

Safety evaluation of heterogeneous traffic

Experiments using different models in SUMO

Master's thesis in Automotive Engineering

WEICHENG XIAO

DEPARTMENT OF MECHANICS AND MARITIME SCIENCES
DIVISION OF VEHICLE SAFETY

MASTER'S THESIS IN AUTOMOTIVE ENGINEERING

Safety evaluation of heterogeneous traffic

Experiments using different models in SUMO

WEICHENG XIAO

Department of Mechanics and Maritime Sciences
Division of Vehicle Safety
CHALMERS UNIVERSITY OF TECHNOLOGY
Göteborg, Sweden 2020

Safety evaluation of heterogeneous traffic
Experiments using different models in SUMO
WEICHENG XIAO

© WEICHENG XIAO, 2020-08-21

Master's Thesis 2020:65
Department of Mechanics and Maritime Sciences
Division of Vehicle Safety
Chalmers University of Technology
SE-412 96 Göteborg
Sweden
Telephone: + 46 (0)31-772 1000

Cover:

A screenshot of the heterogeneous traffic simulation in SUMO (blue: conventional car; green: autonomous car; red: conventional truck; yellow: autonomous truck)

Department of Mechanics and Maritime Sciences
Göteborg, Sweden 2020-08-21

Safety evaluation of heterogeneous traffic
Experiments using different models in SUMO
Master's thesis in Master's programme Automotive Engineering
WEICHENG XIAO
Department of Mechanics and Maritime Sciences
Division of Vehicle Safety
Chalmers University of Technology

Abstract

With the development of self-driving technology, the day when autonomous vehicles share roads with traditional human-operated vehicles seems to be around the corner. This makes the safety evaluation of this so-called heterogeneous traffic particularly important.

In this thesis, by conducting microscopic heterogeneous traffic simulation on the simulation platform SUMO, the impact of autonomous vehicles on traffic safety and efficiency is studied. In order to obtain more accurate results, the initial car-following model and lane-changing model in SUMO need to be modified and calibrated using real-world data before being implemented in the simulation.

As some previous research have shown, the types of vehicles involved on driving situation have an impact on drivers' driving behaviours. The neglect of this impact has led to errors when reproducing the realistic driving behaviour with the existing car-following and lane-changing models. In this thesis, the models are modified by setting appropriate value for some related parameters to reflect this impact. Then the models are calibrated using the data extracted from highD dataset. Three performance indicators, namely number of conflicts, number of lane-changing and a speed performance indicator, are proposed to rate the error in terms of car-following, lane-changing and safety aspect. After the calibration, the best set of parameter values is selected and used to represent those human-operated vehicles in the heterogeneous traffic simulation. As for the autonomous vehicles, both zero-error Intelligent Driver Model (IDM) and Cooperative Adaptive Cruise Control (CACC) model are used to present two different types of autonomous vehicles.

After getting all the models needed, the heterogeneous traffic simulation is conducted in SUMO. Several indicators, such as time to collision and number of lane-changing, etc., are used to evaluate the safety and efficiency of traffic. The results are different when using different autonomous car models. For the zero-error IDM case, the results show that traffic safety and efficiency increase as the penetration rate of autonomous vehicles increases. For the CACC model, the traffic efficiency increases with the increase in the penetration rate, but the traffic safety deteriorates when the penetration rate is low, and it slowly improves only after the penetration rate is higher than 0.5. The simulation results help to understand the impact that autonomous vehicles will bring on heterogeneous traffic.

Key words: heterogeneous traffic, microscopic simulation, intelligent driver model, highD dataset, traffic safety evaluation, mixed traffic

Contents

Abstract.....	I
Contents.....	III
List of figures	V
List of tables	VII
Preface.....	IX
1 Introduction.....	1
1.1 Background.....	1
1.2 Objectives.....	3
1.3 Scope	3
1.4 Research questions.....	4
1.5 Thesis outline.....	4
2 Theory.....	5
2.1 Simulation package and models.....	5
2.1.1 Simulation package SUMO.....	5
2.1.2 Car-following and lane-changing models.....	6
2.2 Impacts of involved vehicle types.....	9
2.3 HighD dataset.....	10
2.4 Surrogate safety measures	11
3 Methodology.....	13
3.1 Modification of models.....	13
3.1.1 Parameter setting for IDM.....	14
3.1.2 Parameter setting for LC2013	19
3.1.3 Other parameter settings.....	19
3.2 Calibration of models.....	20
3.2.1 Driving data.....	20
3.2.2 Performance indicators.....	21
3.2.3 Parameters setting and calibration results	22
3.3 Heterogeneous traffic simulation	25
3.3.1 Simulation settings.....	25
3.3.2 Simulation outputs.....	26
4 Results	27
4.1 Results of zero-error IDM	27
4.2 Results of CACC model.....	28
4.3 Results comparison	30

5	Discussion and conclusion.....	35
5.1	Research questions.....	35
5.1.1	How do the types of vehicle involved in traffic impact drivers' driving behaviours?.....	35
5.1.2	How are these summarized impacts reflected in the models?.....	35
5.1.3	How does the existence of autonomous vehicles influence the safety of heterogeneous traffic?.....	36
5.1.4	How does the existence of autonomous vehicles influence the efficiency of heterogeneous traffic?	36
5.2	Discussion.....	36
5.2.1	Ethical aspect.....	36
5.2.2	Limitations	36
5.2.3	Future work	37
	References	39
	Appendix	45

List of figures

Figure 1 The workflow of this thesis.....	3
Figure 2 A screenshot of SUMO's graphical user interface.....	5
Figure 3 The car-following situation.....	7
Figure 4 The lane-changing situation.....	7
Figure 5 The visualization of highD dataset.....	10
Figure 6 The distribution of minimum car-following-car THW in highD.....	15
Figure 7 The distribution of minimum car-following-truck THW in highD.....	15
Figure 8 The distribution of minimum truck-following-car THW in highD.....	16
Figure 9 The distribution of minimum truck-following-truck THW in highD.....	16
Figure 10 The distribution of minimum THW in highD.....	17
Figure 11 The distribution of mean THW in highD.....	17
Figure 12 The distribution of maximum THW in highD.....	18
Figure 13 The distribution of mean speed in the selected part of highD dataset.....	23
Figure 14 The speed performance indicator (RMSE) for each sets of parameters.....	23
Figure 15 The number of lane-changing for each sets of parameters.....	24
Figure 16 The number of conflicts for each sets of parameters.....	24
Figure 17 A screenshot of the heterogeneous traffic simulation in SUMO (blue: conventional car; green: autonomous car; red: conventional truck; yellow: autonomous truck).....	26
Figure 18 Average number of potential conflicts identified by TTC (zero-error IDM).....	27
Figure 19 Average number of potential conflicts identified by DRAC (zero-error IDM).....	27
Figure 20 Average number of lane-changing (zero-error IDM).....	28
Figure 21 Average traffic flow value (zero-error IDM).....	28
Figure 22 Average number of potential conflicts identified by TTC (CACC).....	29
Figure 23 Average number of potential conflicts identified by DRAC (CACC).....	29

Figure 24 Average number of lane-changing (CACC)	30
Figure 25 Average traffic flow value (CACC).....	30
Figure 26 Box plot and mean value (blue and red lines) of number of potential conflicts identified by TTC.....	31
Figure 27 Box plot and mean value (blue and red lines) of number of potential conflicts identified by DRAC	31
Figure 28 Box plot and mean value (blue and red lines) of number of lane-changing	32
Figure 29 Box plot and mean value (blue and red lines) of traffic flow	32

List of tables

Table 1 Special outputs of simulation in SUMO.	6
Table 2 Control modes of ACC and CACC [34].	9
Table 3 Studies on car-following situation	9
Table 4 Studies on lane-changing situation.....	10
Table 5 Some parameters of IDM and LC2013 [17]	13
Table 6 Setting of some IDM parameters	14
Table 7 Values with highest probability density	18
Table 8 Selected Tau value in IDM	19
Table 9 Parameters in IDM and LC2013	19
Table 10 Parameters of driver state device in SUMO	20
Table 11 Data extracted from highD dataset and corresponding parameters in SUMO	21
Table 12 Parameters used for simulation	22
Table 13 Kolmogorov-Smirnov test result for each set	23

Preface

This thesis project has been conducted at VTI in Gothenburg during the spring and summer of 2020. I would like to thank Maytheewat Aramrattana and Niklas Strand from VTI for giving me the opportunity to participate in this project. They have provided countless help and support to me throughout the process. I would also like to thank my examiner Selpi, who can always enlighten me when I was confused and ensure that I maintain the right direction.

I want to thank Nicklas Pettersson, we helped and encouraged each other during the whole process.

WEICHENG XIAO, Gothenburg, Aug 2020

Nomenclature

All abbreviation occurring in the report are listed alphabetically in the table below.

ACC	Adaptive Cruise Control
AV	Autonomous Vehicle
CACC	Cooperative Adaptive Cruise Control
DHW	Distance Headway
DR	Deceleration Rate
DRAC	Deceleration Rate to Avoid Collision
EV	Ego Vehicle
IDM	Intelligent Driver Model
OD	Origin and Destination
OSM	Open Street Map
PET	Post Encroachment Time
PI	Performance Indicator
PV	Preceding Vehicle
RMSE	Root Mean Square Error
SGAP	Space Gap
SSMs	Safety Surrogate Measurements
SUMO	Simulation of Urban Mobility
TGAP	Time Gap
THW	Time Headway
TTC	Time To Collision
UAV	Unmanned Aerial Vehicle

1 Introduction

For centuries, traffic safety has attracted people's attention. As roads and vehicles are consistently developing, the situations that researchers need to study become more and more varied and complicated. In order to embrace the era of autonomous vehicles, consideration of heterogeneous traffic safety has also been put on the agenda. In this report, the safety is evaluated by means of conducting microscopic traffic simulation.

This thesis is done at Swedish National Road and Transport Research Institute (VTI), and is the final project for a master's degree in Automotive Engineering.

1.1 Background

With the continuous development of autonomous driving technology, autonomous vehicles (AVs) will increasingly appear on the road for the foreseeable future. According to a deployment prediction, approximately 50% of vehicles sold and 40% of vehicles on road could be autonomous in the 2050s [1]. Obviously, in the early phases of AVs' deployment on public roads, they would be sharing roads with other road users such as traditional human-operated vehicles, pedestrians, cyclists, etc. The safety evaluation of this so-called 'Heterogeneous Traffic Situation' becomes particularly important. Nowadays, based on the depth and emphasis of research, several test methods have been proposed for traffic safety evaluation, including field tests, trials and simulation [2]. In this 'Heterogeneous Traffic' case, conducting microscopic simulation is a better way than performing field tests or trials in terms of cost and safety. Therefore, this thesis focuses on using microscopic simulation tool to evaluate the safety of heterogeneous traffic.

Various traffic safety related microscopic simulations have been applied using different simulation tools and their corresponding models. For example, Cunto [3] calibrated the behavioural microscopic models in VISSIM [4] using vehicle tracking data, and then used the calibrated model to investigate the safety implications of two different traffic strategies; Vanderschuren [5] estimated the safety benefits of three types of intelligent transport system measures using the microscopic simulation models based on Paramics[6]; Huguenin, et al. [7] carried out the research with the microscopic traffic simulator AIMSUN [8] to propose a new safety indicator for linear conflicts, which not only considered the probability but also the severity of hypothetical collision. Lopez, et al. [9] depicted the work flow to do the microscopic simulation with the open source simulator Simulation of Urban Mobility (SUMO) [10]. Research has shown that none of the existing simulation tools far exceeds any other for evaluating traffic safety [11][12], but the user-friendly programming interface facilities and convenient model modification make SUMO a good choice. Moreover, the outputs of SUMO contain abundant useful information that can be directly analysed. That is mainly why SUMO is chosen as the simulation tool in this thesis.

As is known to all, the accuracy of the simulation model largely depends on how close it is to the real situation [13], that is why the default models are normally not directly used. In most cases, the models are optimized and calibrated, to some extent, to make them more consistent with the observed data, so as to obtain more accurate simulation results. As researchers learn more and more about the real traffic situation, the models used for traffic simulation are also continuously changed. Currently some research results show that in the car-following and lane-changing situations, the driving behaviours will change with the type of preceding and following vehicle. Aghabayk et

al. [14] compared the driving behaviours when following a passenger car and a heavy truck, then introduced that the impacts of the leading vehicle type are reflected in three aspects, including headways, reaction times, and subject vehicle acceleration. Eagly et al. [15] pointed out the car-following behaviours are also influenced by the following vehicle type, after observing the situations of four types of passenger car and heavy vehicle following combinations. They reached the conclusion that the existence of heavy vehicle means larger space and time headways and longer reaction time. Moridpour et al. [16] believed that a model for lane-changing behaviour of heavy vehicles should be specifically developed because all the existing models fail to represent the behaviour of heavy vehicles, she proposed a framework to capture the lane-changing behaviour of heavy vehicles. However, the original models in SUMO, including the Krauß-model (which is the default car-following model) and the LC2013 model (which is the default lane-changing model) do not reflect the change of driving behaviours [17], e.g. the different headways and reaction time. The Intelligent Driver Model (IDM), which represents the intelligent driver in car-following situations, has been proven that it can replicate driving behaviours better than the Krauß-model [18], but still does not reflect this difference on driving behaviours [19]. All these models take parameters such as vehicle speed, vehicle acceleration and vehicle length into consideration but ignore the change of driving behaviour mentioned above. Therefore, to obtain more realistic and meaningful safety evaluation results, it is necessary to make appropriate modification to IDM car-following model and LC2013 lane-changing model in SUMO.

After the simulation is conducted, the next step is to evaluate the safety of the traffic based on the results. Traditionally, the safety of traffic can be defined as the expected number of crashes [20]. However, the limitations of this definition have been extensively discussed for a long time [20], [21], [22]. Crashes are rare and happen randomly on the road, which makes them hard to observe, and the crash studies or reconstruction could be costly both in terms of time and money. In addition, the number of crashes in the road crash database is quite low, especially for this heterogeneous traffic situation. Because of all these limitations, some indirect measures are proposed for assessing the traffic safety [20]. They are not based on crashes, but make use of other indicators that are related to crashes to represent the safety performance. Those so-called safety surrogate measures (SSMs) are more measurable and able to offer an insight to predict the safety. Time to collision (TTC) and post-encroachment time (PET) are two of the most typical safety surrogate measures. As a matter of fact, those safety surrogate measures have been widely used in simulation-based safety evaluation for a long period. Gettman, et al. [20] derived surrogate measures from existing microscopic traffic simulation models for intersections and then used them to support the safety evaluation for both signalized and unsignalized intersections; Goh et al. [23] implemented another frequently-used surrogate measure—deceleration rate to avoid the crash (DRAC) together with a self-defined crash potential index to investigate the safety impacts of bus priority; Morando, et al. [24] used TTC to identify potential conflicts to study the safety impact of autonomous vehicles. According to those previous examples, it is reasonable to draw the conclusion that in a simulation environment, SSMs are effective tools for safety evaluation. In this study, a method similar to the one mentioned in [23] is applied to do the safety evaluation.

1.2 Objectives

The thesis aims to modify the car-following and lane-changing models in SUMO by taking into account the difference brought about by different preceding and following vehicle types, then use the data from highD dataset to calibrate the modified models. After that, a simulation of heterogeneous traffic will be conducted with those modified models. In the end, safety of this traffic is evaluated using SSMs. The main workflow is shown in Figure 1.

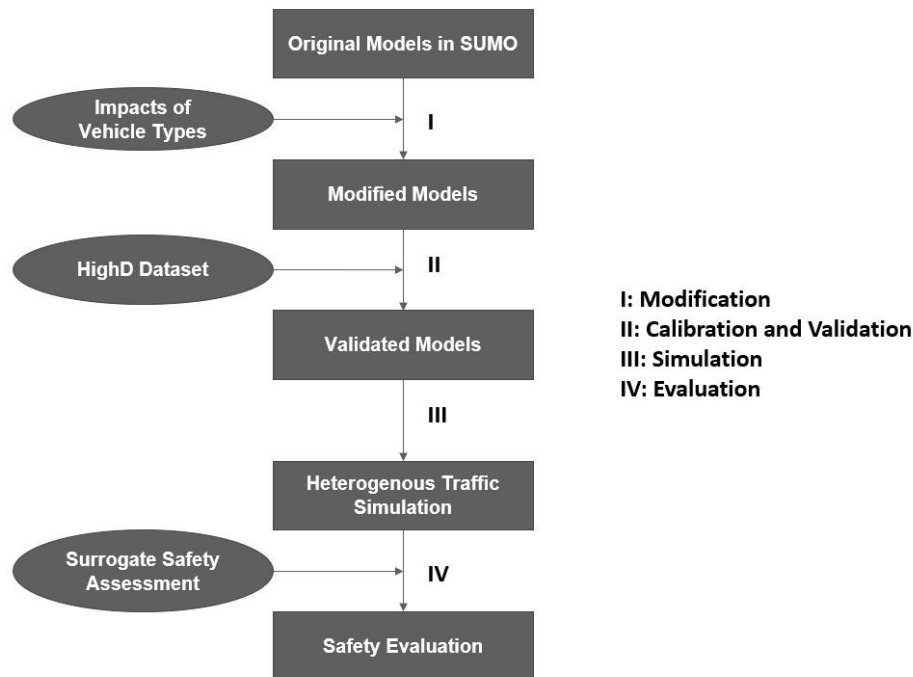


Figure 1 The workflow of this thesis

The following five objectives are addressed in this work:

1. Summarise the known impacts of involved vehicle types on driving behaviour in car-following and lane-changing situation.
2. Modify currently available car-following and lane-changing models in SUMO to address the summarised impacts.
3. Extract necessary vehicles' (i.e., cars' and trucks') data from the highD dataset, and perform the calibration on models basing on the extracted data.
4. Conduct heterogeneous traffic simulation in SUMO taking into account the results of the first three objectives.
5. Evaluate the safety of heterogeneous traffic using results from microscopic simulation in SUMO.

1.3 Scope

In heterogeneous traffic, there are many traffic participants, including cars, trucks, cyclists, pedestrians, etc. However, the scope of this thesis, including both the traffic situation studied and data used for calibration, is limited to two types of vehicles: cars and trucks. No other traffic participants are involved.

To study the impact of involving vehicle types, the data extracted from highD dataset concentrates on four types of car-truck combination, namely the car-following-car (CC), car-following-truck (CT), truck-following-car (TC) and truck-following-truck (TT). But due to the limitation of time and scale, the modification of the models is limited to present the difference between truck drivers and car drivers, that is to say, the difference between CC,CT and TC,TT. Because the difference between car drivers and truck drivers can be expressed by setting different parameters in SUMO, but if the differences in the four combinations mentioned earlier want to be reflected, the source code of SUMO need to be changed accordingly. This choice saves a lot of time while ensuring a certain accuracy, and also avoids many instabilities that may be caused by the changed source code.

1.4 Research questions

By the end of this thesis, the following research questions are answered:

1. Based on existing research, how do the types of vehicle involved in traffic impact drivers' driving behaviours?
2. How are these summarized impacts reflected in the models used for heterogeneous traffic simulation?
3. How does the existence of autonomous vehicles influence the safety of heterogeneous traffic? Do different penetration rates bring different results?
4. How does the existence of autonomous vehicles influence the efficiency of heterogeneous traffic? Do different penetration rates bring different results?

1.5 Thesis outline

The thesis is structured as follows. Chapter 2 describes the main theory for the thesis. The methodology is introduced in Chapter 3. Chapter 4 lists all the results and in Chapter 5 the conclusion and discussion are given.

2 Theory

In this chapter, the theoretical knowledge related to this thesis is introduced, including the simulation package and involved models, the impacts on driving behaviours that have been studied, the dataset used to calibrate and the safety surrogate measures used to evaluate traffic safety.

2.1 Simulation package and models

As mentioned in the introduction part, nowadays various simulation package and corresponding models can be used to do the traffic simulation. In this thesis, SUMO is chosen because it can represent the traffic situation of interest and it can directly output safety related measures utilized for safety evaluation. As a result, the models in SUMO are used as a basis for modification.

2.1.1 Simulation package SUMO

SUMO is an open source microscopic and continuous traffic simulation package, mainly developed by employees of the Institute of Transportation Systems at the German Aerospace Centre. It is freely available and published under the Eclipse Public License V2, and has been widely used in recent years [9]. The word “microscopic” here means each vehicle and its dynamics are individually represented. In traffic simulation, its advantages include simple inputs, useful outputs and convenience to modify the models. In addition, its clear graphical user interface is helpful for observing the traffic situation of vehicles during the simulation.

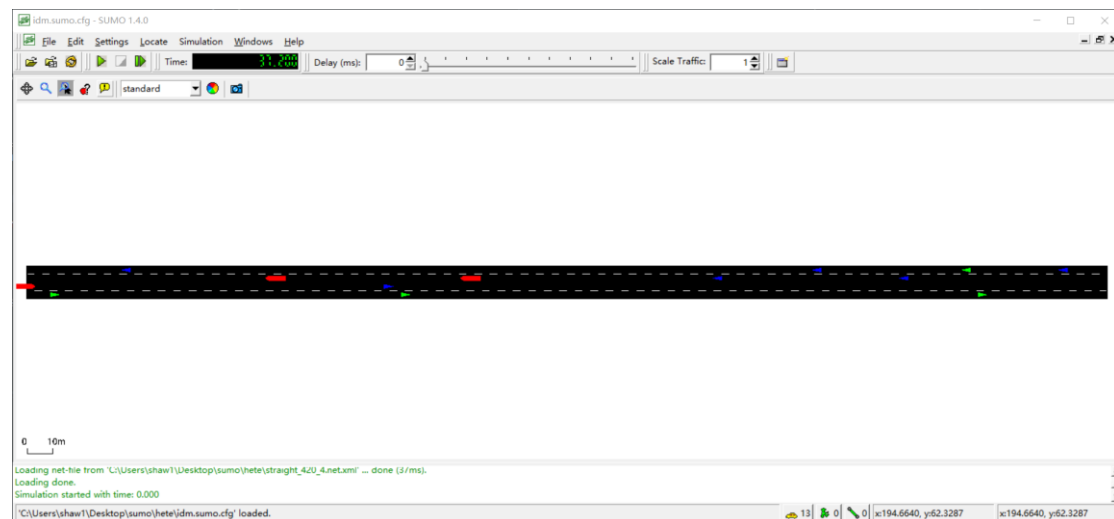


Figure 2 A screenshot of SUMO's graphical user interface

The inputs of SUMO are a network file, a route file and other optional files. The network file basically contains the road information involved in the simulation, which can be designed using the built-in program Nedit or imported directly from digital road map such as Open Street Map (OSM) or from other traffic simulator like Vissim. Such multiple options allow the users to design a simple road to observe the performance of the models first, and then verify the models in a complex road environment. The route file, also called traffic demand file, can be defined as individual

trips, flows or as routes [9]. Several tools are provided in the package to help to generate the route file. For example, the traditional Origin and destination (O-D) matrix, which contains the starting time, the origin, the destination and other details, can be transferred to vehicle trips with the help of the tool OD2TRIPS. In the optional files, more settings can be given and other entities can be loaded. So, in general, it is easy to create or import the concerned road conditions and traffic flows in SUMO.

The useful output is one of the most important reasons why SUMO was chosen. It is known to all that the data presented after the simulation should be abundant and detailed, so that useful ones could be easily selected. Data processing should be as convenient as possible to save time. Just like other simulation tools, there are lots of data sorted by topics after the simulation in SUMO. However, it is special in that the outputs it provides include not only the trajectories output, but also the lane change output, SSMs, and vehicle type probe output listed in Table 1.

Table 1 Special outputs of simulation in SUMO.

Special outputs	Parameters included
Lane change output	Vehicle type, time and place when changing lane, reason of change, current speed, etc.
SSMs	TTC, PET, SGAP, TGAP, etc.
Vehicle type probe	The chosen type and the file path of this specific type.

* TTC: time to collision; PET: post encroachment time; SGAP: space gap; TGAP: time gap

Apparently, these special outputs can be directly used in this thesis and make the data processing part much easier.

In addition to the above two points, the ability to easily modify the models is also an advantage. In SUMO, many different models are provided for car-following, lane-changing and other traffic situations. Some of them are introduced in the next section. It is worth mentioning that when the initial models cannot meet the demand, the users can easily create a new model and describe it with their own parameters by changing the source code. All in all, SUMO is a suitable simulation tool for this thesis.

2.1.2 Car-following and lane-changing models

In this thesis, two of the most common and important traffic situations are considered, namely car-following and lane-changing. The car following situation basically involves two vehicles, the ego vehicle and the preceding vehicle. The longitudinal interaction between them are explored. On the other hand, the lane-changing situation focus on the lateral behaviours.

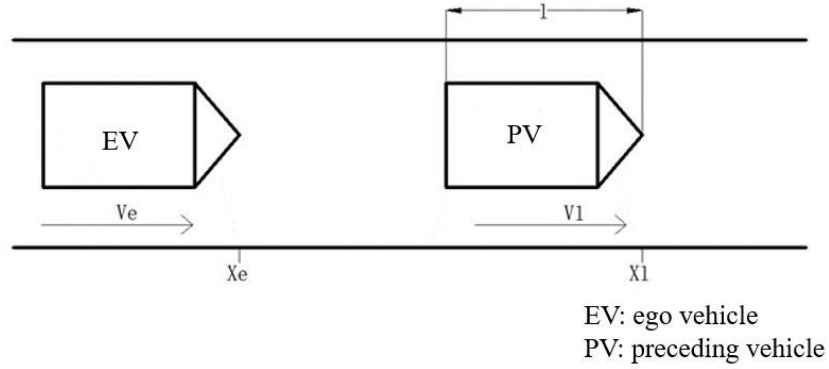


Figure 3 The car-following situation

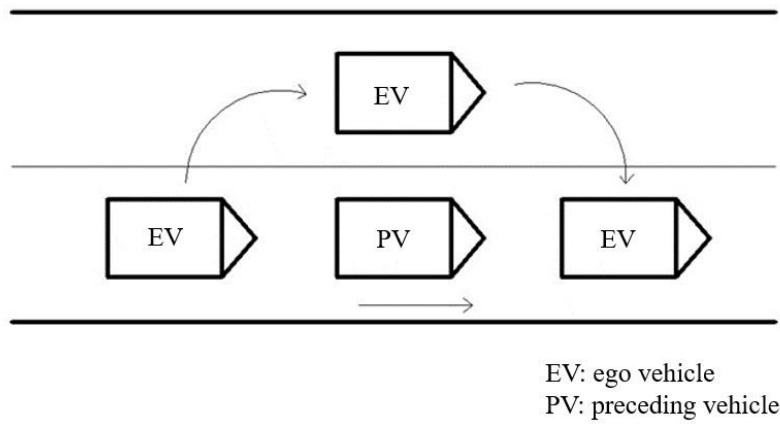


Figure 4 The lane-changing situation

In recent years, various of models have been proposed to simulate these two traffic situations. Luong [25] has compared different car-following models for microscopic simulation in terms of safe distance, stability, and velocity. Zheng et al. [26] did the same thing for lane-changing models. In SUMO alone, there are up to fifteen different car-following models available, based on different principle, using different parameters, e.g., the default Krauss model [19] and the widely used intelligent driver model (IDM) [27]. In addition, there are some more advanced car-following models, such as Adaptive Cruise Control (ACC) model and Cooperative Adaptive Cruise Control (CACC) model [28]. For the lane-changing simulation, there are three models mentioned in SUMO, and the default one is LC2013 [29]. All these named models are going to be briefly introduced and used as basics for modification.

The default Krauss model is the model defined by Stefan Krauß in [19] with some modifications. The principle behind it is quite simple, just drive as fast as you can, as long as you will not get yourself involved in a collision. To fulfil this goal, a concept called safety velocity is proposed, this safety velocity in next time step will be decided by the current speed of both vehicles, the reaction time of the ego vehicle driver, and the deceleration capacity of the ego vehicle. It can be expressed by the following formula.

$$v_{safe}(t) = v_l(t) + \frac{g(t) - v_l(t)\tau}{\frac{v}{b(v)} + \tau} \quad (2.1)$$

$v_{safe}(t)$ is the safety velocity, $v_l(t)$ is the velocity of preceding vehicle, $g(t)$ is the gap to the front vehicle, τ is the reaction time of the following driver, \bar{v} is the velocity of ego vehicle and $b(\bar{v})$ is the deceleration of ego vehicle. After calculating the safety velocity, the driver will choose the minimum value among the safety velocity, the maximum velocity of the ego vehicle and the current velocity of the vehicle ($v_e(t)$) plus the maximum acceleration.

$$v_{des}(t) = \min [v_{safe}(t), v_e(t) + a, v_{max}] \quad (2.2)$$

In order to more realistically simulate human drivers, a disturbance coefficient has been added to represent human error.

While the Krauss model is widely used, it has also been questioned in [30] [31]. Treiber et al. came up with the IDM in [27] to describe the dynamic of a vehicle in car-following situation. The model can be simplified into two ordinary differential equations as below:

$$\dot{x}_e = \frac{dx_e}{dt} = v_e \quad (2.3)$$

$$\dot{v}_e = \frac{dv_e}{dt} = a \left(1 - \left(\frac{v_e}{v_0} \right)^\delta - \left(\frac{s^*(v_e, \Delta v_e)}{s_e} \right)^2 \right) \quad (2.4)$$

With $s_e = x_l - x_e - l$; $\Delta v_e = v_e - v_l$; $s^*(v_e, \Delta v_e) = s_0 + v_e T + \frac{v_e \Delta v_e}{2\sqrt{ab}}$.

s_e is the gap distance, x_l is the position of preceding vehicle and x_e is the position of ego vehicle, v_e is the speed of ego vehicle and v_l is the speed of preceding vehicle, l is the length of preceding vehicle, v_0 is the desired velocity of ego vehicle, s_0 is the minimum spacing ahead, T is the desired time headway, a is the maximum vehicle acceleration, b is the comfortable braking deceleration and δ is an acceleration exponent. Obviously, more parameters are taken into consideration in this model compared to Krauss model, and Pourabdollah et al. [32] have shown that IDM replicates driving behaviours better than Krauss model. In this thesis, the car-following model for human-operated vehicles is modified based on IDM.

In addition to these two models, SUMO also provides other car following models, such as ACC and CACC, which can actively control the speed of the vehicle independently of the driver's behaviour [33]. This allows them to be used to represent autonomous vehicles to some extent in the simulation of SUMO. Their control modes are the same, as shown in Table 2. The difference between CACC and ACC is that CACC added the communication part of V2V and V2I, which makes the traffic safer and more stable.

As mentioned earlier, the default lane-changing model in SUMO is LC2013. It is developed by Erdmann in [29] based on DK2008 model [35]. It divides the lane changes into four types (i.e., strategic change, cooperative change, tactical change and regulatory change) according to the motivation of the manoeuvres, and represents them separately. In addition to some vehicle and traffic parameters, such as vehicle acceleration and lateral gap, parameters about eagerness are also used in this model to reflect the driver's willingness to change lanes.

Table 2 Control modes of ACC and CACC [34].

Control modes	Aim	Active if
Speed control mode	Maintain the pre-defined desired speed	TGAP > 2 s or no preceding vehicle
Gap control mode	Maintain a constant time gap	gap and speed deviations are smaller than 0.2 m and 0.1 m/s
Gap-closing control mode	Enable the smooth transition from speed control mode to gap control mode	TGAP < 1.5 s
Collision avoidance control mode	Prevent rear-end collisions	TGAP < 1.5 s and gap deviation is negative

2.2 Impacts of involved vehicle types

Studies have shown that the involved vehicle types in car-following and lane-changing situations have significant impacts on the drivers' behaviours, which means that if the models want to reflect the real situations, then these impacts should be taken into account. Based on the analysis of naturalistic driving data, many impacts have been proposed. Those involving the four combinations in the situations considered in this thesis are summarized in Table 3 and Table 4.

Table 3 Studies on car-following situation

References	Involved parameters	Conclusions
Sayer et al. [36]	DHW	CC > CT
	THW	CC > CT
Ye et al. [37]	THW	TT > TC > CT > CC
Hoogendoorn et al. [38]	THW	CT > CC TT > TC
Aghabayk et al. [39]	DHW	TT > TC > CT > CC (v > 30Km/h)
		TT > CT > TC > CC (v < 30Km/h)
	THW	Same as DHW
	Reaction time	TT > TC = CT > CC
Sarvi et al. [40]	DHW	CT > TC > CC

From the listed studies on car-following situation, most of them reported that the existence of truck will make time headway (THW) or distance headway (DHW) larger, which makes TT have the largest value. Meanwhile, there are a few studies that hold opposite opinions. The former view says that larger lengths and lower braking capabilities of trucks have led to this result [37]. The latter view argues that the smaller braking rate of the truck will make the following car feel that they have more time to deal with the emergency, thus shortening the following distance [36]. There are also

different opinions regarding the two combinations of CT and TC. Some studies suggest that the situation will change with the speed of ego vehicle as shown in Table 3.

Table 4 Studies on lane-changing situation

References	Conclusions
Moridpour [33]	Heavy vehicles have a larger front space gap than passenger cars.
	Passenger cars do more lane changings to gain speed.
	The heavy vehicle drivers mostly move into the slower lane and the passenger car drivers mainly move into the faster lane during the lane changing maneuver.
Aghabayk et al. [42]	As vehicle length increases, the minimum available gap accepted by vehicles gets larger.

For the lane-changing situation, the current studies focus more on the ego vehicle type instead of the types of all involved vehicles. In terms of lane-changing maneuvers, the truck drivers tend to maintain a larger front space gap than passenger car drivers when changing to the target lane. As for the motivation of lane-changing, passenger car drivers change lanes more often than truck drivers to gain speed, while the truck drivers conduct more cooperative lane-changing to make way for faster vehicles behind.

These impacts mentioned above will be used as a guideline to help to modify the models.

2.3 HighD dataset

After the models are modified, some real-world traffic data is needed for the calibration of the modified models. For this thesis, the highD dataset [44] is selected for the following reasons.

The highD dataset are naturalistic vehicle trajectories recorded on German highway using drone or Unmanned Aerial Vehicle (UAV). It contains a large quantity of vehicles and different traffic situations. Many cases of the four vehicle combinations that this thesis focuses on can be found in the database. A preliminary statistical analysis indicates that it contains 79,000 CC cases, 8,700 CT cases, 8,100 TC cases and 13,000 TT cases.



Figure 5 The visualization of highD dataset

The database consists of 60 recordings, which were recorded at different times and on different roads. But what they have in common is that they were all recorded on the highway, and the traffic flows include two types of vehicles, cars and trucks. Each recording consists of four files. In addition to an image showing the overview of the recording location, it also contains `xx_recordingMeta.csv`, `xx_tracks.csv`, `xx_tracksMeta.csv` (`xx` is the file serial number). The `trackMeta` file contains some general information about the record, such as its serial number, location, start time and duration of the record, the total covered driving distance and time, the total number of cars and the number of trucks. The `track` file records the traffic situation of each vehicle frame by frame, including the vehicle ID, the location of the vehicle in each frame, its real-time lateral and longitudinal speed and acceleration, and its current DHW, THW, TTC, the ID of the vehicle it follows, and the lane number where it is located. From this file, a lot of useful information related to the car-following situation and lane-changing situation can be obtained. The `track meta` file is like a summary of each vehicle in the record. It contains the vehicle ID, the corresponding width and height, the type of vehicle, the first frame in which the vehicle appears, and the last frame in which the vehicle disappears. In addition, the file also presents the minimum, maximum and average longitudinal speed, the minimum DHW, the minimum THW, the minimum TTC and the number of lane changes. Obviously, this file also contains a lot of information that can be directly used in the calibration part.

From the above, it is not hard to find that in addition to a large scale, another good thing about the database is the enriched data. A lot of useful information is pre-extracted, including metrics like THW and TTC, and the driven manoeuvres, for instance, lane changes. Meanwhile, the visualization also makes data extraction and analysis more intuitive.

2.4 Surrogate safety measures

Since SUMO is an accident-free simulator, it is not convenient to count the number of accidents to evaluate the safety, which means some other measurable indicators should be used to indicate traffic safety. In this thesis, it is the number of conflicts that plays this role. The term “traffic conflict” is defined by Amundsen & Hyden in [45]:

“A traffic conflict is an observable situation in which two or more road users approach each other in space and time to such an extent that there is a risk of collision if their movements remain unchanged.”

From the definition it is easy to tell, that number of traffic conflicts can represent the safety of traffic when the accident (or collision) number is not available. By setting an appropriate threshold for Time to Collision (TTC), the potential traffic conflicts can be identified. TTC is defined as the time interval, usually measured in seconds, required for one vehicle to strike another object [46] which is the preceding vehicle in the case of car-following and lane-changing. It can be described in a formula using the parameters in Figure 3 as below:

$$TTC = \frac{x_l - x_e - l}{v_e - v_l}, \text{ when } v_e - v_l > 0 \quad (2.5)$$

In some studies [5] [41], the max below 1.0 or 1.5 seconds is regarded as the trigger to record a conflict, but in [50], 3.0 seconds is used as the threshold for emergency braking. However, if only one measure is used to determine conflict, it may cause some biases,

so another measure, deceleration rate to avoid a crash (DRAC), is also applied. Using the mentioned parameter in Figure 2, DRAC can be represented as below:

$$DRAC = \frac{(v_e - v_l)^2}{2(x_l - x_e - l)}, \text{ when } v_e - v_l > 0 \quad (2.6)$$

A certain value of 3.4 m/s^2 or 3.35 m/s^2 has been widely used as the threshold for DRAC [23].

3 Methodology

In this chapter, the methodology of this master thesis is introduced in detail, including the acquisition of the modified models before the simulation, the heterogeneous simulation with the modified models, and the analysis of the results after the simulation. The acquisition of the model part consists of the modification and verification of models.

3.1 Modification of models

There are many different ways to modify the driving models based on existing models in SUMO, including adding new parameters, modifying the principles behind the models, and setting more reasonable values for current parameters, etc. Out of consideration for the workload and convenience, the third method should be given priority to see if the model can be modified. If not, then adding new parameters or directly changing the principles behind the model should be considered.

It is mentioned in Section 2.1.2 that IDM and LC2013 are selected as the basis for the car-following model and the lane-changing model of human-operated vehicles, respectively. Some of their parameters and corresponding definitions are listed in Table 5.

Table 5 Some parameters of IDM and LC2013 [17]

Model	Parameter	Definition
IDM	minGap	Minimum Gap when standing (m)
	accel	the acceleration ability (m/s ²)
	decel	the deceleration ability (in m/s ²)
	emergencyDecel	The maximum deceleration ability in case of emergency (in m/s ²)
	tau	the driver's desired (minimum) time headway
LC2013	lcStrategic	the eagerness for performing strategic lane changing
	lcCooperative	the willingness for performing cooperative lane changing
	lcSpeedGain	the eagerness for performing lane changing to gain speed
	lcAssertive	Willingness to accept lower front and rear gaps on the target lane

It can be easily found that there are already several parameters that can be used to reflect the summarized impacts of vehicle types. The goals are supposed to be achieved by setting the values of existing parameters reasonably. From the summary in Section 2.2, three main conclusions can be drawn as below:

1. Truck drivers tend to maintain a larger THW/DHW than passenger car drivers in car-following situation.
2. Passenger car drivers tend to change lanes more often to gain speed than truck drivers.
3. The truck drivers tend to maintain a larger front gap distance than passenger car drivers when changing to the target lane.

These three conclusions can be realized by the following settings in IDM and LC2013 :

1. Set a larger tau value for trucks in IDM.
2. Set a smaller lcSpeedGain value for trucks in LC2013.
3. Set a smaller lcAssertive value for trucks in LC2013.

3.1.1 Parameter setting for IDM

The first step is to separately set the mentioned IDM parameters for cars and trucks in SUMO. From [47], the default parameter values corresponding to each vehicle type can be obtained. For cars, the minGap is set to 2.5 m, the acceleration ability is set to 2.9 m/s^2 while the deceleration ability is 7.5 m/s^2 . The emergency deceleration ability is set to 9.0 m/s^2 . As for the trucks, their acceleration ability, deceleration ability, and emergency deceleration ability are set to 1.2 m/s^2 , 4.0 m/s^2 , and 7.0 m/s^2 , respectively.

Table 6 Setting of some IDM parameters

parameter	car	truck
minGap	2.5 m	2.5 m
accel	2.9 m/s^2	1.2 m/s^2
decel	7.5 m/s^2	4.0 m/s^2 .
emergencyDecel	9.0 m/s^2	7.0 m/s^2 .

The next part is the setting of Tau value. In this thesis, two methods are proposed to determine an approximated value, and several groups are set above and below this value to further determine the most suitable value. The first method is to refer to the existing simulation settings. In the online interactive traffic simulation [48] created by Martin Treiber, he also chose IDM to simulate the driving of cars and trucks on the highway. He chose 1.5 seconds as the Tau value of the car and 1.7 seconds as the Tau value of the truck. This setting is used as one of the basic values (set 2 in Table 8).

Another method is to obtain the relevant values from the highD dataset. According to the definition of Tau value, it represents the driver's desired (minimum) time headway. This value can be obtained from the whole dataset through statistical methods. As mentioned in Section 2.3, the track file and trackMeta file contain all the necessary information to identify the car-following situations of each vehicle in the traffic. This makes it possible to analyse the THW distribution during the car-following situations. Considering the impact of involved vehicle types on the drivers' driving behaviours, all car-following situations in highD dataset are first classified according to those four combinations mentioned in Section 1.3, and then their THW data are extracted separately. The distributions of their minimum THW are shown in Figure 6-9.

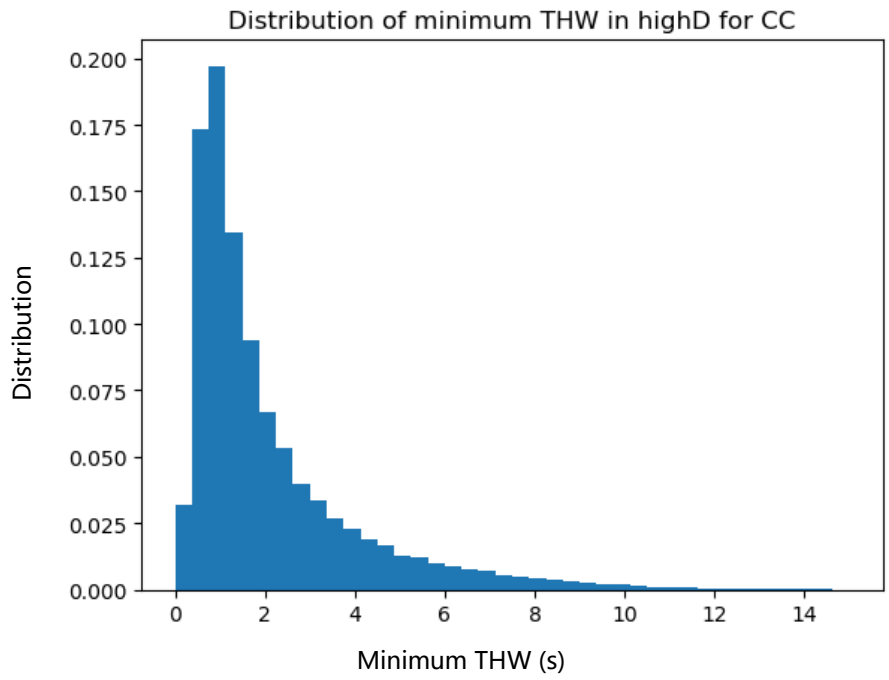


Figure 6 The distribution of minimum car-following-car THW in highD

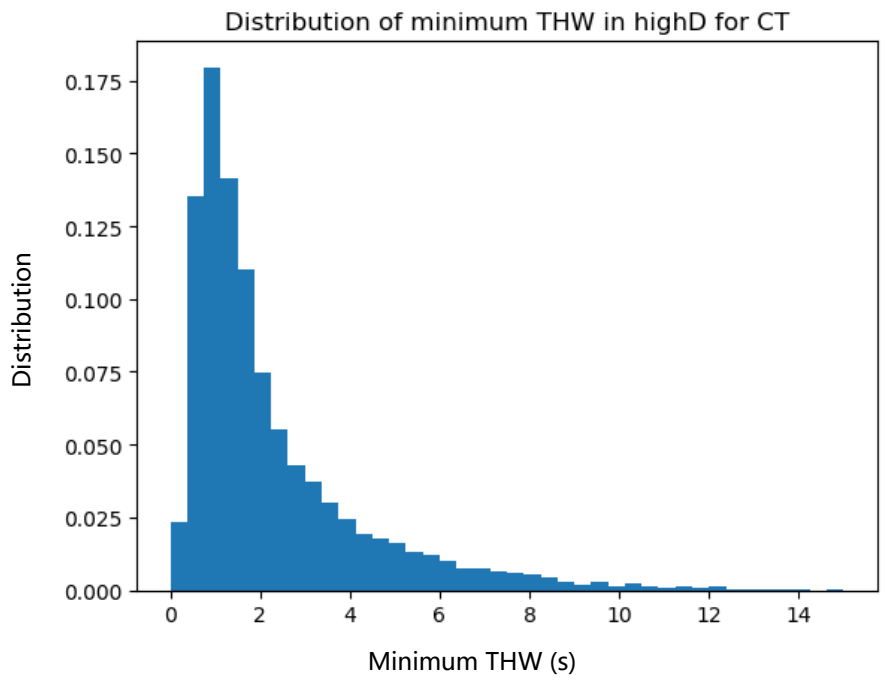


Figure 7 The distribution of minimum car-following-truck THW in highD

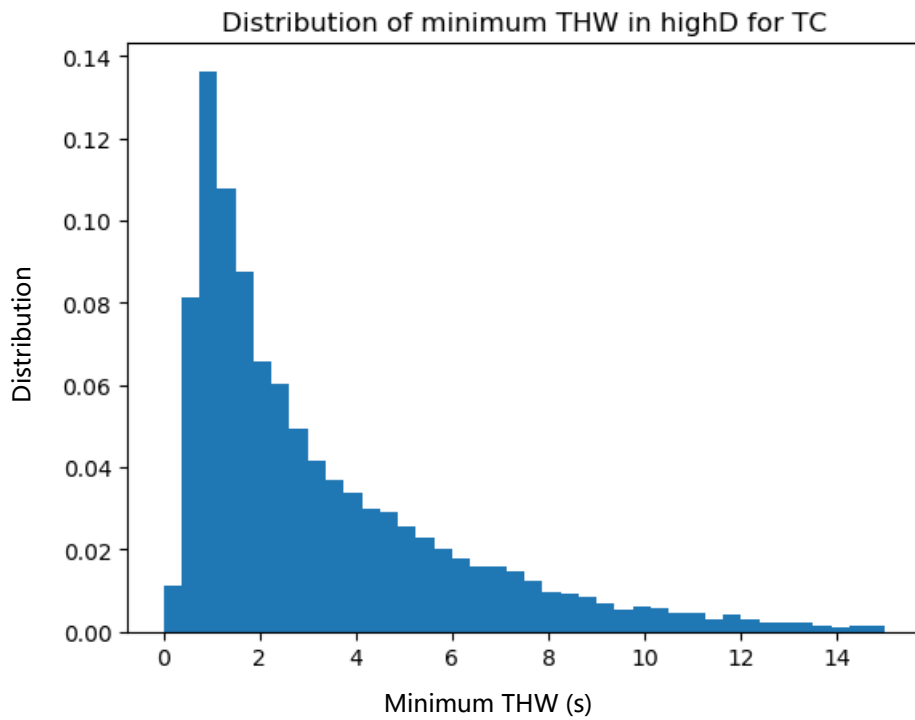


Figure 8 The distribution of minimum truck-following-car THW in highD

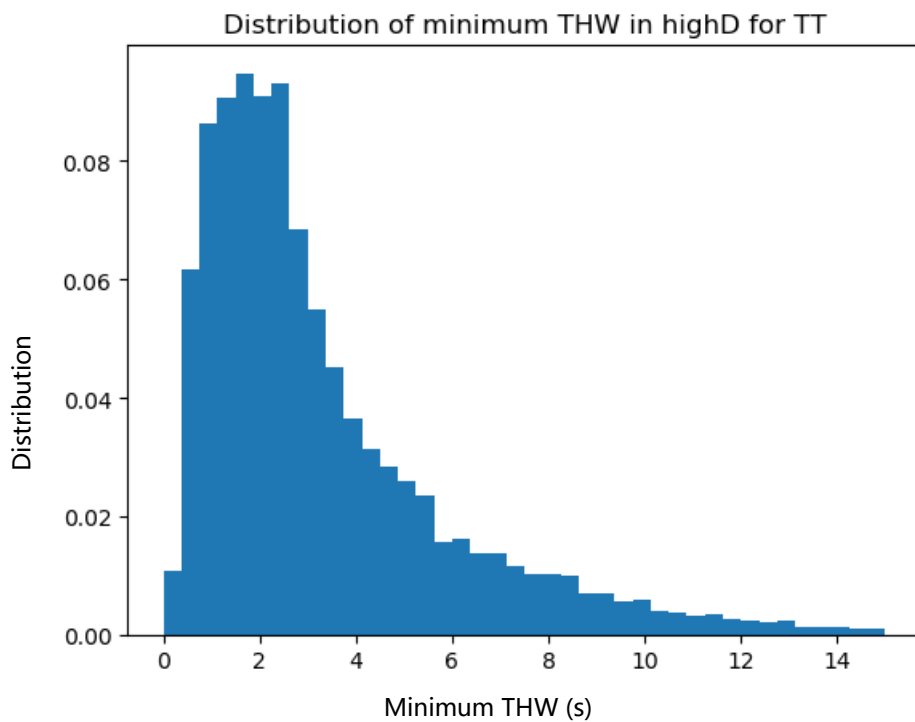


Figure 9 The distribution of minimum truck-following-truck THW in highD

After implementing a kernel density estimation and putting the results together, the distributions of all four combinations are shown in Figure 10.

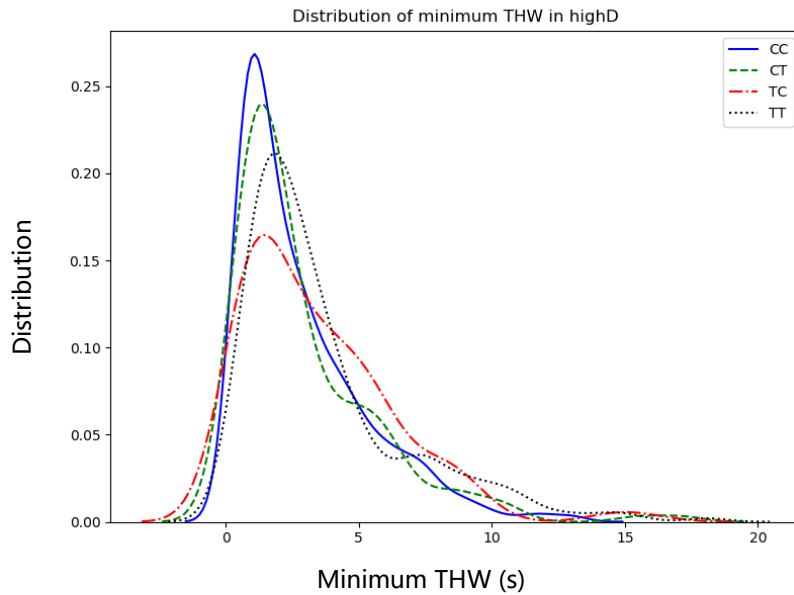


Figure 10 The distribution of minimum THW in highD

In order to better understand the THW selected by drivers in the process of following a vehicle, the same method is applied to study the mean THW and maximum THW. The results are shown in Figure 11 and Figure 12, respectively.

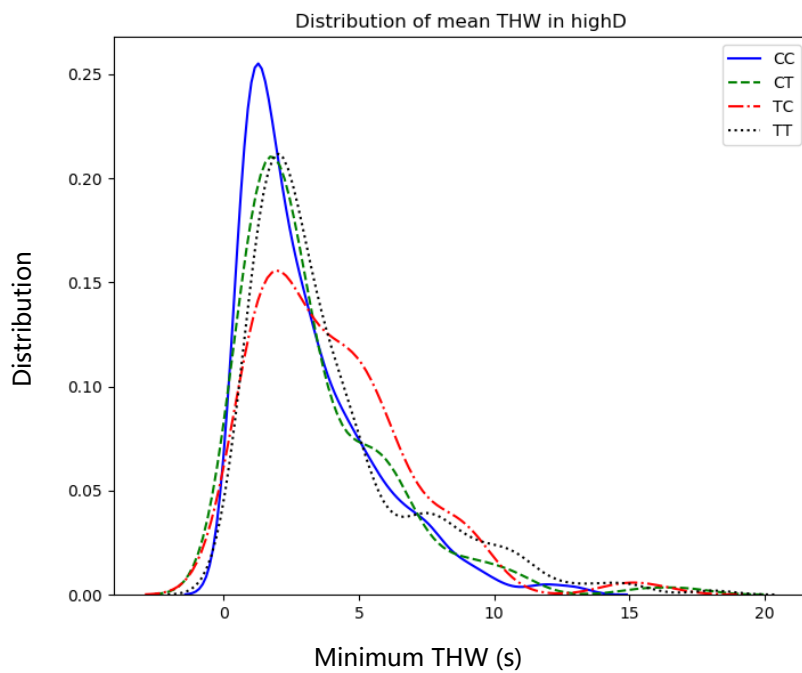


Figure 11 The distribution of mean THW in highD

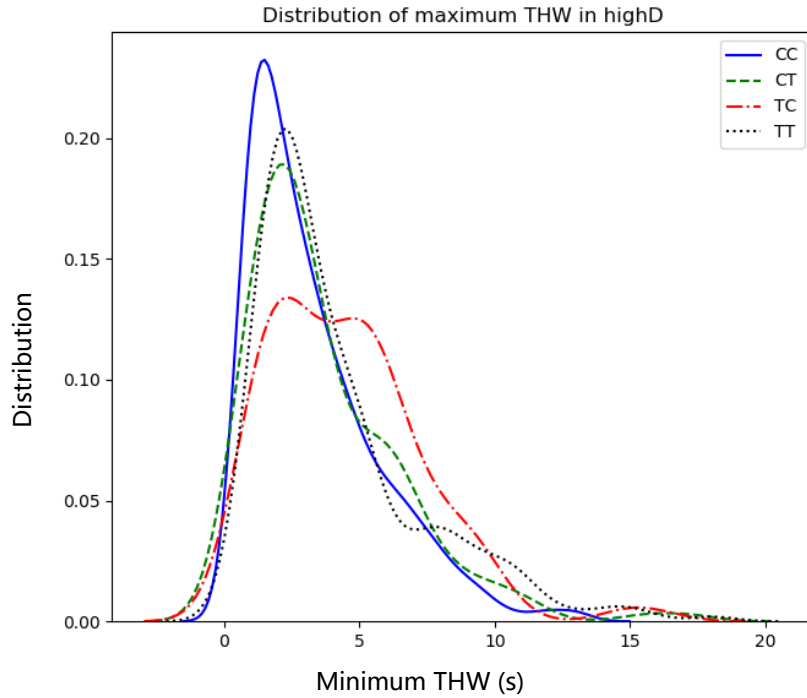


Figure 12 The distribution of maximum THW in highD

It can be drawn from the Figure 10, 11, 12 that values with highest probability density and their corresponding standard deviations in the minimum, maximum and mean values are shown in Table 7.

Table 7 Values with highest probability density

Combination	Maximum	SD	Minimum	SD	Mean	SD
CC	1.50 s	2.42	1.08 s	2.41	1.28 s	2.40
CT	2.24 s	2.81	1.42 s	2.77	1.73 s	2.77
TC	2.33 s	2.85	1.43 s	2.87	1.97 s	2.82
TT	2.33 s	3.12	1.75 s	3.12	2.02 s	3.10

It can be seen from Table 7 that the results obtained from the highD dataset are consistent with the aforementioned research results of [36] [37] [38]. Among the THW values of those four combinations, $TT > TC > CT > CC$. This further verifies that the types of involved vehicle in traffic situation do have an impact on the driver's driving behaviour. After obtaining these data, 1.50 seconds and 2 seconds were selected as the Tau value of the car and the Tau value of the truck in another basic set (set 3 in Table 8). Based on these two sets of basic values, another two sets are expanded upwards and downwards to obtain the initially selected Tau value in IDM parameters, as shown in Table 8.

Table 8 Selected Tau value in IDM

	Tau for car	Tau for truck
Set 1	1.50 s	1.50 s
Set 2	1.50 s	1.70 s
Set 3	1.50 s	2.00 s
Set 4	1.50 s	2.30 s

3.1.2 Parameter setting for LC2013

After setting up the following model, the parameters in the lane-changing model LC2013, such as lcSpeedGain and lcAssertive, also need to be set to obtain a more reasonable lane-changing situation. From the data obtained from highD dataset, there are 14 lane changes per 100 cars and 6 lane changes per 100 trucks. But with the default parameter settings, the number of lane changes will be much smaller than the real data. Therefore, each set of parameters in car-following model need to be correspondingly assigned appropriate parameters in lane-changing model. After many adjustments, the appropriate parameter settings are listed in Table 9.

Table 9 Parameters in IDM and LC2013

	τ_c	τ_t	$lcSG_c$	$lcSG_t$	lcA_c	lcA_t
Set 1	1.5	1.5	3.1	2.0	3.0	1.0
Set 2	1.5	1.7	3.1	2.0	3.8	2.0
Set 3	1.5	2.0	2.4	3.0	4.1	1.5
Set 4	1.5	2.3	3.1	2.0	3.4	1.5

*lcSG: lcSpeedGain; lcA: lcAssertive

3.1.3 Other parameter settings

After setting the parameters of the car-following model and lane-changing model, the new model is expected to reflect the real driving behaviours more realistically. But for a human-operated vehicle, there is another important factor that affects the driving behaviours, that is driving imperfection caused by human drivers. In order to reflect the difference between autonomous vehicles and human-operated vehicles in heterogeneous traffic simulation, driving imperfection needs to be added to the model.

To achieve that, the model is equipped with Driver State Device, which is a generic mechanism provided by SUMO to introduce driving errors to the car-following and lane-changing models. The related parameter settings are shown in Table 10, and their specific meanings are listed in the appendix. This set of parameters has been proved to be effective in representing possible driving errors in [49], so as to better simulate the actual driver driving situation.

The autonomous vehicles in the simulation are represented using two different types of model. The first one is zero-error IDM. It is the default IDM without adding driving errors. The default value of desired (minimum) THW is 1 second, which is lower than the value set for conventional human-operated vehicles. Not only that, in order to reflect the perfect driving of autonomous vehicles, the speed factor in the model is set to 1, and speed deviation is set to zero. The difference in this respect can roughly reflect the

difference between autonomous vehicles and conventional human-operated vehicles in the real situation. This model emphasizes the perfect driving ability of each autonomous vehicle. The second one is CACC model introduced in Section 2.1.2. Different from the zero-error IDM, it not only ensures the excellent driving ability of each autonomous vehicle, but also includes V2V and/or V2I communication. This makes each autonomous vehicle no longer a separate individual, but can exchange information with other autonomous vehicle and influence each other. It can be seen from [51] that when simulating an autonomous vehicle and a human-operated vehicle in the same traffic scene, if the vehicle type is the same, then their acceleration, deceleration, and emergency deceleration can be set to be basically the same, The only difference is that a different model is selected, and the AV driver's imperfection is set to zero. Therefore, when setting the parameters of the CACC model, the acceleration, deceleration and emergency deceleration used by the human-operated vehicle are retained, and the imperfection and speed deviation are set to 0.

Both models are used in heterogeneous traffic simulation separately to observe their impact on traffic safety and efficiency. Their results are also compared with each other to see the difference.

Table 10 Parameters of driver state device in SUMO

Parameter	Value
initialAwareness	1.1
errorTimeScaleCoefficient	100
errorNoiseIntensityCoefficient	0.5
speedDifferenceErrorCoefficient	1
headwayErrorCoefficient	4
speedDifferenceChangePerceptionThreshold	1
headwayChangePerceptionThreshold	1

3.2 Calibration of models

After setting all the parameters and obtaining a new model, an important step is to calibrate it with real driving data. In this section the calibration process of the model is introduced. After this process, a set of parameters closest to the actual situation is selected.

3.2.1 Driving data

As mentioned before, the data used for calibration are from the highD dataset. This database contains real driving data, including 110,000 vehicles and a driving distance of 45,000 km. Taking into account the computation and simulation time, this calibration does not use all the data, but part of the data obtained from it, which are the first two files in this case. They include in total 2160 vehicles. The data obtained from them is mainly divided into two parts, the initial driving states of the vehicle, and the driving statistics. The initial driving states of the vehicle includes the type of vehicle, the initial speed, the initial position, the starting lane and the direction of driving. These data are used as inputs for SUMO simulation to rebuild the driving situation in the database as realistic as possible. The driving statistics are composed of driving speed, the smallest TTC and the number of lane changes of each vehicle, etc. They are used to evaluate

whether the simulation results are close enough to the actual observed data as described below.

Table 11 Data extracted from highD dataset and corresponding parameters in SUMO

Extracted data	Specific information	Parameter in SUMO
Initial driving states	type of vehicle	type
	initial speed	departSpeed
	initial position	departPos
	starting lane	departLane
	driving direction	route
Driving statistics	minimum speed	minXVelocity
	maximum speed	maxXVelocity
	mean speed	meanXVelocity
	minimum TTC	minTTC
	number of lane changes	numLaneChanges

3.2.2 Performance indicators

In the modification section, the method of modifying the model by setting appropriate parameters has been selected. The next task is to solve the problem of how to determine an appropriate set of parameters. As we know, the results of simulation will change with the change of parameters' value, and by evaluating the simulation result, it can be shown whether the model under this set of parameters can simulate the actual situation well. When the errors between the simulated value and the observed value is small enough, this set of parameters can be considered as a good one that can reflect the actual situation. The role of performance indicators is to make errors intuitively clear and comparable.

For this driving model, the car-following situation, lane change situation and safety-related situation during driving all need to be taken into account. Therefore, three performance indicators, consisting of a speed performance indicator, number of lane changes and number of conflicts, are proposed to illustrate the performance in terms of these three aspects. Speed is a significant factor affecting traffic. Hence, its value is expected to be as close as possible to the actual data. If the speed of each vehicle at each time step is extracted and compared, the amount of data to be processed will be extremely large, and it's not necessary to make the model to perform exactly like the vehicle in HighD at every single time step. So, in this thesis, the mean speed, instead of speed at every single time step is extracted and compared and the speed performance indicator is defined as the root mean square error, RMSE, of each vehicle's mean speed, shown in the following formula:

$$PI = RMSE_{mean} \quad (3.1)$$

The root mean square error between the measured value, x_a , and the simulated value, \bar{x}_a , is defined as $RMSE_x = \sqrt{\frac{\sum_{a=1}^n (x_a - \bar{x}_a)^2}{n}}$. The a and n in the RMSE formula refer to the vehicle index and the total number of vehicles respectively. The measured value here is the mean speed of each vehicle extracted from HighD dataset, and the simulated value is the mean speed in the whole simulation run of each vehicle. In the case where the initial speed has been input in advance, this performance indicator can represent to some extent the speed change during the driving process. The lower the value, the closer the speed change in the simulation is to the realistic situation. In addition, by comparing

the simulated and actual number of lane change, one can directly recognize whether the model can reflect the actual situation in terms of lane changes. Similarly, by comparing the number of conflicts, which is defined here as the case where TTC is less than 3.0 seconds, it can be determined whether the model can reflect traffic safety of realistic traffic.

3.2.3 Parameters setting and calibration results

As mentioned earlier, there are a total of six parameters that need to be set, namely the Tau value in IDM for truck (τ_t) and for passenger car (τ_c), and the values of lcSpeedGain and lcAssertive in LC2013 for truck ($lcSpeedGain_t$ and $lcAssertive_t$) and for passenger car ($lcSpeedGain_c$ and $lcAssertive_c$). The tau value represents the driver's desired (minimum) time headway, lcSpeedGain represents the driver's eagerness for performing lane changing to gain speed, and lcAssertive represents the driver's willingness to accept lower front and rear gaps on the target lane. The default value for tau is 1.0 second and the adjustable range is $\tau \geq 0$. The default value for lcSpeedGain and lcAssertive is 1.0 and the range is from 0 to infinity. Based on the conclusions previously summarized, τ_t should be larger than τ_c , $lcSpeedGain_t$ should be smaller than $lcSpeedGain_c$ while $lcAssertive_t$ should be smaller than $lcAssertive_c$. In order to get a better set of parameters, several sets of different parameters are set up for comparison as shown in Table 9.

To more intuitively see the difference between the model with default parameters and the modified model, a set of default parameters (set 5 in Table 8) is also added, so a total of five sets of parameters are used for simulation and their performance indicators are compared.

Table 12 Parameters used for simulation

	τ_c	τ_t	$lcSG_c$	$lcSG_t$	lcA_c	lcA_t
Set 1	1.5 s	1.5 s	3.1	2.0	3.0	1.0
Set 2	1.5 s	1.7 s	3.1	2.0	3.8	2.0
Set 3	1.5 s	2.0 s	2.4	3.0	4.1	1.5
Set 4	1.5 s	2.3 s	3.1	2.0	3.4	1.5
Set 5	1.0 s	1.0 s	1.0	1.0	1.0	1.0

Before calculating the RMSE for each set, there is a quick way to judge whether the speed distributions of cars and trucks in the simulation results are consistent with the actual situation, and that is to use the statistical methods. Firstly, the mean speed for each vehicle in the part of the highD database used for calibration is extracted, and the speed distribution is shown in Figure 13. Then, a Shapiro-Wilk test is used to determine that the two distributions in Figure 13 do not conform to the normal distribution. So after the simulation of each set is completed, Kolmogorov-Smirnov tests between real data and simulation results are conducted. Kolmogorov-Smirnov tests' results are listed in Table 13. From the table it's easy to tell that they all have a very low P value, which means the mean speed distribution in all the simulation results is not from the same distribution of highD dataset. It makes sense because it's always hard to use simulation to fully rebuild the reality. This makes it necessary to use $RMSE_{mean}$ to determine which set can bring better speed simulation results.

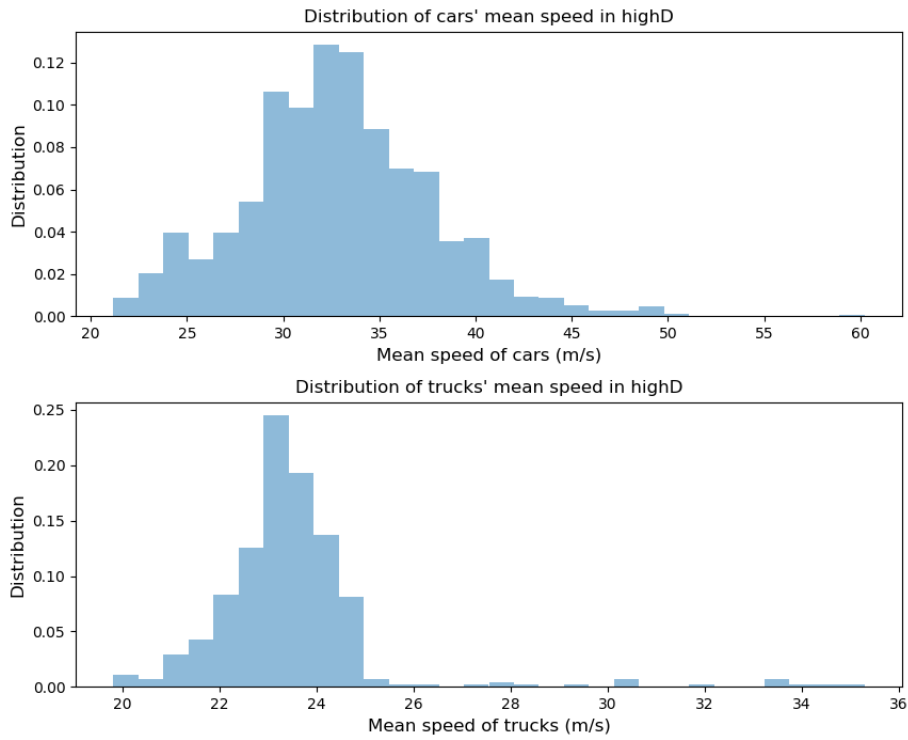


Figure 13 The distribution of mean speed in the selected part of highD dataset

Table 13 Kolmogorov-Smirnov test result for each set

	Set 1	Set 2	Set 3	Set 4	Set 5
Statistic	0.12945	0.29796	0.21283	0.17959	0.06822
P value	5.19E-13	4.14E-67	1.90E-34	1.19E-24	0.0006385

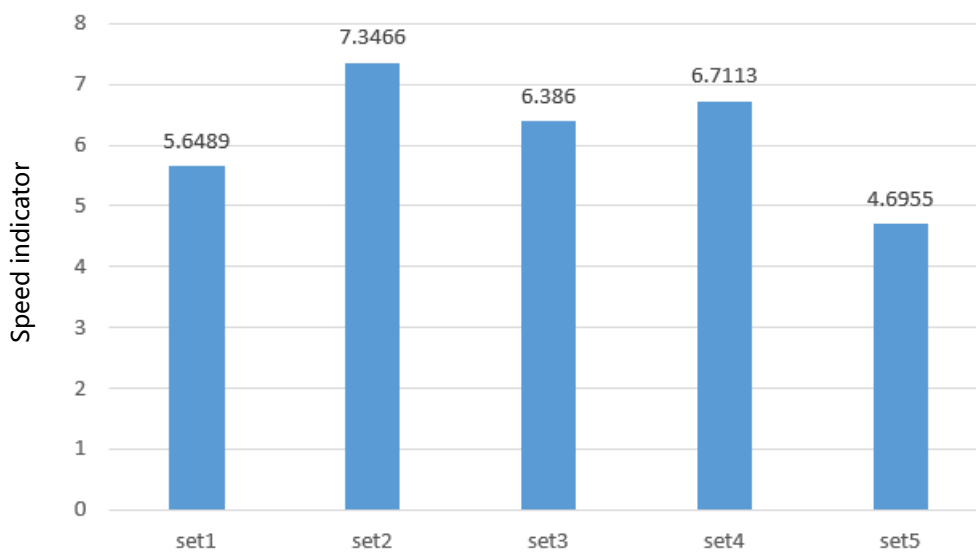


Figure 14 The speed performance indicator (RMSE) for each sets of parameters

After running the simulation in SUMO, their corresponding performance indicators are

shown in Figure 14, 15 and 16. As can be seen from Figure 14, the model with the default values performs well in terms of speed simulation. The selected parameters will more or less increase the error in speed simulation. But among all 5 sets, the difference between them is relatively small.

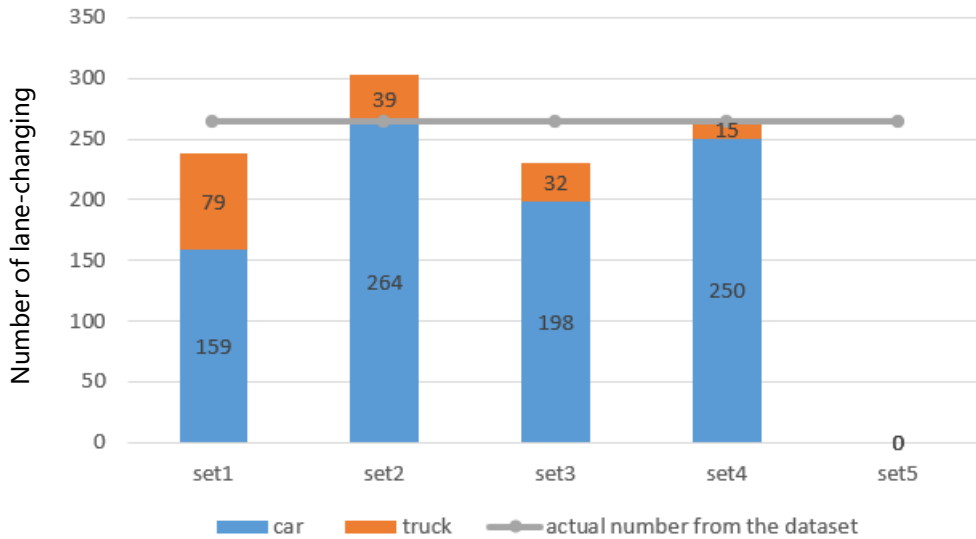


Figure 15 The number of lane-changing for each sets of parameters

As Figure 15 shown, the default model performs very poorly when simulating lane changes. The simulated vehicles did not conduct any lane-changing operations while in the realistic data, there are many lane-changings. In the remaining four sets of parameters, the number of lane changes simulated by set 4 is the closest to the actual number of lane changes, and the other three sets have obvious differences.

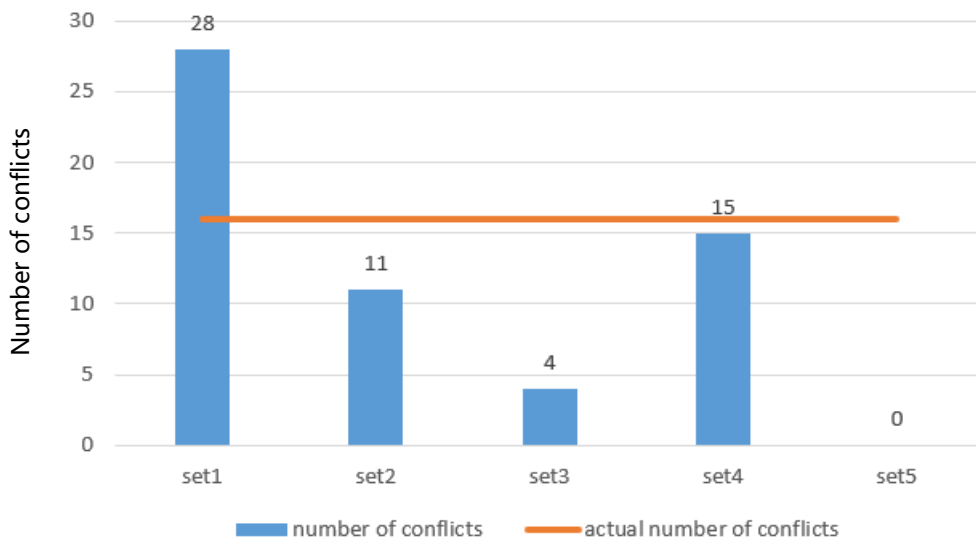


Figure 16 The number of conflicts for each sets of parameters

It can be seen from Figure 16 that in terms of simulated conflicts, the model with default parameters still performs poorly, similar to the situation when simulating lane changes, it does not record any conflicts, which is very inconsistent with the actual situation. The results of other four sets are very different. Set 2 and set 4 are closer to the real situation, while set 1 has too many conflicts and set 3 has too few.

This result proves that if the default model is used for heterogeneous traffic simulation without modification, it will inevitably bring large differences due to the poor performance in terms of lane-changing and safety aspects. Compared with the default model, all of the four sets of selected parameters have reduced the errors in lane-changing and conflicts to a certain extent, and the result of set 4 is most consistent with the realistic data in terms of lane-changing and safety. In other words, set 4 achieves a significant reduction in differences in terms of lane-changing and conflicts without adding large difference in speed simulation.

In general, the calibration results prove that by selecting appropriate parameters, the errors in traffic simulation can be effectively minimized, thus achieving the goal of modifying the model. Through the calibration process, an appropriate set of parameters for simulation human-operated vehicles is obtained and it is going to be used for subsequent heterogeneous traffic simulation.

3.3 Heterogeneous traffic simulation

3.3.1 Simulation settings

After the modified and calibrated model is obtained, it can be used in the heterogeneous traffic simulation to explore the impact of autonomous vehicles on traffic safety and efficiency.

In the simulation, the selected road segment is a straight 4-lane segment with a length of 420 meters, which is consistent with the segment used in calibration. There are two reasons for this choice. One reason is that the modified model has been verified in the calibration process that its performance on this segment is very close to the real situation. The second reason is that although this is a quite simple road scenario, the traffic on it includes all the traffic situations needed to be considered, including car-following situation, lane-changing situation and potential conflicts.

The traffic load in the simulation is set to 4000 vehicles for each run. A python script is written to randomly generate the traffic load. By inputting the desired penetration rate of autonomous vehicles, the script can generate vehicles according to the rate. The output consists of 4000 vehicles, with randomly selected vehicle types, random initial speed, random initial lane and driving direction. These data are output according to a pre-set format, so that they can be directly applied to the simulation inputs. Among all the autonomous vehicles and human-operated vehicles, the ratio of trucks to cars is set to 1:4, and this value is extracted from the highD dataset. For instance, when the input penetration rate of autonomous vehicles is 0.7, the generated traffic load will include 480 human-operated cars, 120 human-operated trucks, 1,120 autonomous cars, and 280 autonomous trucks. In order to eliminate the possible errors caused by randomly generated traffic flow, each penetration rate will go through 100 runs.

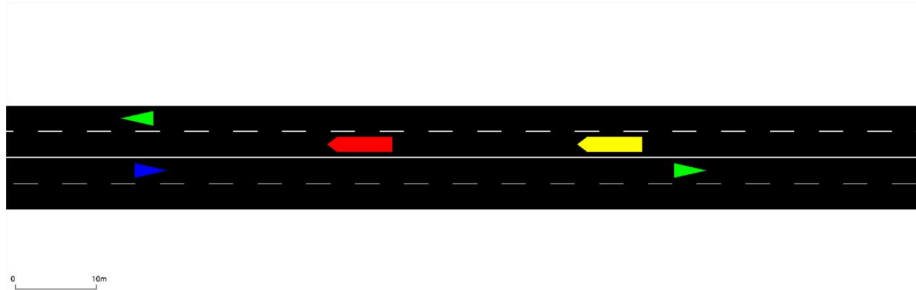


Figure 17 A screenshot of the heterogeneous traffic simulation in SUMO (blue: conventional car; green: autonomous car; red: conventional truck; yellow: autonomous truck)

3.3.2 Simulation outputs

As mentioned in the Section 2.1.1, there are several output files after each run of simulation, including the lane-changing file, SSMs file and a full-output file containing information about every edge, lane, vehicle and traffic light for each time step. From these files, necessary information needs to be extracted to evaluate the safety and efficiency of traffic.

The traffic safety is evaluated using the number of conflicts. In this thesis, TTC and DRAC are the criteria used to identify potential conflicts. The threshold for TTC is 3.0 seconds, and for DRAC 3.4 m/s^2 . Every time the actual value exceeds the threshold, a record is written in the SSMs file. So, it is easy to get the number of conflicts.

While obtaining the number of conflicts, some other information can also be acquired to evaluate the efficiency of traffic. In this case, it is the number of lane-changing and traffic flow value. The number of lane-changing can be directly counted in lane-changing file, and the traffic flow value T can be calculated using the following formula:

$$T = \frac{4000}{\text{simulation time}} \times 3600 \text{ (vehicle per hour)} \quad (3.2)$$

The simulation time can be extracted from the full-output file.

4 Results

In this chapter, the results of the simulation are displayed, including the impacts of the two autonomous vehicle models mentioned above on traffic safety and efficiency at various penetration rates, and the comparison of their results. The possible reasons for the results are also discussed in this chapter.

4.1 Results of zero-error IDM

The heterogeneous traffic simulation requires both traditional human-operated vehicle model and autonomous vehicle model. When the human-operated vehicle model is the modified IDM and the autonomous vehicle model is zero-error IDM, the results are shown as following figures.

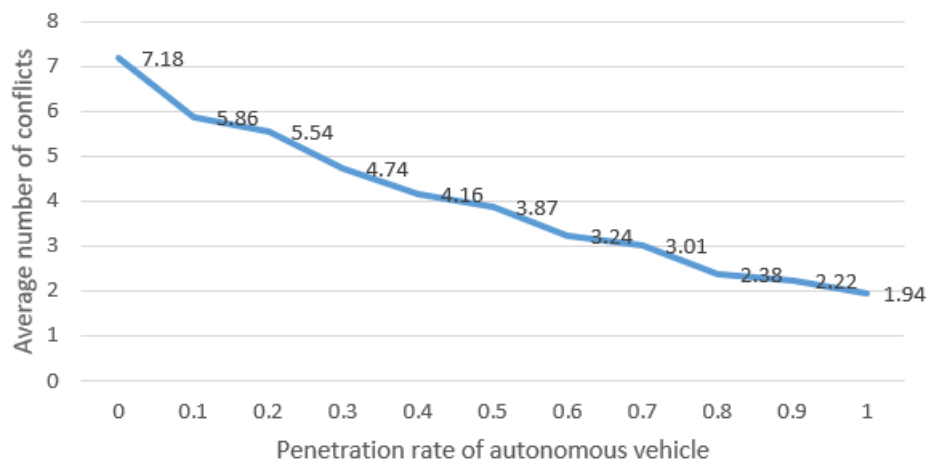


Figure 18 Average number of potential conflicts identified by TTC (zero-error IDM)

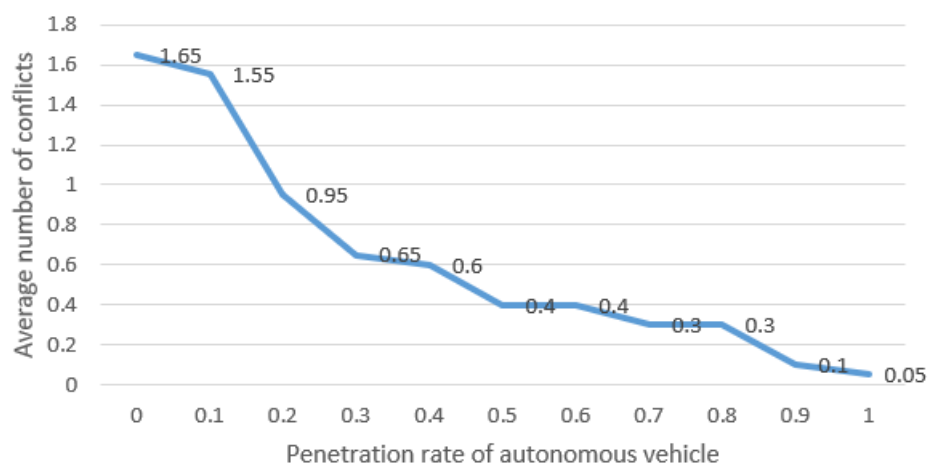


Figure 19 Average number of potential conflicts identified by DRAC (zero-error IDM)

It can be seen from the above two figures that when the applied autonomous vehicle model is zero-error IDM, as the penetration rate of autonomous vehicle increases, the number of potential conflicts is gradually decreasing, whether it is data obtained through TTC or DRAC. It means that with more autonomous vehicles in traffic, traffic safety has been gradually improved.

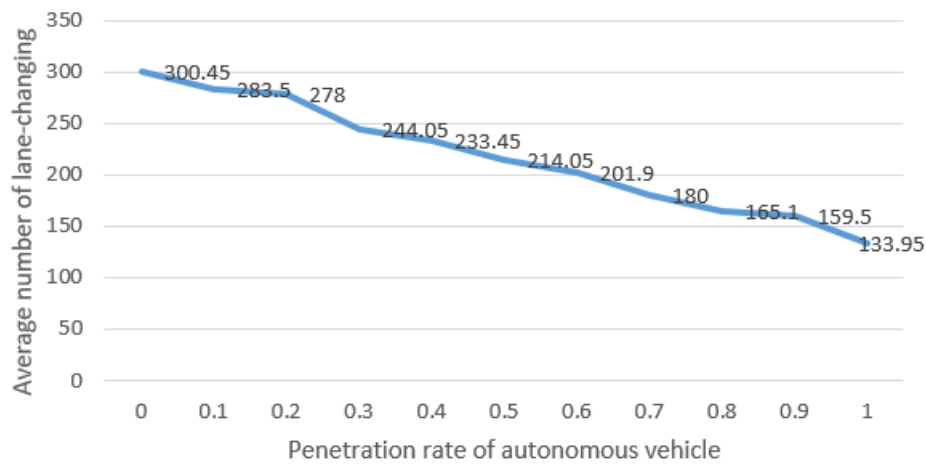


Figure 20 Average number of lane-changing (zero-error IDM)

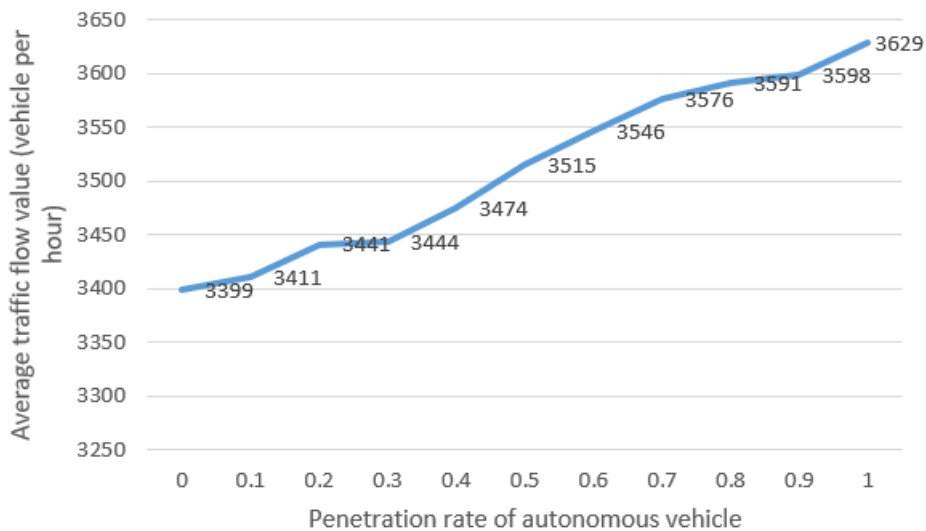


Figure 21 Average traffic flow value (zero-error IDM)

Figure 20 shows that as the penetration rate increases, the number of lane-changing in traffic is gradually reduced. Combined with the continuous increase in traffic flow value shown in Figure 21, it can be found out that when there are more autonomous vehicles in traffic, the traffic efficiency is higher.

4.2 Results of CACC model

When the implemented autonomous vehicle model is changed to the CACC model and run all the simulations again, the results obtained are as follows.

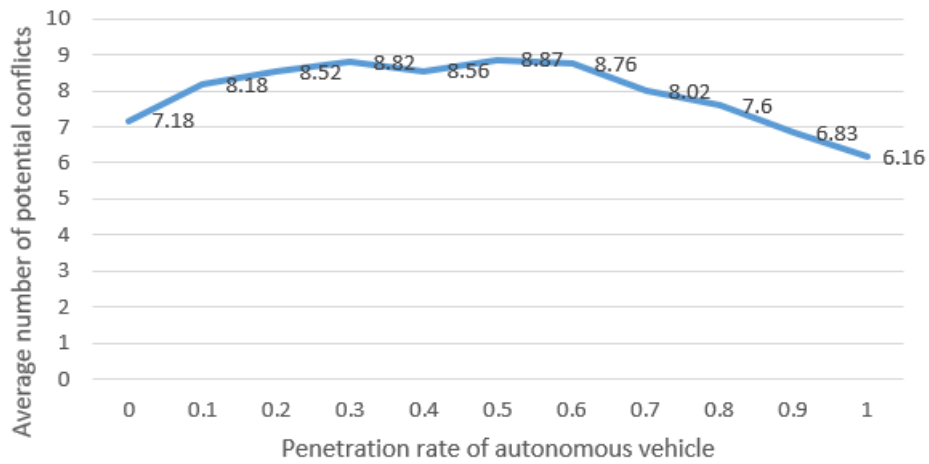


Figure 22 Average number of potential conflicts identified by TTC (CACC)

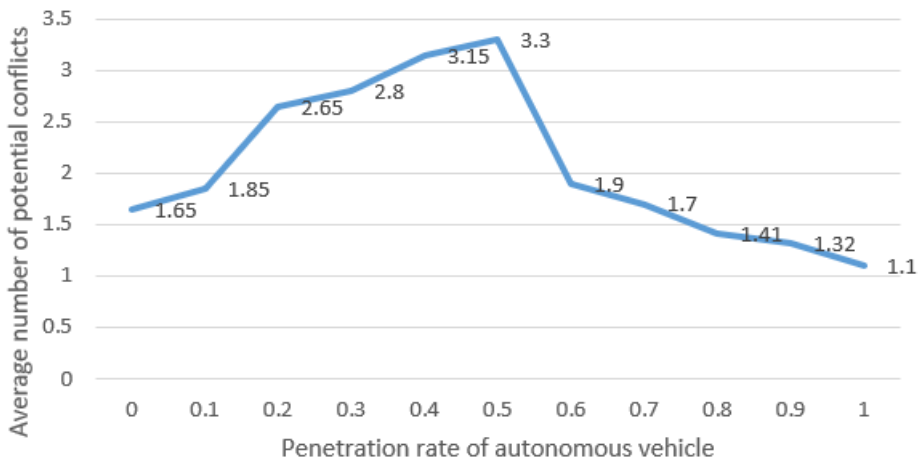


Figure 23 Average number of potential conflicts identified by DRAC (CACC)

It can be seen from Figure 22 and Figure 23 that when the penetration rate is within a low level, the existence of autonomous vehicles does not improve traffic safety, but on the contrary makes it worse. As the penetration rate increases, the degree of this deterioration on safety is still increasing. This deterioration is gradually eliminated after the penetration rate is higher than 0.5. Compared to the initial safety state when there are no autonomous vehicles, traffic safety improves only when the penetration rate exceeds 0.8 or 0.9.

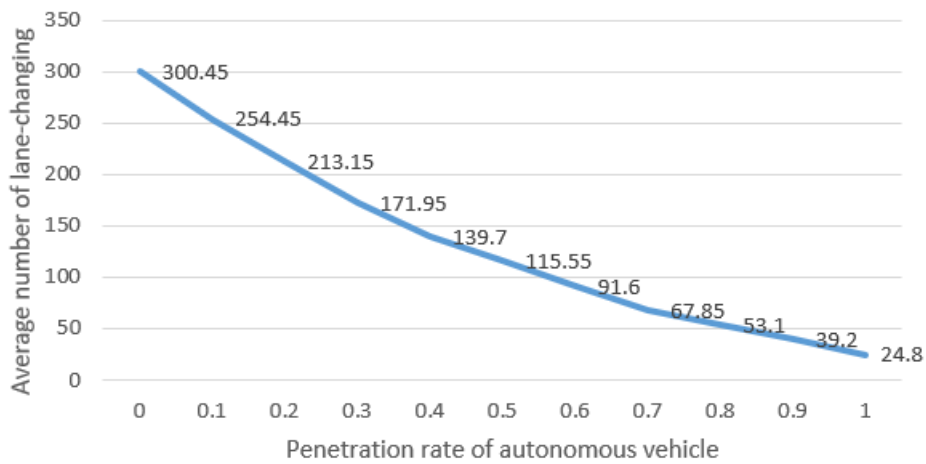


Figure 24 Average number of lane-changing (CACC)

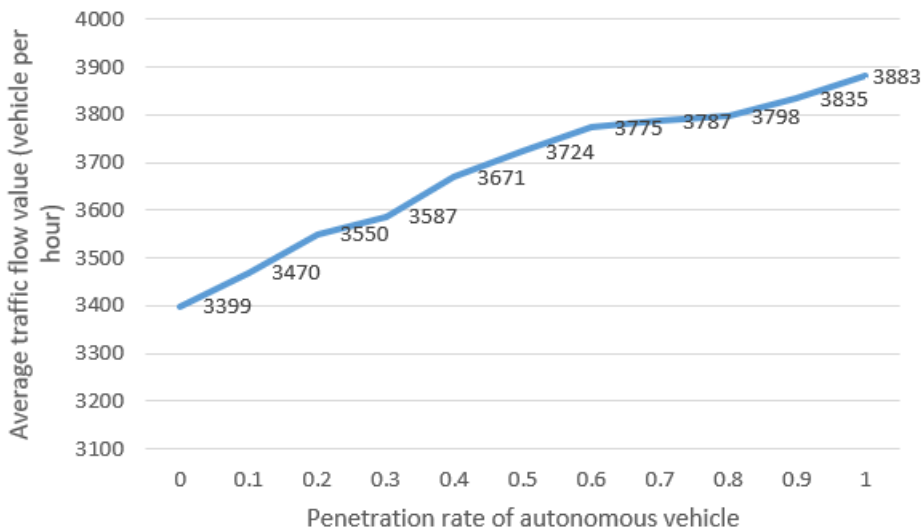


Figure 25 Average traffic flow value (CACC)

Combining the information expressed in Figure 24 and Figure 25, the autonomous vehicles represented by the CACC model can greatly improve traffic efficiency, and as the penetration rate increases, traffic efficiency continues to improve.

4.3 Results comparison

As introduced in Section 3.1.3, The two models zero-error IDM and CACC model can be regarded as representing two types of autonomous vehicles. The former emphasizes the improvement of each vehicle's own independent driving ability without involving communication between autonomous vehicles. The latter not only improves the driving strategy, but also adds V2V and V2I communication, which means that the vehicle will be affected by other autonomous vehicles. When putting the simulation results of the two autonomous vehicle models together, the differences between them can clearly be seen.

In the following figures, each contains two solid lines connecting all the mean values and box plots based on all the result data. The former help us see the trend, and the latter can help us better observe the consistency of the results. The box represents the interquartile range (IQR, 25th percentile to 75th percentile), the middle line in the box represents the median, and the whiskers represent $1.5 \times \text{IQR}$.

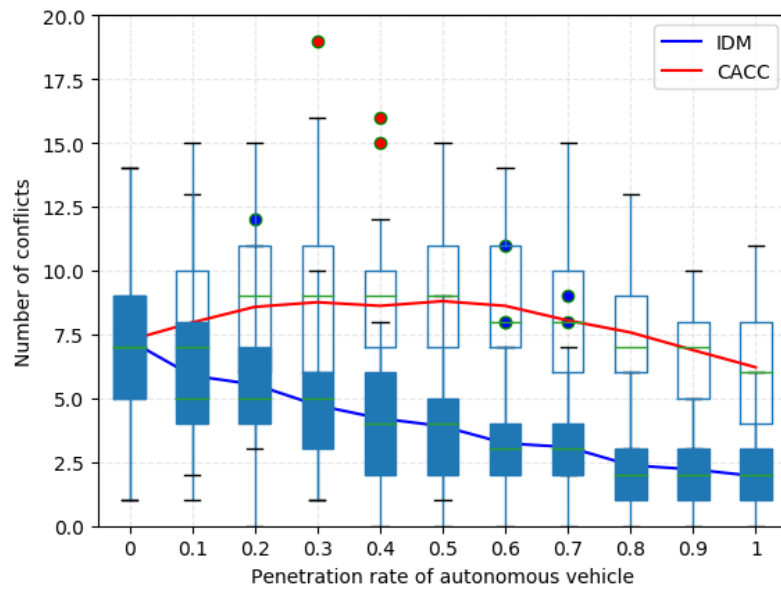


Figure 26 Box plot and mean value (blue and red lines) of number of potential conflicts identified by TTC

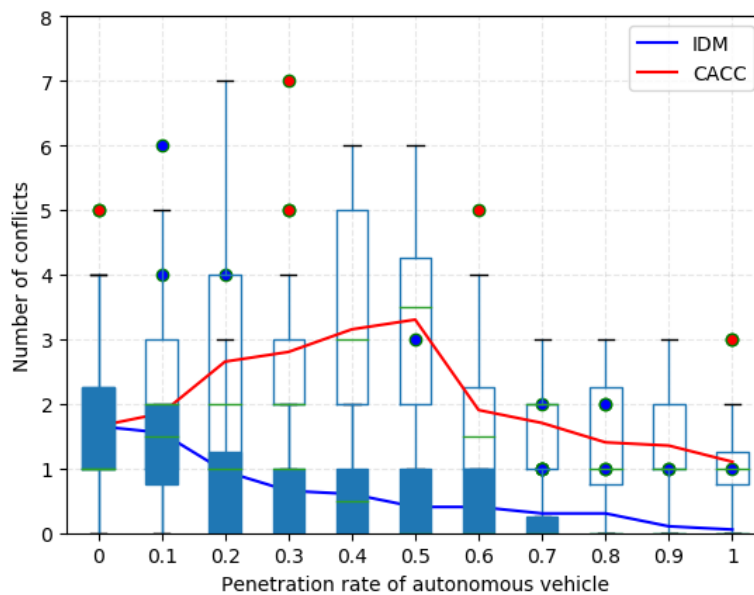


Figure 27 Box plot and mean value (blue and red lines) of number of potential conflicts identified by DRAC

From Figures 26 & 27, it's obvious that the zero-error IDM is much better than the CACC model in improving traffic safety. No matter when the penetration rate is low or high, the traffic safety of IDM is far better than that of CACC model. Especially when the penetration rate is in a high level, the IDM can reduce the number of potential conflicts to a very low level, while the CACC model present a slightly better result than the initial state without any autonomous vehicles. From the box plots, the distributions of both are relatively extensive, which reflects the inconsistency of the results.

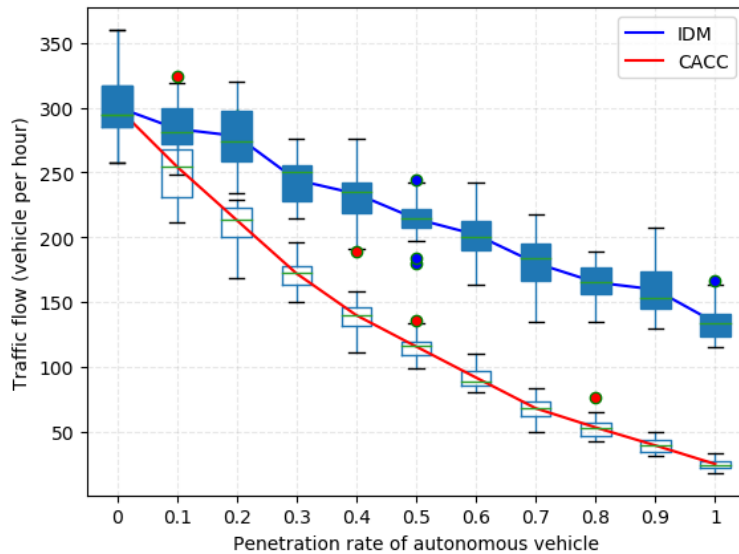


Figure 28 Box plot and mean value (blue and red lines) of number of lane-changing

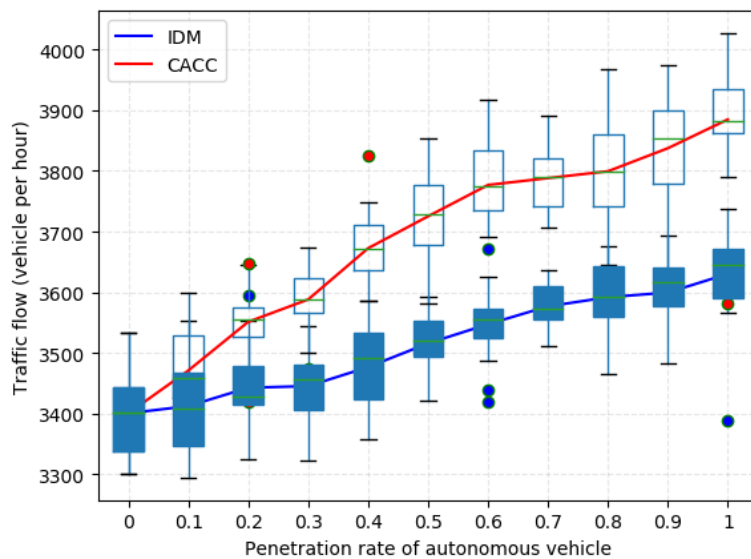


Figure 29 Box plot and mean value (blue and red lines) of traffic flow

Although the CACC model is not effective in improving traffic safety, it is far superior to zero-error IDM in improving traffic efficiency. With the increase in penetration rate, although both models help to continuously improve traffic efficiency, the CACC model always maintains a greater advantage.

From the perspective of the consistency of the results, it can be seen from Figures 28 & 29 that in terms of lane-changing, the performance of the CACC model is much more consistent than that of the zero-error IDM, but from the perspective of traffic flow, both are widely distributed.

One possible reason for the difference in traffic safety and efficiency between the two models is that the CACC model can be influenced by other autonomous vehicles because of the existence of V2V communication. When the penetration rate of autonomous vehicles is low, there are not many autonomous vehicles in traffic, so V2V communication may not be much, and it may be affected by traditional human-operated vehicles in the surrounding environment. This leads to a deterioration on traffic safety. Until the penetration rate of autonomous vehicles exceeded half, V2V communication begins to become more frequent and the interference received is reduced, so it begins to play a positive role in improving traffic safety. The excellent performance of the CACC model in improving traffic efficiency can also be explained by this reason. Due to the existence of V2V communication, autonomous vehicles can obtain additional information from other vehicles, avoiding a lot of unnecessary lane-changing or overtaking manoeuvre, thus improving traffic efficiency.

As for zero-error IDM, because it avoids possible driving errors of human drivers, it can improve traffic safety and efficiency. And it is not affected by other vehicles, so even if the penetration rate of autonomous vehicle is very low, it can still exert a positive effect. The higher the penetration rate, the more significant the positive effect.

5 Discussion and conclusion

This chapter collects all the findings and answers the research questions raised at the beginning of the report. The discussion about the results are also included.

5.1 Research questions

At the beginning of this thesis, four research questions are asked, expected to be answered at the end, the following are the questions:

1. Based on existing research, how do the types of vehicle involved in traffic impact drivers' driving behaviours?
2. How are these summarized impacts reflected in the models used for heterogeneous traffic simulation?
3. How does the existence of autonomous vehicles influence the safety of heterogeneous traffic? Do different penetration rates bring different results?
4. How does the existence of autonomous vehicles influence the efficiency of heterogeneous traffic? Do different penetration rates bring different results?

5.1.1 How do the types of vehicle involved in traffic impact drivers' driving behaviours?

Through sorting out and researching existing related literatures and studies, we found that most researchers can reach a consensus that the types of vehicle involved in traffic will have an impact on the driving behaviours of drivers. But as to how they affect driving behaviours, different researchers hold different views.

In this thesis, the research objects are limited to cars and trucks. Regarding the car-following situation, some researchers observe that truck drivers tend to maintain a larger THW than car drivers, while others hold the opposite view. Instead of directly adopting the view of one side, we use the existing data in the highD dataset for verification. Firstly, the car-following situations are extracted from the dataset and then analysed to check the THW that different drivers tend to maintain. The final conclusion is that truck drivers do tend to have a larger THW than car drivers in car-following situations.

As for lane-changing situation, most researchers observe that truck drivers tend to do fewer lane-changing manoeuvre, and the data extracted from the highD dataset also verifies this conclusion.

5.1.2 How are these summarized impacts reflected in the models?

There are many ways in SUMO to make the default models change according to the user's ideas, including adding new parameters, modifying the principles behind the models, and setting more reasonable values for current parameters, etc. Through the analysis of the existing models and their parameters, we believe that setting reasonable values for the existing parameters is most suitable for us to reflect these summarized impacts, considering the constraint of time. The specific parameter settings include setting a larger Tau value for truck drivers in the car-following model, setting a larger value of lcAssertive and lcSpeedGain for car drivers in the lane-changing model, as mentioned in Section 2.2.

After the modification, we used the data in the highD dataset to perform a calibration to ensure that it reflects these impacts while maintaining its simulation accuracy.

5.1.3 How does the existence of autonomous vehicles influence the safety of heterogeneous traffic?

From the results, autonomous vehicles that do not include V2V communication can improve traffic safety, and as the penetration rate increases, traffic safety can be continuously improved. For autonomous vehicles with V2V communication functions, when they first appear in traffic, they may have a negative effect on traffic safety. As the penetration rate increases, this negative impact first increases. After the rate exceeds 0.5, the negative impact begins to decrease, when the participation rate reaches a high level (over 0.7), it begins to turn into a positive impact on traffic safety.

5.1.4 How does the existence of autonomous vehicles influence the efficiency of heterogeneous traffic?

Both types of autonomous vehicles play a positive role in improving traffic efficiency. It is worth mentioning that autonomous vehicles with V2V communication have a more significant effect on improving traffic efficiency.

5.2 Discussion

Judging from the overall results, the implementation of autonomous vehicles on the road is very anticipated. Whether from the perspective of traffic safety or traffic efficiency, this move will greatly improve today's traffic. The simulation results tell us that as long as the correct type of autonomous vehicle is selected, even if there are only a small number of autonomous vehicles on the road at the beginning, it can still help to improve traffic safety and efficiency.

In addition, these simulation results can also help us when developing plans to put autonomous vehicles on the road step by step. For example, V2V communication function can be putted into use after the proportion of autonomous vehicles in traffic reaches a certain level (for example, over 0.7), to obtain a better effect and avoid possible negative influences.

5.2.1 Ethical aspect

This study provides more realistic models for the simulation of heterogenous traffic and helps to get more accurate safety and efficiency evaluation results. With the help of these results, researchers and companies could design the active safety functions in a better way and these results can also help people to think when is the right time to put autonomous vehicles on the road. But on the other hand, these positive results may make researchers or companies more aggressively want to put autonomous vehicles on the road, which may cause serious safety problems or legal issues. Because even the most perfect simulation cannot fully represent the actual situation, the results of the simulation can only be used as a reference, and should not be used as evidence to prove that people have fully understood the consequences of putting autonomous vehicles on the road.

5.2.2 Limitations

After reviewing the whole process of this thesis, we can find that there are some limitations. First of all, due to time and computation constraints, not all the data in the

highD dataset is used in the calibration phase. Only a part of the data is used as the input of the simulation, and the simulation result is compared with the actual data of this part. Since the setting of parameters is based on the data obtained from the entire highD dataset, errors will inevitably occur if not all the data are used in the calibration process. Another limitation lies in the setting of simulation parameters. We only selected a limited set of parameters as candidates, and then selected the set of parameters with the best results among them. Considering the limited time and energy, this is an effective method to a certain extent, but this method can only allow us to get a set of relatively good parameters, not the best. It is suitable for observing approximate trends when the possible results are not known, but cannot be used for accurate statistics. In addition, the selected method of modifying the model is also one of the limitations. In this case, we chose to modify the models by setting appropriate parameter values. This is a choice made for simplicity and convenience. But this method does not fully reflect the difference between the four combinations (CC, CT, TC and TT). In other words, the models used in the final simulation have only been improved to a certain extent.

5.2.3 Future work

Regarding the limitations, we know there are still several points that can be further improved. First of all, if we use all the data from highD dataset as input of simulation in the calibration phase, We can select a set of parameters that can make the simulation results more consistent with the actual situation. In addition, the method of selecting parameters can also be improved. If there are more time and energy, genetic algorithm could be considered. This algorithm is very suitable for finding the optimal solution, in our case, to find the best set of parameters. Moreover, in the final part, the heterogeneous traffic simulation, this thesis only considers the penetration rate from 0 to 1 with a step length of 0.1. With each rate, 100 runs of simulation are performed, and each run contains 4000 vehicles. If the number of runs and the number of vehicles in each run are further increased, then the accuracy of the results are expected to be improved.

Furthermore, in this thesis, we choose the way of setting more reasonable values for current parameters to reflect those summarized impacts. In this case, we can distinguish the different driving behaviours of the truck driver and the car driver, but we cannot make the driver react differently according to the type of preceding vehicle. To further reflect this difference, the other two ways can be chosen, namely adding new parameters, or directly changing the principle behind the models. Through these two methods, we can set drivers to maintain different THW when following different types of vehicles, so as to simulate the driving behavior of drivers more realistically. In fact, we do make some preliminary attempts based on this idea. A new model called 'heterogeneous', that has been proposed outside this thesis project, was explored during the thesis project. In the new model, we can set the four new parameters of `heterogeneous_cc_thw`, `heterogeneous_ct_thw`, `heterogeneous_tc_thw` and `heterogeneous_tt_thw`, so that different drivers can give more realistic reactions when following different types of vehicles. The subsequent steps are the same as the method used in this report. After calibration, the new model can be used for heterogeneous traffic simulation. But in the process of doing so, many new problems have also appeared, such as the problem of how to choose the appropriate values of these four mentioned parameters. Therefore further work to solve such problems is needed before further exploration with the new model can be done.

References

- [1] Litman, T. (2017). Autonomous vehicle implementation predictions. Victoria, Canada: Victoria Transport Policy Institute.
- [2] Roesener, C., Sauerbier, J., Zlocki, A., Fahrenkrog, F., Wang, L., Varhelyi, A., & de Gelder, E. (2017). A Comprehensive Evaluation Approach for Highly Automated Driving. 25th Enhanced Safety of Vehicles 2017, 1–13.
- [3] Cunto, F. (2008). Assessing Safety Performance of Transportation Systems using Microscopic Simulation. *Conflict* (p. 190).
- [4] VISSIM 7 User Manual, PTV AG, Karlsruhe, Germany, 2015.
- [5] Vanderschuren, M. (2008). Safety improvements through Intelligent Transport Systems: A South African case study based on microscopic simulation modelling. *Accident Analysis and Prevention*, 40(2), 807–817.
- [6] Duncan, G. I. (1995). PARAMICS wide area microscopic simulation of ATT and traffic management. In *Proceedings of the 28th International Symposium on Automotive Technology and Automation (ISATA)*.
- [7] Huguenin, F., Torday, A., & Dumont, A. (2005, March). Evaluation of traffic safety using microsimulation. In *Proceedings of the 5th Swiss Transport Research Conference–STRC, Ascona, Swiss*.
- [8] TSS-Transport Simulation Systems. Aimsun. version 6.1 ed. J. Barceló, Ed. Barcelona. p. Traffic simulation software (2010)
- [9] Lopez, P. A., Behrisch, M., Bieker-Walz, L., Erdmann, J., Flotterod, Y. P., Hilbrich, R., Wiebner, E. (2018). Microscopic Traffic Simulation using SUMO. In *IEEE Conference on Intelligent Transportation Systems, Proceedings, ITSC (Vol. 2018-h*
- [10] German Aerospace Center, Institute of Transportation Systems. sourceforge.net. [Online].; 2011 [cited 2012 February 16. Available from: http://sumo.sourceforge.net/doc/current/docs/userdoc/SUMO_User_Documentation.html.
- [11] Mahmud, S. M. S., Ferreira, L., Hoque, M. S., & Tavassoli, A. (2019, April 1). Microsimulation modelling for traffic safety: A review and potential application to heterogeneous traffic environment. *IATSS Research*. Elsevier B.V.
- [12] Ronaldo, A., Ismail, M. (2012). Comparison of the two micro- simulation software Aimsun & Sumo for highway traffic modelling. (Master's thesis, Linköping University, Department of Science and Technology)

- [13] Rao, P. R. V. (2011). Increasing Accuracy of Simulation Modeling via a Dynamic Modeling Approach (Master's Theses, University of Tennessee)
- [14] Aghabayk, K., Sarvi, M., Forouzideh, N., & Young, W. (2013). New Car-Following Model considering Impacts of Multiple Lead Vehicle Types. *Transportation Research Record: Journal of the Transportation Research Board*, 2390(1), 131–137.
- [15] Eagely, S. K. A., Sarvi, M., Young, W., & Wang, Y. (2012). Investigating heavy vehicle interactions during the car following process. In *Transportation Research Board (USA) Annual Meeting 2012* (pp. 1-13). Transportation Research Board.
- [16] Moridpour, S., Sarvi, M., & Rose, G. (2007). Modelling lane changing behaviour of heavy commercial vehicles. *30th Australasian Transport Research Forum*, 1–15.
- [17] Definition of Vehicles, Vehicle Types, and Routes. (n.d.). Retrieved Feb 1, 2020, from SUMO website: http://sumo.sourceforge.net/userdoc/Definition_of_Vehicles_Vehicle_Types,_and_Routes.html
- [18] Pourabdollah, M., Bjarkvik, E., Furer, F., Lindenberg, B., & Burgdorf, K. (2018). Calibration and evaluation of car following models using real-world driving data. *IEEE Conference on Intelligent Transportation Systems, Proceedings, ITSC, 2018-March (October)*, 1–6.
- [19] Krauß, S. (1998). *Microscopic modeling of traffic flow: Investigation of collision free vehicle dynamics* (Doctoral dissertation).
- [20] Gettman, D., & Head, L. (2003). Surrogate Safety Measures from Traffic Simulation Models. In *Transportation Research Record* (pp. 104–115). National Research Council.
- [21] Laureshyn, A., & Varhelyi, A. (2018, April 4). The Swedish Traffic Conflict Technique: observer's manual. Retrieved from <http://lup.lub.lu.se/record/fc968207-ed2b-49cc-9346-57d8c74cac41>
- [22] Tarko, A. P. (2018). Surrogate measures of safety. In *Transport and Sustainability* (Vol. 11, pp. 383–405). Emerald Group Publishing Ltd.
- [23] Goh, K., Currie, G., Sarvi, M., & Logan, D. (2014). Experimental microsimulation modeling of road safety impacts of bus priority. *Transportation Research Record*, (2402), 9–18.
- [24] Morando, M. M., Tian, Q., Truong, L. T., & Vu, H. L. (2018). Studying the Safety Impact of Autonomous Vehicles Using Simulation-Based Surrogate Safety Measures. *Journal of Advanced Transportation*, 2018.

- [25] Luong, E. (2018). Microscopic Car Following Models Simulation Study. UC Irvine. ProQuest ID: Luong_uci_0030M_15394. Merritt ID: ark:/13030/m56m85fc. Retrieved from <https://escholarship.org/uc/item/8tx731bq>
- [26] Zheng, Z. (2014). Recent developments and research needs in modeling lane changing. *Transportation Research Part B: Methodological*, 60, 16–32. <https://doi.org/10.1016/j.trb.2013.11.009>
- [27] Treiber, M., Hennecke, A., & Helbing, D. (2000). Congested traffic states in empirical observations and microscopic simulations. *Physical Review E - Statistical Physics, Plasmas, Fluids, and Related Interdisciplinary Topics*, 62(2), 1805–1824. <https://doi.org/10.1103/PhysRevE.62.1805>
- [28] Xiao, L., Wang, M., & van Arem, B. (2017). Realistic Car-Following Models for Microscopic Simulation of Adaptive and Cooperative Adaptive Cruise Control Vehicles. *Transportation Research Record: Journal of the Transportation Research Board*, 2623, pp. 1–9. <https://doi.org/10.3141/2623-01>
- [29] Erdmann, J. (2015). SUMO's lane-changing model. In *Modeling Mobility with Open Data* (pp. 105-123). Springer, Cham.
- [30] Saifuzzaman, M., & Zheng, Z. (2014). Incorporating human-factors in car-following models: A review of recent developments and research needs. *Transportation Research Part C: Emerging Technologies*, 48, 379–403.
- [31] Song, J., Wu, Y., Xu, Z., & Lin, X. (2015). Research on car-following model based on SUMO. *Proceedings of 2014 IEEE 7th International Conference on Advanced Infocomm Technology, IEEE/ICAIT 2014*, January, 47–55.
- [32] Pourabdollah, M., Bjarkvik, E., Furer, F., Lindenberg, B., & Burgdorf, K. (2018). Calibration and evaluation of car following models using real-world driving data. *IEEE Conference on Intelligent Transportation Systems, Proceedings, ITSC, 2018-March (October)*, 1–6. <https://doi.org/10.1109/ITSC.2017.8317836>
- [33] Milanés, V., & Shladover, S. E. (2014). Modeling cooperative and autonomous adaptive cruise control dynamic responses using experimental data. *Transportation Research Part C: Emerging Technologies*, 48, 285–300.
- [34] Car-Following-Models/CACC (n.d.). Retrieved Feb 1, 2020, from SUMO website: <https://sumo.dlr.de/docs/Car-Following-Models/CACC.html>
- [35] Krajzewicz D. (2010) Traffic Simulation with SUMO – Simulation of Urban Mobility. In: Barceló J. (eds) *Fundamentals of Traffic Simulation*. International Series in Operations Research & Management Science, vol 145. Springer, New York, NY

- [36] Sayer, J. R., Mefford, M. L., & Huang, R. W. (2005). The Effects of Lead-Vehicle Size on Driver Following Behavior: Is Ignorance Truly Bliss? (pp. 221–225). The University of Iowa.
- [37] Ye, F., & Zhang, Y. (2009). Vehicle type-specific headway analysis using freeway traffic data. *Transportation Research Record*, (2124), 222–230.
- [38] Hoogendoorn, S. P., & Bovy, P. H. L. (2000). A Closer Examination of Overtaking Prohibitions for Trucks. *IFAC Proceedings Volumes*, 33(9), 227–234.
- [39] Aghabayk, K., Sarvi, M., & Young, W. (2012). Understanding the dynamics of heavy vehicle interactions in car-following. *Journal of Transportation Engineering*, 138(12), 1468–1475.
- [40] Sarvi, M. (2013). Heavy commercial vehicles-following behavior and interactions with different vehicle classes. *Journal of Advanced Transportation*, 47(6), 572–580.
- [41] Morando, M. M., Tian, Q., Truong, L. T., & Vu, H. L. (2018). Studying the Safety Impact of Autonomous Vehicles Using Simulation-Based Surrogate Safety Measures. *Journal of Advanced Transportation*, 2018.
- [42] Moridpour, S., Rose, G., & Sarvi, M. (2010). The effect of surrounding traffic characteristics on lane changing behaviour. *Journal of Transportation Engineering*, 136(11), 1-12.
- [43] Aghabayk, K., Moridpour, S., Young, W., Sarvi, M., & Wang, Y.-B. (2011). Comparing Heavy Vehicle and Passenger Car Lane-Changing Maneuvers on Arterial Roads and Freeways. *Transportation Research Record: Journal of the Transportation Research Board*, 2260(1), 94–101.
- [44] Krajewski, R., Bock, J., Kloeker, L., & Eckstein, L. (2018). The highD Dataset: A Drone Dataset of Naturalistic Vehicle Trajectories on German Highways for Validation of Highly Automated Driving Systems. *IEEE Conference on Intelligent Transportation Systems, Proceedings, ITSC, 2018-Novem (November)*, 2118–2125.
- [45] Amundsen, F. H., & Hyden, C. (1977). *Proceedings of first workshop on traffic conflicts*. Oslo, TTI, Oslo, Norway and LTH Lund, Sweden.
- [46] Society of Automotive Engineers (2013). *Operational Definitions of Driving Performance Measures and Statistics (SAE Recommended Practice J2944)*, Warrendale, PA: Society of Automotive Engineers.
- [47] Vehicle Type Parameter Defaults (n.d.). Retrieved Feb 1, 2020, from SUMO website: https://sumo.dlr.de/docs/Vehicle_Type_Parameter_Defaults.html

- [48] Treiber, M., & Kesting, A. (2010). An Open-Source Microscopic Traffic Simulator. *IEEE Intelligent Transportation Systems Magazine*, vol. 2, no. 3, pp. 6-13.
- [49] Andreotti, E., Baykas, P.E., Selpi, S. (2020). Safety-centred analysis of transition stages to traffic with fully autonomous vehicles. *IEEE Conference on Intelligent Transportation Systems, Proceedings, ITSC*.
- [50] Mintsis, E., Luecken, L., Porfyri, K., Koutras, D., Zhang, X., Rondinone, M., & Mitsakis, E. (2019). TransAID Deliverable 3.1-Modelling, simulation and assessment of vehicle automations and automated vehicles' driver behaviour in mixed traffic. Retrieved July 30,2020, from www.transaid.eu/deliverables/
- [51] European Commission. (2012). implementing Regulation (EC) No 661/2009 of the European Parliament and of the Council with respect to type-approval requirements for certain categories of motor vehicles with regard to advanced emergency braking systems. (Commission Regulation (EU) No 347/2012 of 16 April 2012). Retrieved from <https://op.europa.eu/en/publication-detail/publication/ac59c0ef-09ee-4e38-b1d4-1f3521a28a13>
- [52] Driver state. (n.d.). Retrieved Feb 1, 2020, from SUMO website: https://sumo.dlr.de/docs/Driver_State.html

Appendix

Table A. Parameters of the Driver State Device in SUMO. The information are taken directly from [52]

Parameter	Type	Description
initialAwareness	float	The initial awareness assigned to the driver state.
errorTimeScale-Coefficient	float	Time scale constant that controls the time scale of the underlying error process.
errorNoiseIntensity-Coefficient	float	Noise intensity constant that controls the noise intensity of the underlying error process.
speedDifferenceError-Coefficient	float	Scaling coefficient for the error applied to the speed difference input of the car-following model.
headwayErrorCoefficient	float	Scaling coefficient for the error applied to the distance input of the car-following model.
speedDifferenceChange-PerceptionThreshold	float	Constant controlling the threshold for the perception of changes in the speed difference
speedDifferenceChange-PerceptionThreshold	float	Constant controlling the threshold for the perception of changes in the distance input.

DEPARTMENT OF MECHANICS AND MARITIME SCIENCES
DIVISION OF VEHICLE SAFETY
CHALMERS UNIVERSITY OF TECHNOLOGY
Gothenburg, Sweden
www.chalmers.se



CHALMERS
UNIVERSITY OF TECHNOLOGY