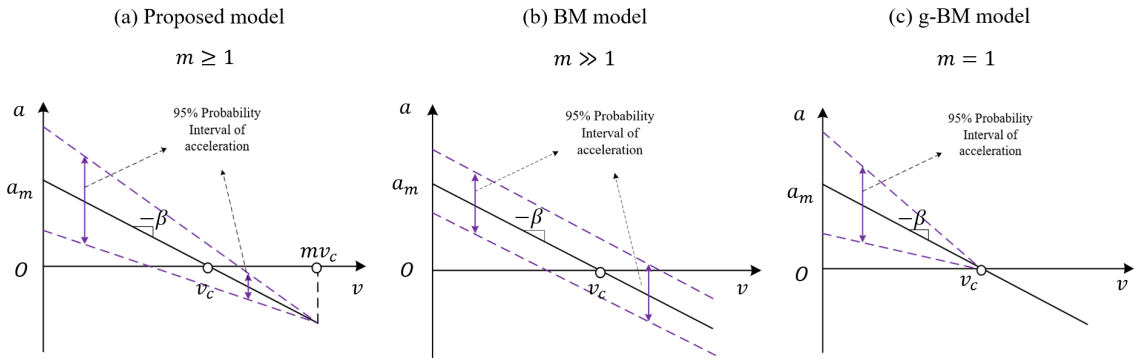

Figure 3.4: Desired acceleration models. In the proposed model the standard deviation of the desired acceleration decrease linearly with the vehicle speed, but does not vanish at v_c. The region between two purple lines represents the 95% probability interval of accelerations.

contrast, the g-BM model cannot because the SD $[a(v(t))]$ is assumed to be zero at the desired speed v_c.

2. speed-capacity relationship: the g-BM model reproduces this relationship because SD $[a(v(t))]$ is a decreasing function of the speed. In contrast, the BM model cannot because SD $[a(v(t))]$ is assumed constant.

Based on these observations, we conjecture that the acceleration error SD $[a(v(t))]$ has to be nonzero at v_c and be a decreasing function of the speed. Here, we impose that SD $[a(v(t))]$ vanish at mv_c and the stochastic differential equation for the proposed acceleration model becomes:

$$\begin{cases} d\xi(t) = v(t)dt, \quad \xi(0) = 0, & \text{(3.10a)} \\ dv(t) = (v_c - v(t))\beta dt + (mv_c - v(t))\sigma dW(t), \quad v(0) = v_0, & \text{(3.10b)} \end{cases}$$

and has analytical solution. For the car-following implementation we simply generate $\xi(t)$ as a normal random variable with mean $\mathrm{E}[\xi(t)]$ given by (3.4) and $\mathrm{Var}[\xi(t)]$ shown in Appendix A.

As mentioned previously, the parameter m can be interpreted as a knob that regulates

how close the model is to the BM or g-BM models. When $m = 1$, it can be seen immediately from (3.10b) that the proposed model is equivalent to the g-BM car-following model. When $m >> 1$, the model is equivalent to the BM model. To see this, notice that

$$\frac{\text{SD}\left[a(v)\right]}{\text{SD}\left[a(0)\right]} = \frac{(mv_c - v)\sigma dW(t)}{(mv_c - 0)\sigma dW(t)} = 1 - \frac{v}{mv_c}. \tag{3.11}$$

A reasonable vehicle speed v is usually greater than 0 and smaller than $2v_c$, so we have $0 < \frac{v}{v_c} < 2$. Along with $m >> 1$, we have $1 - \frac{v}{mv_c} \approx 1 - 0 = 1$, and thus we can say that the driver error, $\text{SD}\left[a(v(t))\right]$, is constant in terms of vehicle speed, and therefore the acceleration process is a BM.

Physical bounds on the parameters $m, \tilde{\sigma}$

As with any normally distributed random variable there is always a probability that it will be negative. Following Laval, Toth, and Zhou (2014), for the worst case scenario $v(0) = 0$, we need $P\{\tilde{v}(\beta\tau) \leq 0\} \leq \alpha$, for a small prescribed α. For typical values $\beta \approx 0.07\text{s}^{-1}$ and $\tau \approx 1.2$s and $\alpha = 0.05$, one has:

$$m\tilde{\sigma} \leq 0.2$$

is a secure range of parameter values for the acceleration process starting from a complete stop. If a negative speed is realized in the simulation, it should be set to 0.

Here, $m\tilde{\sigma} \leq 0.2$ is the secure range of parameters during an acceleration process from a complete stop, i.e. $v(0) = 0$. For usual cases, for example when $v(0) = 60$ km/h, the bound of the product of two parameters could be much higher because $v(\beta\tau) \leq 0$ is very unlikely to happen.

Formulation with roadway upgrade

To take the upgrade of the roadway into account, we replace (3.3) by

$$\mathrm{E}\left[a(v)\right] = \left(u - v(t)\right)\beta - \alpha g \max\{0, G\}, \tag{3.12}$$

where u is the free-flow speed, i.e. the desired speed on a flat road segment, $g = 9.81$ m/s^2 is the acceleration of gravity, G is the roadway grade expressed as a decimal and α is a dimensionless parameter. Notice that in the literature $\alpha = 1$, which is consistent with the assumption that the acceleration due to gravity in the direction of movement, $g \max\{0, G\}$, is subtracted, in its entirety, from the acceleration the driver would impose to the vehicle on a flat segment, $(v_c - v(t))\ \beta$. Here, we add the parameter α to relax this strong assumption in the literature: $0 < \alpha < 1$ indicates they compensate for the upgrade, i.e. that they press the gas pedal harder than they would on a flat segment, while $\alpha > 1$ implies a softer than usual pressing. Values of $\alpha < 0$ are unlikely as it would indicate that acceleration increases with the upgrade

Notice that the desired speed becomes a function of the upgrade, and can be expressed as:

$$v_c = u - \alpha g \max\{0, G\}/\beta. \tag{3.13}$$

Dimensionless formulation

In this section we formulate the proposed model in a dimensionless form to identify a the key parameters that drive the model. One possibility is the following transformations:

$$\tilde{t} = \beta t, \qquad \tilde{v}(\tilde{t}) = v(\tilde{t})/v_c, \qquad \tilde{\xi}(\tilde{t}) = \beta\xi(\tilde{t})/v_c, \qquad \tilde{\sigma}^2 = \sigma^2/\beta. \tag{3.14}$$

where dimensionless variables are indicated with a tilde. It follows that the dimensionless form of (3.10) is:

$$\begin{cases} \tilde{\xi}(\tilde{t}) = v(\tilde{t})dt, \quad \tilde{\xi}(0) = 0, & \text{(3.15a)} \\ \tilde{v}(\tilde{t}) = (1 - \tilde{v}(\tilde{t}))d\tilde{t} + (m - \tilde{v}(\tilde{t}))\tilde{\sigma}dW(\tilde{t}), \quad \tilde{v}(0) = v_0/v_c, & \text{(3.15b)} \end{cases}$$

It can be seen that besides the initial conditions $\tilde{\xi}(0)$ and $\tilde{v}(0)$, the only two non-observable parameters that drive this model are m and $\tilde{\sigma}$.

3.4 Model Properties

In this section, we undertake a detailed analysis of the unified model and show that it is able to replicate virtually all types of traffic instabilities consistently with empirical data, including oscillations, capacity drop, hysteresis and the concave growth pattern of vehicle speeds along a platoon recently reported in Jiang et al. (2014).

3.4.1 The free-flow acceleration process

In this section we analyze the variability of the free-flow acceleration process (3.10). To eliminate irrelevant parameters from this analysis, we can formulate the key variables of this model in dimensionless form. One possibility is the following transformations:

$$\tilde{t} = \beta t, \qquad \tilde{v}(\tilde{t}) = v(\tilde{t})/v_c, \qquad \tilde{\sigma}^2 = \sigma^2/\beta. \tag{3.16}$$

where dimensionless variables are indicated with a tilde.

Figure 3.5 shows 5 realizations of the free-flow acceleration process (3.10) along with the 90% probability bounds, for different values of model parameters. All three combinations of parameter values capture the exponential growth of the vehicle speed during the acceleration process, but they all exhibit different variabilities. As a reference, the figure also includes the empirical data from Laval2014.

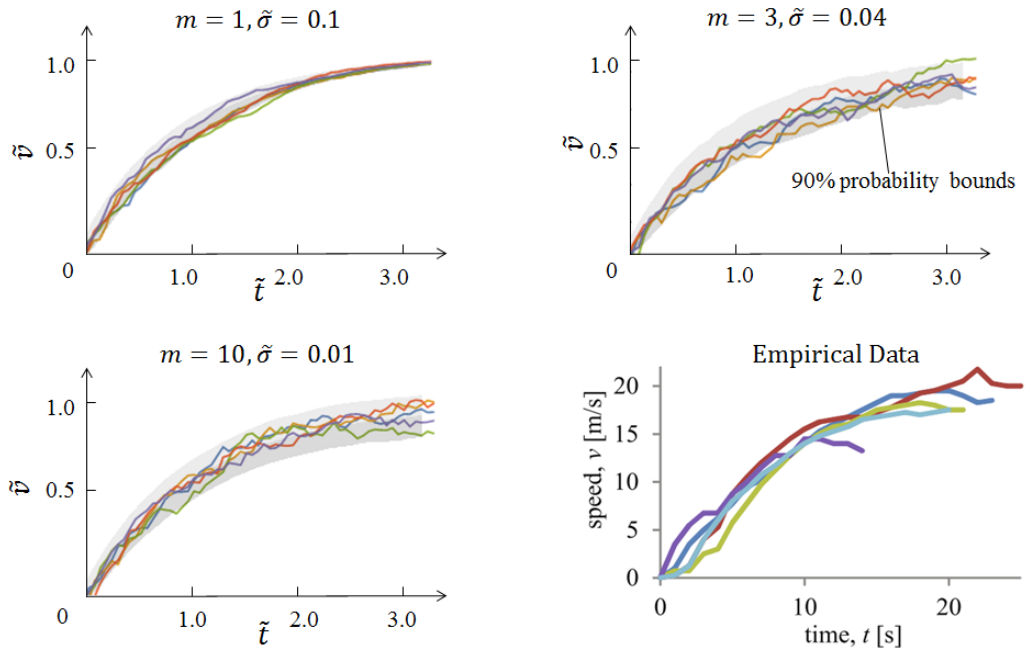

Figure 3.5: Five realizations along with the 90% probability bounds for the acceleration process starting from $\tilde{v_0} = 0$ along with empirical vehicle acceleration data in Laval, Toth, and Zhou (2014). In the figures, each line represents each acceleration process and the gray area represents the 90% probability bounds.

.

The speed variability can be measured by the coefficient of variation of the dimensionless speed: $C[\tilde{v}(\tilde{t})] = V[\tilde{v}(\tilde{t})]^{1/2}/E[\tilde{v}(\tilde{t})]$. we present $C[\tilde{v}(\tilde{t})]/(m\tilde{\sigma})$ in Figure 3.6, which shows that most of the speed variability occurs at the beginning of an acceleration process, i.e. for $\tilde{v_0} < 0.5$ but converges for all values of $\tilde{v_0}$. It can be shown that this value is almost independent of $\tilde{\sigma}$ and also independent of m for large ms. One can obtain that $C[\tilde{v}(\tilde{t})]$ converges to $\frac{(m-1)\tilde{\sigma}}{\sqrt{2-\sigma^2}}$. For a g-BM model, i.e. $m = 1$, the variability of speed converges to 0. For a BM model, i.e. $m >> 1, \tilde{\sigma} \leq \frac{0.23}{m} \approx 0$, the value $C[\tilde{v}(\tilde{t})]/(m\tilde{\sigma})$ converges to $\frac{(m-1)\tilde{\sigma}}{m\tilde{\sigma}\sqrt{2-\sigma^2}} \approx \sqrt{0.5}$, which corresponds well with Laval, Toth, and Zhou (2014). A similar analysis shows that the position variability converges to 0 regardless of the initial speed.

These results indicate that regardless of the parameter values, the speed variability is maximal at the beginning of the acceleration process and that it slowly converges to a value between 0 and $\sqrt{1/2}$ as time passes.

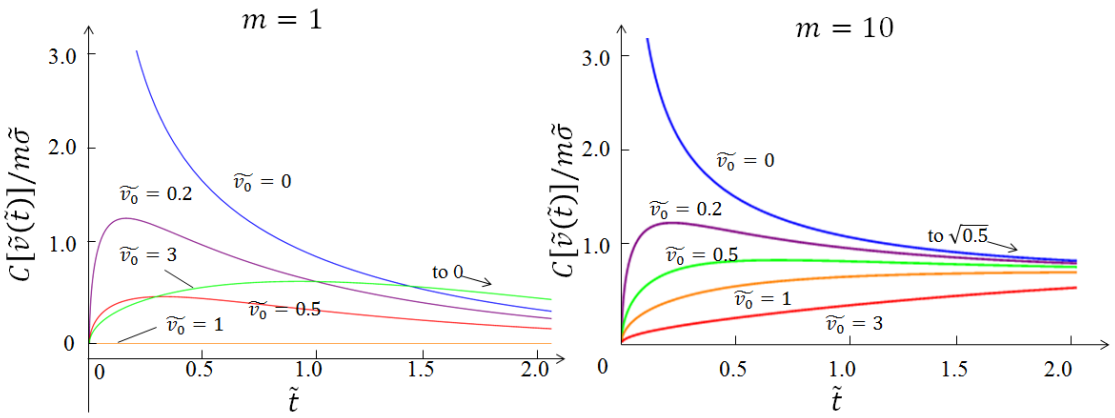

Figure 3.6: Rescaled coefficient of variation for the dimensionless speed $\tilde{v}(\tilde{t})$ for several values of the initial speed $\tilde{v_0}$.

.

3.4.2 Concave growth of platoon oscillation

The concave growth of the vehicle platoon oscillation was first observed by Jiang et al. (2014). Tian et al. (2016) find that the growth pattern in different oscillations collapses into a single concave curve, which indicates the law of the oscillation growth is universal. Treiber and Kesting (2017) shows that the Brownian motion model in Laval, Toth, and Zhou (2014), and the IDM model with noise and indifference regions are all able to reproduce the concave growth pattern, indicating strongly that a stochastic component is vital for replicating this important phenomena.

Table 3.1 shows the estimated values of only a part of parameters of the proposed model, where for simplicity we assumed $\rho = 0$ in the estimation. For more information of the estimation of parameter values, readers are referred to Chapter 4.

Here we show that the model analyzed in this thesis is also able to reproduce the concave growth of the platoon oscillation; see Figure 3.7. The left column in the figure corresponds to the model results with parameters from Table 3.1, where it can be seen that the fit deteriorates for low lead vehicle speeds. The right column in the figure correspond to the same parameter values except for m and $\tilde{\sigma}$ who were chosen to optimize the fit with

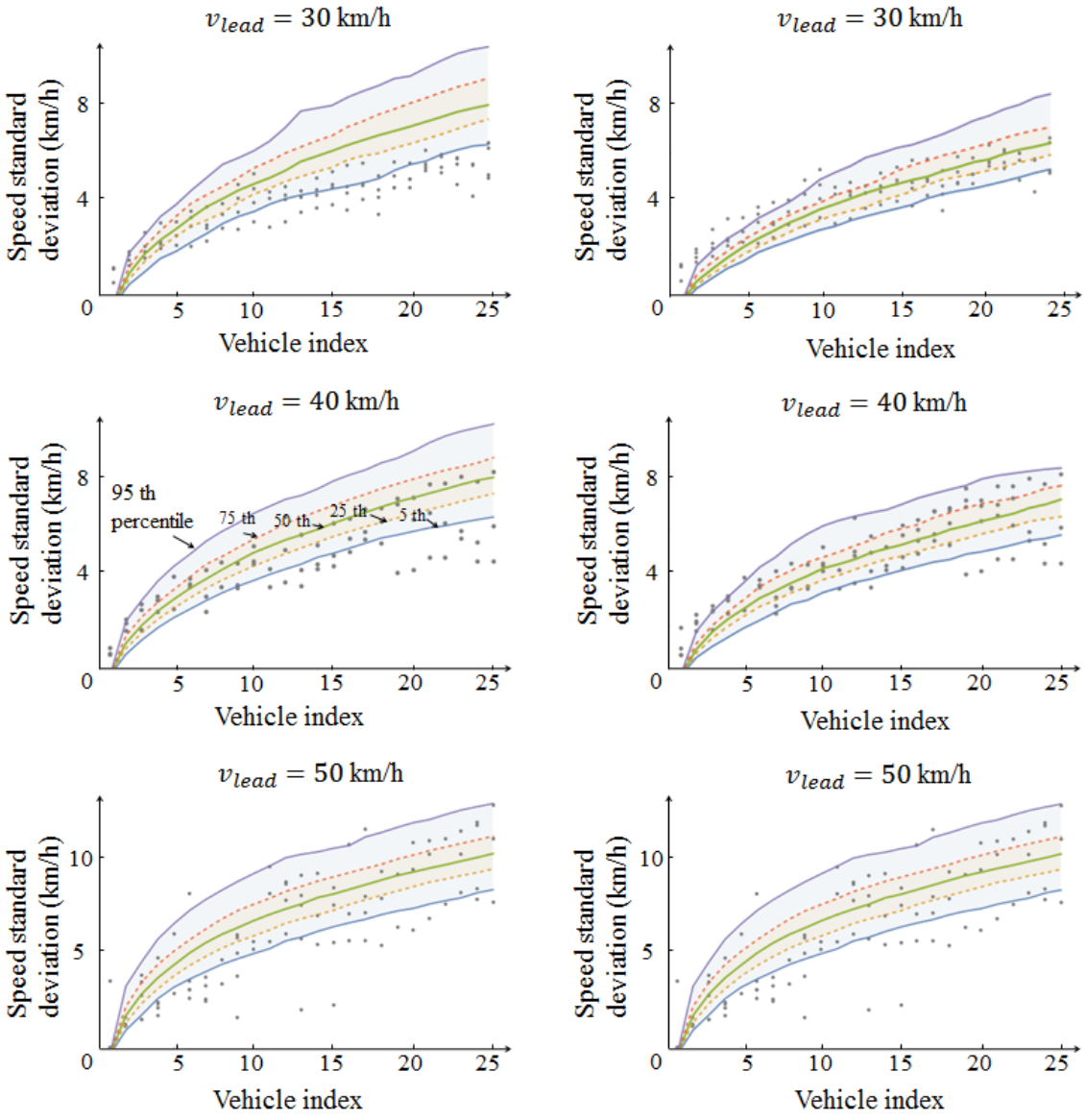

Figure 3.7: Simulated speed standard deviations for 25 vehicles in a vehicle platoon. The leading vehicle drives at $v_{lead} = 30$ km/h, $v_{lead} = 40$ km/h and $v_{lead} = 50$ km/h. The gray dots are obtained from Jiang's car-following experiments (Jiang et al., 2014). The probability bounds in the left column are generated with the estimated parameter values in Table 3.1 while those in the right column are generated with the selected parameter values. The selected values are $m = 1.25$, $\tilde{\sigma} = 0.165$ for $v_{lead} = 30$ km/h and $v_{lead} = 40$ km/h, and the selected values are the same as the estimated values for $v_{lead} = 50$ km/h.

.

the data. This shows that the proposed model reproduces the concave growth of platoon oscillation, and that the parameters m and $\tilde{\sigma}$ dictate the shape of the growth.

Table 3.1: Parameter value estimation results from the car-following experiment data (Jiang et al., 2014)

Parameter	Estimated value	t-stat
μ_τ	0.62	3.3
μ_δ	4.94	3.8
β	80	96.8
m	9.8	6.1
$\tilde{\sigma}$	0.020	5.8
σ_τ	0.24	0.6
σ_δ	2.27	1.0

We also conducted 300-vehicle car-following simulation experiment that shows that the speed standard deviation first increases with vehicle index in a concave way and then stabilizes; see Figure 3.8. This result is consistent with Li and Ouyang (2011) and Yuan et al. (2018).

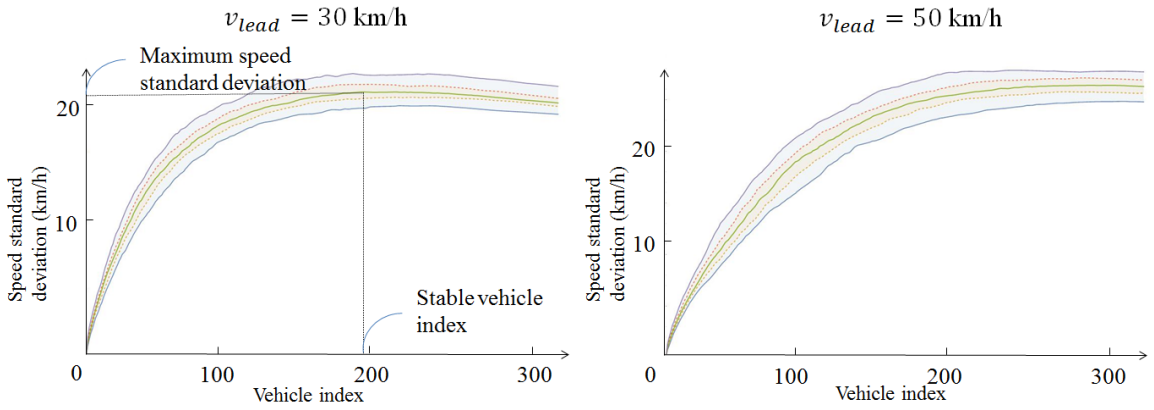

Figure 3.8: Simulated speed standard deviations for 25 vehicles in a vehicle platoon. The leading vehicle drives at $v_{lead} = 30$ km/h (left column) and $v_{lead} = 50$ km/h (right column). Parameter values are $m = 1.25$, $\tilde{\sigma} = 0.176$. The lines in the figures represents the 5th, 25th, 50th, 75th and 95th percentile of the realizations.

.

It is interesting to investigate the maximum oscillation amplitude and also the number of vehicles (stable vehicle index) it takes to stop the oscillation growth; see Figure 3.8. We conducted simulations with different lead vehicle speeds and model parameters. It is shown that the maximum oscillation amplitude is an increasing function of the lead vehicle speed, almost independent of model parameters. In the meantime, the stable vehicle index

increases with v_{lead} and decreases with the product of $m\tilde{\sigma}$. This relationship is visualized in Figure 3.9.

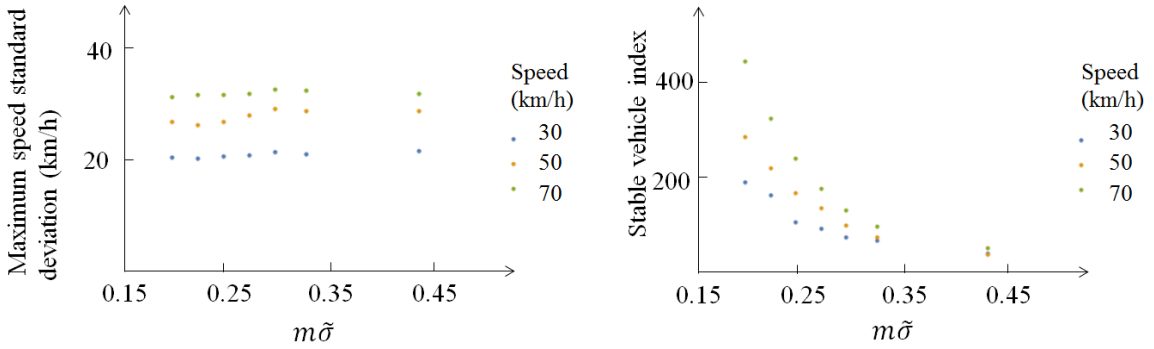

Figure 3.9: Relationship between the maximum speed variation (left column), the stable vehicle index (right column) and the v_{lead}, $m\tilde{\sigma}$. Here the speed process does not start with $\tilde{v}(0) = 0$ so we can accept a larger range of $m\tilde{\sigma}$ than we mentioned in the previous section.

.

3.4.3 Periodic oscillation at uphill bottlenecks

In this section, we study the effects of the parameters m and $\tilde{\sigma}$ on the period and amplitude of the oscillations produced by the model, and compare them with empirical data. We conducted simulations on a 600m one-lane flat road except for a 100m uphill segment starting at location x =400 m.

To take roadway grade into account, one can express the desired speed as $v_c = u - gG/\beta$, where $g = 9.81$ m/s^2 is the acceleration of gravity and G is the grade expressed as a decimal. We used a free-flow speed $u = 120$km/h for the simulations and varying the parameters β, G, m and $\tilde{\sigma}$ within their typical ranges. Then we used Fourier spectrum analysis to estimate the period and amplitude of the speed series oscillations 100m upstream of the bottom of the upgrade, as suggested by Li, Peng, and Ouyang (2010) and Laval, Toth, and Zhou (2014).

Figure 3.10 shows the distribution of dimensionless periods (period times β) and dimensionless amplitudes (amplitude divided by v_c) as a function of $m\tilde{\sigma}$ and the average

speed at the start of the uphill segment, V_{avg}. It can be seen that: (i) the dimensionless period is a decreasing function of $m\tilde{\sigma}$ and the dimensionless amplitude is an increasing function of $m\tilde{\sigma}$ (ii) the distribution of dimensionless period and dimensionless amplitude as a function of V_{avg} accords well with the empirical data in Treiber and Kesting (2012) and Knoop et al. (2012), and (iii) all distributions exhibits a large variance. These results are similar to those in Laval, Toth, and Zhou (2014) except that the relationship between the dimensionless period and amplitude are correlated with $m\tilde{\sigma}$ but not simply $\tilde{\sigma}$.

The relationship between V_{avg} and $m\tilde{\sigma}$ is presented in Figure 3.11. V_{avg} decreases with $m\tilde{\sigma}$. Again, this result is similar to Laval, Toth, and Zhou (2014) except that V_{avg} is correlated with $m\tilde{\sigma}$ but not only $\tilde{\sigma}$. This indicates that the average speed at the bottleneck decreases with the the randomness of driver accelerations.

3.4.4 Speed-capacity relationship in congestion

Recent empirical findings by Yuan, Knoop, and Hoogendoorn (2015) show that the queue discharge rate out of a moving jam increases with the speed in the queue. On average, they observed that the capacity drop magnitude is around 26% when the vehicular speed in congestion is 0 km/h and decreases as the vehicular speed increases. The mechanisms behind this "speed-capacity" relationship was recently unveiled in Yuan et al. (2018): the key ingredients are bounded vehicle acceleration and an error term that decreases with the speed. As noted earlier, the geometric Brownian motion model in Yuan et al. (2018) reproduces this phenomenon for the first time but loses the ability to generate realistic traffic oscillations. The model analyzed in this thesis does not have this drawback, as shown next.

We perform the simulation of a 25-vehicle car-following experiment on a one-lane road with no passing, where the lead vehicle drives initially at a constant low speed, then accelerates to free-flow speed in the resulting discharge rate is measured at that location; see Figure 3.12.

Figure 3.14 plots the results, where it can be seen that the model is able to successfully

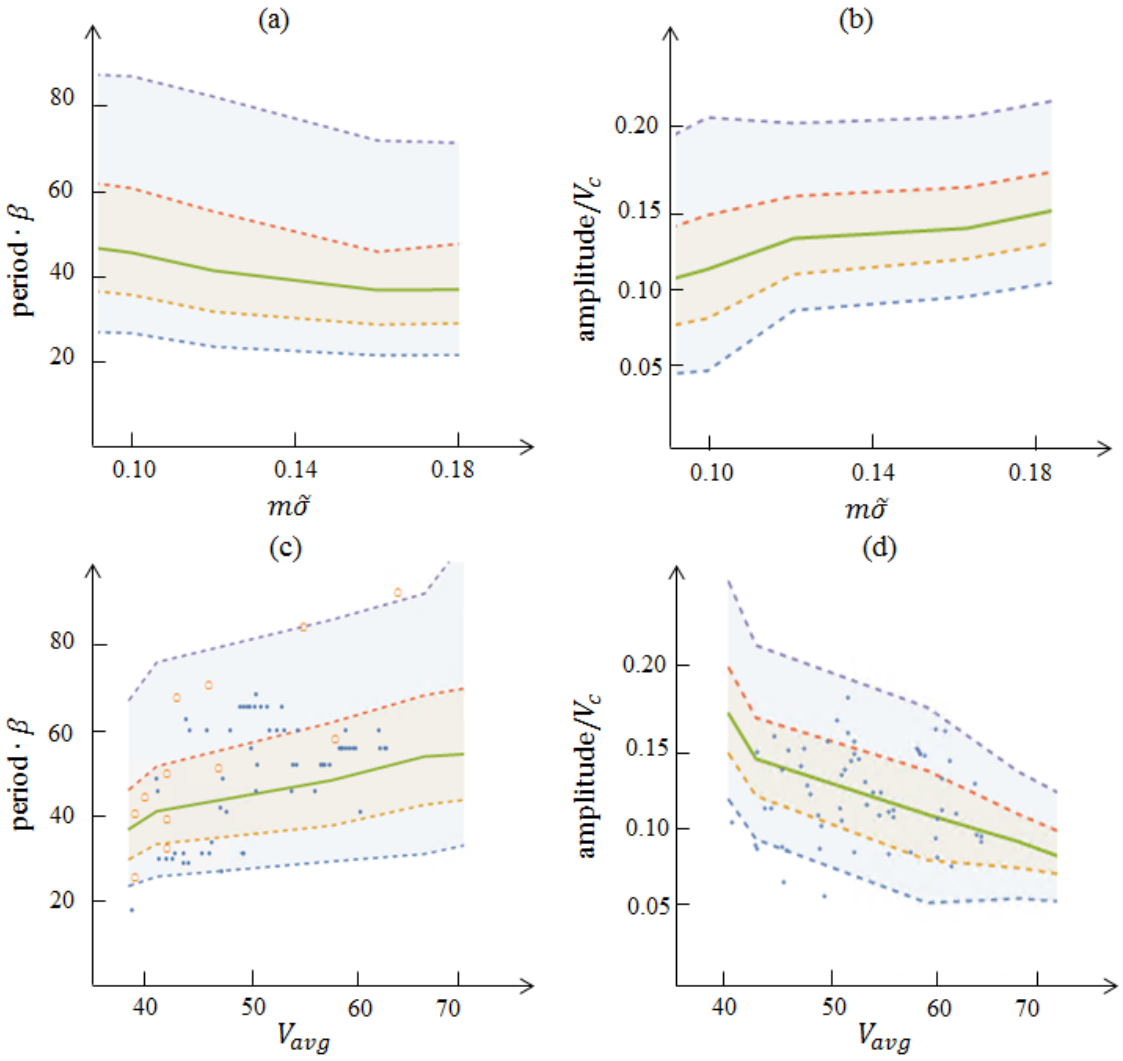

Figure 3.10: The dimensionless period and the dimensionless amplitude as a function of model parameter $m\tilde{\sigma}$ and V_{avg}. The lines represents the 5th, 25th, 50th, 75th and 95th percentile of the simulation results. The orange circles in part (c) are empirical data originally taken from Treiber and Kesting (2012) and the blue dots in part (c) and (d) are taken from Knoop et al. (2012). To convert the empirical data in dimensionless form we used $w = 18$km/h, $\beta = 0.07\text{s}^{-1}$ and the following linear regression on the simulated data: $v_c = 3.6V_{avg} - 71.4$.

reproduce the increasing speed-capacity relationship and that the slope of this increasing relationship is inversely proportional to m. This is as expected because small values of m means that the model is closer to the geometric Brownian motion model.

The important point here, is that a value of $m \approx 1.2$ is able to replicate both the speed-

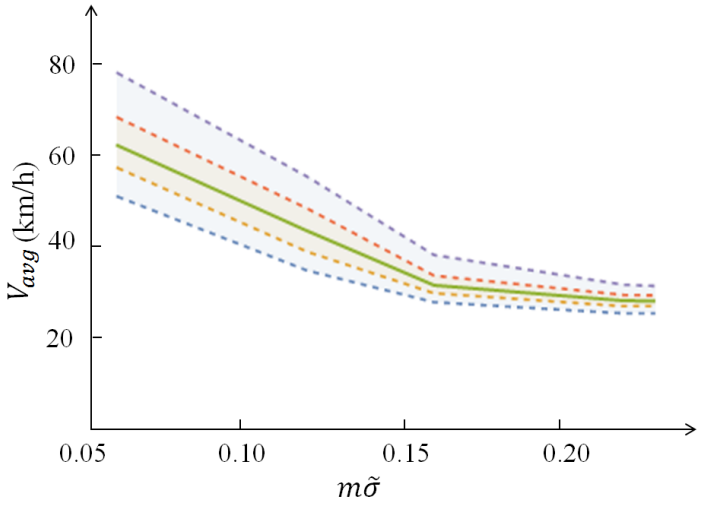

Figure 3.11: Average speed at the bottom of the uphill segment as a function of $m\tilde{\sigma}$. The lines represents the 5th, 25th, 50th, 75th and 95th percentile of the simulation results.

.

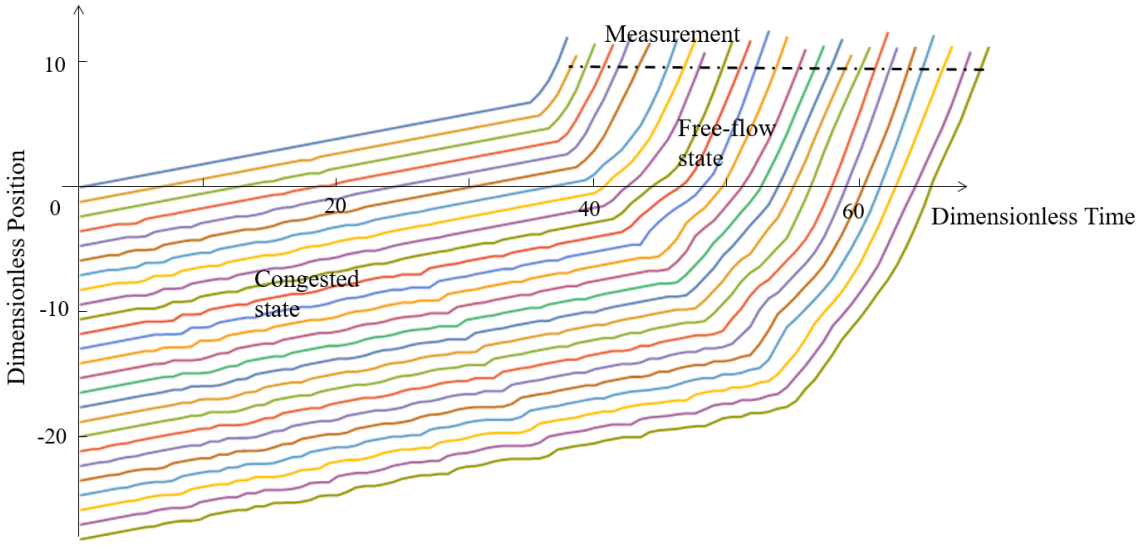

Figure 3.12: A sample trajectory of the queue discharge experiment with $\tilde{v_0} = 0.1$, $m = 1.25$ and $\tilde{\sigma} = 0.35$. The initial spacing of the vehicles is calculated based on the triangular fundamental diagram $\rho = \delta + \tilde{v_0}\tau$. Discharge rate is measured at the black dash line where the vehicle speeds reach free flow speed.

capacity relationship and realistic traffic oscillations. To see this, the speed-capacity relationship is shown in Figure 3.14 and 3.15, and an example of the realization of traffic oscillations is provided in Figure 3.13.

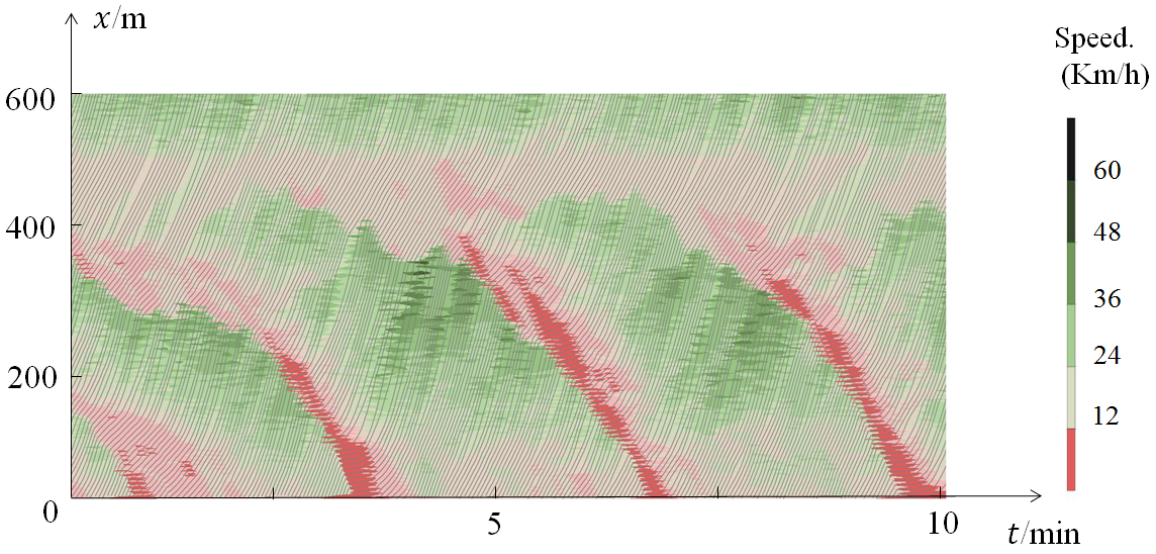

Figure 3.13: An example of the realistic traffic oscillations generated with the proposed model. The model parameters are $m = 1.2$ and $\tilde{\sigma} = 0.16$.

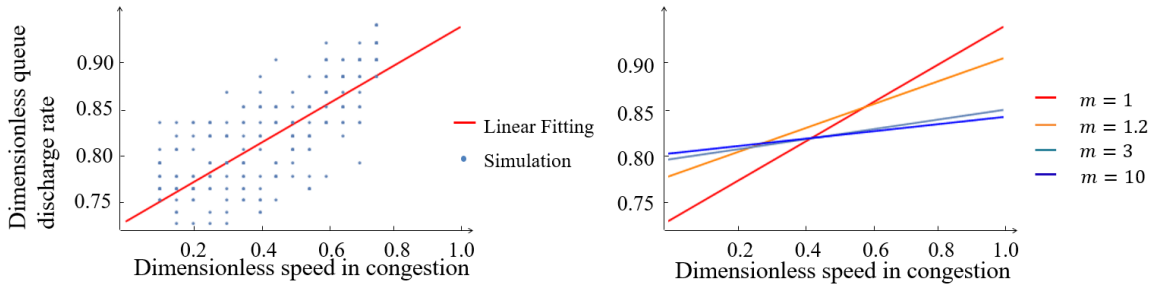

Figure 3.14: (a) Simulation results and a linear fitting of the desired model with parameter $m = 1$, which shows a positive relationship between queue discharge rate and speed in congestion. (b) Dimensionless queue discharge rate as a function of dimensionless speed in congestion for $m \in \{1, 1.2, 3, 10\}$. The model gradually loses its ability to catch the positive relationship between speed and queue discharge rate with the increase of the value of m. The dimensionless queue discharge rate is the ratio of the queue discharge rate to the roadway capacity.

We also find that for the queue discharge rate at decreases with $\tilde{\sigma}$ for the same values of m and $\tilde{v}$, two examples are provided in Figure 3.15. As we can see, for $m = 1.25$ and $\tilde{v} = 0.6$, the corresponding dimensionless queue discharge rates are approximately 0.95, 0.91 and 0.87 for $\tilde{\sigma} = 0.15, 0.25$, and 0.35.

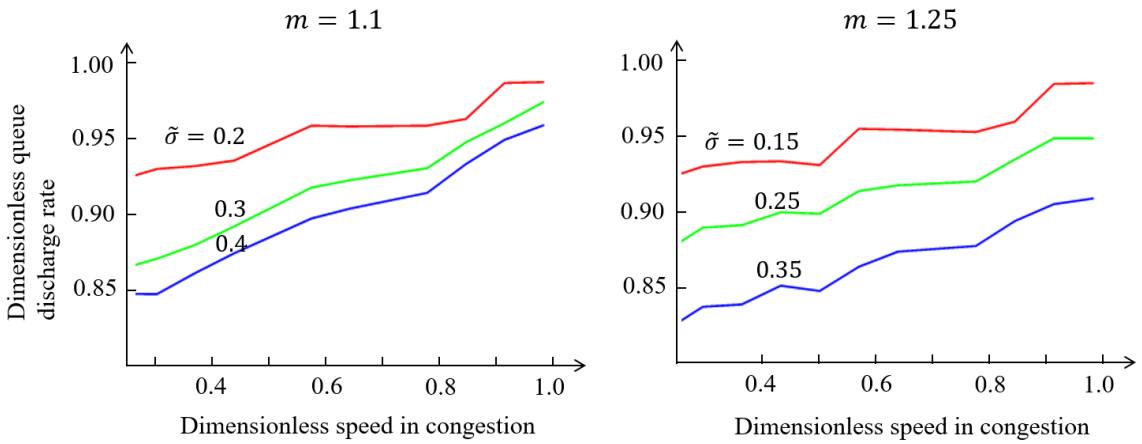

Figure 3.15: Dimensionless queue discharge rate as a function of speed for different values of $\tilde{\sigma}$. In the figures, we can observe that for the same speed, queue discharge rate decreases with $\tilde{\sigma}$.

3.4.5 Vehicle speed distributions

Here we use repeated simulations of the same scenario to approximate the speed distribution of each vehicle $n = 1, 2, \ldots$ in a platoon, where the trajectory of the leader $n = 1$ is taken from empirical data. We use two datasets: (i) from Jiang et al. (2014) who perform a car-following experiment with 25 vehicles on a two-way 3.5 km suburban road, and (ii) from Laval, Toth, and Zhou (2014) who perform a similar car-following experiment on a two-lane road near Georgia Tech with six vehicles. The grade data is also provided. For both two datasets, the realizations of the follower speeds (15th vehicle for dataset 1, and 6th vehicle for dataset 2) fall into the probability bands and the width of the probability band is a function of $m\tilde{\sigma}$. Examples are provided in Figure 3.16. In the predictions, the trajectories of the platoon leader and the initial conditions of all followers are provided. Then we run simulations repeatedly to get the distributions of the predicted vehicle speed.

3.4.6 Oscillations propagating downstream

Here we conduct simulations with the proposed model. A comparison between the simulation results and field data is also presented.

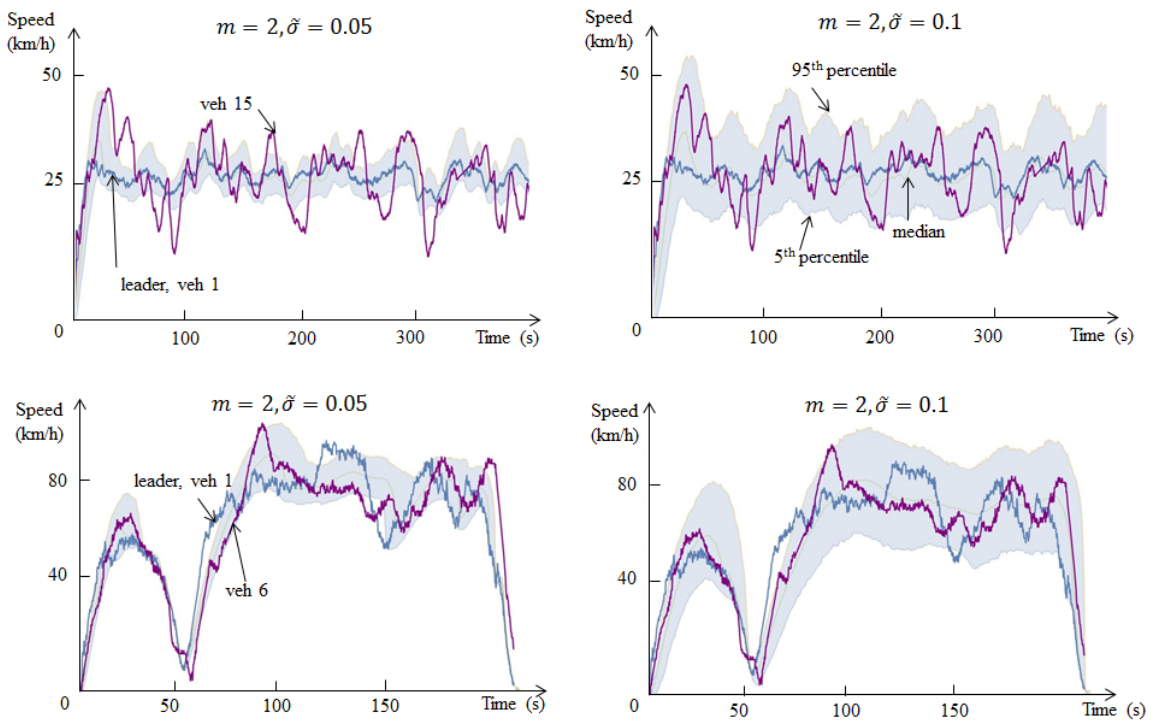

Figure 3.16: Results of car-following experiments with initial data from Jiang et al. (2014) (top row) and Laval, Toth, and Zhou (2014) (bottom row). The observed speeds for the leader and the vehicles are shown in solid lines. The shaded area represents the 90%-probability band of the predicted vehicle speed. The vehicle speed curve has been shifted to the left by the wave trip time between the two vehicles.

For each vehicle in the simulation, we randomly generate τ_j and δ_j for 25 cars based on the estimated mean and standard error values from dataset 1. In the simulation, we input the leader trajectory from dataset 1 and then generate initial conditions of the platoon based on parameters τ_j and δ_j . Given the estimated parameter values from dataset 1 and the initial conditions, we generate vehicle trajectories one by one with the model (3.7) - (4.2). The results are shown in Fig. 3.17.

In Fig. 3.17, we can see that the simulation results and the real data show high resemblance. The high-speed regions in the experiments occur at similar places in the simulations. Besides, we can observe oscillation propagating downstream and trajectories that ”bounce” at the end of the vehicle-platoon in both field data and simulation results, as explained next.

It is well-known that traffic oscillations in congested traffic propagate upstream at the wave speed of approximately 20 km/hr. To our surprise, however, dataset 1 exhibits oscillation that travel downstream, and this phenomenon also occurs in our simulation results; see Fig. 3.17.

To explain this, recall that the driver acceleration process is a stochastic process, where some drivers would appear to accelerate more aggressively than others. As a result, large spacing will appear between leaders that are more aggressive than their followers, which can lead to low-density traffic states in the platoon. Let "A" denote these free-flow states where vehicles have the possibility of accelerating to speeds close to the free-flow speeds; see Figs. 3.17 and 3.18. Notice that these "pockets" of free flow inside congestion have been observed empirically in Kim and Cassidy (2012). Eventually, these vehicles will catch up to the leader traveling at a nearly constant speed and inducing congested traffic states "U" in Fig. 3.18. As a result, and according to kinematic wave theory, a shock is produced between states A and U which travels at the speed s_{AU} shown in the figure. That the speed s_{AU} is positive means that the flow in state A tends to be lower than in state U because of the large spacings mentioned earlier. After the first shock, the process starts over leading to periodic oscillations moving downstream.

We also notice that in Fig. 3.17 (c) and (d), there is a slight difference in terms of vehicle spacings. But what is important here is that the speed-contour plots in (c) and (d) match well, which means the simulation captures the main macroscopic features of the field trajectories.

3.5 Discussion

This chapter presents a generalization of the works conducted by Laval, Toth, and Zhou (2014) and Yuan et al. (2018), each producing different features of traffic instabilities, by adding a dimensionless parameter m. This parameter regulates the type of driver free acceleration error on a scale from Brownian motion (BM, $m >> 1$) to geometric Brownian

(g-BM, $m = 1$) acceleration processes. When $m >> 1$, this model captures proper periodicity of traffic oscillations at bottlenecks and demonstrates that human factor alone can cause traffic instabilities. This result accords with the work conducted by Laval, Toth, and Zhou (2014) well. When $m = 1$, this model is equivalent to the model in (Yuan et al., 2018), which successfully reproduces the speed-capacity relationship observed empirically.

To understand why $m >> 1$ leads to oscillations, recall that for oscillations to develop first they have to form. From equation (11b) note that the error added to the vehicle trajectory decreases with vehicle speed. When $m = 1$ this error goes to zero when traveling at the desired speed, which prevents any initial perturbation from occurring, let alone developing into a fully grown oscillation. When $m >> 1$ the error is greater than zero and increasing with m, which allows the initial perturbation to occur (and possibly grow into a mature oscillation). To understand why $m \to 1$ leads to speed-capacity relationship in bottleneck, consider a (n+1)-vehicle car-following experiment on a one-lane road with no passing, as shown in Fig. 3.19. In this experiment the lead vehicle drives initially at a constant low speed v_0, then accelerates to free-flow speed in the resulting discharge rate is measured at that location $x = 0$. We can express the discharge flow as:

$$Q = \frac{n}{t_n - t_0}, \tag{3.17}$$

where t_n is the time that vehicle n crosses the measurement location, i.e. $x_n(t_n) = L$ which is also equal to $n\delta$ and $t_0 = -n\tau$ in our case. Using the law of total probability we can write:

$$\begin{aligned} P(x_n(t) \geq L) &= P(x_n(t) \geq L \mid t_n \leq t)P(t_n \leq t) + P(x_n(t) \geq L \mid t_n > t)P(t_n > t) \\ &= 1 \times P(t_n \leq t) + 0 \times P(t_n > t) \\ &= P(t_n \leq t). \end{aligned} \tag{3.18}$$

The second equality in (3.18) follows from $x_n(t)$ being a monotonic increasing function,

and thus $x_n(t) \geq x_n(t_n) = L$ only if $t_n \leq t$. Consequently, $P(x_n(t) \geq L \mid t_n \leq t) = 1$ and $P(x_n(t) \geq L \mid t_n > t) = 0$. Combining (3.17) and (3.18) we have that the cumulative distribution function of the discharge rate Q can be written as:

$$\begin{aligned} P(Q < q) &= P(t_n > \frac{n}{q} + t_0) \\ &= P(x_n(\frac{n}{q} + t_0) < L) \\ &= 1 - \frac{1}{4}\text{erfc}\left(\frac{n\delta - \text{E}\left[\xi_n(\frac{n(1-\tau q)}{q})\right]}{\sqrt{2}\text{SD}\left[\xi_n(\frac{n(1-\tau q)}{q})\right]}\right)\text{erfc}\left(\frac{\delta q - u}{\sqrt{2}q\sqrt{\sigma_\delta^2 + \sigma_\tau^2 u^2}}\right) \end{aligned} \tag{3.19}$$

which is analytical. Fig. 3.20 shows the expected value of the discharge flow, $E(Q)$, against the speed in congestion for different model parameters. The left panel in the figure shows that the slope of the line decreases with m, and the model loses its ability to generate speed-capacity relationship in the Brownian motion limit.

The right panel in the figure strongly suggests that the product $m \cdot \tilde{\sigma}$ is the main parameter of the model, while the bottom panel illustrates how the discharge flow decreases with $m \cdot \tilde{\sigma}$. We verified that these results hold for a wide range for the remaining parameter values.

We find that if we select $m \approx 1.2$, the model could do both the oscillation and speed-capacity relationship (Xu and Laval, 2019). In the future, we plan to do statistical tests on the model homogeneity between different drivers and different vehicle classes (regular vehicles and trucks). If we find the model parameters are not homogeneous between different drivers or vehicle types, we will consider assigning different values of ms to different groups of drivers or vehicles to build hybrid models. Such hybrid models have good potential to reproduce different types of traffic instabilities.

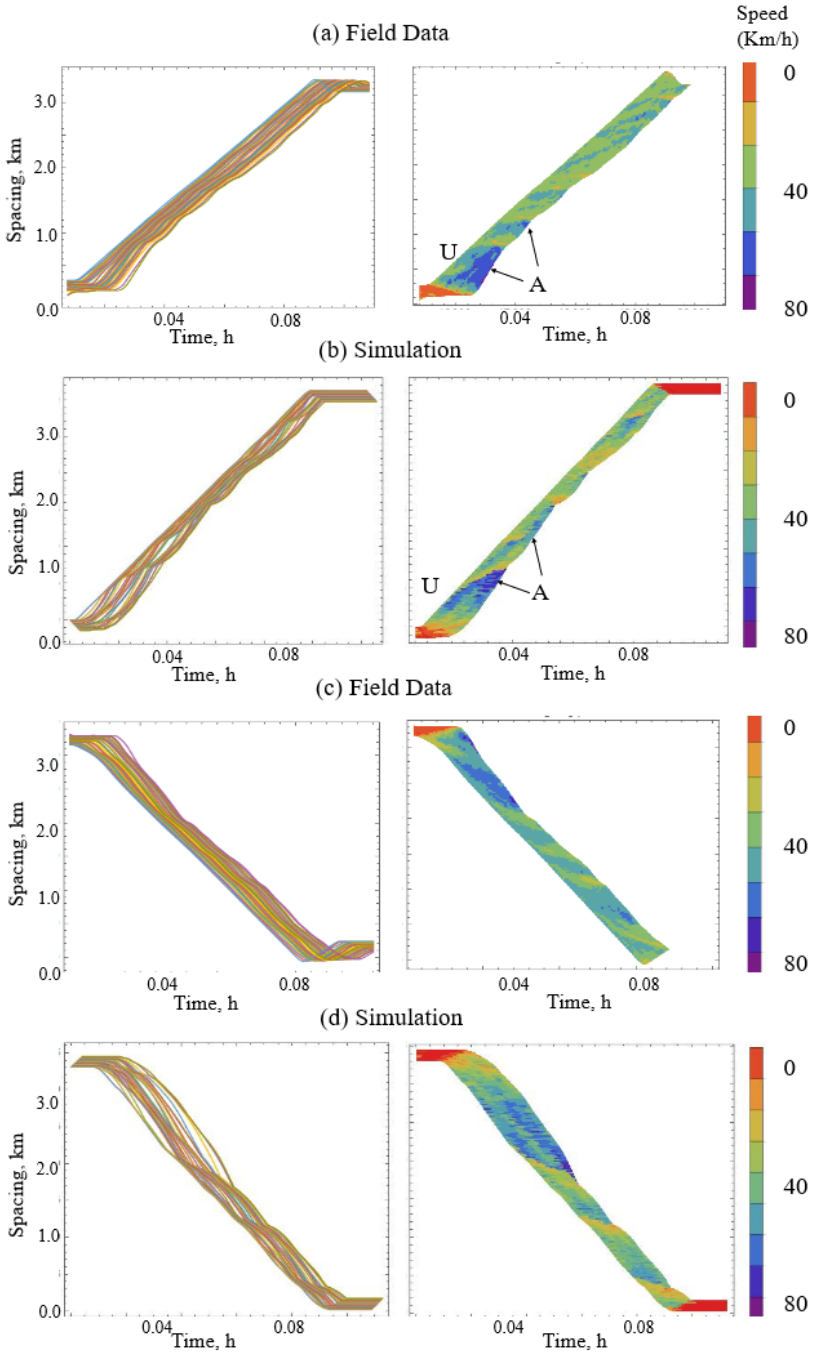

Figure 3.17: A comparison between field trajectory data and simulation results. (a) Trajectories of a northbound car-following experiment; (b) simulation of the experiment in (a); (c) trajectories of a southbound car-following experiment; (d) simulation of the experiment in (c). The low-speed areas are identified in red or yellow, while high-speed areas are identified in purple or blue. We can observe the high resemblance between real trajectories and simulation results. In the figures, we can observe different traffic states U and A which will be explained next and shown in Figure 3.18.

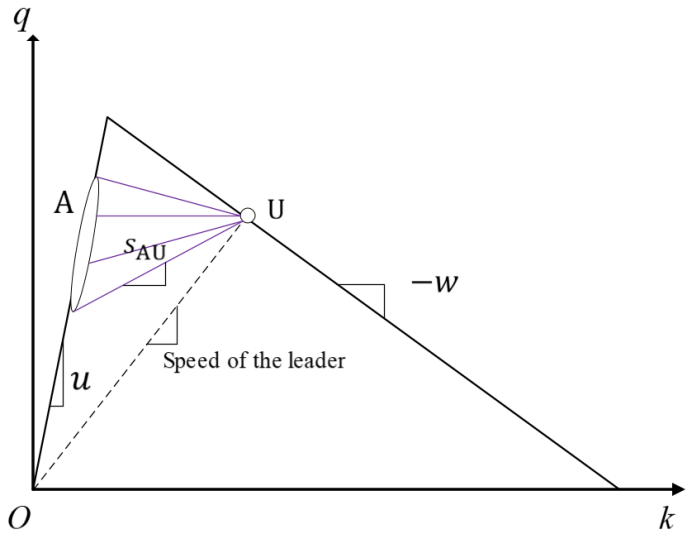

Figure 3.18: The triangular fundamental diagram used to analyze the unusual wave speed and the trajectory bounces. The ellipse represents the possible positions of traffic state A. The purple lines represents the possible traffic shock waves with negative or positive wave speed.

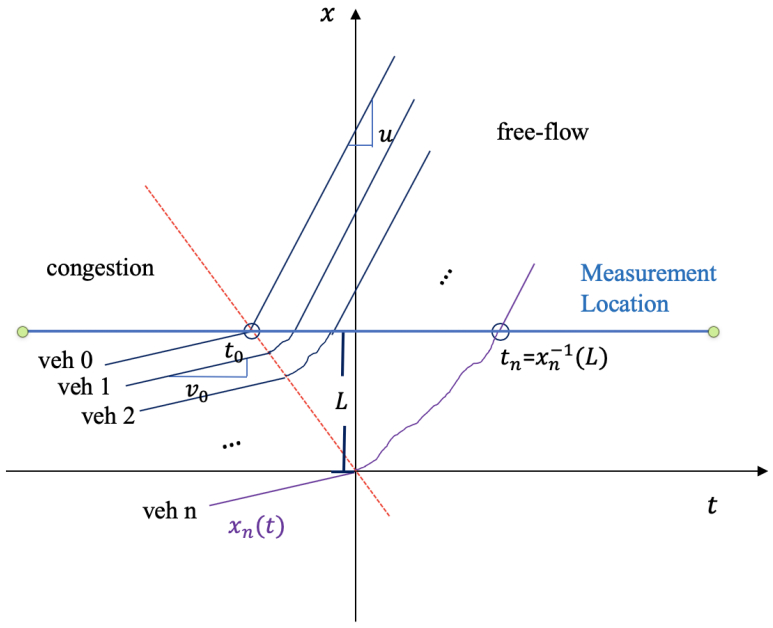

Figure 3.19: The discharge flow experiment: vehicles are travelling at a constant low speed v_0 at the beginning and then accelerate to free-flow speed. The trajectory of nth vehicle $x_j(t_n)$ is shown in the figure and t_n is the time that the nth vehicle crosses the measurement location.

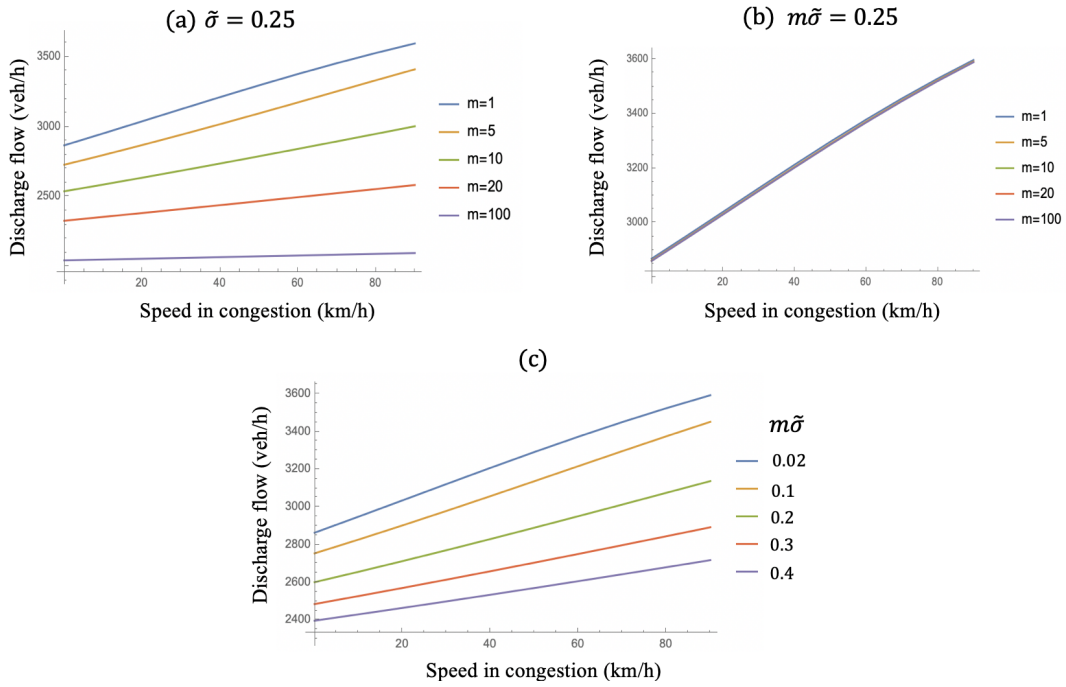

Figure 3.20: The plot of discharge flow against speed in congestion. We used $n = 50$, $\mu_\delta = 6$m, $\mu_\tau = 0.75$s, $u = 100$km/h, $\beta = 200\text{h}^{-1}$, $\sigma_\delta = 1$m and $\sigma_\tau = 0.4$s.

CHAPTER 4

PARAMETER ESTIMATION AND STATISTICAL INFERENCE

This chapter is modified from the published paper:

Xu, T., & Laval, J. A. (2020). Statistical inference for two-regime stochastic car-following models. Transportation Research Part B. 134, pp. 210-228.

4.1 Introduction

The estimation of stochastic car-following model parameters has a short history because the vast majority of car-following models existing in the literature are deterministic. To the best of our knowledge, the only relevant reference would be Ahmed (1999) and Hoogendoorn and Ossen (2005), who use MLE for stochastic car-following models of the type:

$$x_j(t) = F(\mathbf{x}; \Theta) + \epsilon \tag{4.1}$$

where $x_j(t)$ is the vehicle position time t (or speed or acceleration), F is a deterministic car-following model with parameters $\Theta = (\theta_1, \theta_2, \ldots)$, **x**=$\{x_j(t_i)\}$ are the trajectory data points for vehicle $j = 1, 2, \ldots$ at times $t_i, i = 1, 2, \ldots$, and ϵ is a normal random variable with mean of zero and constant standard deviation. The problem with this formulation is that the additive error may not be appropriate for car-following models, which are generally accepted to be *two-regime* models, such as the one analyzed in this thesis. These models recognize that the free-flow and congestion components obey completely different laws, which are reconciled under the minimum operator. It follows that a stochastic extension would *not be* an additive model. If random noises are to be added in such models, each regime should include errors of different nature and magnitudes, which would not be consistent with (4.1). In the proposed model of this thesis work, the probability density of

vehicle positions turns out to be analytical. This means that we are not constrained by the additive error assumption of existing formulations, and model parameters will have better statistical properties.

4.2 Formulation

In this section, we apply maximum-likelihood estimation (MLE) to estimate the parameters in our proposed model (3.9). Based on the formulation in the previous section, it is not difficult to show that the mean and variance of the free-flow term ($Y = x_j(t-\tau') + \xi_j(\tau')$) and congestion term ($Z = x_{j-1}(t-\tau) - \delta$) are given by:

$$\begin{cases} \mu_Y = x_j(t-\tau') + \mathrm{E}\left[\xi(\tau')\right], & \text{(4.2a)} \\ \sigma_Y^2 = \mathrm{Var}\left[\xi(\tau')\right], & \text{(4.2b)} \\ \mu_Z = x_{j-1}(t-\mu_\tau) - \mu_\delta - a_{j-1}(t-\mu_\tau)\sigma_\tau^2/2, & \text{(4.2c)} \\ \sigma_Z^2 = v_{j-1}^2(t-\mu_\tau)\sigma_\tau^2 + \sigma_\delta^2 + 2\rho v_{j-1}^2(t-\mu_\tau)\sigma_\tau\sigma_\delta, & \text{(4.2d)} \end{cases}$$

which is all we need to evaluate (3.8). In (4.2a) and (4.2b) recall that $\tau' = 1.2$s is a constant, and that $\mathrm{E}\left[\xi(t)\right]$ is given in (3.4) and $\mathrm{Var}\left[\xi(t)\right]$ in Appendix A. For (4.2c) and (4.2d) we used second- and first-order approximations (Ang and Tang, 2007) for $\mathrm{E}\left[x_{j-1}(t-\tau)-\delta\right]$ and $\mathrm{Var}\left[x_{j-1}(t-\tau)-\delta\right]$, respectively, and $a_{j-1}(t)$ is the acceleration of vehicle $j-1$ at time t.

Before we can use MLE to estimate the parameters:

$$\Theta = (\mu_\tau, \mu_\delta, u, \beta, m, \tilde{\sigma}, \rho, \sigma_\tau, \sigma_\delta, \alpha, \rho_0)$$

we perform three additional steps. First, we constrain the parameter search space within the upper and lower bounds shown in Table 4.1, to ensure that the estimation results remain physically meaningful and to avoid spurious local maxima.

Table 4.1: Upper and lower bounds for parameter estimation

Parameter	Lower bound	Upper bound
μ_τ	0.4	2.0
μ_δ	3	20
u	60	90
β	50	350
m	1	10
$\tilde{\sigma}$	0	0.3
ρ	-1	1
σ_τ	0	1
σ_δ	0	5
α	-2	4
ρ_0	-1	1

Secondly, we perform a third-order interpolation on the discrete trajectory data so that the $x_j(t)$, its derivative $v_j(t)$ and its second derivative $a_j(t)$ become continuous functions of time. This is illustrated in Fig. 4.1, where at times $\{t_1, t_2, ...t_i, ...\}$, we have $\mathbf{x} = \{x_j(t_1), x_j(t_2), ...x_j(t_i), ...\}$ represented by circles, the $\{x_j(t_1 - \tau'), x_j(t_2 - \tau'), ...x_j(t_i - \tau'), ...\}$ in (4.2a) represented by triangles and the $\{x_{j-1}(t_1 - \mu_\tau), x_{j-1}(t_2 - \mu_\tau), ...x_{j-1}(t_i - \mu_\tau), ...\}$ in (4.2c) represented by squares. The randomness of parameter τ is shown by the uncertainty of the position of the squares on Fig. 4.1.

4.2.1 Data

We have two independent datasets for the parameter estimation.

Dataset 1: Jiang et al. (2014) conduct a car-following experiment with 25 vehicles all equipped with data collection devices. In this experiment, the vehicle platoon traveled northbound and southbound repeatedly, on a two-way 3.5 km suburban road segment in Hefei, China. An example of trajectory data is provided in Fig. 4.2. The grade data is also contained in the dataset.

Dataset 2: The second dataset is from (Laval, Toth, and Zhou, 2014) who perform a similar car-following experiment on a two-lane road near Georgia Tech with six vehicles. The grade data is also provided.

In both two sets of experiments, the lead vehicle was instructed to maintain a given constant speed while other vehicles were expected to follow its leader naturally. There was no interference from traffic signals or other vehicles in the experiments.

Fig. 4.2 shows a typical car-following experiment, with the lead vehicle maintaining a constant low speed. Notice that the followers travel at speeds higher than the lead vehicle speed occasionally, as seen by the blue region in the figure, which is possible because the leader speed is much lower than the free-flow speed u.

4.3 Estimation results and statistical inference

Here we present the MLE parameters using all the data within each data set, and also for each particular experiment within each data set.

We randomly choose five car-following experiments from dataset 1. For each car-following experiment, we select a data point $x_j(t_i)$ each $dt = 12$ sec during a five-minute period for the first 15 cars such that 25 data points are used for each car and 375 data points are used in each car-following experiment. We perform MLE for each car-following experiment and also for the overall dataset ($375 \times 5 = 1875$ data points are used in total). For dataset 2, we used the same method for the six-vehicle platoon with a smaller $dt = 8$ sec during a 140-second period. 18 data points are used for each vehicle and $18 \times 5 = 90$ data points are used in each car-following experiment. $90 \times 5 = 450$ data points are used for the estimation of the overall dataset. We decreased the value of dt for dataset 2 in order to obtain more sample points without sacrificing independence. We tested different values of dt and find that the correlation among data points are negligible with the chosen dt. The parameter evaluation results for the overall datasets are provided in Table 4.2 and the results for each car-following experiments are listed in Table 5.3 and Table 5.2.

Parameters u and β are two major parameters in the free-flow branch, which determine the mean of the vehicle acceleration and ξ. The estimation results of u and β for both datasets have high statistical significance. These two variables are both significant. The

Table 4.2: MLE parameter values for the overall datasets

Parameter	Dataset 1 mean value	t-stat	Dataset 2 mean value	t-stat
$\widehat{\mu_\delta}$	5.78	12.4	4.28	3.8
$\widehat{\mu_\tau}$	0.54	7.	0.46	6.6
$\widehat{u}$	60.14	40.3	70.	24.9
$\widehat{\beta}$	94.78	66.9	210.11	7.4
$\widehat{m}$	6.13	9.1	4.21	3.3
$\widehat{\tilde{\sigma}}$	0.04	8.2	0.05	2.7
$\widehat{\rho}$	-0.49	-1.4	-0.46	-4.7
$\widehat{\sigma_\delta}$	1.63	2.1	4.88	4.6
$\widehat{\sigma_\tau}$	0.32	3.2	0.25	7.8
$\widehat{\alpha}$	0.55	9.6	0.56	5.8
$\widehat{\rho_0}$	-0.85	-29.3	-0.83	-12.6

Table 4.3: MLE parameter values for dataset 1

Experiment	$\widehat{\mu_\delta}$		$\widehat{\mu_\tau}$		$\widehat{u}$		$\widehat{\beta}$		$\widehat{m}$		$\widehat{\tilde{\sigma}}$		$\widehat{\rho}$		$\widehat{\sigma_\delta}$		$\widehat{\sigma_\tau}$		$\widehat{\alpha}$		$\widehat{\rho_0}$		log likelihood
	mean	t-stat	mean	t-stat	mean	t-stat	mean	t-stat	mean	t-stat	mean	t-stat	mean	t-stat	mean	t-stat	mean	t-stat	mean	t-stat	mean	t-stat	
1	5.24	6.2	0.65	3.5	60.	19.	66.7	15.4	10.	1.4	0.02	1.4	-0.93	-2.	1.45	0.5	0.66	1.2	0.51	15.3	-0.83	-10.2	2546
2	6.03	6.3	0.74	6.2	66.93	5.9	88.76	34.8	4.46	4.2	0.05	3.5	-0.71	-3.8	2.33	3.5	0.56	3.2	0.91	3.5	-0.56	-2.6	2447
3	5.53	108.4	0.59	14.9	60.52	3.6	86.62	13.8	7.31	1.4	0.04	1.1	-0.71	-0.1	0.12	1.6	0.19	1.7	0.52	1.3	-0.8	-9.7	2416
4	3.44	2.	0.84	2.8	61.52	44.8	142.22	20.8	6.49	3.7	0.04	3.4	-0.93	-30.	4.23	8.9	0.74	4.9	0.96	4.7	-1.	-2572.1	2393
5	6.13	6.6	0.5	3.8	60.11	7.9	70.69	9.9	8.76	2.3	0.03	2.2	-0.86	-4.4	2.66	1.6	0.41	1.7	0.51	2.8	-0.84	-12.	2484
Overall	5.78	12.4	0.54	7.	60.14	40.3	94.78	66.9	6.13	9.1	0.04	8.2	-0.49	-1.4	1.63	2.1	0.32	3.2	0.55	9.6	-0.85	-29.3	12259
Likelihood-ratio test statistic $\Lambda(\mathbf{x}) = 54$, the P-Value is 0.144																							

free-flow speed u in dataset 2 is slightly higher than in dataset 1. The maximum acceleration $u\beta$ represents the desired acceleration when is at rest. A quick calculation tells us $u\beta$ in dataset 2 are approximately twice that in dataset 1, indicating the vehicles in dataset 2 have much larger acceleration at rest than those in dataset 1. This result might be explained because in dataset 1 the lead vehicle is controlled to be traveling at a low speed near 30 km/h while leaders in dataset 2 are traveling at higher speeds. Parameters m and $\tilde{\sigma}$ determine the stochastic term in the free-flow branch. We have $m >> 1$ for both two datasets with high t-statistic and this indicates the driver acceleration process is closer to BM according to (3.11).

Parameters τ and δ are the key parameters controlling the congested branch in the

Table 4.4: MLE parameter values for dataset 2

Experiment	$\widehat{\mu_\delta}$		$\widehat{\mu_\tau}$		$\widehat{u}$		$\widehat{\beta}$		$\widehat{m}$		$\widehat{\tilde{\sigma}}$		$\widehat{\rho}$		$\widehat{\sigma_\delta}$		$\widehat{\sigma_\tau}$		$\widehat{\alpha}$		$\widehat{\rho_0}$		log likelihood
	mean	t-stat	mean	t-stat	mean	t-stat	mean	t-stat	mean	t-stat	mean	t-stat	mean	t-stat	mean	t-stat	mean	t-stat	mean	t-stat	mean	t-stat	
1	6.45	7.8	0.56	6.2	65.01	5.8	268.75	2.4	4.44	3.9	0.05	3.2	-0.59	-1.7	3.16	2.5	0.41	0.8	0.57	3.5	-1.	-40.	541
2	6.43	5.5	0.57	3.9	68.88	13.5	203.43	3.9	3.88	1.6	0.05	1.3	-0.62	-1.3	2.54	1.5	0.3	1.6	0.91	3.5	-0.63	-1.6	565
3	4.54	0.9	0.44	1.2	72.64	8.	223.17	3.3	8.84	2.9	0.02	2.6	-0.61	-4.9	2.85	0.4	0.4	13.3	0.56	1.4	-0.64	-1.7	555
4	4.16	1.3	0.42	1.4	65.44	4.4	222.04	1.7	6.37	1.4	0.03	1.3	-0.61	-0.7	2.68	0.6	0.53	1.4	0.54	2.5	-0.82	-6.1	539
5	6.87	3.8	0.5	5.8	80.49	17.2	260.93	4.9	3.46	2.7	0.04	1.9	-0.67	-0.6	2.5	2.	0.35	1.3	0.87	3.1	-0.49	-1.1	550
Overall	4.28	3.8	0.46	6.6	70.	24.9	210.11	7.4	4.21	3.3	0.05	2.7	-0.46	-4.7	4.88	4.6	0.25	7.8	0.56	5.8	-0.83	-12.6	2734
Likelihood-ratio test statistic $\Lambda(\mathbf{x}) = 32$, the P-Value is 0.911																							

Table 4.5: 95%-CIs of m for each car-following experiment

Experiment	Dataset 1		Dataset 2	
	lower bound	upper bound	lower bound	upper bound
1	-4.0	24.0	2.2	6.7
2	-1.8	10.7	-0.8	8.6
3	-2.9	17.5	2.9	14.8
4	3.1	9.9	-2.4	15.1
5	1.3	16.2	1.0	6.0
Overall	4.8	7.5	1.7	6.7

model. Both μ_τ and μ_δ are statistically significant variables for the overall model given the high t-statistic. A typical value of τ would be around 1.2s and that of δ would be 7m. However, the estimated values of τ and δ are much smaller than the typical values, which means the vehicles in both datasets follow its leader closer than that in a typical vehicle platoon. We also find that τ and δ are negatively correlated because the estimated values of ρ are both less than 0 for both datasets and with high t-statistic. As can be seen in Fig.4.3, this is as expected because all feasible (τ,δ) pairs give points on segment AB, and moving along this segment one can see that if τ increases then δ decreases and vice-versa.

Next, we take advantage of the statistical inference tools made possible by MLE to test several interesting hypotheses.

4.3.1 BM or g-BM?

Recall that we obtained high values of m in the estimation. Here we present the 95%-confidence intervals (CIs) of parameter m to test the hypotheses that $m = 1$. As we can see from Table 4.5, $m = 1$ is contained in five of these CIs so we cannot reject the null hypothesis $H_0 : m = 1$ in these five cases. However, the CIs of parameter m in most experiments strongly indicate that $m > 1$. Given the estimated mean value and CIs of m, one can see that the acceleration process is closer to BM than g-BM.

It can be seen that these CIs are fairly wide, and could be narrowed by increasing the sample size.

4.3.2 Are parameters different across datasets?

Here we test parameter homogeneity across datasets by performing likelihood-ratio tests. The methodology is summarized in Appendix B. If we estimate the parameters for each dataset, we have $n_1 = 2 \times 11 = 22$ parameters. If we estimate the parameters for two datasets combined, we have $n_2 = 10$ parameters. We have $\Lambda(\mathbf{x}) = 2\left[\ell(\mathbf{x}; \widehat{\Theta_1}) - \ell(\mathbf{x}, \widehat{\Theta_2})\right] = 2((12259+2734)-13808)) = 2370$, resulting in a p-value of 0.000 with degrees of freedom of 20. This means we cannot accept the null hypothesis that the model with n_1 parameters has similar goodness-of-fit as the model with n_2 parameters, with a significance level of 0.05. In other words, we cannot use unified parameter values for these two datasets. This result might be explained given that the values of β are significantly different across the two datasets.

4.3.3 Are parameters different across experiments?

Here we test parameter homogeneity across car-following experiments by performing likelihood-ratio tests. If we estimate parameter values for each experiment, we have $n_1 = 5 \times 11 = 55$ parameters. If we estimate parameter values for the overall dataset, we have $n_2 = 11$ parameters. Therefore, the degrees of freedom is 44. We have $\Lambda_1(\mathbf{x}) = 54$ (p-value is 0.14) for dataset 1 and $\Lambda_2(\mathbf{x}) = 32$ (p-value is 0.91) for dataset 2. This means we accept the null hypothesis that the model with n_1 parameters has similar goodness-of-fit as the model with n_2 parameters with a significance level of 0.05.

4.3.4 Should grade data be included in the model?

Here we discuss the importance of the grade in the model by testing the significance of parameter α in the model. The estimation results of α were shown in Table 5.3 for the overall models and in 5.2 for the individual experiments, giving the strong indication that this parameter is very significant in all cases. The fact that in all cases $0.5 \leq \hat{\alpha} \leq 1$ reveals that drivers tend to compensate for the upgrade, as explained in section 3.2.2, and that the

Table 4.6: 95%-CIs of α for each car-following experiment

Experiment	Dataset 1		Dataset 2	
	lower bound	upper bound	lower bound	upper bound
1	0.2	1.2	0.3	0.9
2	0.3	1.0	0.4	1.4
3	0.7	1.2	-0.2	1.4
4	0.1	1.2	0.1	1.0
5	0.3	0.9	0.3	1.4
Overall	0.5	0.7	0.4	0.8

current assumption in the literature that $\alpha = 1$ might not be warranted.

To take a closer look, the 95%-confidence intervals (CIs) of α are shown in Table 4.6. It can be seen that the lower bounds are almost all positive, which is reassuring of the fact that upgrades have a detrimental effect on acceleration. That some upper bounds are > 1 is problematic, as it indicates that some drivers might press the gas pedal softer than they would on level terrain, which is counterintuitive. However, this only happens in individual experiments where the reduced sample size might be to blame. In fact, the overall models exhibit upper and lower bounds that are very reasonable. In Chapter 5, we study the impact of roadway grade on vehicle accelerations more in detail with the NGSIM field trajectory data.

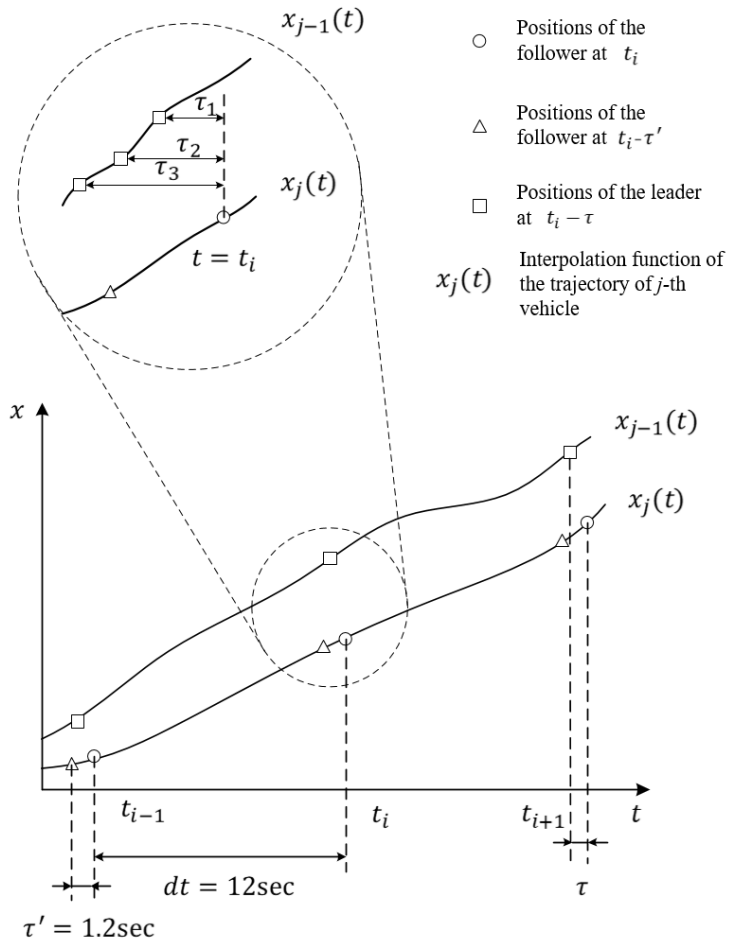

Figure 4.1: Key aspects for maximum likelihood estimation. In this figure, $x_j(t)$ represents the interpolated trajectory of vehicle j. To perform the maximum likelihood estimation, we first choose discrete times t_1, t_2, ..., t_i, ... as our data **x**. Then for each time t_i, the circle represents the observed position of vehicle j at t_i, the triangle represents the observed position of vehicle j at time $(t_i - \Delta t)$, and the square represents the observed positions of vehicle $j-1$ at time $(t_i - \tau)$. The zoomed in figure shows that the positions of squares are not fixed but depend on the value of the random variable τ. The uncertainty of the position of the squares indicates the randomness of parameter τ.

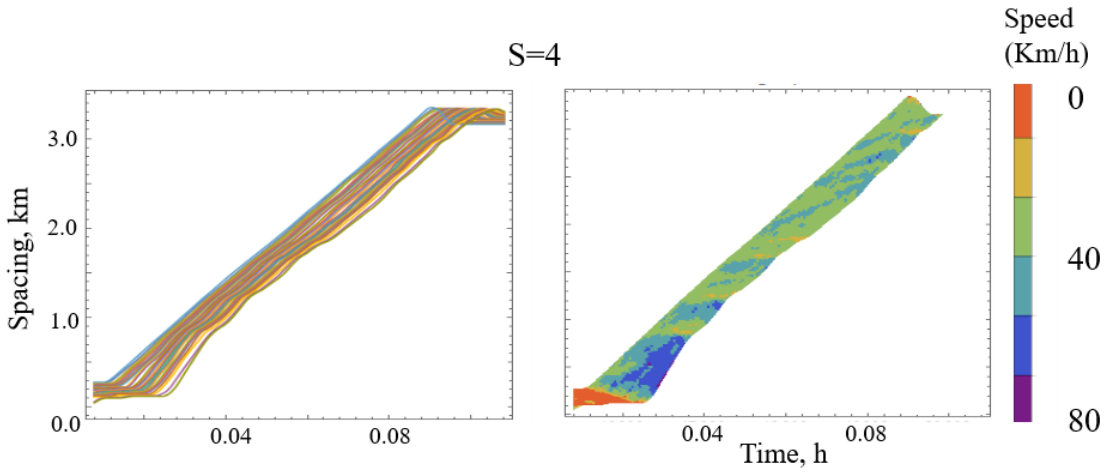

Figure 4.2: An example of the trajectory data collected in a series of car-following experiments (Jiang et al., 2014). S=4 represents the ordinal number of the experiment.

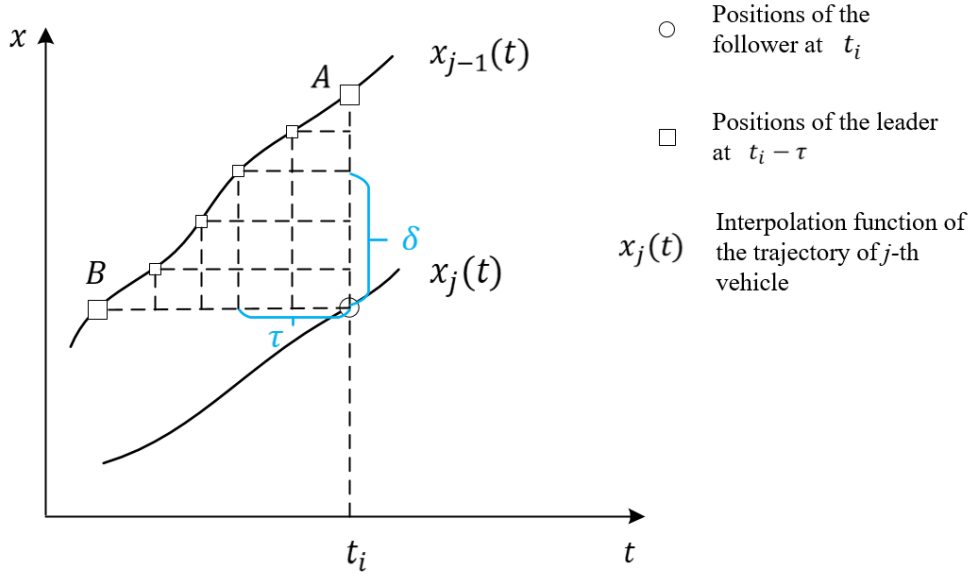

Figure 4.3: This figure helps us understand the negative correlation between τ and δ. Note that all feasible (τ,δ) pairs give points on segment AB. Moving along this segment one can see that if τ increases then δ decreases and vice-versa.

CHAPTER 5
DRIVER REACTIONS TO UPHILL GRADES

This chapter is modified from the following paper:
Xu, T., & Laval, J. A. (2020). Driver Reactions to Uphill Grades: Inference from a Stochastic Car-following Model. Accepted by Transportation Research Record.

5.1 Introduction

Modeling truck performance with roadway grade has a long history. Back in the 1970s, researchers collected speed and weight data for trucks operating on highway grades (Walton and Lee, 1975). Based on the data, they find that trucks experience great acceleration loss on uphill segments and the magnitude of acceleration loss depends on the approach speed. Later, the previous study was extended in Gillespie (1985) and the authors develop a methodology to predict truck acceleration loss on grades for different classes of trucks. These studies for designing climbing lanes only consider the tangent vertical profiles so Yu (2005) focuses on the impact of vertical curvature on truck performance.

The impact of the roadway grade on passenger cars has long been neglected. In the AASHTO report (Officials, 2011), the authors point out that passenger cars can generally negotiate with a grade of 5 percent or less without significant acceleration loss. Modeling work to incorporate roadway grade with vehicle acceleration can be found in Laval (2004) where the author argues that roadway grade is a key ingredient to produce traffic instabilities as observed empirically. In this study, the crawl speed for both passenger cars and trucks is calculated based on the free motion model in McLean (1989). Later, the author successfully replicates periodic oscillations at uphill segments with a car-following model that includes the effects of gravity (Laval, Toth, and Zhou, 2014). They assume that on an upgrade segment, the acceleration due to gravity in the direction of movement, gG, is

subtracted from the acceleration a driver would have imposed on a vehicle on a flat segment. However, this assumption implies that drivers do not press the gas pedal harder on uphill segments, which may not be the actual case, as reported in Bernat Goñi Ros (2016). The study in Bernat Goñi Ros (2016) presents a novel car-following model along with a lane-changing model and find that the influence of roadway grade on vehicle longitudinal motion is more realistic than that predicted by existing models. However, this study did not investigate the difference among vehicle classes and more data are needed for calibration and validation.

The purpose of this chapter is to test the impact of roadway grade on vehicles and find the difference between regular vehicles and trucks. We use the US101 dataset to estimate model parameters. Then statistical tests are performed. In the end, a discussion is provided.

5.2 Dataset

5.2.1 Trajectory data

The Next Generation SIMulation (NGSIM) datasets contain detailed and accurate field data for traffic research and development. In this thesis, we use the US Highway 101 (US 101) dataset, which contains detailed vehicle trajectory data on southbound US 101, also known as the Hollywood Freeway, in Los Angeles, CA, collected on June 15th, 2005. The length of the study segment is approximately 640 meters (2,100 feet). There are five mainline lanes throughout the section and one auxiliary lane through a portion of the corridor between the on-ramp at Ventura Boulevard and the off-ramp at Cahuenga Boulevard (See Fig. 5.1). This dataset provides the precise location of each vehicle area every one-tenth of a second during 45 minutes in the morning rush hour. This period includes uncongested and congested traffic states and the transition between these two states. (James Colyar, 2007)

A time-space diagram of the vehicle trajectories are given the Fig. 5.2. In Fig. 5.2, the distance of each vehicle is drawn against time. the slope of the trajectories indicates the vehicles' speed. The green color represents a free-flow state and the red color represents a

congestion state. We can observe periodic traffic congestion from the visualization of the dataset.

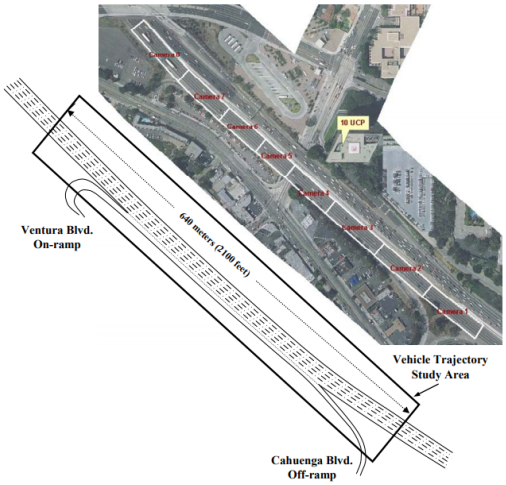

Figure 5.1: The site for US 101 data collection. (James Colyar, 2007)

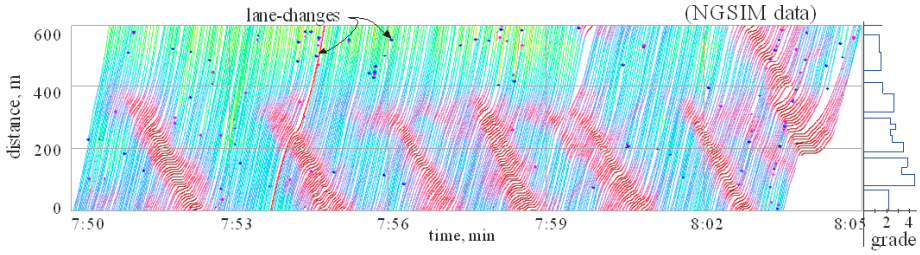

Figure 5.2: A visualization of the US 101 dataset. (Laval, Toth, and Zhou, 2014)

5.2.2 Grade Data

Modeling vehicle acceleration with roadway grade requires high-accuracy grade data. We used Google Earth to obtain the elevation profile along the study corridor, and derived the grade data using a second-order interpolation; see Fig. 5.2. One can see that the segment has an average upgrade around 2%, with larger grades near the beginning of the segment. Also note in this figure that the oscillations are not caused by the grade, but by a crew of workers located near the 400m mark, as reported in Chen et al. (2014).

5.2.3 Sample size

We use the data from five mainline lanes during the period of 7:50 to 8:05. Lane 1 is the leftmost lane and Lane 5 is the rightmost lane. We removed vehicles that performed lane changing during the period. Remaining vehicles are used in the parameter estimation.

For vehicle i, if it entered the study corridor at t_{enter} and leaved the study corridor at t_{leave}, we can generate 50 uniform random variables $t_1, t_2, ...t_{50}$ between t_{enter} and t_{leave} as our sample time stamp, such that we will have 50 sample points for one vehicle. In Table 5.1, we summarize the sample size for the parameter estimation for each lane, each vehicle class.

Table 5.1: Number of sample points for parameter estimation (50 sample points each vehicle)

Lane	Motorcycles	Cars	Trucks	Total
1	300	13,450	50	13,800
2	150	10,750	0	10,900
3	0	9,150	200	9,350
4	100	6,650	100	6,850
5	50	5,800	100	5,950
Total	600	45,800	450	46,850

Data for each vehicle are randomly divided into five groups, each with a size of 10. Four groups (80% of the data) are used as training data and the remainder group (20% of the data) is used as validation data (James et al., 2013). For the four groups of training data, we use maximum likelihood to estimate parameter values, then we use the results from the four training groups separately on the validation data. The estimation result that gives the biggest log-likelihood value on the validation data is selected as the final estimation result.

5.3 Estimation results and statistical inference

5.3.1 Homogeneity among vehicle classes

Table 5.2 summarizes the estimation results for different vehicle classes. The first three rows in this table give the results for models estimated with data only involving the corre-

sponding vehicle class as the follower, while the last row corresponds to a model estimated with the overall data set, without distinguishing vehicle types. It can be seen that we cannot reject the hypothesis that **model parameters are different for different vehicle classes**.

Fig. 5.3 shows the 95% confidence intervals (CI) of the estimated values of β, $\tilde{\sigma}$, and α, which can be used to test the equality of these parameter values across vehicle types simply by observing if these intervals overlap or not. The half width of the error bars equals the standard deviation of the corresponding parameter. It can be seen that the intervals for motorcycles and trucks are much wider than that of passenger cars, which is the consequence of the sample sizes available to us. We can see that the main differences between cars and trucks are in the following parameters:

1. β: Trucks have a higher value than cars for this parameter, which implies that trucks tended to accelerate more aggressively than cars.

2. $\tilde{\sigma}$: Trucks exhibit a higher value, which implies that truck drivers have a larger variation when accelerating, which might be related to the number of gears that truck engines use when accelerating.

3. α: Again, this parameter is larger for trucks, and the implications are discussed below.

Of the findings above, we believe that the last one is the most significant in terms of the impact of current modeling frameworks. Recall that the parameter α represents the proportion of the gravitational forces due to upgrades that are absorbed by the motion of the vehicle, and has been assumed to be 100% until now. Our estimation results show that for regular vehicles, $\widehat{\alpha} = 0.51$ and for heavy vehicles $\widehat{\alpha} = 0.96$. This means that car drivers demand more power from the engine to compensate for gravitational forces such that their actual motion loses only about 50% of what they would have lost had they continued to drive without compensating.

In the case of trucks, we can see that the standard assumption in the literature is appropriate, as they seem to compensate for only 4% of the gravitational force. For motorcycles, $\widehat{\alpha} = 0.79$ is significantly higher than for regular vehicles, which may seem counter-intuitive since motorcycles have a much higher power-to-weight ratio. A possible explanation is that, precisely for this reason, motorcycles are almost always in the car-following model with little slack for extra acceleration.

Table 5.2: MLE parameter values for different vehicle types

Vehicle Class	$\widehat{\mu_\delta}$		$\widehat{\mu_\tau}$		$\widehat{u}$		$\widehat{\beta}$		$\widehat{m}$		$\widehat{\tilde{\sigma}}$		$\widehat{\rho}$		$\widehat{\sigma_\delta}$		$\widehat{\sigma_\tau}$		$\widehat{\alpha}$		$\widehat{\rho_0}$		log likelihood
	mean	t-stat	mean	t-stat	mean	t-stat	mean	t-stat	mean	t-stat	mean	t-stat	mean	t-stat	mean	t-stat	mean	t-stat	mean	t-stat	mean	t-stat	
motorcycle	5.59	3.4	0.4	3.7	70.09	3.	90.07	2.6	4.06	3.1	0.23	3.	-0.48	-5.2	1.69	9.8	1.	15.2	0.79	1.8	0.32	2.8	549
regular vehicle	7.76	65.8	0.4	25.3	76.92	144.7	97.58	336.3	4.1	74.9	0.11	97.3	0.17	3.3	1.43	21.1	0.35	30.6	0.51	87.	-0.1	-5.4	49,638
heavy vehicle	4.54	1.4	0.43	1.4	73.45	17.5	116.26	369.1	3.38	44.2	0.22	12.6	-0.48	-1.5	1.46	3.	0.57	3.9	0.96	14.	-0.39	-3.	488
Overall	8.06	72.8	0.4	28.	85.23	21.9	118.27	36.3	4.48	34.	0.07	36.5	0.26	11.7	1.89	27.8	0.22	29.7	0.61	44.3	-0.14	-9.5	50,419
Likelihood-ratio test statistic $\Lambda(\mathbf{x}) = 512$, the P-Value is 0.000																							

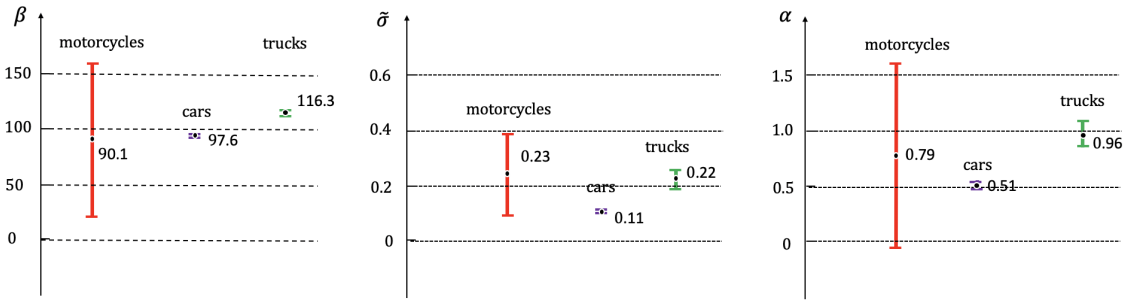

Figure 5.3: 95% Confidence Intervals of the estimated values of β, $\tilde{\sigma}$, and α for different vehicle classes.

The estimation of ρ_0 indicates that the correlation between Y and Z for this dataset is very weak ($|\widehat{\bar{\rho}_0}| < 0.4$), especially for cars ($|\widehat{\bar{\rho}_0}| = 0.1$). For the rest of this chapter, we assume that Y and Z are independent, which help us to reduce one parameter ρ_0 in the model.

5.3.2 Homogeneity among lanes

Table 5.3 and 5.4 includes the estimation results for separate lanes, and Fig. 5.4 shows the CIs of estimated parameter values. It can be seen from Table 5.3 that the p-value of the likelihood ratio test is larger than 0.01, which indicates that the additional complexity of

the model for the three middle lanes is not justified with a significance level of 0.01. As we can see from the results, the only slight difference between the vehicles on three middle lanes are in τ and the way they deal with upgrades, which is implied by the different values of α.

These results accord well with Fig. 5.4, where we can verify that most parameter values are the same for Lane 2, 3 and 4 since their confidence intervals overlap. It can be seen that the main difference between the middle three lanes compared to the median and shoulder lane resides in the jam spacing, where in the middle lanes is substantially higher. This can be explained by the higher proportion of trucks and the smaller proportion of motorcycles usually observed in the middle lanes. Another significant difference among lanes can be seen in the middle part of the figure, where middle lanes exhibit a significantly smaller diffusion coefficient $\tilde{\sigma}$. This implies that these lanes driver acceleration tends to have less variation than on other lanes. Unfortunately, the explanation that is due to the higher proportion of trucks does not apply here since we have seen the trucks tend to have a higher $\tilde{\sigma}$. Another possibility is that these lanes exhibit less stop-and-go activity compared to the other lanes and therefore provide fewer opportunities for accelerating in free flow, which is when $\tilde{\sigma}$ can be estimated.

Table 5.3: MLE parameter values for different lanes

Lane ID	$\widehat{\mu_\delta}$		$\widehat{\mu_\tau}$		$\widehat{u}$		$\widehat{\beta}$		$\widehat{m}$		$\widehat{\tilde{\sigma}}$		$\widehat{\rho}$		$\widehat{\sigma_\delta}$		$\widehat{\sigma_\tau}$		$\widehat{\alpha}$		log likelihood
	mean	t-stat	mean	t-stat	mean	t-stat	mean	t-stat	mean	t-stat	mean	t-stat	mean	t-stat	mean	t-stat	mean	t-stat	mean	t-stat	
1	4.39	5.5	0.4	3.9	74.0	16.9	92.8	87.6	4.04	14.2	0.16	29.7	- 0.29	- 6.3	2.	42.	0.67	44.5	0.82	17.1	14,429
2	7.83	36.4	0.4	11.4	76.9	10.2	97.6	5.4	5.08	28.6	0.07	10.4	- 0.03	- 0.1	1.07	6.6	0.53	16.7	0.47	2.6	12,062
3	7.82	19.5	0.93	41.2	76.9	58.	97.6	19.9	5.08	10.2	0.07	10.5	- 0.03	- 1.5	1.07	7.9	0.51	57.1	0.45	8.4	10,003
4	7.78	84.6	0.4	30.7	76.9	14.6	97.6	4.6	5.08	6.6	0.07	5.6	- 0.05	- 0.2	1.08	2.7	0.18	4.6	0.26	2.1	7,844
5	5.35	6.5	0.4	6.7	70.0	28.9	90.0	91.9	2.93	124.5	0.25	18.	- 0.55	- 0.8	1.79	3.1	0.57	5.1	0.71	17.2	6,430
Overall	7.06	48.4	0.4	22.5	80.9	53.7	107.4	101.1	4.82	31.6	0.08	30.9	0.21	13.3	1.9	15.4	0.31	19.3	0.53	58.3	50,389
Likelihood-ratio test statistic $\Lambda(\mathbf{x}) = 758$, the P-Value is 0.000																					

Table 5.4: MLE parameter values for three middle lanes

Lane ID	$\widehat{\mu_\delta}$		$\widehat{\mu_\tau}$		$\widehat{u}$		$\widehat{\beta}$		$\widehat{m}$		$\widehat{\tilde{\sigma}}$		$\widehat{\rho}$		$\widehat{\sigma_\delta}$		$\widehat{\sigma_\tau}$		$\widehat{\alpha}$		log likelihood
	mean	t-stat	mean	t-stat	mean	t-stat	mean	t-stat	mean	t-stat	mean	t-stat	mean	t-stat	mean	t-stat	mean	t-stat	mean	t-stat	
2	7.83	36.4	0.4	11.4	76.9	10.2	97.6	5.4	5.08	28.6	0.07	10.4	- 0.03	- 0.1	1.07	6.6	0.53	16.7	0.47	2.6	12,062
3	7.82	19.5	0.93	41.2	76.9	58.	97.6	19.9	5.08	10.2	0.07	10.5	- 0.03	- 1.5	1.07	7.9	0.51	57.1	0.45	8.4	10,003
4	7.78	84.6	0.4	30.7	76.9	14.6	97.6	4.6	5.08	6.6	0.07	5.6	- 0.05	- 0.2	1.08	2.7	0.18	4.6	0.26	2.1	7,844
three lanes together	7.81	44.7	0.7	35.1	76.91	207.7	97.57	3102.8	5.08	468.4	0.07	154.4	-0.03	-0.2	1.07	10.	0.51	25.3	0.43	209.5	29,891
Likelihood-ratio test statistic $\Lambda(\mathbf{x}) = 36$, the P-Value is 0.015																					

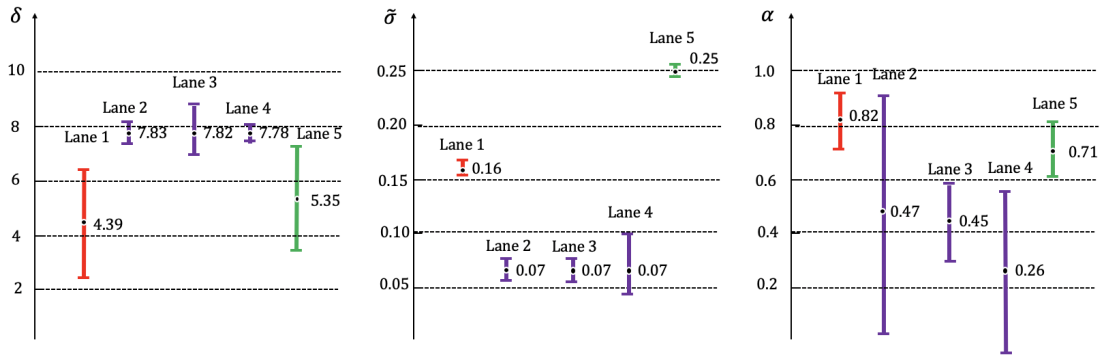

Figure 5.4: 95% Confidence Intervals of the estimated values of δ, $\tilde{\sigma}$, and α based on data from five lanes.

5.4 Discussion

In this chapter, we have used statistical inference based on the maximum likelihood estimation of a two-regime stochastic car-following model using NGSIM data. Our main finding pertains to the impact of uphill grades on the acceleration drivers choose to impose on their vehicles. We found evidence that the current assumption in the existing free-motion acceleration models (α=1) does not apply to car drivers, who tend to overcome half of the gravitational effects by using more engine power. Truck drivers only compensate for 4% of the loss, possibly because of limited engine power; however, more truck data is needed to increase our confidence in these results.

This finding is important for current applications because it means not only that current models are severely overestimating the operational impacts that uphill grades have on regular vehicles, but also underestimating their environmental impacts. For example, the VISSIM model (PTV, 2015) uses $\alpha = 1$ and emissions calculations do not consider driver compensation as found here. We calculated the emission and energy consumption for a Honda Pilot 2004 with the MOVES model (USEPA, 2019). In the simulation experiment, the car is accelerating from a complete stop on a 5% uphill segment. The free-flow speed $u = 100$ km/h and the inverse relaxation time $\beta = 0.07\text{s}^{-1}$. The results are shown in

Table 5.5: CO2 emission and energy consumption comparison for different values of α

Distance Traveled (m)	CO2 Emission (gram)			Energy Consumption (kJ)		
	$\alpha = 0$	$\alpha = 0.5$	$\alpha = 1$	$\alpha = 0$	$\alpha = 0.5$	$\alpha = 1$
500	287.7	254	210.1	4003	3535	2924
1000	484.5	418.9	366.5	6742	5816	5387
2000	857.4	744.7	673.1	11931	10362	9366

Table 5.5. The results showing that the current free-motion acceleration models ($\alpha = 1$) are underestimating the environmental impacts by 10 % in an acceleration process at uphill grades.

We also find that car-following model parameters related to the acceleration process are significantly different among shoulder, median and middle lanes. We saw that this can be partially explained by the greater proportion of trucks the middle lanes, but more data is needed to complete the analysis.

Parameters δ and τ are the key parameters controlling the congested branch in the model. In the model, we assume they follow a bi-variate normal distribution, as suggested in Ahn, Cassidy, and Laval (2003). From the estimation results, we find the correlation between δ and τ is zero because the parameter ρ has a low t-statistic. This result is consistent with Ahn, Cassidy, and Laval (2003) and indicates that traffic waves propagate as a random walk.

CHAPTER 6
CONCLUSIONS AND FUTURE WORK

This thesis presents a generalization of the works conducted by Laval, Toth, and Zhou (2014) and Yuan et al. (2018), each producing different features of traffic instabilities by adding a dimensionless parameter m. The model in this thesis regulates the type of driver free acceleration error on a scale from Brownian motion (BM, $m >> 1$) to geometric Brownian (g-BM, $m = 1$) acceleration processes. When $m >> 1$, this model captures proper periodicity of traffic oscillations at bottlenecks and demonstrates that human factor alone can cause traffic instabilities. This result accords with the work conducted by Laval, Toth, and Zhou (2014) well. When $m = 1$, this model is equivalent to the model in (Yuan et al., 2018), which successfully reproduces the speed-capacity relationship observed empirically. The main contribution of the thesis is that if we select a proper value for m ($m \approx 1.2$) in the proposed model, the model can explain both the oscillation and speed-capacity relationship.

The model is so flexible that it can be combined with lane changing models (Laval and Daganzo, 2006) in microscopic-level traffic simulation (Chilukuri, Laval, and Guin, 2014). The interference of traffic signals could also be accomplished by simply changing the desired vehicle speed in the model. That is, with this empirically validated model, we can simulate almost every scenario in traffic streams without too much additional work.

Another major contribution of this thesis lies in the estimation of the parameter values. The probability density function of the vehicle position in the proposed model turns our to be analytical such that model parameters can be easily estimated with MLE without an additive error term. This idea could provide reference for other studies outside of the traffic flow area, that it is beneficiary to model a process controlled by two random processes such that model parameters can be estimated with MLE without an additive error term. In

addition, we combine cross-validation method with the maximum likelihood estimation by dividing the data into training data and validation data. Using validation data could help us avoid over-fitting such that the results are more reliable.

The framework of the two-regime stochastic car-following models also allows us to perform statistical tests on various cases, such as the homogeneity of the driver/vehicle population and the statistical significance of the impacts of roadway geometry or weather. With the help with this tool and the NGSIM datasets, we find that current free-motion models are not only severely overestimating the operational impacts that uphill grades have on regular vehicles, but also underestimating their environmental impacts.

The study in this thesis also has the following limitations and these limitations should be the focus of future research. One of the limitations of this study is that we assumed a linear model between vehicle acceleration and roadway grade. However, the impact of roadway grade on the vehicle acceleration may be nonlinear. In future research, we will estimate different functional forms for $f(G)$ for different values of G, including downgrades. The function $f(G)$ could be piece-wise linear, polynomial or more complex. Some options for $f(G)$ is shown in Fig. 6.1. We will carefully compare different models and choose the best fit of $f(G)$ based on various datasets. We also plan to investigate the impact of gradient sequence on vehicle acceleration behavior as an extension of the current work. Also, increasing the sample size of trucks and motorcycles is helpful in the future. The data gives a strong indication that the value of m is larger than 3, which means that the acceleration processes of drivers are closer to BM than g-BM. Unfortunately, it also means that the model loses its ability to explain the speed-capacity relationship. This could be explained by a number of reasons and should be the focus of future research. One possibility is that m might be different for different traffic states and vehicles may have values of $m \approx 1$ when the queue is discharging. This assumption needs to be tested in the future.

Finally, we conjecture that the proposed model is well suited to represent the operations of automated vehicles. To see this, (i) manufacturers strive to implement car-following

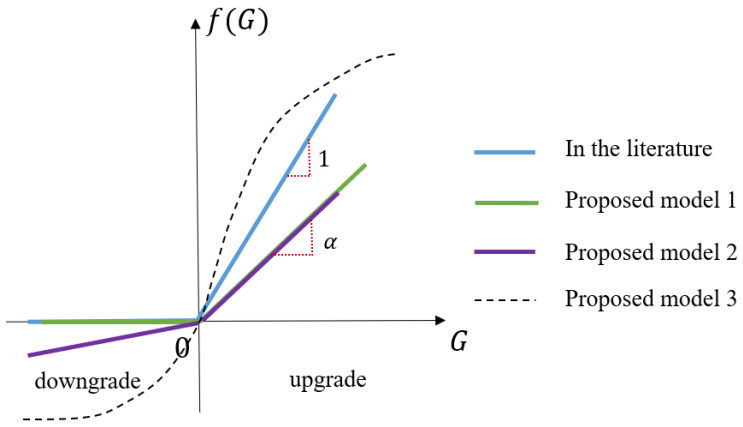

Figure 6.1: Possible relationships between $f(G)$ and roadway grade in the model.

behaviors that are similar to human drivers, (ii) AVs also have to overcome gravity and are subject to limited engine power, which implies that a bounded acceleration model is in order for AVs as well, (iii) existing single-regime models for AV's can be incorporated in the congestion term in (3.9) by expressing them in terms of vehicle positions rather than acceleration. In the future, the impacts of AVs on this model might be tested with statistical inference methods. Preliminary results indicates that there is a significant difference among the current commercial AV models.

Appendix A

$$
\begin{aligned}
\text{Var}\left[\xi(t)\right] \;=\; & \frac{1}{\beta^2(2\beta^2-3\beta\sigma^2+\sigma^4)^2}e^{-t(\beta+\sigma^2)}(e^{t(-\beta+\sigma^2)}(-1+e^{\beta t})((v_c-v_0)^2 \\
& +e^{\beta t}((5+2(-4+m)m)v_c^2+2(-1+2m)v_cv_0-v_0^2))\sigma^8+4\beta^4e^{t(-\beta+\sigma^2)} \\
& ((-1+e^{t\sigma^2})(v_c-v_0)^2+e^{2\beta t}(-1+m)^2tv_c^2\sigma^2+4e^{\beta t}(-1+m)tv_c(-v_c+v_0)\sigma^2 \\
& 2\beta^3e^{-\beta t}\sigma^2(6e^{t\sigma^2}(v_c-v_0)^2-e^{2t\sigma^2}((2+m)v_c-3v_0)(mv_c-v_0)+4e^{t(\beta+\sigma^2)} \\
& (-2mv_c^2+m^2v_c^2+2v_cv_0-v_0^2+4(-1+m)tv_c(v_c-v_0)\sigma^2)+e^{t(2\beta+\sigma^2)} \\
& ((2-4m)v_cv_0+v_0^2+v_c^2(-6+10m-3m^2-5(-1+m)^2t\sigma^2))) \\
& 2\beta\sigma^6(3e^{t(-\beta+\sigma^2)}(v_c-v_0)^2+e^{t\sigma^2}(-5v_0^2+2v_cv_0(1+4m-(-1+m)t\sigma^2) \\
& +v_c^2(8-18m+5m^2+2(-1+m)t\sigma^2))-e^{t(\beta+\sigma^2)}(4(-1+2m)v_cv_0-2v_0^2 \\
& +v_c^2(11-18m+5m^2+(-1+m)^2t\sigma^2)))+\beta^2e^{-\beta t}\sigma^4(-13e^{t\sigma^2}(\beta-v_0)^2 \\
& +2e^{2t\sigma^2}(-mv_c+v_0)^2-4e^{t(\beta+\sigma^2)}(4e^{t(\beta+\sigma^2)}(4(1+(-3+m)m)v_c^2 \\
& +4(1+m)v_cv_0-4v_0^2+5(-1+m)tv_c(v_c-v_0)\sigma^2)+e^{t(2\beta+\sigma^2)}(10(-1+2m)v_cv_0 \\
& -5v_0^2+v_c^2(29-48m+14m^2+8(-1+m)^2t\sigma^2))))
\end{aligned}
$$

Appendix B: MLE

We use MLE for the estimation of parameters in our proposed model. In a nutshell, MLE consists in finding the value of Θ that maximize the log-likelihood $\ell(\mathbf{x};\Theta) = \sum_{i,j} \ln[f(x_j(t_i);\Theta)]$.

This gives the MLE estimator of Θ, and is denoted $\widehat{\Theta}$. MLE is appealing because for large samples one can perform statistical inference analysis to answer important questions in car-following behavior. For large samples the distribution of MLE estimators tends to the multivariate normal distribution, i.e. $\widehat{\Theta} \overset{dist}{\rightarrow} N(\Theta, \text{SD}\,[\Theta])$, where the covariance matrix SD $[\Theta]$ can be approximated using the Cramer-Rao lower bound: SD $[\Theta] \approx J(\widehat{\Theta})^{-1}$ where $J(\Theta) = -\frac{\partial^2 \ell(\Theta)}{\partial \Theta^2}$ is the observed Fisher's information for the sample. This important result allows us to use the statistical inference toolbox, in particular confidence intervals and hypothesis testing.

A likelihood-ratio test can be used to compare the goodness-of-fit of different model specifications. For example, if we have two models with number of parameters n_1 and n_2 ($n_1 > n_2$), respectively, the likelihood-ratio test statistic is $\Lambda(\mathbf{x}) = 2\left[\ell(\mathbf{x};\widehat{\Theta_1}) - \ell(\mathbf{x},\widehat{\Theta_2})\right]$ which follows a chi-square distribution with $n_1 - n_2$ degrees of freedom. For more information on MLE, the reader is referred to Hubbert (2012).

REFERENCES

Bilbao-Ubillos, Javier (2008). "The costs of urban congestion: Estimation of welfare losses arising from congestion on cross-town link roads". In: *Transportation Research Part A: Policy and Practice* 42.8, pp. 1098 –1108.

Ahmed, Kazi Iftekhar (1999). "Modeling Driver's Acceleration and Lane Changing Behavior". PhD thesis.

Kilpeläinen, Markku and Heikki Summala (2007). "Effects of weather and weather forecasts on driver behaviour". In: *Transportation Research Part F: Traffic Psychology and Behaviour* 10.4, pp. 288 –299.

Li, Hu et al. (2007). "Impact of Traffic Conditions and Road Geometry on Real World Urban Emissions Using a SI Car". In: *SAE World Congress & Exhibition*. SAE International.

Chen, Danjue et al. (2012b). "Microscopic traffic hysteresis in traffic oscillations: A behavioral perspective". In: *Transportation Research Part B* 46.10, pp. 1440 –1453.

Laval, J. A. and L Leclercq (2010). "A mechanism to describe the formation and propagation of stop-and-go waves in congested freeway traffic". In: *Philosophical Transactions of The Royal Society A* 368.1928, pp. 4519–4541.

Yuan, Kai et al. (2018). "A Geometric Brownian Motion car-following model: Towards a better understanding of capacity drop". In: *Transportmetrica B*. Forthcoming.

Laval, Jorge A., Christopher S. Toth, and Yi Zhou (2014). "A parsimonious model for the formation of oscillations in car-following models". In: *Transportation Research Part B: Methodological* 70, pp. 228 –238.

Jiang, Rui et al. (Apr. 2014). "Traffic Experiment Reveals the Nature of Car-Following". In: *PLOS ONE* 9.4, pp. 1–9.

James Colyar, John Halkias (2007). *US Highway 101 Dataset Fact Sheet*. Tech. rep.

Xu, Long et al. (2013). "Comparative Study of Fuel Consumption Estimations for Car-following Models Based on VSP Distributions". In: *Journal of Highway and Transportation Research and Development* 30.6, pp. 126–147.

Pipes, Louis A (1953). "An operational analysis of traffic dynamics". In: *Journal of applied physics* 24.3, pp. 274–281.

Chandler, Robert E., Robert Herman, and Elliott W. Montroll (1958). "Traffic Dynamics: Studies in Car Following". In: *Operations Research* 6.2, pp. 165–184. eprint: https://doi.org/10.1287/opre.6.2.165.

Herman, Robert et al. (1959). "Traffic Dynamics: Analysis of Stability in Car Following". In: *Operations Research* 7.1, pp. 86–106. eprint: https://doi.org/10.1287/opre.7.1.86.

Gazis, Denos C, Robert Herman, and Renfrey B Potts (1959). "Car-following theory of steady-state traffic flow". In: *Operations research* 7.4, pp. 499–505.

Gazis, Denos C, Robert Herman, and Richard W Rothery (1961). "Nonlinear Follow-the-Leader Models of Traffic Flow". In: *Operations Research* 9.4, pp. 545–567.

Michaels, R.M (1963). "Perceptual Factors in Car Following". In: *Proceedings of the 2nd International Symposium on the Theory of Road Traffic Flow*, pp. 44–59.

Bando, M. et al. (1995). "Dynamical Model of Traffic Congestion and Numerical Simulation". In: *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics* 51.2, pp. 1035–1042.

Treiber, Martin, Ansgar Hennecke, and Dirk Helbing (2000). "Congested Traffic States in Empirical Observations and Microscopic Simulations". In: pp. 1805–1824.

Newell, G F (2002). "A Simplified Car-Following Theory : a Lower Order Model". In: *Transportation Research Part B* 36.3, pp. 195–205.

Dougherty, Mark (1995). "A review of neural networks applied to transport". In: *Transportation Research Part C: Emerging Technologies* 3.4, pp. 247–260.

Newell, G F (1961). "Non-Linear Effects in the Dynamics of Car Following". In: *Opns. Res.* 2.9, pp. 209–229.

Wiedemann, Rainer and U Reiter (1992). "Microscopic Traffic Simulation: the Simulation System MISSION, Background and Actual State". In: *Project ICARUS (V1052) Final Report* 2.

Fritzsche, Hans-thomas and Daimler-benz Ag (1994). "A Model for Traffic Simulation". In: *Traffic Engineering & Control* 35.5, pp. 317–321.

Gipps, P.G. (1981). "A Behavioural Car-following Model for Computer Simulation". In: *Transportation Research Part B: Methodological* 15.2, pp. 105–111.

Kesting, Arne, Martin Treiber, and Dirk Helbing (2010). "Enhanced intelligent driver model to access the impact of driving strategies on traffic capacity". In: *Philosophical Trans-*

actions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences 368.1928, pp. 4585–4605. eprint: http://rsta.royalsocietypublishing.org/content/368/1928/4585.full.pdf.

Helbing, Dirk and Benno Tilch (1998). "Generalized force model of traffic dynamics". In: *Physical review E* 58.1, p. 133.

Jiang, Rui, Qingsong Wu, and Zuojin Zhu (2001). "Full Velocity Difference Model for a Car-following Theory". In: *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics* 64.1, pp. 1–4.

Mauch, M and M J Cassidy (2002). "Freeway Traffic Oscillations: Observations and Predictions". In: *15th Int. Symp. on Transportation and Traffic Theory*. Ed. by M.A.P. Taylor. Pergamon-Elsevier, Oxford,U.K.

Laval, J A (2005). "Linking synchronized flow and kinematic wave theory". In: *Traffic and Granular Flow '05*. Ed. by T. Schadschneider A.and Poschel et al. Springer, pp. 521–526.

Laval, J A and C F Daganzo (2006). "Lane-changing in traffic streams". In: *Transportation Research Part B* 40.3, pp. 251–264.

Ahn, S and M Cassidy (2007). "Freeway Traffic Oscillations and Vehicle Lane-Change Manoeuvres". In: *17th International Symposium on Transportation and Traffic Theory*. Ed. by B Heydecker, M Bell, and R Allsop. Elsevier, New York, pp. 691–710.

Zheng, Zuduo et al. (2011). "Applications of wavelet transform for analysis of freeway traffic: Bottlenecks, transient traffic, and traffic oscillations". In: *Transportation Research Part B: Methodological* 45.2, pp. 372 –384.

Sugiyama, Yuki et al. (2008). "Traffic jams without bottlenecks-experimental evidence for the physical mechanism of the formation of a jam". In: *New Journal of Physics* 10, p. 033001.

Treiber, Martin, Ansgar Hennecke, and Dirk Helbing (1999). "Derivation, Properties, and Simulation of a Gas-Kinetic-Based, Non-Local Traffic Model". In: *Physical Review E* 59 (1), p. 239.

Wilson, R. E. (2008). "Mechanisms for spatio-temporal pattern formation in highway traffic models". In: *Phil. Trans. R. Soc. A* 1872, pp. 2017–2032.

Wilson, R.E. and J.A. Ward (2011). "Car-following models: fifty years of linear stability analysis - a mathematical perspective". In: *Transportation Planning and Technology* 34.1, pp. 3–18. eprint: http://dx.doi.org/10.1080/03081060.2011.530826.

Treiber, Martin and Arne Kesting (2012). "Validation of traffic flow models with respect to the spatiotemporal evolution of congested traffic patterns". In: *Transportation Research Part C: Emerging Technologies* 21.1, pp. 31 –41.

— (2017). "The Intelligent Driver Model with Stochasticity -New Insights Into Traffic Flow Oscillations". In: *Transportation Research Procedia* 23.Supplement C. Papers Selected for the 22nd International Symposium on Transportation and Traffic Theory Chicago, Illinois, USA, 24-26 July, 2017., pp. 174 –187.

Chen, Danjue et al. (2012a). "A behavioral car-following model that captures traffic oscillations". In: *Transportation Research Part B* 46.6, pp. 744 –761.

Delpiano, Rafael et al. (2015). "The kinematic wave model with finite decelerations: A social force car-following model approximation". In: *Transportation Research Part B: Methodological* 71, pp. 182 –193.

Hall, F L and K Agyemang-Duah (1991). "Freeway Capacity Drop and the Definition of Capacity". In: *Transportation Research Record* 1320, pp. 91–98.

Coifman, Benjamin and Seoungbum Kim (2011). "Extended bottlenecks, the fundamental relationship, and capacity drop on freeways". In: *Transportation Research Part A: Policy and Practice* 45.9. Select Papers from the 19th International Symposium on Transportation and Traffic Theory (ISTTT), pp. 980 –991.

Leclercq, Ludovic, Jorge A. Laval, and Nicolas Chiabaut (2011). "Capacity Drops at Merges: an endogenous model". In: *Procedia - Social and Behavioral Sciences* 17.0. Papers selected for the 19th International Symposium on Transportation and Traffic Theory, pp. 12 –26.

Yuan, Kai, Victor L. Knoop, and Serge P. Hoogendoorn (2015). "Capacity Drop: Relationship Between Speed in Congestion and the Queue Discharge Rate". In: *Transportation Research Record: Journal of the Transportation Research Board* 2491, pp. 72–80. eprint: `https://doi.org/10.3141/2491-08`.

Chen, Danjue and Soyoung Ahn (2018). "Capacity-drop at extended bottlenecks: Merge, diverge, and weave". In: *Transportation Research Part B: Methodological* 108, pp. 1 –20.

Jin, Wen-Long (2017). "A first-order behavioral model of capacity drop". In: *Transportation Research Part B: Methodological* 105, pp. 438 –457.

Kontorinaki, Maria et al. (2017). "First-order traffic flow models incorporating capacity drop: Overview and real-data validation". In: *Transportation Research Part B: Methodological* 106, pp. 52 –75.

Oh, Simon and Hwasoo Yeo (2015). "Impact of stop-and-go waves and lane changes on discharge rate in recovery flow". In: *Transportation Research Part B: Methodological* 77, pp. 88 –102.

Wong, GCK and SC Wong (2002). "A multi-class traffic flow model : an extension of LWR model with heterogeneous drivers". In: *Transportation Research Part A* 36, pp. 827–841.

Xu, Tu and Jorge Laval (2018). "Parameter Estimation of a Stochastic Microscopic Car-Following Model". In: *97th Annual Meeting of Transportation Research Board*. Washington. ISBN: 4048942360.

Tian, Junfang et al. (2016). "Empirical analysis and simulation of the concave growth pattern of traffic oscillations". In: *Transportation Research Part B: Methodological* 93, pp. 338 –354.

Lighthill, M J and GB Whitham (1955). "On kinematic waves. I Flow movement in long rivers. II A theory of traffic flow on long crowded roads". In: *Proceedings of the Royal Society of London* 229.A, pp. 281–345.

Zheng, Zuduo, Soyoung Ahn, and Christopher M. Monsere (2010). "Impact of traffic oscillations on freeway crash occurrences". In: *Accident Analysis & Prevention* 42.2, pp. 626 –636.

Leclercq, L et al. (2007). "Relaxation Phenomenon After Changing Lanes: Experimental Validation with NGSIM Data Set". In: *Transportation Research Record* 1999, pp. 79–85.

Leclercq, L and J A Laval (2007). "A multiclass car-following rule based on the "Lighthill-Whitham-Richards" model." In: *Traffic and Granular Flow '07 (Forthcoming)*. Ed. by C Appert-Rolland et al. Springer.

Laval, J A and L Leclercq (2008). "Microscopic modeling of the relaxation phenomenon using a macroscopic lane-changing model". In: *Transportation Research Part B* 42.6, pp. 511–522.

Treiterer, J and J A Myers (1974). "The Hysteresis Phenomenon in Traffic Flow". In: *6th Int. Symp. on Transportation and Traffic Theory*. Ed. by D. J. Buckley. A.H. and A.W. Reed, London, pp. 13–38.

Ahn, Soyoung, Sravani Vadlamani, and Jorge Laval (2013). "A method to account for non-steady state conditions in measuring traffic hysteresis". In: *Transportation Research Part C* 34, pp. 138–147.

Laval, Jorge A. (2011). "Hysteresis in traffic flow revisited: An improved measurement method". In: *Transportation Research Part B: Methodological* 45.2, pp. 385 –391.

Nadarajah, S. and S. Kotz (2008). "Exact Distribution of the Max/Min of Two Gaussian Random Variables". In: *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, pp. 210–212.

Ahn, S, M Cassidy, and J A Laval (2003). "Verification of a simplified car-following theory". In: *Transportation Research Part B* 38.5, pp. 431–440.

Li, Xiaopeng and Yanfeng Ouyang (2011). "Characterization of Traffic Oscillation Propagation under Nonlinear Car-Following Laws". In: *Procedia - Social and Behavioral Sciences* 17. Papers selected for the 19th International Symposium on Transportation and Traffic Theory, pp. 663 –682.

Li, Xiaopeng, Fan Peng, and Yanfeng Ouyang (2010). "Measurement and estimation of traffic oscillation properties". In: *Transportation Research Part B: Methodological* 44.1, pp. 1–14.

Knoop, V.L. et al. (2012). "Number of Lane Changes Determined by Splashover Effects in Loop Detector Counts". In: *Transactions on Intelligent Transportation Systems* 13(4), pp. 1525–1534.

Kim, Kwangho and Michael J. Cassidy (2012). "A capacity-increasing mechanism in freeway traffic". In: *Transportation Research Part B: Methodological* 46.9, pp. 1260 –1272.

Xu, Tu and Jorge Laval (2019). "Analysis of a two-regime Stochastic Car-Following Model: explaining capacity drop and oscillation instabilities". In: *Transportation Research Record*. Accepted.

Hoogendoorn, Serge and Saskia Ossen (Jan. 2005). "Parameter Estimation and Analysis of Car-following Models". In: *Transportation and traffic theory: flow, dynamics and human interaction: proceedings of the 16th international symposium on transportation and traffic theory University of Maryland*, pp. 245–265.

Ang, A. H. and W.H. Tang (2007). *Probability Concepts in Engineering*. Second. John Wiley & Sons.

Walton, Michael C and Clyde E Lee (1975). *Speed of vehicles on grades*. Tech. rep. University of Texas at Austin, Center for Highway Research.

Gillespie, Thomas D. (1985). *Methods for Predicting Truck Speed Loss on Grades*. Tech. rep. The University of Michigan Transportation Research Institute.

Yu, Bin (2005). “Modeling Truck Motion along Grade Sections”. PhD thesis. Virginia Polytechnic Institute and State University.

Officials, Transportation (2011). *A Policy on Geometric Design of Highways and Streets, 2011*. AASHTO.

Laval, J A (2004). “Hybrid models of traffic flow: impacts of bounded vehicle accelerations”. PhD thesis. Dept. of Civil Engineering, Univ. of California, Berkeley.

McLean, J R (1989). “Two-Lane Highway Traffic Operations (TWOPAS)”. In: *Transportation Studies*. Ed. by N Ashford and WG Bell. Vol. 11. Amsterdam: Gordon and Breach Science Publishers.

Bernat Goñi Ros Victor L. Knoop, Yasuhiro Shiomi Toshimichi Takahashi Bart van Arem Serge P. Hoogendoorn (2016). “Modeling traffic at sags”. In: *International Journal of Intelligent Transportation Systems Research*. Accepted.

Chen, Danjue et al. (2014). “On the periodicity of traffic oscillations and capacity drop: The role of driver characteristics”. In: *Transportation Research Part B: Methodological* 59, pp. 117 –136.

James, Gareth et al. (2013). *An Introduction to Statistical Learning with Applications in R*. Stanford University, Stanford, CA, USA: Springer.

PTV (2015). *PTV VISSIM 8 USER MANUAL*.

USEPA, U.S. Environmental Protection Agency (2019). *MOVES and Other Mobile Source Emissions Models*.

Chilukuri, Bhargava Rama, Jorge A Laval, and Angshuman Guin (2014). “MicroSimulation-Based Framework for Freeway Travel Time Forecasting”. In: *Transportation Research Record: Journal of the Transportation Research Board* 2470, pp. 34–45.

Hubbert, S (2012). *Essential Mathematics for market risk management*. 2nd. John Wiley & Sons.