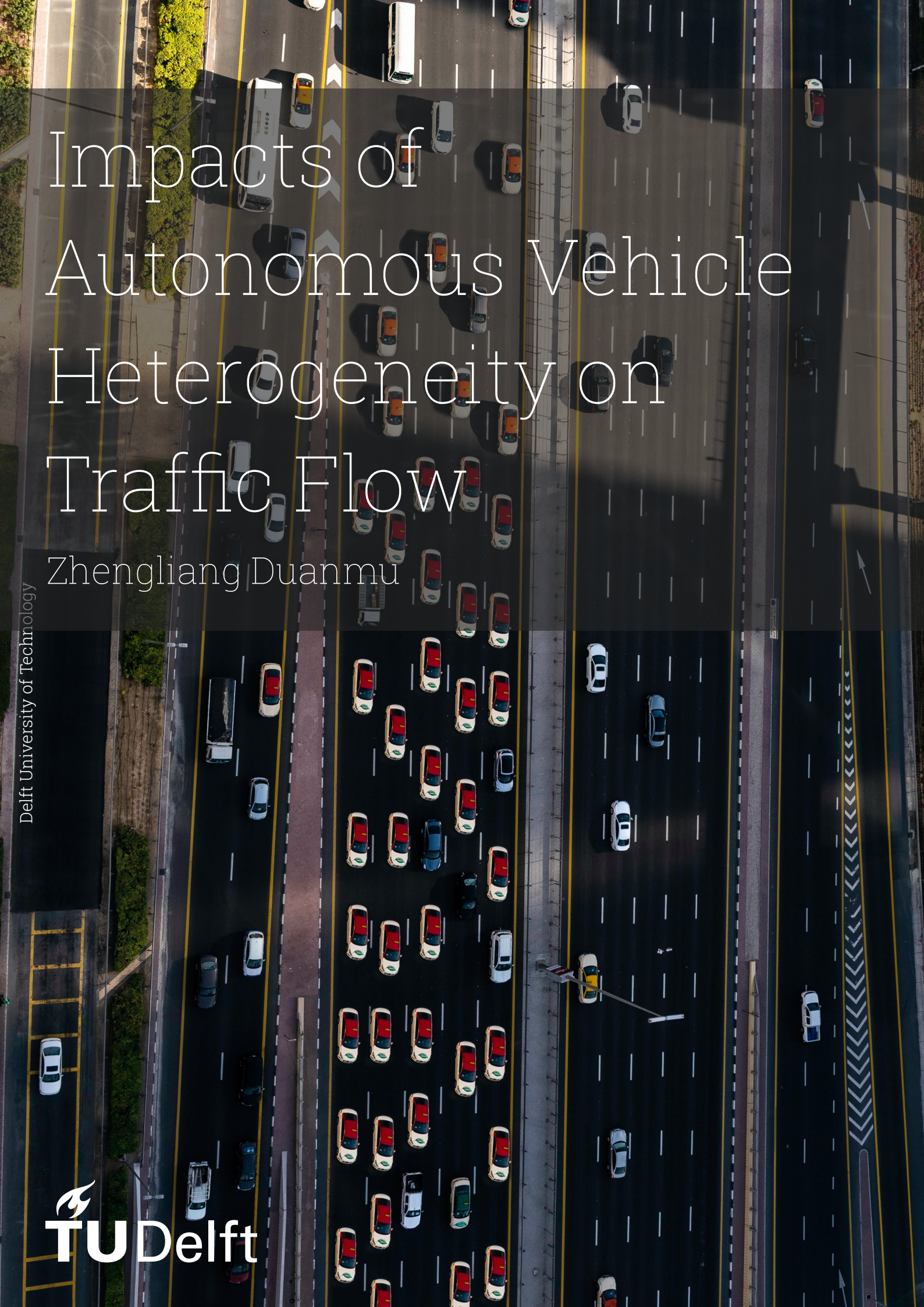


Impacts of Autonomous Vehicle Heterogeneity on Traffic Flow

Zhengliang Duanmu

Delft University of Technology

TU Delft

Impacts of Autonomous Vehicle Heterogeneity on Traffic Flow

Zhengliang Duanmu

Student number: 5458498
Project duration: January 17, 2023 – November 5, 2023
Thesis committee: Dr. ir. Simeon Calvert, TU Delft, chair
Xue Yao, TU Delft, daily supervisor
Dr. ir. Marco Rinaldi, TU Delft, supervisor

TU Delft

Preface

It has been one year since I embarked on my graduation project, a journey that has been both challenging and rewarding. Throughout this process, I have had the privilege of learning, growing, and evolving as a researcher. The completion of this master's thesis marks a significant milestone in my academic journey, and it is with gratitude and humility that I reflect on the support and guidance that have made this achievement possible.

Firstly, I would like to express my gratitude to my daily supervisor, Xue Yao. Your support and expertise have been invaluable in shaping the trajectory of my research. You have taught me the importance of precision in every aspect of the research process and provided me with continuous, constructive feedback that has propelled me forward. I must also extend my appreciation to Dr. Simeon Calvert for the insightful feedback and guidance. Your critical analysis of my results has pushed me to strive for excellence and develop a more comprehensive understanding of my field. Your mentorship has been instrumental in my growth as a researcher. I am grateful to Dr. Marco Rinaldi for the invaluable perspectives and feedback, which have broadened my horizons and challenged my thinking. Your insights have encouraged me to explore my research from different angles, leading to a more robust thesis.

To all those who have supported me behind the scenes, especially my parents, I owe an immeasurable debt of gratitude. Your unwavering belief in me, your constant encouragement, and your unconditional support have been the foundation upon which I've built my academic pursuits. Without your love and guidance, this journey would not have been possible.

Lastly, I want to express my appreciation to the people who accompanied me, Lu, Jiayi, Ximing, and others who have stood by my side. Your assistance and companionship have been a source of strength and motivation throughout this memorable experience.

As I look back on the path that has led me to this moment, I am filled with a sense of accomplishment and gratitude. This thesis represents not only the culmination of my academic endeavors but also the embodiment of the support and mentorship that have shaped me as a researcher. It is my hope that I will contribute to the ever-expanding field of traffic and transportation and this work will serve as a testament to the power of perseverance, mentorship, and support.

Zhengliang Duanmu

Delft, November 2023

Summary

With the development of Autonomous Vehicles (AVs), a promising future for their implementation becomes increasingly apparent. However, it is essential to acknowledge that the effects of AV deployment are not straightforward, particularly when considering scenarios involving AVs from different manufacturers and various levels of automation. The current research predominantly concentrates on human-driven vehicles(HDVs) and AVs. However, the assumption that AVs exhibit homogeneous behavior is a simplification that does not reflect the actual diversity within this category. The potential variations within AVs are not adequately addressed in the current study.

This research is undertaken to evaluate the impact of autonomous vehicle heterogeneity on traffic flow. To assess AV heterogeneity, the initial step involves an examination of the manifestations of AV heterogeneity through data analysis. This analysis will serve as compelling evidence of the existence of heterogeneity among AVs. The data source of the data analysis includes two parts, the Adaptive Cruise Control (ACC) data, and the high-level AV data. For the ACC data, the open ACC dataset is used. As for the high level, the processed Waymo and Lyft 5 datasets are used. These datasets encompass essential information, including the position, acceleration, and speed of the vehicles within the platoon, which is instrumental in identifying and characterizing heterogeneous driving behaviors. The analysis focuses on analyzing parameters such as Time-to-Collision (TTC), time gap, and acceleration/deceleration patterns. As for the time gap, the investigations include the distribution of time gaps under different speed ranges and different acceleration conditions. The results of the analysis contribute to the conclusion that heterogeneity among AVs is evident, not only across various automation levels but also within the same level of AVs. Recognizing the presence of intra-level heterogeneity among AVs enables us to draw further conclusions regarding their behavior and characteristics.

Given the presence of heterogeneity, characterized by the same or different automation levels with differing behavioral patterns among AVs, the car-following models are employed to capture this heterogeneity. Therefore, these parameters are calibrated using a genetic algorithm and maximum likelihood estimation is applied to determine the best-fit distributions of desired time gaps and maximum accelerations. Calibrated car-following models are then employed to represent the longitudinal behaviors of AVs. The parameters are drawn from

distributions, it is expected that AVs will exhibit slightly varying behaviors.

To assess the impact of heterogeneous traffic on traffic flow, various scenarios are constructed and evaluated. The scenarios encompass ACC vehicles, HDVs, and combinations of ACC, highly automated vehicles(HAVs), and HDVs. The first scenario aims to assess the impact of heterogeneity among AVs of the same automation level, so the different shares of ACC vehicles are involved. In contrast, the second scenario involves HAVs and ACC vehicles to evaluate the influence of heterogeneity arising from various AV automation levels. The ultimate conclusion drawn from this study suggests that heterogeneity negatively impacts traffic efficiency. Specifically, the efficiency gains afforded by vehicles equipped with ACC are offset by the presence of heterogeneous traffic at low penetration rates. Furthermore, the results obtained from simulation scenario 2 indicate that the introduction of multi-level AVs may have a detrimental effect on traffic efficiency and stability. These findings underscore the need to validate and improve AV performance comprehensively before embarking on large-scale implementation efforts.

Contents

List of Figures

List of Tables

1

Introduction

In this chapter, the background of this research is discussed in section 1.1, and the research outline is introduced in section 1.2.

1.1. Background

Traffic heterogeneity typically refers to variations caused by non-identical vehicle/driver pairs. Those differences could be different car-following behaviors or maneuvering behaviors. According to Makridis et al.[22], the studies of heterogeneity could be divided into three categories: intra-driver heterogeneity(the variability in the behavior of the same driver), inter-driver heterogeneity(the variability in the behavior of different drivers), and combination of both.

Considering its profound impacts, studies on traffic heterogeneity are of great significance. According to Yuan et al.[40], the stochastic desired accelerations to a large extent caused the capacity drop. Traffic heterogeneity also has potential effects on traffic and brings instability and uncertainty to traffic. Another reason to study traffic heterogeneity is that heterogeneity brings difficulties to the microscopic traffic simulation. According to Ossen and Hoogendoorn[26], the level of heterogeneity modeled strongly impacts the traffic flow predictions made by microsimulation tools. This difficulty leads to the failure of traffic management measures when addressing highly heterogeneous traffic.

The study of heterogeneity could help us to learn the traffic better and get more precise simulation results. The CCF(classified car-following) models considered the heterogeneity proposed by Sun et al.[35] can provide a more accurate description of the car-following behaviors compared with the unified car-following model. Therefore, it is essential to gain comprehen-

sive knowledge about traffic heterogeneity, which could arise from various disturbances or from unique vehicle/driver combinations.

In the realm of autonomous vehicles, heterogeneity also exists. Current studies predominantly focus on the mixture of AVs and human-driving vehicles, largely overlooking the inherent heterogeneity within AVs themselves. However, this oversight is significant, as variations in behavior may manifest even within AVs of the same level. The heterogeneous AV behavior (such as distinct desired headways at identical speeds) can be attributed to both hardware and software discrepancies. On the hardware front, factors such as vehicle mechanical limits and sensor characteristics contribute to the heterogeneity. Similarly, on the software front, variations in training inputs and the implementation of different algorithms also compound the heterogeneous behavior. The heterogeneity within autonomous vehicles is a problem that needs to be addressed prior to their widespread implementation since there is no unified standard for AVs' behavior in real life. The heterogeneous nature of AVs may diminish or even negate the efforts of traffic measures intended to improve traffic efficiency and traffic safety since the current management measures do not consider the appearance of different levels of AVs. With the development of AVs, the impact of heterogeneous AVs is becoming more and more severe. According to the estimation done by Calvert et al.[6], the total share of AVs will be around 18 percent in 2030. And this number will rise to 26 percent in 2035. Despite the pressing need for addressing this issue, the implications of AV heterogeneity remain largely uncharted territory. Therefore, further research is needed to elucidate the impact of such diversity on the broader framework of autonomous transportation systems.

1.2. Thesis organization

The structure of this research is shown in Figure 1.1. The first step is to analyze the ACC data.

According to SAE[13], ACC vehicles are defined as level 1 autonomous vehicles. And there are several real-life experiments that could provide available data sets. Thus, the ACC vehicle could be an ideal starting point for studying autonomous heterogeneity. The ACC data subject to analysis is sourced from the openACC database. The objective of the openACC database is to furnish comprehensive data on ACC behavior to the broader scientific community. This initiative aims to enhance the understanding of the characteristics of ACC-equipped vehicles and their potential impact on traffic dynamics. It also facilitates the anticipation of potential challenges associated with the widespread adoption of ACC systems. With the available data, the first step is to identify the heterogeneity from the data sets. The exploratory data analysis could help identify the heterogeneous driving style in the level 1 AV data sets. Then the next step is to describe the driving behaviors. A comprehensive examination of previous

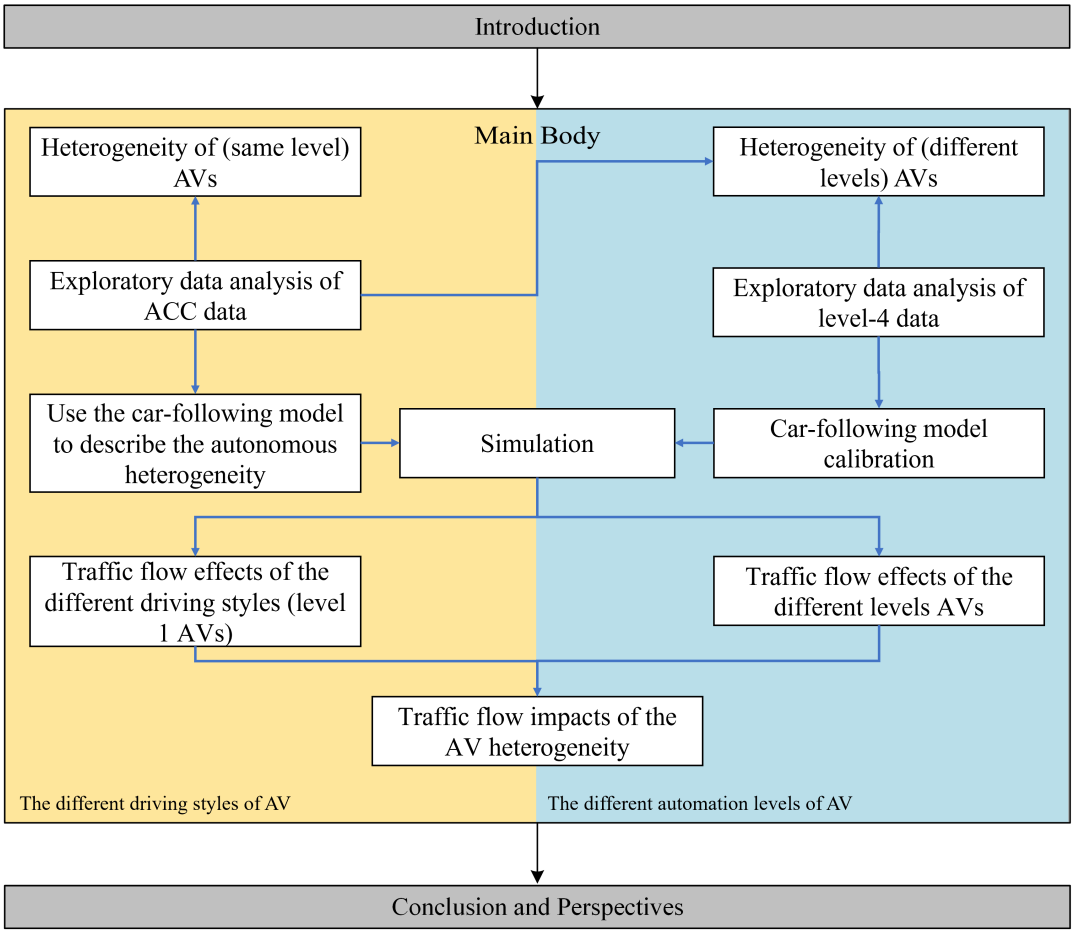


Figure 1.1: Research structure

research in the literature review part indicates that the method employed to characterize the heterogeneity of AVs could be the calibration of car-following models. Due to the fact that the data sets of ACC and CACC vehicles are mostly about longitudinal behaviors, then the traffic flow impact is mostly about the longitudinal traffic impact. The driving behavior model will only be the car-following model. After the calibration of the car-following model, the next step is to define the driving heterogeneity using the calibration result. The calibrated parameters will be fitted to different distributions to find the best fit distribution. Then the traffic flow impact of the ACC heterogeneity could be derived by simulating scenarios with the defined parameter distribution. Then further conclusions on the AV heterogeneity can be drawn based on the simulation results. Another aspect of this research is the traffic flow impact of vehicles with heterogeneous automation levels. The level 4 data processed from Waymo and Lyft 5 are used to define the parameters of level 4 autonomous vehicle longitudinal behavior. The time gap distributions and other performance indicators could help to identify the performance differences of vehicles of different levels. Then the simulation with mixed levels of AVs will be conducted to identify the traffic flow effect of the AV heterogeneity. The simulation of AVs with

different automation levels will include both the level 4 vehicles and level 1 vehicles. With the simulation result of the two parts, one can draw a conclusion on the traffic flow effect of the AV heterogeneity.

2

Literature review

In this chapter, a comprehensive examination is conducted on the existing research pertaining to traffic heterogeneity and simulations involving autonomous vehicles. The primary objectives of this review are to meticulously scrutinize the intricacies of traffic heterogeneity and to establish a precise definition of the primary research focus, which is the heterogeneity of autonomous vehicles. This review is further directed towards delving into the underlying mechanisms of AVs and conducting a comparative analysis of car-following models suitable for representing autonomous vehicles. Additionally, the existing research gaps is elucidated while refining our research objectives.

2.1. Driving heterogeneity

Ossen and Hoogendoorn[26] gave a definition of heterogeneity, in their research heterogeneity is defined as the car-following behaviors differences of driver/vehicle combinations under the same conditions. They also distinguish the two kinds of driver heterogeneity: the inter-driver and the intra-driver. The inter-driver heterogeneity denotes the variation in driving behaviors exhibited by different drivers under similar conditions. A multitude of research endeavors have attempted to characterize this inter-driver heterogeneity, often categorizing behaviors into distinct driving styles, such as aggressive or normal. On the other hand, intra-driver heterogeneity pertains to the inconsistencies in a driver's behavior when subjected to various driving environments. A substantial body of research exists addressing the fluctuations in driver behavior consequent to changes in the driving environment such as freeway or non-freeway.

The methods of characterizing the heterogeneity are shown in Table 2.1.

Table 2.1: Methods of characterizing the heterogeneity

Year	Author	heterogeneity type	method
2013	Sundbom et al.	inter-driver heterogeneity (normal driving and aggressive driving)	ARX models
2018	Berthaume et al.	intra-driver heterogeneity (different road type, different congestion level)	Psychophysical Car-Following Framework
2022	Makridis et al.	inter-driver heterogeneity (mild, normal and dynamic) and intra-driver heterogeneity (use IQR to quantitatively describe)	vehicle- and speed-independent acceleration-based metric: Independent Driving Style (IDS)
2018	Wang et al.	Inter-driver heterogeneity	a Bayesian nonparametric approach based on a hidden semi-Markov model
2015	Taylor et al.	Inter-driver heterogeneity	Dynamic Time Warping (DTW) algorithm and its application for calibrating this microscopic simulation model
2021	Yarlagadda and Pawar	Intra-driver heterogeneity and inter-driver heterogeneity	PCA and K-means clustering

From Table 2.1, it can be observed that when related to the inter-heterogeneity, a commonly used method is to character the driving behavior into different styles by clustering or other statistical methods, Sundbom et al.[36] developed a driver behavior classification method based on online estimation of a PrARX model to identify the two different driving styles: aggressive and normal. And when it comes to intra-heterogeneity, the researchers mainly focus on the factor that mainly influences driving behavior, especially car-following behavior. Berthaume et al.[4] verified the heterogeneity in car-following acceleration behaviors by applying the Psychophysical Car-Following Framework. Yarlagadda and Pawar [38] employed Principal Component Analysis (PCA) and K-means clustering to discern that the level of driver aggression varies across different driving regimes. They further established that characterizing drivers based on abstract driving features does not adequately capture the diversified nature of driving behavior. Ossen et al.[27] have found the existence of the two kinds of heterogeneity based on the trajectory data. The result shows that two kinds of heterogeneity are present in traffic at the same time and the intra-driver heterogeneity is less strong compared to the inter-driver heterogeneity.

2.2. AV heterogeneity

The inter-driver and intra-driver heterogeneity does not fully suit the autonomous vehicle since (part of) driving behaviors are not implemented by humans. The AVs also have the extra character that HDVs do not have, the different automation levels. Within the same automation level, the AVs still have different driving behaviors due to the different algorithms or different training data sets. In the analogy between the HDV and AV, AV heterogeneity could also be categorized into two levels: intra-automation level heterogeneity and inter-automation level heterogeneity. Thus, the definition of AV heterogeneity in this research could be given: differences between the car-following behaviors of different levels of AVs and different driving behaviors of the same level AVs.

The studies that related to the AV heterogeneity mainly focus on the mixture of the different automation levels of vehicles (usually level 0 and level 1 AV) since there are clear behavior differences among the AV and HDV. However, there are still some studies that explore the heterogeneity within the same level of AVs though the concept of heterogeneity is not clearly proposed. The difference in the AV could mainly be defined as the following aspects: desired time gap, acceleration/deceleration, and reaction time.

The first aspect under consideration pertains to the desired time gap, a fundamental element in car-following behavior. The significance of the desired time gap in modeling vehicle behavior has been widely acknowledged in the literature. Numerous ACC algorithms have been specifically engineered to uphold a consistent desired time gap during vehicular operations. ACC-equipped vehicles have gained recognition for their ability to effectively maintain smaller time gaps while in motion, thus contributing to enhanced operational efficiency. In simulations involving ACC vehicles, the desired time gap is often configured to adopt either a fixed value or a selection from among a limited set of predefined values. Historical research has frequently relied upon subjective preferences to establish these fixed values. For instance, Nowakowski et al. [25] conducted a study wherein subjective opinions regarding ACC/CACC systems were collected and subsequently utilized to inform time gap settings. However, in this research, they constrained the options for ACC vehicle time gap settings to just three values: 1.1 seconds, 1.6 seconds, and 2.2 seconds. These values correspond to the three primary levels of gap settings as provided by vehicle manufacturers, namely, short, intermediate, and long. Subtle differentiations within each level of the gap setting have been disregarded for simplicity. It is noteworthy that the desired time gaps can vary significantly among different vehicles. In the case of HDVs, these gaps are typically categorized into clusters denoting aggressive or normal driving behaviors. Even within these clusters, some degree of variation may persist. It is conceivable that such variations also exist within ACC-equipped vehicles

and HAVs.

In conclusion, the ACC vehicles in the past simulation are modeled very homogeneously and these methods neglect the difference between the shortest desired time gap setting difference of vehicles from different manufacturers.

The desired time gap may, to some extent, be influenced by user preferences. while the determination of maximum acceleration and deceleration for ACC vehicles and HAVs during car-following behavior is not under the control of the drivers. Consequently, the strategies for acceleration and deceleration in AVs might exhibit a greater degree of diversity compared to the desired time gap. Previous research has shown that ACC systems tend to deliver smoother acceleration profiles. In fact, many ACC-equipped vehicles incorporate mechanisms that limit jerk, thereby enhancing the overall driving experience. Conversely, high-level AVs designed to operate across a wide range of scenarios may have more relaxed acceleration constraints to accommodate challenging conditions. It's important to emphasize that due to variations in algorithms and the distinct objectives of different AV systems, the acceleration patterns exhibited by HAVs can differ markedly.

Additional factors, such as the accuracy of sensors, may indeed influence vehicle behavior. Nevertheless, these factors have not been taken into account in our analysis, as they do not exert a direct impact on traffic flow. Rather, their influence is more indirect, affecting key features like desired time gap and acceleration, which in turn can influence traffic flow. Therefore, only the main behavior difference of the AV will be included in this research. The behavioral differences among the desired time gap and acceleration will be detailed examined.

2.3. AV simulation review

The study of AV traffic flow heterogeneity is motivated by the significant impact that AVs have on traffic flow dynamics. The structure of autonomous vehicles is inherently complex, characterized by a sequential process involving the sensing, recognition, decision, and action loop[39]. While the output of AVs can be translated into acceleration or braking actions, the inputs to the AV system typically consist of raw data collected from various sensors. Consequently, modeling autonomous vehicles poses a challenging task. Accurate modeling of AVs is crucial for addressing design issues or evaluating their performance. However, when examining the traffic flow impact of AVs, such a method may be unnecessary. Table 2.2 shows the AV modeling of recent years.

Table 2.2: Review of the recent AV model methods

Author	Year	Modelling method	Simulation scenario
Calvert et al.[6]	2017	LMRS-IDM+ model	motorway corridor with three lanes
So et al.[31]	2020	Enhanced Intelligent Driver Model	Frankfurter Ring road section
Fang et al.[9]	2022	automatic lane change algorithm, ACC model based on IDM	The upstream part of a motorway in Hungary
Berrazouane et al.[3]	2019	LC2013, ACC car-following model	motorway segment
Lu et al.[20]	2019	Krauss Model	Grid network and urban traffic network
Kavas-Torris et al.[14]	2023	Wiedemann 99 (CoExist parameters)	Four different scenarios contains freeway and urban road
Guo et al.[10]	2023	cellular automata model	single-lane scenario with different communication range
Zheng et al.[41]	2020	A stochastic Lagrangian model	single-lane straight highway section

The prevalent approach for microscopically modeling AVs is the integration of Car-Following and Lane-Changing models, as evident from existing research. This method could provide accuracy that exploring the traffic flow impact of AVs could be used and reduce the complexity at the same time. The used car-following models include the IDM, EIDM, Krauss Model, Wiedenmann 99 model, and ACC/CACC model. The parameters of those models came from the literature or from the calibration of experiment data.

There are studies that provide the recommended parameters of the car-following model to represent the behavior of AVs. The first project is the CoExist project. CoExist project using the Wiedemann 99 car-following model to represent the longitudinal behavior of AVs[33]. The recommended parameters are shown in Table 2.3. The AVs are specified into four different categories: Rail safe, Cautious AV, Normal AV, and Aggressive AV. The results are validated by comparing the car-following model result to the autonomous vehicle driving logic in different simulation scenarios.

Table 2.3: CoExist W99 model parameters

Parameters	rail safe	cautious	normal	all knowing
CC0	1.5	1.5	1.5	1
CC1	1.5	1.5	0.9	0.6
CC2	0	0	0	0

CC3	-10	-10	-8	-6
CC4	-0.1	-0.1	-0.1	-0.1
CC5	-0.1	0.1	0.1	0.1
CC6	0	0	0	0
CC7	0.1	0.1	0.1	0.1
CC8	2	3	3.5	4
CC9	1.2	1.2	1.5	2

The CoExist project does not specify the level of AVs. In the report given by Atkins[2], the different parameters of the W99 model are applied to represent the different levels of AVs. The suggested parameters are shown in Table2.4.

Table 2.4: W99 model parameters suggested by Atkins[2]

Capability Levels		**CC0 (m)**	**CC1 (s)**	**CC7 (m/s2)**	**CC8 (m/s2)**	**CC9 (m/s2)**
Level 2		1.5	0.9	0.25	3.5	1.5
Level 3	Cautious	2.5	1.8	0.1	3.2	1.2
	Normal cautious	2	1.2	0.2	3.4	1.4
	Normal assertive	1	0.8	0.3	3.6	1.6
	Assertive	0.5	0.6	0.4	3.8	1.8
Level 4		0.5	0.6	0.4	3.8	1.8

The predominant approach for assessing the impact of AVs is through simulation. Within the realm of simulation scenarios, the most used method is to consider different penetration rates of the AVs and evaluate the simulation result to determine the impact of the AVs under different penetration rates. One common view on the traffic flow impact of the AV is that when the penetration rate is low, the AV will have a negative impact or limited positive impact on the traffic flow, when the penetration rate is high (50% or larger), the AV will have positive impact on the traffic flow.

Calvert et al. [6] conducted simulations of a motorway segment and used the LMRS-IDM model to represent ACC vehicles, revealing that low-level automated vehicles in mixed traffic scenarios tend to yield a marginal adverse effect on traffic flow and road capacities. Improvements are observed only when the penetration rate exceeds 70%. Similarly, Fang et al. [9] arrived at comparable results, where the introduction of HAVs and CAVs negatively impacted network performance. However, high CAV penetration rates led to a notable reduction in traffic delays. Guo et al.'s research[10] suggested that a large AV penetration rate ensures a low level of traffic congestion in mixed traffic flows.

It is important to note that divergent viewpoints exist within the research community. For instance, Berrazouane et al. [3] conducted simulations of a motorway segment, revealing a decrease in the maximum traffic flow rate with an increasing penetration rate of AVs. Meanwhile, Lu et al.[20] contend that AV penetration has a positive impact on enhancing road network capacity in a quasi-linear fashion.

Though the opinions on the traffic flow impact of the introduction of AVs are quite different due to the fact researchers are applying different methods and making different assumptions, the introduction of the AV heterogeneity will surely help to improve the AV simulation method and make the simulation result more reliable.

2.4. Research gap

The future of Autonomous Vehicles is becoming increasingly evident, with the consequence of their implementation becoming increasingly apparent. Among the various impacts, the influence of AVs on traffic flow in mixed traffic environments emerges as a critical consideration, given its intrinsic connection to both efficiency and safety.

The second research gap is the current research has not sufficiently addressed the diverse driving behavior of autonomous vehicles at the same level of automation, as well as the variations across different levels of automation for AVs. The current research on the traffic flow impact of autonomous vehicles is limited, as studies often consider only a single level of automation as it is shown in Table 2.2. The AVs in the simulation only contain different shares of ACC (level 1 AVs) and human-driving vehicles. In the investigation of AV traffic flow impact, it is often assumed that the AVs are highly homogeneous, or even exhibit identical behavior (Same car-following model parameter). Thus, the concept of AV heterogeneity is introduced to explore the traffic flow impact of AVs. In this research, autonomous vehicle heterogeneity comprises the different driving behavior of the same level of AVs and the different automation levels of AVs.

2.5. Research objective and research questions

The primary objective of this research is to evaluate the traffic flow impact of autonomous vehicle heterogeneity. This study leverages empirical data to systematically observe and analyze the heterogeneity within a given context. Subsequently, the simulation techniques will be employed to comprehensively assess the implications and consequences of this heterogeneity. This research can yield valuable insights into autonomous vehicle heterogeneity and its external manifestations, thereby informing the development of future autonomous vehicle technologies. To achieve the research objective, the following research questions are pro-

posed.

- Does empirical data exhibit evidence of autonomous vehicle heterogeneity?
- What methodology is most suitable for modeling the inherent heterogeneity in autonomous vehicles?
- What are the traffic flow impacts of autonomous vehicle heterogeneity?

To model the AV heterogeneity, the first step is to identify the heterogeneity in the empirical data. In the literature review section, the definition of autonomous vehicle heterogeneity is given: differences between the car-following behaviors of different levels of AVs and different driving behaviors of the same level AVs. After giving the definition of AV heterogeneity, one can identify the AV heterogeneity in the dataset. Thus, the first research question includes two sub-questions:

- Does the same level of AVs show different driving behaviors?
- Do different levels of AVs show different behaviors?

It seems there is a plain fact for the answer. The variations in sensors, decision-making algorithms, and overall vehicle design will result in performance differences. Nonetheless, an empirical dataset will be subjected to analysis to substantiate the presence of heterogeneity. The same-level heterogeneity will be analyzed based on the ACC data set. Starting with the low-level autonomous vehicle could be a great option due to the data sets availability. The analysis will focus on the time gap, and the time-to-collision and acceleration will also be included. Regarding the second sub-question, the level 4 data set obtained from Waymo and Lyft5 will be used. The objective is to compare the performance of a level 4 AV with that of a level 1 AV.

The second research question is how to model heterogeneous autonomous vehicles. This question not only includes how to model the AV in the further simulation but also is about how to obtain the heterogeneity in the simulation. The answer to this question will based on the existing literature and the comparison between different models.

With the measured AV heterogeneity, it comes to the third question, What could be the traffic flow impact of ACC vehicles? The impact of the heterogeneous low-level autonomous vehicle will be derived using microscopic simulation. The simulation scenario will include ACC vehicles with different driving behaviors. The impact of different level AVs will include different shares of ACC vehicles and level 4 AVs. Then the conclusion on heterogeneity could be drawn by analyzing the simulation result and the explanation of the AV heterogeneity could be given. Then the main result question could be answered.

2.6. Research methodology

As illustrated in Figure 1.1, this research can be divided into two distinct components: data analysis and simulation. The data analysis segment primarily centers around the characterization of AV heterogeneity. This characterization encompasses an evaluation of heterogeneity concerning the factors discussed in the literature review, specifically, desired time gap and acceleration/deceleration patterns. Furthermore, an essential surrogate safety metric, the time-to-collision of ACC vehicles and Highly Automated Vehicles, is scrutinized.

Following the elucidation of heterogeneity, it becomes imperative to examine the ensuing traffic flow impact within the context of heterogeneous traffic conditions. To fulfill the objectives of this research, two simulation tools are considered: PTV Vissim and Eclipse SUMO which are widely recognized as two of the most commonly employed traffic simulation software. PTV Vissim stands as a commercially available microscopic traffic simulation software developed by PTV Group, ideally suited for practical traffic engineering and analysis. Conversely, Eclipse SUMO is an open-source microscopic traffic simulation software, developed collaboratively by the Eclipse Foundation and the German Aerospace Center, predominantly utilized for research and development endeavors in the realm of transportation and mobility. The choice between these tools hinges on specific project requirements and the extent of customization and extensibility needed. Simulation work involving heterogeneous vehicles necessitates software that offers a convenient means of customizing the behavior of each vehicle. In PTV Vissim, such customization is achieved by defining behavior through GUI or PTV's scripting language. In contrast, Eclipse SUMO allows for straightforward behavior modification through the editing of XML configuration files. It's worth noting that PTV Vissim exclusively supports the W99 and W74 car-following models while integrating other models requires additional effort. On the other hand, SUMO supports a broader range of car-following models, simplifying the modeling process. In consideration of the user-friendly nature and enhanced customization capabilities, Eclipse SUMO is selected as the preferred simulation tool.

As depicted in Table 2.2, a plethora of methods are available for representing AVs within the simulation scenario. In this research, the approach entails employing a car-following model coupled with a lane-changing model. Specifically, the car-following models utilized are IDM and EIDM, complemented by the lc2013 lane-change model. The reason behind selecting these models is discussed in Section 4.3.

To capture the heterogeneity among AVs, a distribution-based method is employed. Within this approach, distributions for each parameter of the car-following model are constructed. Subsequently, this distribution is utilized for the resampling of vehicles within the simulation environment. It is essential to underscore that this process necessitates both car-following

model calibration and the identification of the best-fitting distribution.

For the calibration of the car-following model, the genetic algorithm is chosen as the method of choice. In the context of car-following model calibration, a typical GA workflow involves encoding potential parameter sets as chromosomes, assessing their fitness by comparing model output to observed data, selecting the fittest chromosomes for the production of offspring via crossover and mutation operations, and iteratively repeating this process across multiple generations until a satisfactory solution is achieved. The selection of the genetic algorithm for car-following model calibration is well-founded, supported by prior research and optimization algorithm assessments, as demonstrated by Punzo et al. [29]. Their results indicate that genetic algorithms are capable of rediscovering the true values of model parameters. Even when a genetic algorithm becomes trapped in local minima, it consistently converges towards the known values of the most sensitive parameters. This underscores the effectiveness of calibrating the model based on the spacing between the lead vehicle and the follower, yielding acceptable results in terms of vehicle speed. Furthermore, a substantial body of existing car-following model calibration work[15][28] has effectively employed GA as a robust optimization approach. Additionally, there are also works that calibrated the car-following models of AV using the GA[19]. For the identification of the best-fit distribution, the Python package Fitter is applied. The fitting process applied the MLE and the selection of the best-fit distribution will based on the multi-criterion which includes: SSE, AIC, and BIC.

3

Analysis on AV heterogeneity

In this chapter, the heterogeneity of autonomous vehicles is analyzed by comparing the empirical data from different data sets and the conclusions on the presence of AV heterogeneity will be drawn. The result helps us to know the heterogeneity better and serves as the foundation for subsequent simulations.

3.1. Data pre-processing

3.1.1. ACC data sets

The analysis of the ACC system heterogeneity is based on the ACC platoon empirical data. The descriptions of the data sets are shown in table 3.1. The data set includes six data sets and the data source is the Open ACC database[7]. The openACC database is uploaded to provide comprehensive data on ACC behavior to the scientific community. The database contains all essential longitudinal information on the ACC vehicles in platoons of different sizes. The dataset includes vehicles from various manufacturers and experiments conducted in multiple locations. Part from the ring road data set and the low-speed data set, other subsets in the open data set are utilized in this research. The data set contains the position, speed, and inter-platoon spacing of the ACC vehicle platoon. The platoon sizes vary from two to eleven. The lengths of the data segment are different, most of the segments are longer than 1000s.

Table 3.1: ACC data sets description

Pliot location	Vehicle	Driving type	Description
Sweden	Audi (A6), Tesla (Model 3), Mercedes (A Class)	ACC/ Manual driving	Speed, Inter Vehicle Spacing, GPS Position
Italy	Hyundai Nexo, SsangYong REXTON	ACC/ Manual driving	Speed, Inter Vehicle Spacing, GPS Position
Italy	Fiat (500X), Volvo (XC40), VW (Polo), Hyundai (Ioniq hybrid)	ACC/ Manual driving	Speed, Inter Vehicle Spacing, GPS Position
Italy	Hyundai Nexo, Ford S-Max	ACC/ Manual driving	Speed, Inter Vehicle Spacing, GPS Position
Italy	Ford S-Max, KIA (Niro), Mitsubishi (SpaceStar), Mitsubishi (Outlander PHEV), Peugeot (5008 GT Line), VW (Golf E), Mini (Cooper)	ACC/ Manual driving	Speed, Inter Vehicle Spacing, GPS Position
Hungary	Tesla (model X), Tesla (model 3), Tesla (model S), Mercedes-Benz (GLE 450 4Matic), Jaguar (I-Pace), BMW (I3 s), Audi (E-tron), Toyota (Rav 4), Mazda (3), Audi (A4 Avant)	ACC/ Manual driving	Speed, Inter Vehicle Spacing, GPS Position

The first step of pre-processing is data smoothing. The simple moving average over 3 data points is used to reduce the noise of data. After the smoothing, two filters are applied to the data. For the ACC data sets, only the driving data under the ACC driving mode are taken into account. The second filter is the ACC setting. As shown in Figure 3.1, the two peaks of the time headway distribution correspond to the two different settings of the ACC system. Apart from the effect of the heterogeneous time-gap setting, other heterogeneity effects are explored under the minimum gap setting due to data availability. Additionally, for all the ACC data utilized in this research, the ego vehicles are under the car-following condition.

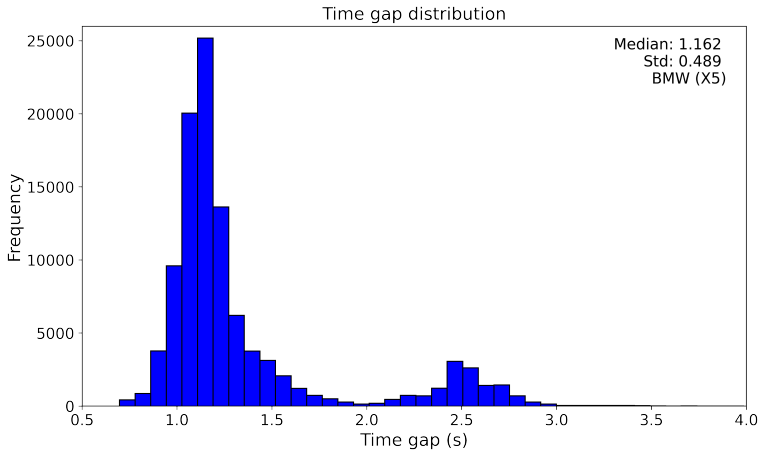


Figure 3.1: BMW (X5) time gap distribution

3.1.2. Level 4 autonomous vehicle data set

The datasets processed by Hu et al.[12] and Li et al.[18] are utilized in this study. The data descriptions are presented in Table 3.2.

Hu's dataset is derived from the Waymo dataset, which is one of the most well-known AV datasets. It contains information about the movement of the AV and the data extracted from its surrounding environment. In their work, Hu et al. process the data by extracting all car-following pairs, removing outliers, and filtering noise. Subsequently, they reformat the data to align with the typical datasets collected by the traffic flow research community. The car-following pairs in this dataset consist of Level 4 vehicles and human-driven vehicles, specifically focusing on AV-HDV(AV following HDVs) pairs. Out of 1000 pairs, 195 are selected for analysis. This dataset contains comprehensive information related to the behavior of autonomous vehicles and the behavior of the leading vehicle. Quality assessments of the data indicate that the overall quality surpasses that of the NGSIM dataset[12]. In terms of size and scope, this dataset aligns with the Waymo open dataset[34], as it is derived from it. It consists of 1000 segments, each lasting 20 seconds, which is considerably shorter than the ACC data.

The dataset processed by Li et al.[18] is derived from the Lyft 5 dataset[11]. The Lyft 5 dataset was collected by 20 autonomous cars following a predetermined route in Palo Alto, California for four months. Li et al. first selected car-following pairs based on specific criteria. They further evaluated the raw data for anomalies to assess its quality. To enhance the data, they applied motion planning, Kalman filtering, and wavelet denoising techniques to correct and improve the raw car-following data. In this research, only the AV-HDV pairs are selected as well. The size of the processed Lyft dataset is larger than the processed Waymo dataset. The processed Lyft 5 dataset contains 8262 AV-HDV pairs and the length of each segment is around 25s. Lyft include urban and non-urban environment while the Waymo data set contains

mostly the urban environment.

Table 3.2: High-Level autonomous vehicle dataset

Author	Number of HV-AV pairs	Segment length
Hu et al.	195	20s
Li et al.	8262	around 25s

3.2. Exploratory data analysis

3.2.1. Time to collision

Various surrogate safety measures have been introduced, with TTC standing out as the most well-known and frequently utilized indicator. This concept is instrumental in the domains of traffic safety and collision avoidance systems, as it quantifies the time required for a vehicle to reach a specific point or object if it maintains its current trajectory. TTC serves as a valuable parameter for the assessment and prediction of potential collisions and is widely used in a range of automotive safety systems, including ADAS and high-level AVs. The Time-to-Collision (TTC) calculation is defined as follows:

$$TTC_i = \frac{X_i}{\Delta V_i}$$

Where the X_i is the distance between the two vehicles at time i and the ΔV_i is the speed difference of two vehicles at time i.

The time-to-collision cumulative distribution functions are shown in Figure 3.2, where the

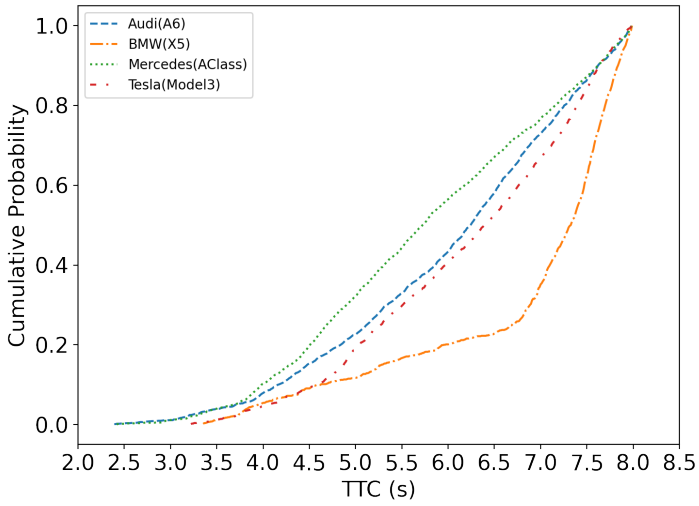


Figure 3.2: Time-to-collision cumulative distribution function

TTCs in this figure are the TTCs of the ego vehicle to the leading vehicle in the platoon. Only the TTCs less than 8s are considered. From the figure, it can be observed that the cumulative distribution function of different vehicles has different trends. Among the ACC vehicle groups, the Audi(A6) and the Mercedes(AClass) have a clearly higher percentage of small TTCs(TTCs < 3s). This fact indicates that the current ACC systems are taking different driving strategies and some vehicles behave more dangerously compared to others.

As for the level 4 Waymo data, the TTC cdf is shown in Figure 3.3. The start point of the CDF line is 2.5 seconds which indicates that there might be risky behaviour during the car-following process. The increase rate of Waymo CDF between 3.3s and 5s is large and the overall time-to-collision of Waymo vehicles is pretty large compared to ACC vehicles.

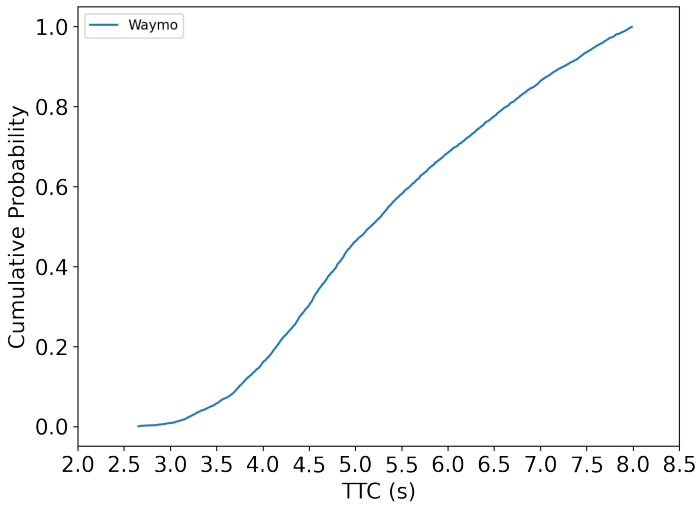
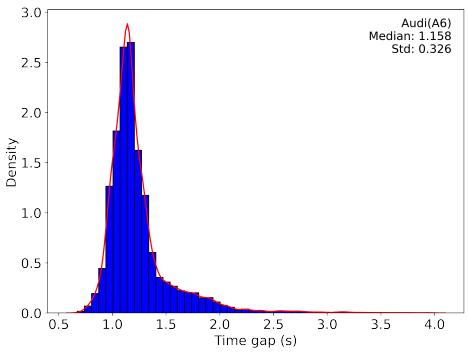


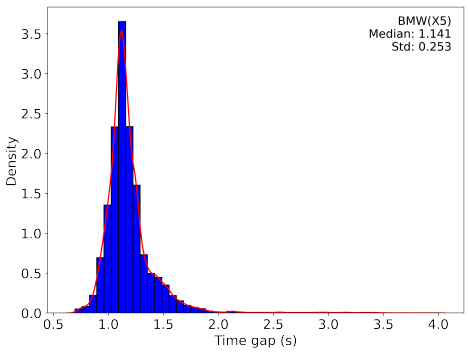
Figure 3.3: Waymo time-to-collision CDF

3.2.2. Time gap

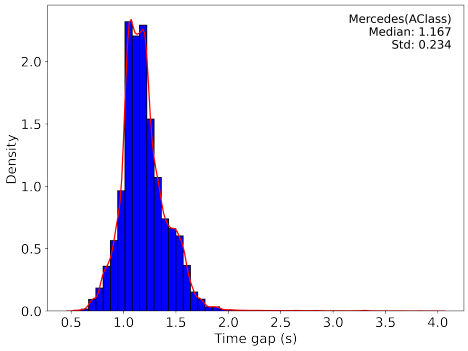
For all the time gaps considered in this research, the vehicles are under the car-following condition and the time gap settings of the ACC vehicles are the minimum gap setting. The distributions of the time gap are shown in Figure 3.4. For these four ACC vehicles, the median of the time gap distribution is different and the standard deviations of the time gaps are different. The Tesla Model 3 has the largest median of the time gap, indicating that Tesla might have the largest desired headway.



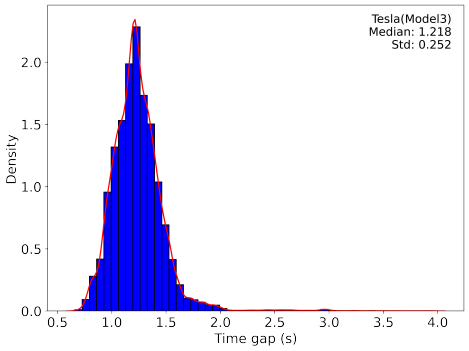
(a) Time gap distribution Audi(A6)



(b) Time gap distribution BMW(X5)



(c) Time gap distribution Mercedes(AClass)



(d) Time gap distribution Tesla(Model3)

Figure 3.4: Time gap distribution of different vehicles

The distribution of time gaps for the Waymo vehicle is illustrated in Figure 3.5. The Level 4 AV exhibits a large desired headway and a large standard deviation in the time gap since the median of the time gap is 3.514s and the standard deviation of the time gap distribution is 2.926. This could be attributed to the algorithm implemented in the Level 4 vehicle, which is designed to handle more complex environments. An increased desired time gap may contribute to improved emergency response capabilities.

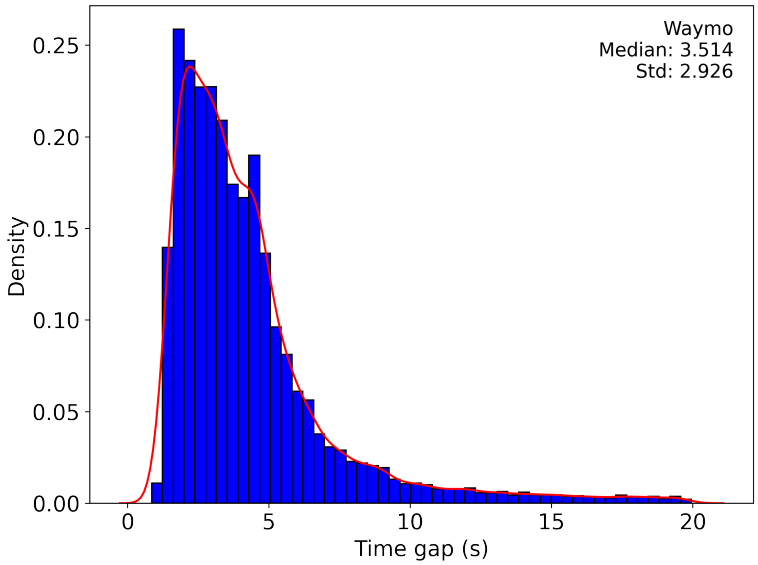


Figure 3.5: Time gap distribution Waymo

To further explore the heterogeneity of the autonomous vehicle, the time-gap distributions of ACC vehicles under different acceleration conditions are shown in Figure 3.6, 3.7, and 3.8. There are three distinct conditions in this context: acceleration, deceleration, and stability. Each condition is determined based on specific criteria, as outlined in the reference [32]. More specifically, a vehicle is classified as being in acceleration condition if its acceleration exceeds $0.5m/s^2$. Conversely, if the acceleration falls below $-0.5m/s^2$, the vehicle is categorized under deceleration condition. Otherwise, if the acceleration is between these two thresholds, the vehicle is considered to be in stable condition. It can be observed that when the vehicles are under acceleration condition, the distribution have the largest difference.

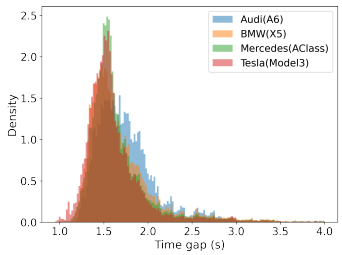


Figure 3.6: Headway distribution (acceleration)

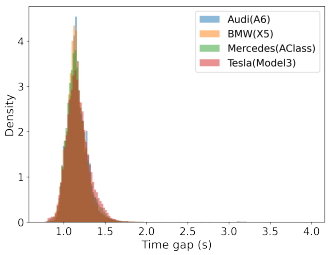


Figure 3.7: Headway distribution (stable)

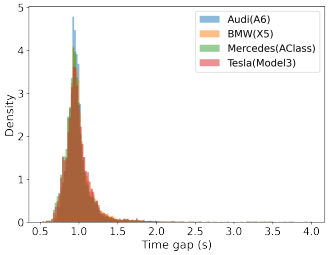


Figure 3.8: Headway distribution (deceleration)

The K-S test is applied to identify whether the two samples are from the same distribution. The K-S test checks whether the two distributions are similar or not by generating the cumulative figures and finding the largest distance along the y-axis. The statistical value of K-S test

is:

$$D_{m,n} = \max_x |F_m(x) - F_n(x)|$$

Where $F_m(x)$ is the first sample and the size of the sample is m. $F_n(x)$ is the second sample, the size is n. The $c(\alpha)$ is the inverse of the Kolmogorov distribution at α. The null hypothesis is that both samples are from the same distribution. The $D_{m,n,\alpha}$ is calculated by the following equation:

$$D_{m,n,\alpha} = c(\alpha)\sqrt{\frac{m+n}{mn}}$$

And if the statistical value $D_{m,n}$ is bigger than $D_{m,n,\alpha}$, we can reject the null hypothesis at a significant level α. If not, we can not reject the null hypothesis which means that the two distributions are probably the same. According to the result of the K-S test, the time gaps are likely from two different distributions. The result of the K-S test are shown in Table 3.3. If the p-value is smaller than 0.05, the null hypothesis (datasets are from the same distribution) is rejected.

Table 3.3: K-S test result

Sample 1	Sample 2	Statistic value	P-value
Audi(A6)	Audi(A6)	0	1.0000
Audi(A6)	BMW(X5)	0.153164741	0.0000
Audi(A6)	Mercedes(Aclass)	0.209611841	0.0000
Audi(A6)	Tesla(Model3)	0.255334544	0.0000
BMW(X5)	Audi(A6)	0.153164741	0.0000
BMW(X5)	BMW(X5)	0	1.0000
BMW(X5)	Mercedes(Aclass)	0.066837009	0.0000
BMW(X5)	Tesla(Model3)	0.108462778	0.0000
Mercedes(Aclass)	Audi(A6)	0.209611841	0.0000
Mercedes(Aclass)	BMW(X5)	0.066837009	0.0000
Mercedes(Aclass)	Mercedes(Aclass)	0	1.0000
Mercedes(Aclass)	Tesla(Model3)	0.076607449	0.0000
Tesla(Model3)	Audi(A6)	0.255334544	0.0000
Tesla(Model3)	BMW(X5)	0.108462778	0.0000
Tesla(Model3)	Mercedes(Aclass)	0.076607449	0.0000
Tesla(Model3)	Tesla(Model3)	0	1.0000

The time gap distributions of the Waymo vehicle under different conditions are shown in Figure 3.9, Figure 3.10 and Figure 3.11. Since the overall time gap distribution of Waymo is

within a wider boundary, the time gap distributions under different acceleration conditions are different.

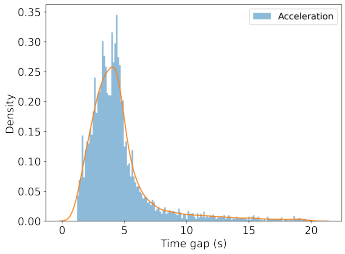


Figure 3.9: Waymo headway distribution (acceleration)

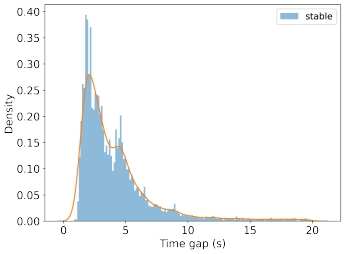


Figure 3.10: Waymo headway distribution (stable)

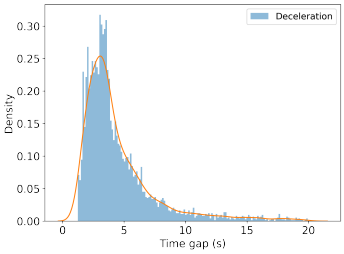
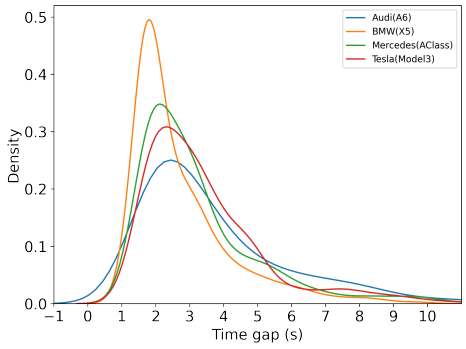
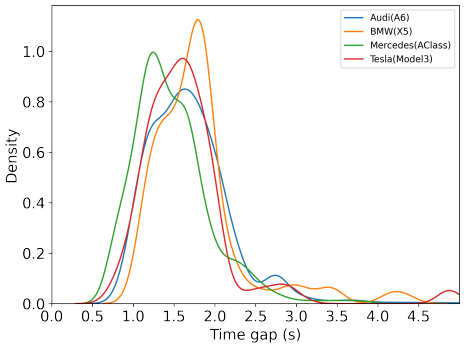


Figure 3.11: Waymo headway distribution (deceleration)

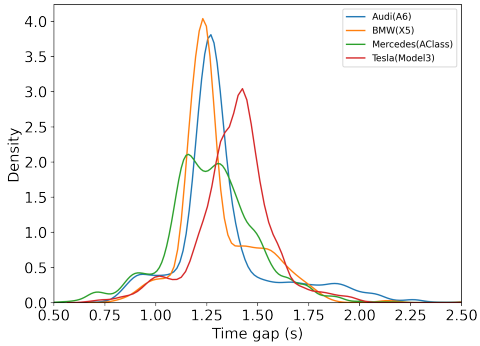
The speed filter could also be applied to the time gap distribution. The speed-specified time gap distributions are shown in Figure 3.12. Four speed ranges have been specified. For



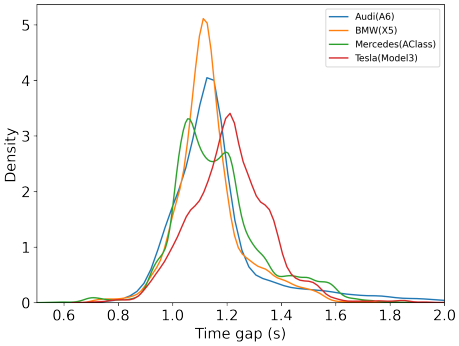
(a) Speed range 0.5-5.5



(b) Speed range 5.5-10.5



(c) Speed range 10.5-15.5



(d) Speed range 15.5-20.5

Figure 3.12: Speed specific time gap distribution

the range of 0.5m/s to 5.5m/s, the time gap distributions of ACC vehicles are similar. However,

the peak of the BMW is less than 2 seconds, while for the other vehicles in the first data set, the peak is larger than 2 seconds. This indicates that BMWs exhibit riskier behavior under low-speed conditions. Additionally, the time gap distribution for other speed ranges shows ubiquity in heterogeneity.

As for the Waymo dataset, the speed-specific time gap distributions are shown in Figure 3.13. The overall trends of the distributions are similar to the ACC distribution, low-speed condition has a larger time gap.

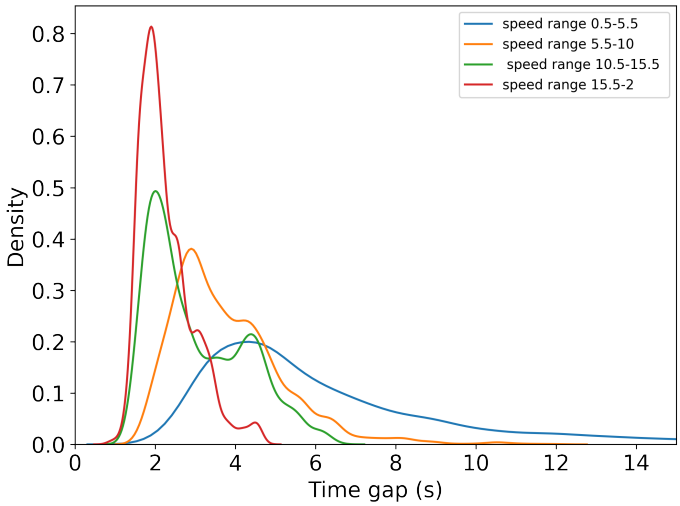


Figure 3.13: Speed-specific time gap distribution of Waymo

3.2.3. Acceleration and deceleration

The cumulative distribution functions of autonomous vehicles' acceleration are shown in 3.14 and 3.15. The increase rate of the CDF curve at a given point reflects how rapidly the probability is accumulating in that particular region. If the increase rate of the CDF curve is high at a certain point, it indicates that a large portion of the data is concentrated around that value. Conversely, a low increase rate suggests that the data is more dispersed or less frequent around that value. Commercial ACC systems are designed to keep the maximum acceleration within certain limits to ensure a pleasant driving experience. However, the cumulative distribution ratio of acceleration and deceleration indicates that there are differences in these limits. The percentage of stable acceleration conditions is high which indicates that the car-following condition of Waymo vehicles is quite stable. And the maximum acceleration is around $2.8m/s^2$ and the largest deceleration is around $2.9m/s^2$. The CDF of level 4 AV is lower than the curve of ACC vehicle when the acceleration (deceleration) is small which indicates that high-level AVs are more likely to take extreme acceleration or break.

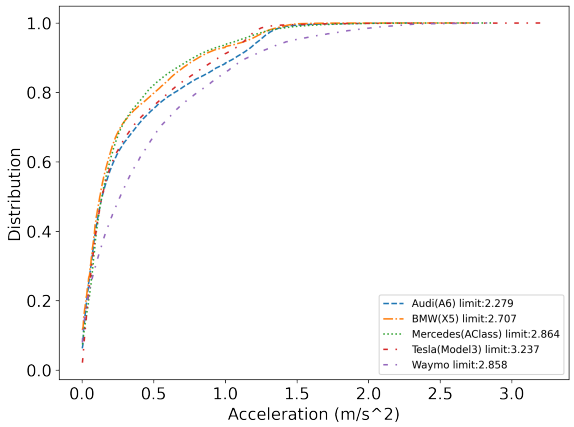


Figure 3.14: Acceleration cumulative distribution function

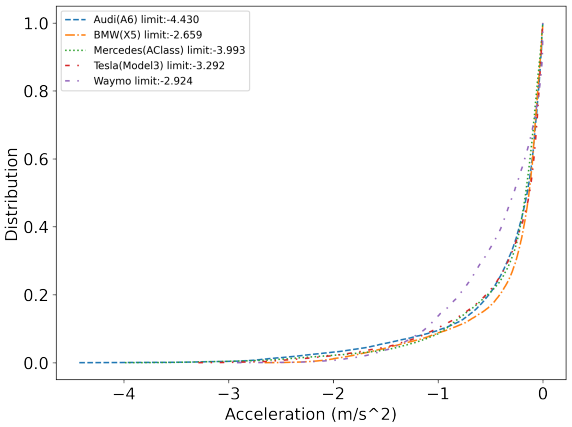


Figure 3.15: Deceleration cumulative distribution function

Another thing that need to be mentioned is the string stability of ACC vehicles. Within the openACC database, the string stability of ACC vehicles exhibits variability, with both string-stable and string-unstable platoons present. Research conducted by Makridis et al.[21], has identified the coexistence of stable and unstable platoons in the dataset. According to their research, platoon stability is primarily dependent on the time gap setting. There exists a threshold time headway setting, above which a platoon can exhibit string stability, and below which the platoon may become string-unstable. This observation highlights the heterogeneity of ACC vehicles with respect to string stability.

3.2.4. Summary

From the above analysis, conclusions on the manifestations of AV heterogeneity could be drawn. The heterogeneity of same-level AVs are shown in different aspects. Some ACC systems are applying more risky driving strategies than other ACC systems. And the ACC vehicles have different time headway distributions under comparable conditions. The desired time intervals for various ACC systems exhibit variation, even when all settings are configured to their minimum values. Throughout the entire frame considered in this research, the vehicles are consistently under car-following conditions, with the time gap setting for the ACC vehicles consistently configured to the minimum gap setting. The comparison of the data reveals that the time gap distributions of ACC vehicles exhibit variations even when the acceleration conditions are identical. Furthermore, this conclusion remains consistent across the four speed ranges examined. The analysis in this chapter proves that heterogeneity within the same automation level exists. The level 4 Waymo vehicle also has different longitudinal characteristics from the ACC vehicle. According to the time gap distribution, the Waymo vehicle takes more cautious behavior and could be considered a cautious driving style.

4

Model Calibration

In this chapter, the candidate car-following models are introduced and subsequently calibrated. Following the calibration process, the calibration results are briefly explained. Additionally, the model to be used in simulations is selected mainly based on the calibration results.

4.1. Car-following models

The Eclipse SUMO encompasses various car-following models, including the ACC/CACC model, W99 model, IDM, and EIDM. For the first simulation scenario, the three models are considered candidate car-following models of the ACC vehicles: the ACC model developed by Milanés and Shladover[23], the EIDM, and the W99 model. For the HAVs, the W99 and EIDM are taken into consideration. The candidate models are selected based on two aspects: accessibility and capability. For Highly Automated Vehicles (HAVs), we take into consideration the W99 and EIDM models.

The selection of candidate models hinges on two primary criteria: accessibility and the model's capacity for accurately representing the longitudinal behavior of vehicles. Since the SUMO is selected as the simulation tool, the selected model must be accessible in the SUMO. Another standard is its capability to model the longitudinal behavior of AVs and HDVs. This selection of the candidate car-following model is predicated on the extensive utilization of the IDM within numerous prior investigations as the designated car-following model for AVs/CAVs. It is also noteworthy that employing the same car-following model for vehicles with different automation levels is advantageous, as it simplifies the comparison of model parameters and fosters a more consistent and scientifically robust evaluation.

As for the lane-changing model, the default model lc2013[8] is selected as the lane-changing model of the vehicle in the simulation. ACC-equipped vehicles and human-driven vehicles adhere to the default parameters, whereas high-level autonomous vehicles have modified parameters to align with their distinctive characteristics. To make the calibration result more comparable, the same car-following model EIDM is utilized.

4.1.1. W99

The Wiedeman 99 model is one of the most widely recognized psycho-physical models in traffic studies and serves as the default car-following model in PTV Vissim software. The W99[1] model encompasses four distinct conditions: free-flowing, approaching, following, and emergency. These conditions are determined by the ”perceptual threshold” boundary. The Wiedemann 99 car-following model is visually represented in Figure 4.1. The figure displays various boundaries that play an essential role in defining the model's behavior, including:

- AX: The desired distance between a vehicle and the one it is following.
- ABX: The desired minimum following distance, setting the lower limit for the following regime.
- SDX: The maximum following distance, establishing the upper limit for the following regime.
- SDV: The approaching point where the following driver perceives that they are closing in on a slower leading vehicle.
- CLDV: The decreasing speed difference between the following and leading vehicles.
- OPDV: The increasing speed difference between the following and leading vehicles.

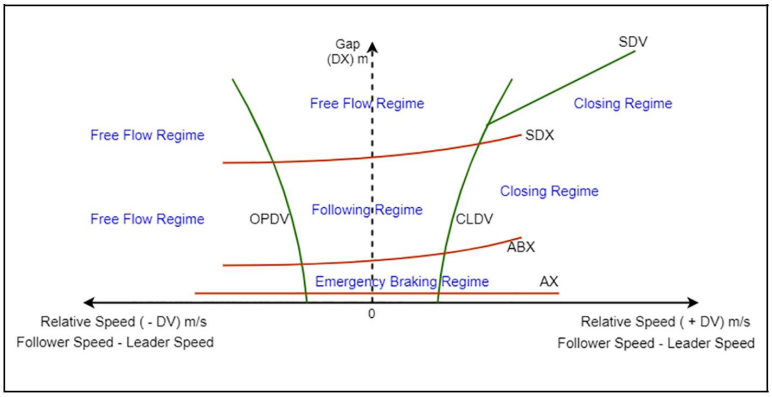


Figure 4.1: Schematic representation of Wiedemann model [1]

The parameters that will be calibrated are shown in Figure 4.1.

Table 4.1: Wiedemann 99 parameters

Parameters	Description
CC0	The desired gap between two vehicles in a stopped condition
CC1	Time gap following the driver keeps in for a safety in moving state
CC2	Range of gap between vehicles in the following regime
CC3	The time between the beginning of deceleration after perceiving of slow-moving leader to start the unconscious-following behavior
CC4	Speed difference during the following process. CC4 controls speed differences during the opening process (Negative relative speed),
CC5	Speed difference during the following process. CC5 controls speed differences in the closing process (Positive relative speed).
CC6	Influence of distance on speed oscillation during the following condition
CC7	Actual acceleration during oscillation in the unconscious-following regime
CC8	Desired acceleration when the vehicle starting from the standing condition
CC9	Desired acceleration at 80km/hr. However, it is limited by maximum acceleration for the vehicle type.

4.1.2. ACC/CACC model

Another model is the ACC (CACC) model developed by Milanés et al.[23]. Since the vehicles in the platoon are always under the following condition defined by the model (the clearance distance is smaller than the minimum threshold), then only the car following part will be calibrated. The car following part of the model is shown in Equation 4.1.

$$a = k_2(d - Tv - L) + k_3(v_l - v) \tag{4.1}$$

where k_2 is the gain in positioning difference between the preceding vehicle and the subject vehicle, and k_3 is the gain in speed difference between the preceding vehicle and the subject vehicle.

4.1.3. Enhanced IDM model

The EIDM is an advanced version of the Intelligent Driver Model developed by Treiber and Kesting. The EIDM builds upon the IDM's foundational principles and incorporates several model extensions to enhance its accuracy and applicability. These extensions primarily focus on reducing the jerk, especially during maneuvers like lane changes and accelerating from a standstill. The EIDM takes inspiration from various existing model extensions of the IDM and integrates them into a comprehensive framework. One of the notable features of the EIDM provided by SUMO is its flexibility, allowing individual extensions to be selectively activated or deactivated by adjusting their corresponding parameters. This adaptability ensures that the model can be tailored to specific driving scenarios and conditions. The IDM has been widely used as a reference for ACC car-following models [23]. The model is shown in Eqution 4.2 and Eqution 4.3.

$$a_{IDM} = a[1 - (\frac{v}{v_0}^{\delta} - (\frac{s^*(v, \Delta v)}{s})^2] \tag{4.2}$$

$$s^*(v, \Delta v) = s_0 + vT + \frac{v\Delta v}{2\sqrt{ab}} \tag{4.3}$$

The enhanced version of IDM is developed by Kesting et al.[16]. The model is shown in Function 4.4.

$$a_{ACC} = \begin{cases} a_{IDM} & if\ a_{IDM} \geq a_{CAH}, \\ (1-c)a_{IDM} + c[a_{CAH} + btanh(\frac{a_{IDM} - a_{CAH}}{b})] & otherwise. \end{cases} \tag{4.4}$$

where a_{IDM} is the acceleration calculated by the IDM, and a_{CAH} is the constant-acceleration heuristic (CAH) of the safe acceleration.

4.2. Calibration method

Before the calibration, the calibration method needs to be specified. The genetic algorithm is used in this project since the genetic algorithm is well used in the car-following model calibration and according to the optimization algorithms assessment performed by Punzo et al.[29], the genetic algothrim can get a more satisfactory result compared to other methods when dealing with a car following model calibration problem. The Python package Scikit-opt is used to realize the genetic algorithm.

4.2.1. Objective function

Before comes to the details of the calibration settings, it is necessary to determine the measure of performance (MoP) and the goodness of fit (GoF) need to be decided. As calibration is only a part of simulation preparation and not the core focus of this research, the choice of objective function is based on prior research. In the car-following model calibration, the spacing, speed, and acceleration are commonly selected as the MoP(measure of performance). Punzo et al.[30] compared various combinations of measures of performance and goodness of fit. Their findings indicated that using RMSE and acceleration as a measure of performance is the best option when using a single MoP. This calibration approach is more Pareto efficient than other combinations of single MoP. Thus, the acceleration is used as the MoP, and the GoF is selected as RMSE. The objective function is shown in Equation 4.5.

$$RMSE_{spacing} = \sqrt{\frac{\Sigma_{i=1}^{N}(s_i^{sim} - s_i^{real})^2}{N}} \tag{4.5}$$

Where s^{sim} and s^{real} represent simulated and ground truth spacing, respectively, and N is the number of data points.

4.2.2. Calibrator settings

The configuration parameters employed for the Genetic Algorithm calibrator are provided in Table 4.2. These settings are the default values of the calibrator, and it's important to note that the maximum number of iterations has been set to a sufficiently high value to achieve satisfactory results.

Table 4.2: Calibrator settings

Parameter	Value
Maximum number of iterations	500
Population size	50
Dimension of objective function	3
Mutation rate	0.001
Precision	1.00E-04

The default values of EIDM parameters are determined based on the simulation parameters utilized in the research conducted by Kesting[16] et al. The boundaries of these parameters are established according to the previous work on IDM model calibration.

Table 4.3: EIDM parameters bounds

Parameters	Description	Bounds	Default value
v_0	desired speed	[1, 70]	33.4
T	desired time gap	[0.1, 5]	1.1
s_0	jam distance	[0.1, 8]	2
a	maximum acceleration	[0.1, 6]	1.4
b	desired deceleration	[0.1, 6]	2
c	coolness factor	[0, 1]	0.99

As for the ACC model, the bounds and default values of the parameters are shown in Table 4.4. The default value of k_2 and k_3 are the simulation result from Milanés et al.[23].

Table 4.4: Bounds and default values of ACC/CACC model

Parameter	Desicription	Bounds	Default value
k_2	gain in positioning difference between the preceding vehicle and the subject vehicle	[0, 1]	0.23
k_3	gain in speed difference between the preceding vehicle and the subject vehicle	[0, 1]	0.07
t	Desired time gap	[0, 3]	1.1

The parameter settings of W99 model are shown in Table 4.5. The initial values are set with reference to Zhu et al.[42]. And the default value is the recommanded value for Intermediate ACC vehicle accodring to Bierstedt et al.[5].

Table 4.5: W99 car following model parameters

Parameters	Desicription	Bounds	Default value
CC0	tandstill Distance	[0, 20]	1.25
CC1	Spacing time	[0, 5]	0.8
CC2	Following Variation	[0, 10]	3
CC3	Threshold for Entering ”Following”	[-20, 0]	-12
CC4	Negative ”Following” Threshold	[-5, 0]	-0.35
CC5	Positive ”Following” Threshold	[0.1, 5]	0.35
CC6	Speed dependency of oscillation	[0.1, 20]	0
CC7	Oscillation acceleration	[-1, 1]	0.25
CC8	Standstill acceleration	[0, 8]	3.5

CC9	Acceleration at 80 km/h	[0, 8]	1.5
VDES	Desired speed of following vehicle	[72, 108]	90

4.3. Model selection

In this section, the car-following model for the simulation is chosen following a thorough calibration process. The average RMSE values obtained from the calibration results are presented in Table 4.6. Assessing the overall calibration results, it is evident that the ACC model and the EIDM model exhibit a notably superior fit when compared to the W99 model, as indicated by the significantly larger RMSE values for the W99 model. Consequently, the selection of the car-following model will be made from the EIDM and ACC models.

Table 4.6: Average RMSE of different models

	EIDM	W99	ACC model
ACC dataset	0.545	5.209	0.239
High level AV dataset	0.057	0.043	

The comparisons between the acceleration generated by the calibrated car-following models and the ground truth data are illustrated in Figure 4.2 and Figure 4.3. Notably, the trajectory produced by the ACC model exhibits a closer alignment with the ground truth data. However, it is essential to acknowledge that the ACC model falls short in capturing the extrema of acceleration and deceleration during the driving process. This discrepancy may be attributed to the fact that the calibrated values of parameters k_2 and k_3 significantly deviate from the recommended values. Specifically, the recommended values are 0.23 for k_2 and 0.07 for k_3," whereas the calibration results yield a higher value for k_3 and a substantially lower value for k_2. Consequently, the calibrated results exhibit a superior conformity to the observed acceleration when compared to the EIDM model. However, it is important to note that the ACC model does not effectively capture extreme acceleration and deceleration events.

To further make the decision between the ACC model and EIDM, a pilot simulation is conducted to evaluate the capability of the ACC model and EIDM to represent the ACC-equipped vehicles. The simulation scenario is a simplified version of the same one introduced in Section 5.2. The number of errors is shown in Table 4.7. The number of collisions and emerging brake incidents associated with the EIDM are deemed acceptable, with no collisions and only 1 emerging brake event recorded during the two-hour simulation. In contrast, the ACC model generated a significantly higher number of undesirable outcomes, which is deemed unacceptable. This model recorded 322 collisions and 203 emerging brake events, highlighting a sub-

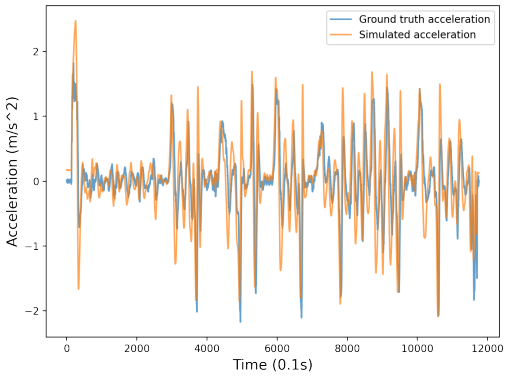


Figure 4.2: Comparison of ground truth acceleration vs. ACC Model generated acceleration

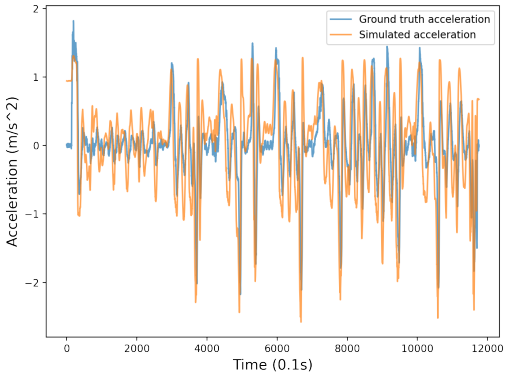


Figure 4.3: Comparison of ground truth acceleration vs. EIDM generated acceleration

stantial performance gap in dealing with merging problems when compared to the EIDM model. Analysis of the simulation logs reveals that scenarios employing the ACC model consistently generate emerging brake actions and collisions at specific locations, specifically at edges E2 and E4. These edges correspond to positions following the onramp, where vehicles are in the process of merging into the main traffic stream. This recurrent issue can be attributed to the underlying data used for calibrating the ACC model, which were collected under stable car-following conditions. Consequently, the ACC model is calibrated with conservative values for speed gain and space gain, making it ill-suited for simulations involving vehicle merging scenarios. While one potential solution could involve the introduction of human takeover, implementing this in the SUMO environment is not straightforward. SUMO only offers Gap-closing control mode and Collision avoidance control mode for the ACC model. Even with both modes activated, only collision incidents are mitigated, leaving an unacceptably high number of emerging brake events due to the absence of appropriate values for speed gain and space gain. As a result, the EIDM emerges as the most suitable car-following model for accurately simulating heterogeneous ACC-equipped vehicles in scenarios involving vehicle merging.

Table 4.7: Collisions and brake applications of different models

	Number of collisions	Number of emergency brake applications
EIDM	0	1
ACC model	322	203

Regarding high-level automated vehicles, the EIDM model is the chosen car-following model for representing longitudinal behaviors. This selection is predicated on the extensive utilization of the IDM within numerous prior investigations as the designated car-following model

for Connected and Automated Vehicles and using the same car-following model for vehicles with different automation levels will make it easier to compare the parameters. Given that the EIDM effectively represents AVs within the model, it becomes the logical choice for maintaining uniformity. In a pilot simulation, it is worth noting that both the EIDM car-following models can be utilized without leading to unacceptable levels of collisions or emerging brake incidents.

4.4. Calibration result analysis

4.4.1. ACC vehicle calibration result

The mean, median, and coefficient of variation of the calibrated parameters are shown in Table 4.8.

Table 4.8: ACC vehicle EIDM calibration results

	v_0	**T**	**s_0**	**a**	**b**	**c**	**RMSE**
Mean	38.46	1.01	2.30	0.72	2.47	0.87	0.55
Coefficient of variation	0.46	0.37	0.81	0.87	0.84	0.33	1.53
Median	34.98	1.10	2.00	0.61	1.96	1.00	0.28

The box plots of the calibrated results are shown in Figure 4.5 and Figure 4.4. Observations reveal that while there are a few outliers, the majority of the recorded time gaps fall within the range of 1 to 2 seconds. The calculated mean and median time gaps are 1.01 seconds and 1.1 seconds, respectively. These figures closely align with the anticipated values for the desired time gap in ACC-equipped vehicles. In terms of minimum distance, the prevalent range lies between 1 meter and 3 meters. This range aligns well with our expectations. In the context of acceleration, the ACC-equipped vehicles demonstrate relatively modest acceleration profiles. The majority of recorded accelerations fall below 1.2 m/s². This behavior can be attributed to two factors. Firstly, commercial ACC systems intentionally restrict the maximum acceleration to enhance passenger comfort. Secondly, it's important to note that the dataset primarily comprises platoon driving scenarios, where instances of rapid acceleration, such as going from a standstill, constitute a small percentage or may even be absent. Regarding comfortable deceleration, the distribution exhibits significant variability. The level of comfortable deceleration is notably influenced by the frequency of sharp deceleration events within the trajectory. In terms of the desired speed, the calculated median speed is 34.98 m/s, a value consistent with typical highway speed limits.

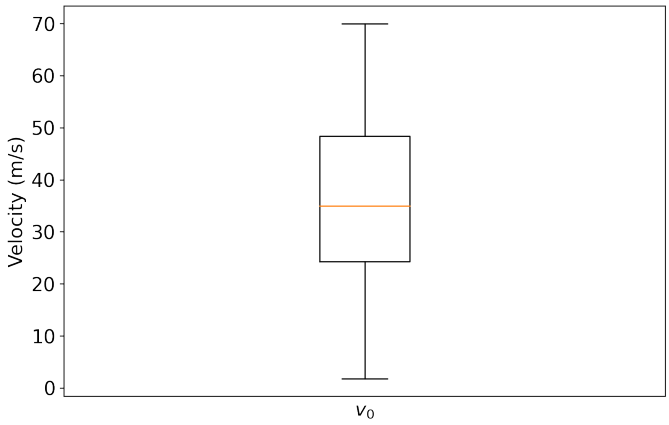


Figure 4.4: Boxplot of ACC vehicles EIDM desired speed

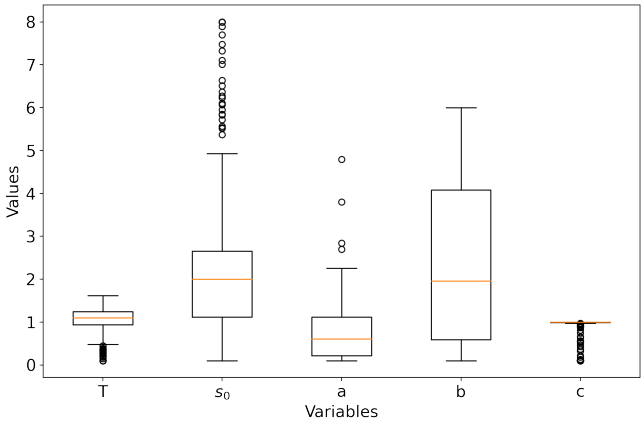
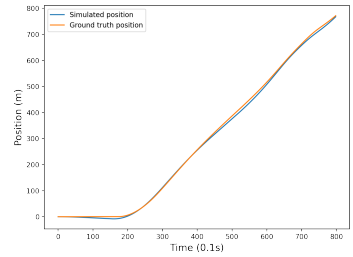


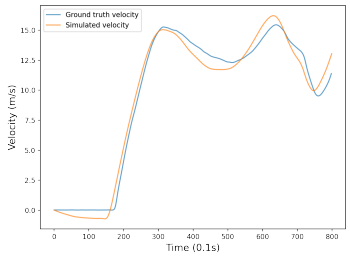
Figure 4.5: Boxplot of ACC vehicles EIDM parameters

To further validate the quality of the calibration, the trajectories of the vehicles are plotted to make a comparison between the ground truth and the car-following model simulated trajectories. The acceleration, velocity, and position of the simulated result and ground truth data are shown in Figure 4.6. Due to the objective function of the calibration being selected as the RMSE of the spacing, both the position, velocity, and acceleration have a relatively good fit. Both the short and long trajectories are displayed. For shorter trajectories, the alignment of position, velocity, and acceleration is notably high, though the peaks of the acceleration are not perfectly captured. However, for longer trajectories, while acceleration exhibits a strong correspondence, some disparities emerge in both velocity and position. It can be observed that there is a significant duration during which the simulated velocity consistently remains lower than the corresponding ground truth velocity. Overall, the trajectory analysis suggests that the EIDM effectively represents the car-following behavior of the ACC-equipped vehicles. It aptly

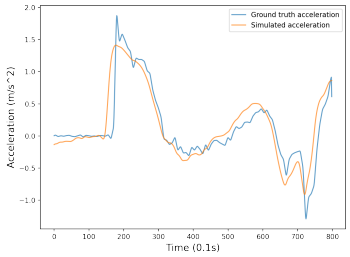
captures the trend and variations in acceleration, despite some discrepancies in simulated velocity and position.



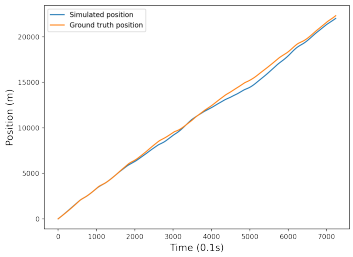
(a) Simulated and real position (short trajectory)



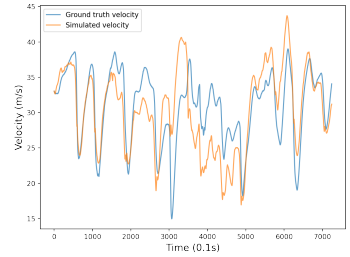
(b) Simulated and real velocity (short trajectory)



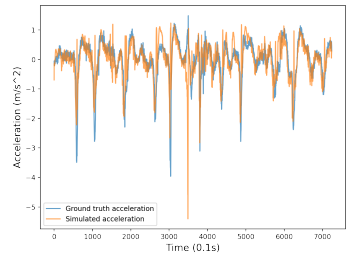
(c) Simulated and real acceleration (short trajectory)



(d) Simulated and real position (long trajectory)



(e) Simulated and real velocity (long trajectory)



(f) Simulated and real acceleration (long trajectory)

Figure 4.6: Simulation of ACC vehicles trajectories by GA-calibrated EIDM acceleration, velocity, and position

4.4.2. ACC string stability

Treiber and Kesting[37] introduced a simple explicit stability criterion for evaluating the string stability of car-following models. The simplified criterion for the IDM is expressed as follows:

$$a \geq \frac{s_0}{T^2} \tag{4.6}$$

Where a denotes the maximum acceleration, s_0 represents the standstill distance, and T signifies the desired time gap. Treiber and Kesting also confirm that the EIDM, when subjected to the constant-acceleration heuristic constraint, generally exhibits behavior closely aligned with that of the IDM, thus preserving its well-established and desirable properties. Consequently, this criterion is applied to the calibration results of the EIDM to conduct a preliminary assessment of the string stability of the modeled vehicles. The results are shown in Figure 4.7.

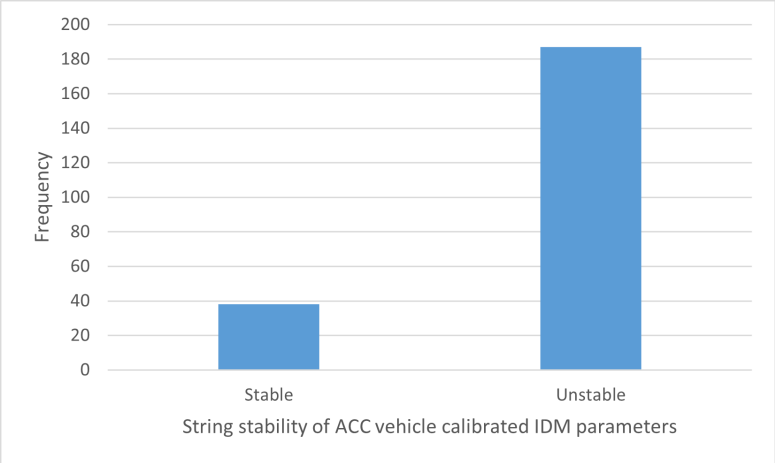


Figure 4.7: ACC vehicle string stability

The outcomes of this analysis reveal that while a majority of vehicles exhibit instability, there are still some sets of parameters that result in stable behavior. Importantly, these findings are consistent with the results obtained from the data analysis.

4.4.3. HAV calibration result

The overall review of the high-level autonomous vehicles calibration result are shown in Table 4.9.

Table 4.9: HAV EIDM calibration result

	v_0	T	s_0	a	b	c	RMSE
Mean	26.91	2.47	4.99	0.94	1.34	0.60	0.04
Coefficient of variation	0.81	0.58	0.51	0.77	1.29	0.66	0.62
Median	16.03	2.23	5.90	0.79	0.58	0.75	0.04

The box plots of the calibrated results are shown in Figure 4.9 and Figure 4.8. The mean desired speed for the autonomous vehicle is 26.86 m/s, considerably lower than the average desired speed exhibited by the ACC vehicles. This contrast is attributed to the inclusion of urban driving scenarios within the dataset of level 4 autonomous vehicles, where lower desired speeds are typical due to the urban environment's characteristics. The desired time gap of the HAV is mainly distributed between 1.5s and 3.8s, as evidenced by the boxplot representation. Notably, the coefficient of variation for this parameter is higher when compared to ACC-equipped vehicles which indicates the HAV vehicle is more heterogeneous than ACC-equipped vehicles. Regarding the minimum distance, both the mean and median values are calculated at 5 meters and 5.9 meters, respectively. These findings regarding the desired time

gap and minimum distance collectively suggest that HAVs adopt a more cautious driving style compared to their ACC-equipped counterparts. Furthermore, when examining acceleration behavior, it is observed that HAVs exhibit smaller maximum acceleration values when compared to human-driven vehicles. This conservative acceleration pattern is likely attributed to the overall cautious driving behavior adopted by the HAVs.

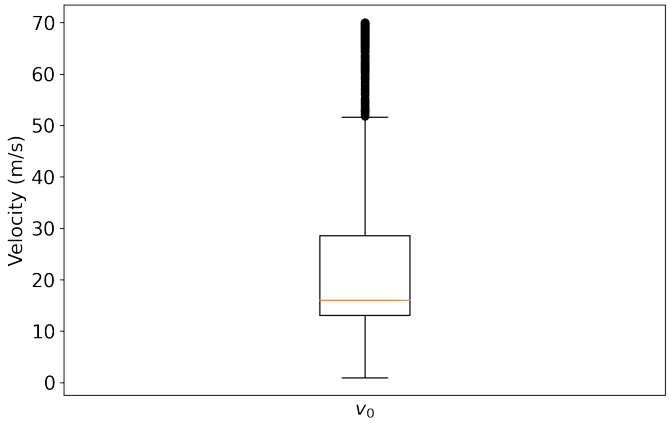


Figure 4.8: Boxplot of HAV calibration results

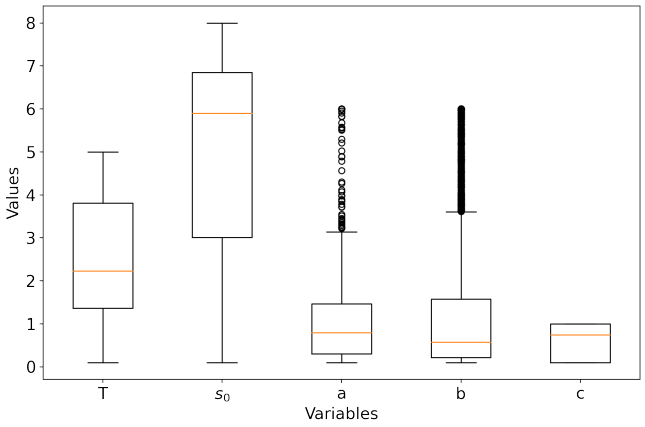
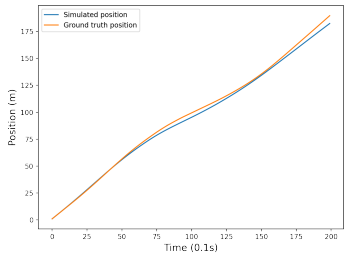


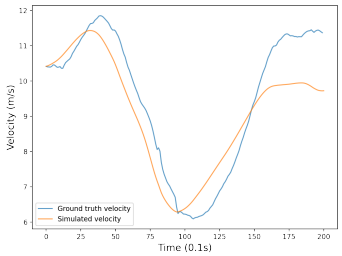
Figure 4.9: Boxplot of HAV calibration results

The simulated trajectories and ground truth data pertaining to Highly Automated Vehicles are presented in Figure 4.10. In the context of representing HAV behavior, there are some unfavorable calibration results, primarily attributable to the inherently more intricate nature of HAV dynamics. For instance, in trajectory two, it becomes evident that although the EIDM captures the general trends satisfactorily in acceleration, disparities emerge in the estimation of position and velocity. Notably, as demonstrated in Figure 4.10e, the velocity profile of the simulated data exhibits shortcomings, primarily attributed to the model's inability to accurately

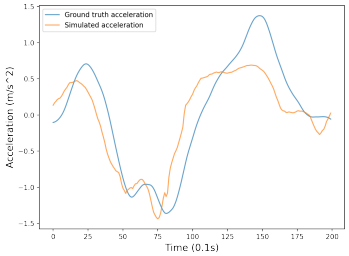
replicate the peak acceleration values. In a broader assessment, it becomes apparent that the overall efficacy of employing the EIDM to replicate car-following behavior in HAVs falls short when compared to its performance in the context of ACC-equipped vehicles. However, using the EIDM to reproduce the HAVs could still be a reasonable idea because of EIDM's capability of capturing the trends of acceleration.



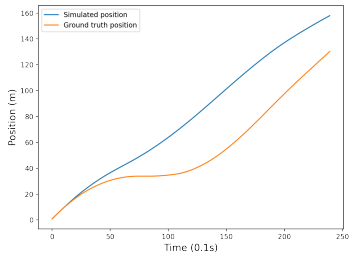
(a) Simulated and real position (data segment 1)



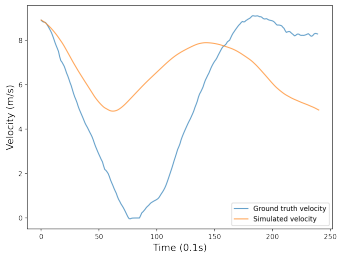
(b) Simulated and real velocity (data segment 1)



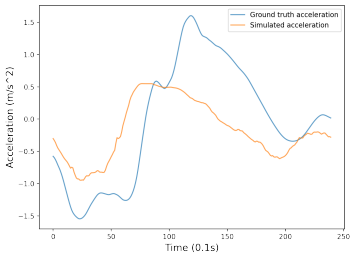
(c) Simulated and real acceleration (data segment 1)



(d) Simulated and real position (data segment 2)



(e) Simulated and real velocity (data segment 2)



(f) Simulated and real acceleration (data segment 1)

Figure 4.10: Simulation of HAV trajectories by GA-calibrated EIDM acceleration, velocity, and position

4.5. Summary

The objective function is selected as the RMSE of spacing, guided by existing literature and research objectives. Based on the calibration results, the car-following model for use in simulations is determined. The EIDM was selected for its accurate simulation of position, velocity, and acceleration, which aligns with the research goals. To maintain consistency, the EIDM is also selected as the car-following model for HAVs and HDVs. The calibration results indicate that the EIDM can effectively represent ACC vehicles in both short and long trajectories. Although most desired time gaps hover around 1.1 seconds, with a small coefficient of variation, there are some outliers. Nonetheless, this is consistent with expectations. It has been observed that many calibrated parameter sets for ACC vehicles exhibit poor string stability, which is consistent with the data analysis findings. However, calibration results for HAV data are different. Some trajectories align well with the model, while others exhibit less fitting due

to the limited duration of each HAV data segment (20 seconds at a 10Hz frequency). The short segment duration leads to significant parameter variations. The calibrated desired time gaps for ACC vehicles are reliable, but the reliability of desired time gaps is contingent on the quality of the data. Nevertheless, the larger volume of HAV data contributes to the increased reliability of the calibration results.

5

Simulation setup

In this chapter, two simulation scenarios are established to explore the impact of heterogeneity within AVs. These scenarios involve variations in AV levels, with one focusing on the heterogeneity among AVs at the same automation level, and the other examining heterogeneity across different levels of AVs. The setup process encompasses the determination of vehicle behaviors within the simulation. Furthermore, the chapter provides a detailed introduction to each scenario, which includes the geometric design of the motorway section, specifications regarding input flows into the main section and the onramp, and the placement of loop detectors. Additionally, the composition of traffic flows within both scenarios is presented.

5.1. Parameter settings

Within the simulation framework, two parameters extracted from the calibrated values are embedded within the enhanced Intelligent Driver Model (IDM): the desired time gap and the maximum acceleration.

In the case of several other parameters, the desired speed is significantly influenced by the prevailing speed limit in the given location. Conversely, for the standstill distance, no discernible specific trends are identified. In the context of the coolness factor, the distribution exhibits a relatively tight concentration, with the majority of values clustering around 1. The rationale behind selecting these two parameters is also to emphasize the primary factors that exert the most significant influence on the study. The omission of these parameters is intended to underscore the impact of the desired time gap and acceleration, which aligns with the primary focus of the data analysis. In the data analysis segment, the emphasis is predomi-

nantly on assessing the influence of time gap and acceleration, rather than speed or headway distance. The selection of the parameters also acts as the simplification of the research. Furthermore, it is important to note that the method for modeling heterogeneous traffic relies on random parameter selection. Increasing the number of randomly chosen parameters raises the likelihood of vehicles displaying unusual or erratic behavior.

The desired time gap and the parameters related to acceleration constitute the primary manifestations of heterogeneity in external factors. The decision to omit deceleration from this context arises from the intricate nature of calibrating comfortable deceleration. This calibration process significantly hinges upon the dataset segment. In situations where the segment predominantly comprises stable data, occurrences of abrupt deceleration are scarce. Consequently, the calibrated deceleration values derived from such segments might not accurately represent the magnitude necessary for faithful simulation outcomes.

Then the best-fit distribution of the desired time gap and acceleration are selected from all distributions supported by the Scipy library. The fit method of the Scipy will return the maximum likelihood estimates of the parameters of each distribution, then the best-fit distribution of the calibrated parameters will be determined based mainly on the sum of the squared estimate of errors (SSE). Additionally, the AIC and BIC of each fitted distribution are also calculated to help determine the selection of the distribution. The AIC is computed using $aic = 2*k-2*log(Lik)$, and the BIC is computed as $k*log(n)-2*log(Lik)$. Both AIC and BIC provide ways to evaluate and compare models based on a combination of how well they fit the data and how complex they are. AIC strikes a balance between fit and complexity, while BIC leans towards simpler models by imposing a stronger penalty for complexity.

5.1.1. HDV parameters

The parameters of the IDM for human-driven vehicles are sourced from existing literature. The research also takes into account the heterogeneity within human-driven vehicles. However, it's important to note that the primary focus of this research does not revolve around the heterogeneity of human-driven vehicles. Therefore, the distribution of parameters for human-driven vehicles is drawn from established literature.

The shape of the parameter distribution is derived from the work of Kim and Mahmassani[17]. In their study, they calibrated trajectories using data from the Generation Simulation trajectory data and subsequently obtained the shape of the lognormal distribution that best fits the car-following model parameters. The specific parameters of the distribution for the car-following model of human-driven vehicles are presented in Table 5.1.

Table 5.1: EIDM parameter best Fit distribution (HDV)

	Distribution	Distribution parameters
T	lognormal	{'μ': 1.27, 'σ': 0.50}
a	lognormal	{'μ': 1.41, 'σ': 1.01}

As the final parameters will be sampled from a distribution, specific limits are imposed on the desired time gap and acceleration to prevent the emergence of impractical values. These limits are defined in accordance with the boundaries established by Treiber and Kesting[37]. In practice, realistic values for the desired time gap typically range from 2 seconds to 0.8 seconds or even lower. Similarly, the practical range for maximum acceleration usually falls within the interval of $0.8m/s^2$ to $2.5m/s^2$.

5.1.2. ACC vehicle parameters

For the desired time gap, the five distributions with the smallest SSE are shown in Table 5.2. The AIC value for the gamma distribution is 136.03, surpassing those of both the Weibull and Laplace distributions. However, it remains smaller than the values observed for other distributions. Concurrently, the BIC value for the gamma distribution is 145.77. Both the AIC and BIC values exhibit relatively modest magnitudes, whereas the SSE is significantly lower in comparison to alternative distributions. Consequently, the gamma distribution is identified as the optimal choice for modeling the heterogeneous desired time gap among ACC-equipped vehicles.

Table 5.2: Desired time gap distributions fitting result (ACC)

	sumsquare_error	aic	bic
dgamma	83.12	136.03	145.77
gennorm	92.58	137.15	146.89
dweibull	98.25	118.85	128.59
laplace_asymmetric	104.21	141.62	151.36
laplace	108.81	121.25	127.75

As for the acceleration, the five distributions with the smallest SSE are shown in Table 5.3. The selection of the Burr distribution is substantiated by the significantly lower values of the AIC and BIC compared to alternative distribution choices while the differences in SSE are very small.

Table 5.3: Acceleration distributions fitting result (ACC)

	sumsquare_error	aic	bic
burr	12.82	770.23	780.14
hypsecant	12.78	1126.90	1131.85
genlogistic	12.84	987.33	994.77
gumbel_r	12.84	984.45	989.41
logistic	12.86	1247.97	1252.92

Within this context, the optimal distribution model for the desired time gap is the dgamma distribution, whereas the burr distribution characterizes the best-fit distribution for acceleration patterns. The most probably distribution parameters associated with EIDM are presented in Table 5.4. Parameters ”a”, ”c” and ”d” govern the shape of the distribution, while ”loc” and ”scale” serve as parameters utilized to adjust the distribution’s location and scale.

Table 5.4: EIDM parameter best fit distribution (ACC Vehicle)

	Distribution	Distribution parameters
T	dgamma	a’:0.78, ’loc’: 1.10, ’scale’: 0.20
a	burr	c’: 3.50, ’d’: 0.49, ’loc’: 0.76, ’scale’: 0.63

Illustrations of the desired headway distributions are shown in Figure 5.1 and Figure 5.2. The empirical data are the desired time gap distribution that calibrated from the database, the theoretical density is from the best-fit distribution.

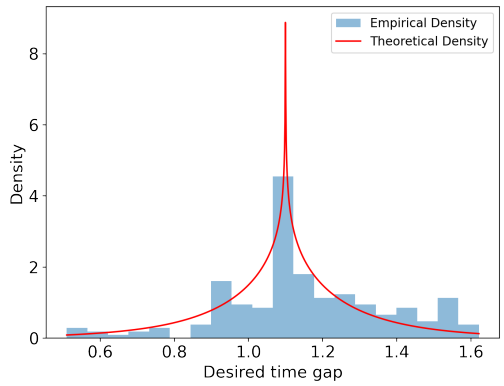


Figure 5.1: Empirical and theoretical density of desired time gap distribution

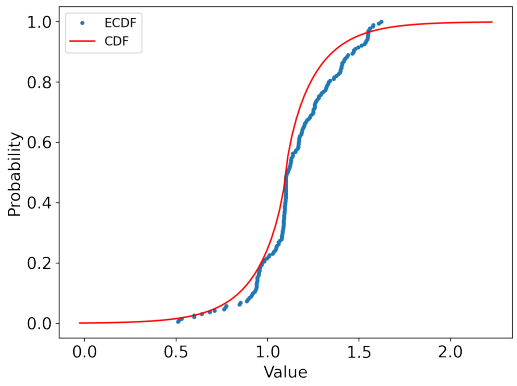


Figure 5.2: Empirical and theoretical CDF of desired time gap distribution

It can be observed that the distribution of desired time gaps exhibits a pronounced concentration of around 1.1 seconds. Furthermore, there is a difference between the desired time gap

exceeding 1.1 seconds and those not meeting this threshold. The empirical and theoretical distributions exhibit overall trends; however, a notable divergence exists: the theoretical cdf commences at 0 seconds, whereas the smallest observed desired time gap within the empirical dataset approximates 0.5 seconds. Furthermore, the maximum desired time gap among ACC-equipped vehicles registers at approximately 1.6 seconds. These findings suggest that the fitted distribution may effectively represent the distribution of time gaps within a specific range, specifically from 0.5 seconds to 1.5 seconds.

In order to enhance the validation of the best-fit distribution, a probability plot of the empirical data and the selected distribution is provided. The probability plot effectively communicates the degree of fit along the x-axis. As shown in Figure 5.3, the observed distribution closely aligns with the theoretical distribution within the range of 0.5 seconds to 1.5 seconds. However, notable discrepancies become apparent when the desired time gap falls below 0.5 seconds or exceeds 1.5 seconds. This outcome corroborates earlier analysis.

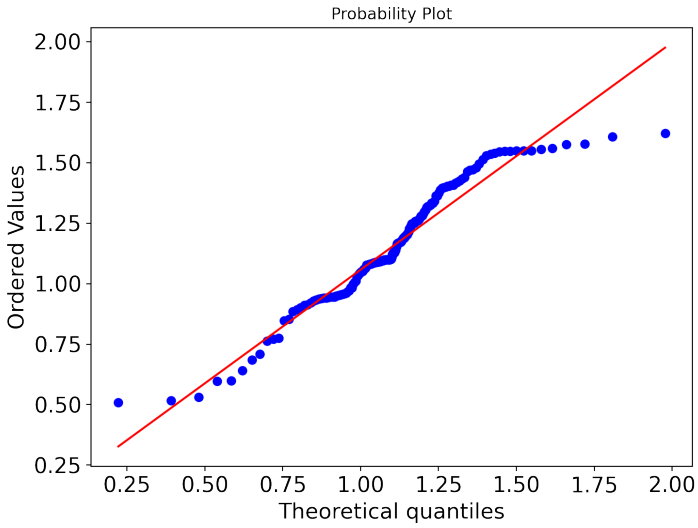


Figure 5.3: Probability plot of ACC vehicle desired time gap

In terms of acceleration characteristics, the empirical and theroratic distribution are shown in Figure 5.4 and Figure 5.5. The majority of maximum accelerations are below 2m/s^2. And a sharp decline is observed after crossing the 1.2m/s^2 threshold, and almost disappearing after reaching 3m/s^2. The trend of the empirical data closely resembles that of the theoretical data before the $1.5m/s^2$, although there is a notably steeper decline in the empirical dataset. Specifically, a substantial divergence becomes evident at an acceleration rate of $1.75m/s^2$. The disparity might be because of the limited size of the calibrated parameters.

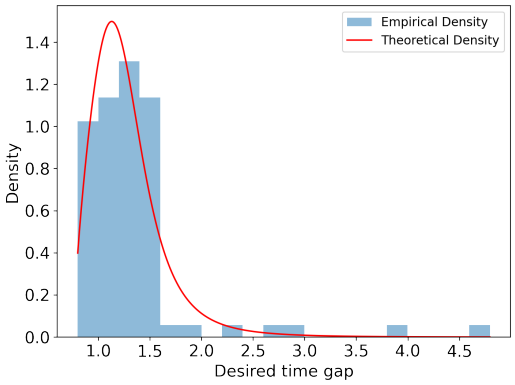


Figure 5.4: Empirical and theoretical density of acceleration distribution

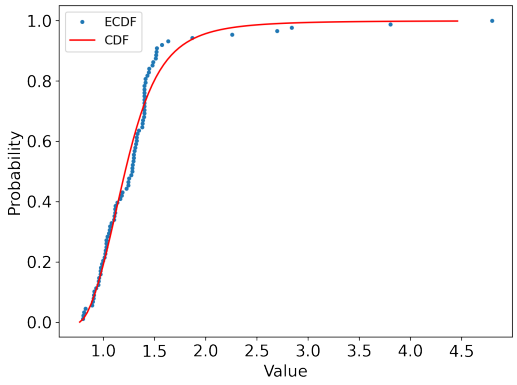


Figure 5.5: Empirical and theoretical cdf of acceleration distribution

The probability plot of ACC acceleration and best-fit distribution is shown in Figure 5.6. The degree of fit shown in the probability plot is less satisfactory; nevertheless, it remains adequate for conveying the overarching pattern of acceleration, which is characterized by uniformity between $0.8m/s^2$ and a sharp decrease after reaching $1.5m/s^2$.

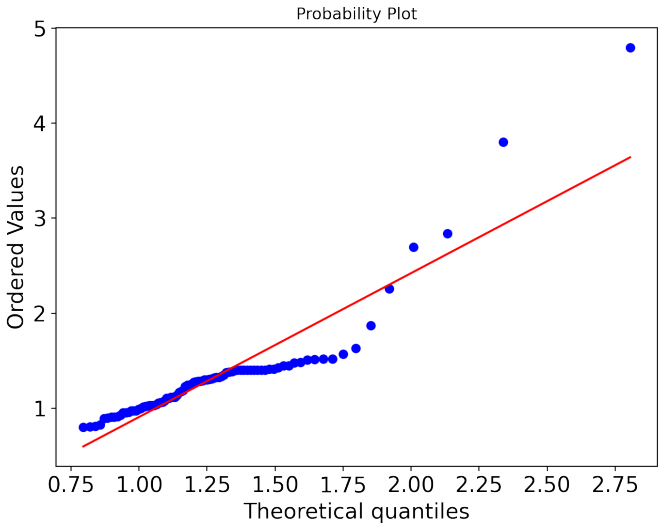


Figure 5.6: Probability plot of ACC vehicle acceleration

Drawing from the empirical and theoretical cumulative distribution functions, along with the empirical and theoretical probability density functions, and considering the information gleaned from the probability plot, we can identify the most suitable distribution for the acceleration and desired time gap variables. This best-fitted distribution can then be employed for resampling purposes.

5.1.3. HAV parameter setting

Differing from the desired time gap distribution of the ACC vehicles and HDV, from Figure5.7 it can be clearly seen that there are two clear clusters in the desired time gap distribution. The median of the first cluster is around 1.5s and the median of the second cluster is about 4 seconds. The assumption is made that the desired time gap of HAV is the mixture of two different distributions. This assumption is reasonable since the traffic context of the lyft5 data includes both the expressway and urban road. The cluster with a smaller desired time gap could represent a reasonable simulation time gap. The Bayesian Gaussian mixture model will be applied to filter the data that are not suitable for the simulation. Bayesian Gaussian mixture models constitute a form of unsupervised learning and can be useful in fitting multi-modal data for tasks such as clustering, data compression, outlier detection, or generative classifiers. Therefore, the GMM could be used to cluster the one-dimensional data. Another thing that could be seen is that the too many values around 0s and 5s, this is because of the boundary. This phenomenon also appears in many other car-following model calibration works[28][12]. So before the clustering, the time gaps that are smaller than 0.1s or larger than 5s are filtered out. These values are considered as noise of GA. The scikit-learn library of Python is used to divide the calibrated parameters into two clusters, the clustering result is in Figure 5.7. The boundary of the two clusters is 3 seconds, for the desired time gaps are larger than the 3 seconds are assumed as the data that is not suitable for the simulation since the context is different.

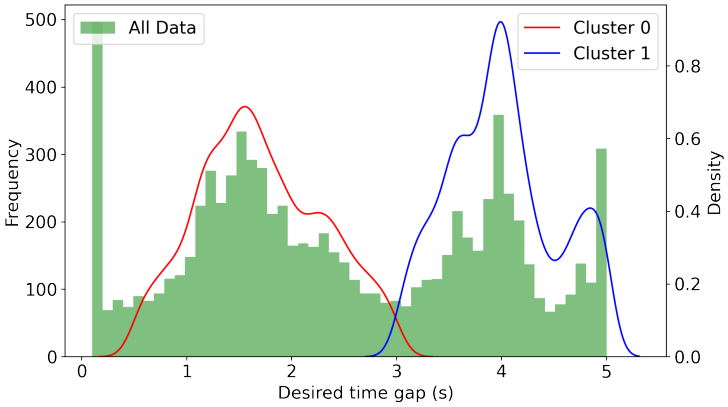


Figure 5.7: Clustering result of desired time gap

Then apply the same process as the ACC vehicle parameter distribution fitting. Then the best fitting result of the HAV vehicle is shown in Table 5.5 and Table 5.6. The Maxwell distribution emerges as the most suitable fit for the desired time gap distribution. This conclusion is supported by the observation that the SSE for the Maxwell, Chi, and Nakagami distributions

are comparable, while the AIC and BIC values for the Maxwell distribution are lower. These AIC and BIC values suggest a superior model fit when accounting for model complexity. The Maxwell distribution exhibits lower complexity while maintaining an equivalent level of fitting quality.

Table 5.5: Desired time gap distributions fitting result (HAV)

	sumsquare_error	aic	bic
maxwell	0.85	229.74	242.50
chi	0.85	232.14	251.28
nakagami	0.85	232.14	251.28
burr12	0.87	231.88	257.40
chi2	0.90	235.88	255.02

The Maxwell distribution is also identified as the optimal fit for the acceleration data, primarily due to its significantly lower SSE. However, the AIC and BIC values for the Maxwell distribution are relatively higher. This outcome implies that the Maxwell distribution offers a superior fit while being associated with a higher level of complexity.

Table 5.6: Acceleration distributions fitting result (HAV)

	sumsquare_error	aic	bic
maxwell	0.35	1945.03	1957.29
rayleigh	0.39	1778.29	1790.55
norm	0.40	2635.32	2647.58
logistic	0.48	1359.28	1371.54
dweibull	0.54	1465.51	1483.91

The parameters of the best fitting distribution are shown in Table 5.7.

Table 5.7: EIDM parameter best fit distribution (HAV)

	Distribution	Distribution parameters
T	maxwell	{'loc': 0.27, 'scale': 0.90}
a	maxwell	{'loc': 0.52, 'scale': 0.65}

The optimal-fit distribution for the desired time gap is depicted in Figure 5.8 and Figure 5.9. In comparison to the desired time gap distribution observed in ACC-equipped vehicles, the distribution of desired time gaps for highly autonomous vehicles exhibits greater variability. Additionally, the HAVs display an overall larger desired time gap in comparison to ACC-equipped

vehicles. This finding concurs with the conclusions drawn during the data analysis phase. From the CDF plot, it can be seen that the overall fitted situation of the maxwell distribution is pretty well.

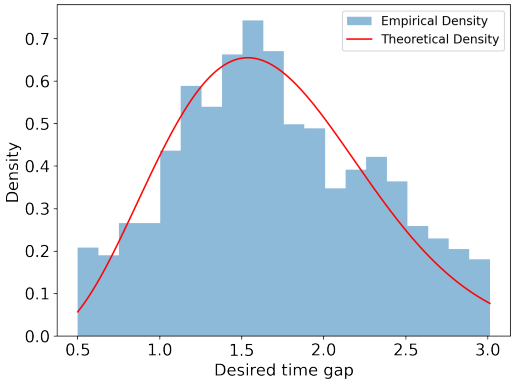


Figure 5.8: Empirical and theoretical density of desired time gap distribution

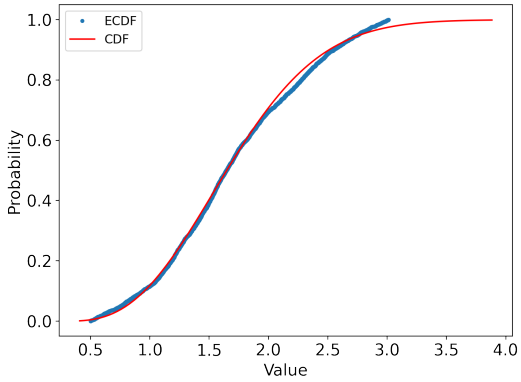


Figure 5.9: Empirical and theoretical CDF of desired time gap distribution

The probability plot of the best-fitted distribution is shown in Figure 5.10. It is evident that the Maxwell distribution provides a satisfactory fit for the data within the time interval ranging from 0.5 seconds to 3 seconds. The result of the probability plot suggested that the Maxwell distribution could be used to re-sample the desired time gap of the HAVs.

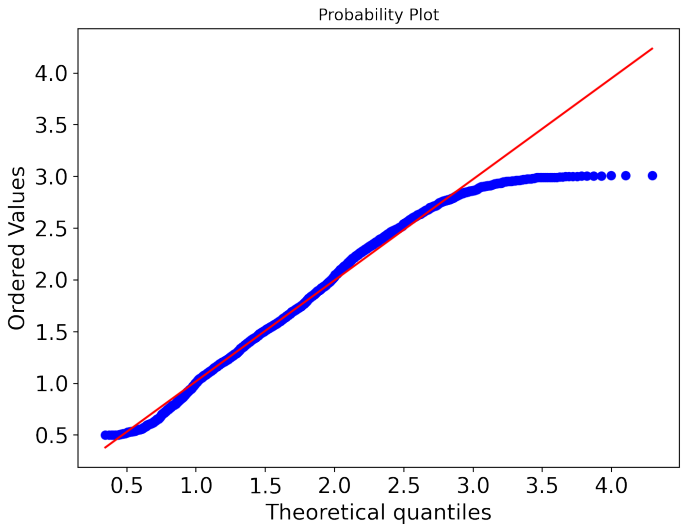


Figure 5.10: Probability plot of ACC vehicle acceleration

In terms of the maximum acceleration, the illustration of the distribution and CDF are shown in Figure 5.11 and Figure 5.12. The accelerations are main between the $0.8m/s^2$ and the $2.5m/s^2$, then there is a sharp decline before $3m/s^2$. The probability reduces to almost 0

after $3m/s^2$. The cdf plots show that the empirical distribution and the theoretical distribution overall fit pretty well. HAVs exhibit a higher overall maximum acceleration when contrasted with ACC-equipped vehicles.

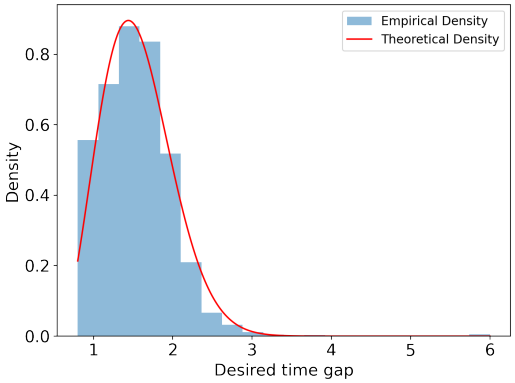


Figure 5.11: Empirical and theoretical density of desired time gap distribution

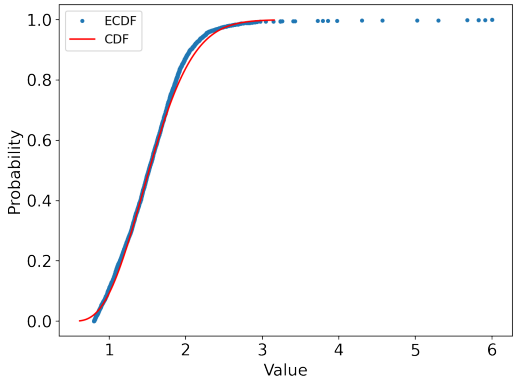


Figure 5.12: Empirical and theoretical CDF of desired time gap distribution

The probability plot is shown in Figure 5.13. The data almost fit the line within a certain scope. The probability plot of the HAV maximum acceleration suggested a similar result with the HAV desired time gap: the Maxwell provides a satisfactory fitted result and could be used as the maximum acceleration of the highly autonomous vehicles.

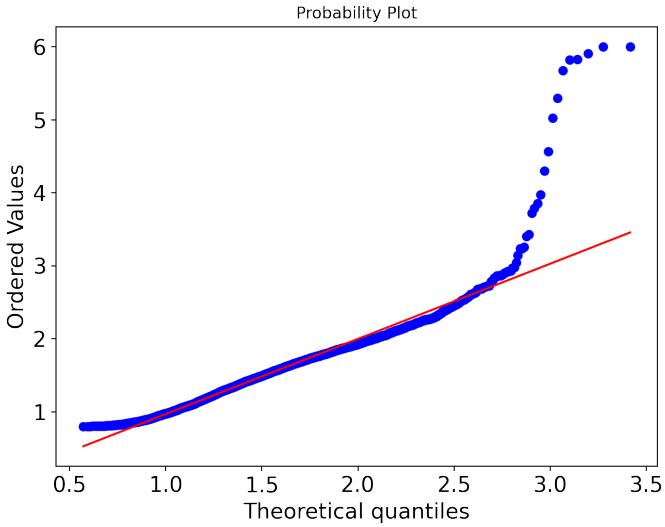


Figure 5.13: Probability plot of ACC vehicle acceleration

5.2. Simulation scenario

The simulation scenario in this research is a motorway section. The sketch of the motorway is shown in Figure 5.14. The road section contains a main section and an on-ramp. The function of the main section is to let the vehicles accelerate to a suitable speed before the on-ramp. The main section also keeps the congestion not spilling back to the start position. The on-ramp is used to invoke the merging disturbance.

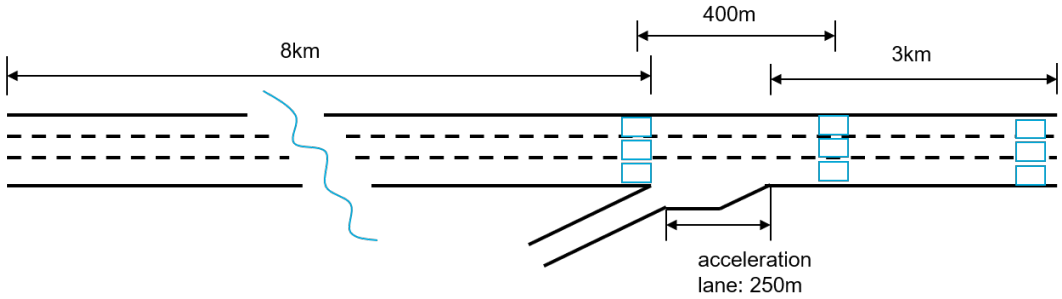


Figure 5.14: Road section of the simulation

The loop detectors are employed to produce simulation outputs. The positions of these detectors are strategically chosen: before the on-ramp, after the on-ramp, and at the end of the bottleneck. The first set of loop detectors is dedicated to monitoring traffic flow before the on-ramp and plotting the fundamental diagrams. Subsequently, the second set of detectors is positioned 150 meters after the on-ramp to capture the merging capacity. The third set of detectors is situated after the bottleneck to collect data on the queue discharging rate.

5.3. Heterogeneity towards HDVs and AVs

For the same level of AV heterogeneity, the traffic will include human-driving vehicles and heterogeneous level 1 autonomous vehicles (ACC vehicles). The ACC vehicles and human-driving vehicles in the simulation are represented by EIDM. The onramp input flow and the main section input flow are shown in Figure 5.15. The flow of the main section increased from 2700veh/h to 5400veh/h, then maintained the 5400veh/h for ten minutes. The 5400veh/h is 80 percent of the capacity, this input flow is sensitive enough to the onramp flow. The input flow of 5400veh/h and 10 minutes lasting time could invoke the congestion that could be solved before the end of the simulation. After reaching 80 percent of the capacity, the main section flow starts to drop to zero. The increasing and decreasing are used to gently generate and solve the congestion. The trend of the onramp flow is similar, however, the max flow of the onramp is 800 to invoke the disturbance. The start point of the input flow is half of the maximum input flow, to make the congestion start at a preferable time. The two-hour simulation time is

adopted by AV simulations of existing research.

For both the onramp flow and the main section flow, the traffic flow will include different shares of the ACC vehicle and human-driving vehicle.

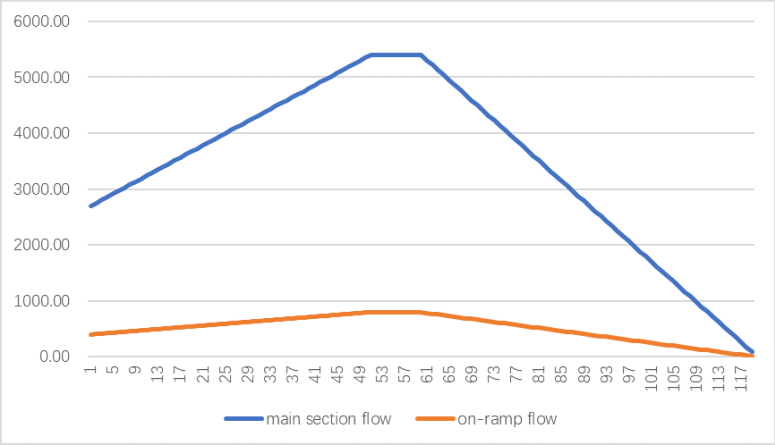


Figure 5.15: Input flow of the motorway section

Both the autonomous vehicles and the human-driving vehicle in the experiments are represented by a car-following model and a lane-changing model. The detailed settings of the simulation are shown in Table 5.8.

Table 5.8: Vehicles settings for simulation 1

	Human-driving vehicle	ACC vehicle
Car-following model	EIDM	EIDM
Lane-changing model	LC2013	LC2013
Share	0%, 20%, 40%, 60%, 80%, 100%	0%, 20%, 40%, 60%, 80%, 100%

The parameters of the EIDM are shown in Table 5.9. The only two differences are the desired time gap distributions and the maximum comfortable acceleration distributions. The desired time gap of human-driving vehicles sampled from a lognormal distribution with a mean of $1.5s$ and the maximum acceleration sampled from a lognormal distribution with a mean of $1.4m/s^2$. The selection of the distribution is based on the work of Kim and Mahmassani [17]. As for the ACC vehicles, the desired time gap and the maximum acceleration are sampled from the best-fit distribution of the calibrated parameters. The parameters for dgamma distribution and burr distribution are shown in Table 5.4.

Table 5.9: EIDM parameters for HDVs and ACC vehicles

	Human driving vehicles	ACC vehicles
desired speed	120 km/h	120 km/h
free acceleration exponent	4	4
desired time gap	lognormal distribution(mean 1.2s)	dgamma distribution
jam distance	2 m	2 m
maximum acceleration	lognormal distribution(mean: 1.4 m/s^2)	burr distribution
desired deceleration	4.5 m/s^2	4.5 m/s^2
coolness factor	0.99	0.99

As for the lane-changing model, the parameters are the default parameters for the LC2013 model for both the ACC vehicle and HV since the lane-changing behaviors of ACC vehicles are also performed by human drivers.

5.4. Heterogeneity towards different levels of AVs

The second simulation is used to identify the traffic flow impact of different levels of autonomous vehicles. In this scenario, the traffic will include human-driving vehicles, ACC vehicles, and level 4 autonomous vehicles. The share of each type of vehicle is fixed, and the input flow is the same with simulation 1. The level 4 autonomous vehicles are represented by a car-following model(EIDM) and a lane-changing model. The simulation settings are shown in Table 5.10.

Table 5.10: Vehicles settings for simulation 2

	Human-driving vehicle	ACC vehicle	High level AV
car-following model	EIDM	EIDM	EIDM
lane-changing model	LC2013	LC2013	LC2013
Share	82%(2030),74%(2035)	15%(2030),20%(2035)	3%(2030),6%(2035)

Simulation 2 includes two distinct scenarios: one for the year 2030 and another for 2035. The primary distinction between these scenarios lies in the proportion of AVs. The determination of these years and proportions draws from precedent research as documented in the work of Calvert et al. [6]. 2030 is the assumed year that the high-level AV appeared on the public road. The next year chosen for simulation is 2035. For subsequent years beyond 2035, the behavior of HAVs might exhibit alterations or enhancements due to the increasing penetration rate of HAVs. Consequently, these additional years have not been included in the current selection for study.

As for the car-following model parameters, the ACC vehicle and human-driving vehicles are similar to simulation 1. As for the high-level autonomous vehicle, the car-following parameters are shown in Table 5.11. The shape of the maxwell distribution is shown in Table 5.7. And the car-following model parameters of the ACC vehicles and HDVs in this simulation scenario are shown in Table 5.9.

Table 5.11: HAV parameters

	Highly autonomous vehicle
desired speed	120 km/h
free acceleration exponent	4
desired time gap	Maxwell distribution
jam distance	2 m
maximum acceleration	Maxwell distribution
desired deceleration	4.5 m/s^2
coolness factor	0.99

The lateral dynamics of high-level autonomous vehicles are modeled using the lc2013 lane-changing model. However, it is essential to acknowledge that the behavioral characteristics of HAVs substantially differ from those of human-driven vehicles. Consequently, a critical parameter within the lc2013 model, namely ”lcAssertive” (indicating the degree of willingness to tolerate reduced front and rear gaps within the target lane), is established at 0.9. This specific value is recommended by the Transaid project as outlined in the work of Mintsis et al. [24], and it accurately describes the lateral conduct of high-level automated vehicles. The chosen value of 0.9 for lcAssertive indicates the conservative lateral behavior of HAV and accord with the cautious longitudinal behavior that has been observed. This value ensures an accurate representation of HAVs' lateral dynamics while accounting for differences between automated and human-driven vehicles.

5.5. Assessment of heterogeneity

The outcomes of the simulations are used to ascertain the influence of AVs on traffic flow when introducing AV heterogeneity. Additionally, the results of Simulation 1 provide insights into the continued validity of the assertion that ACC vehicles contribute to enhanced road capacity when introducing AV heterogeneity. The outcome of the second simulation could draw a conclusion about how autonomous vehicles will influence the highway traffic flow in a more realistic future situation. To achieve the experiment objective, the following evaluation indicators are applied: the fundamental diagram, queue discharging rate, merging capacity,

and average time loss. The average time loss could indicate the overall traffic efficiency of the different scenarios. The merging capacity and the queue discharging rate show how the capacity changed when the share of AV changed. The fundamental diagrams will be plotted based on loop detectors before the on-ramp. This set of loop detectors could capture the spillback effect of the congestion. The fundamental diagram could offer a clue about how the macroscopic will be effect when introducing AV heterogeneity. These findings also contribute to establishing a quantitative relationship between the proportion of heterogeneous AVs and the resulting traffic capacity. Such results can serve as inputs for macroscopic demand models, predicting network performance changes due to varying AV penetration rates.

6

Simulation result

In this chapter, the results of the two simulation scenarios are analyzed. In both scenarios, statistical results are presented initially. Subsequently, the analysis centers on the speed contour and fundamental diagram. Additionally, the merging rate, as well as queue discharge are analyzed. These results are presented to address the traffic flow impact of AV heterogeneity.

6.1. Heterogeneity of same level AVs

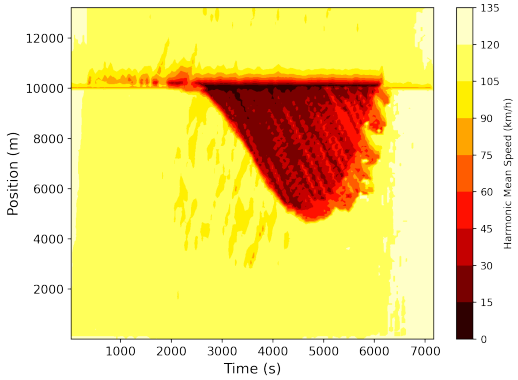
The average time loss and average speed for different ACC penetration rates are shown in Table 6.1. The average time loss and average speed of each vehicle indicate the overall traffic flow efficiency. It can be observed that with the increase in the ACC vehicle penetration rate, the average speed first decreases and then increases. The trend of the average time loss is similar to the average speed. The trend of the indicators indicates that traffic efficiency initially decreases when the ACC penetration rate is low. However, when the proportion of ACC vehicles increases to 60%, traffic efficiency improves.

The speed contour is shown in Figure 6.1. The analysis will mainly focus on the synchronized flow and wide moving jam that can be observed clearly in the congestion pattern. When the penetration rate of ACC is low(0% and 20%), the wide-moving jam appears more frequently compared to other penetration rate. However, the distance that the low-speed wave propagates upstream is shorter than the distance in a high ACC penetration rate scenario. In 40%, 60%, 80% scenarios, the low-speed area becomes larger. However, the severity of the congestion decreases since the black area(speed smaller than 15m/s) is much smaller. When the ACC penetration rate is 100 percent, the wide-moving jam is less presented. But

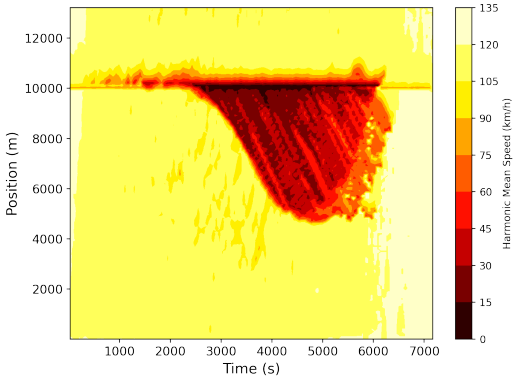
Table 6.1: Statistics of ACC vehicle heterogeneity simulation

Vehicle Share		Simulation Result	
HDV	ACC	Average Time Loss (s)	Average Speed (m/s)
100%	0%	191.19	23.82
80%	20%	200.97	23.35
60%	40%	196.28	23.51
40%	60%	175.78	24.27
20%	80%	170.93	24.49
0%	100%	110.41	26.93

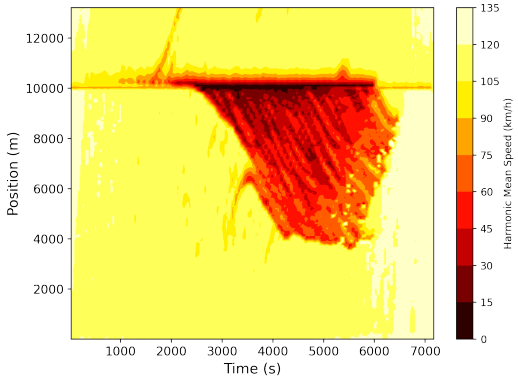
the low-speed area propagates longer upstream. The low-speed area is smaller. Thus, the outcomes of congestion patterns align with the indicator results, the overall congestion severity first increases and then decreases with the increase of the ACC penetration rate.



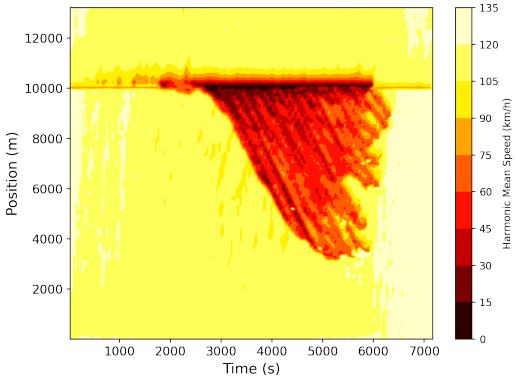
(a) 0% ACC vehicle



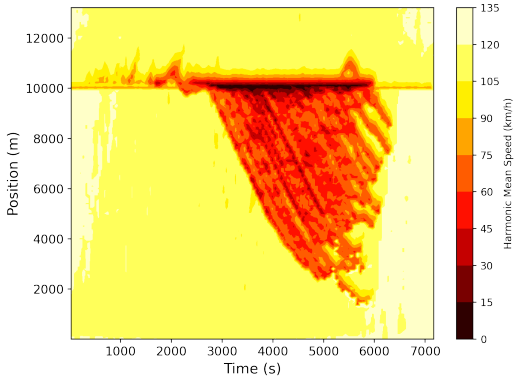
(b) 20% ACC vehicle



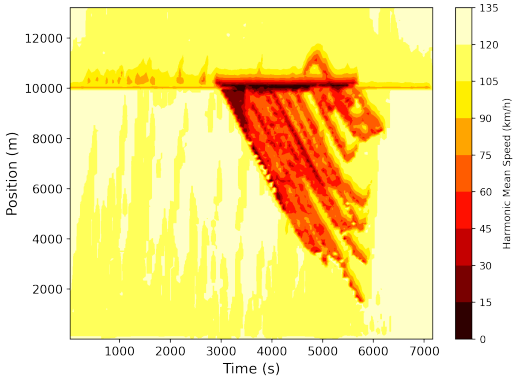
(c) 40% ACC vehicle



(d) 60% of ACC vehicle



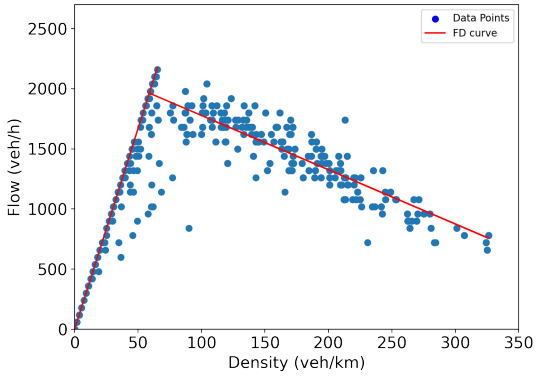
(e) 80% ACC vehicle



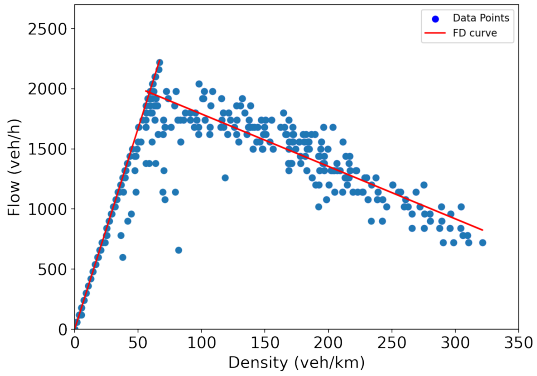
(f) 100% ACC vehicle

Figure 6.1: Speed contour of different ACC vehicle penetration rate

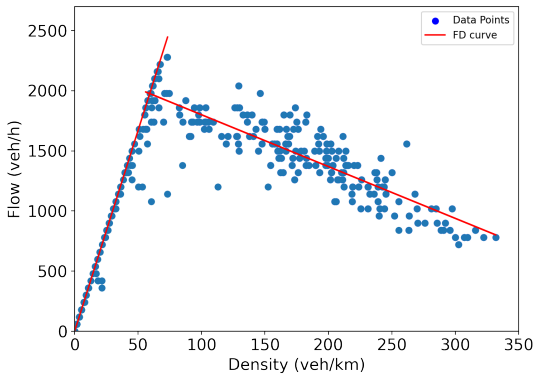
Figure 6.2 shows the fundamental diagram of the different penetration rates of ACC vehicles. The two red lines represent the estimated fundamental diagram, wherein the left segment corresponds to the free-flow regime, while the right segment pertains to the saturated flow regime. The apex of the flow-density fundamental diagram represents the maximum capacity. The corresponding maximum capacity and capacity drop are shown in Table 6.2. This capacity peak increases with higher ACC penetration rates, due to the smaller desired time gap among ACC-equipped vehicles. Maximum capacity rises from around 2160 vehicles/hour (0% ACC) to 2640 vehicles/hour (100% ACC). Another observation is the more pronounced capacity drop at higher ACC penetration rates. This indicates that ACC traffic outflow is high during non-congested times but reduces substantially during congestion.



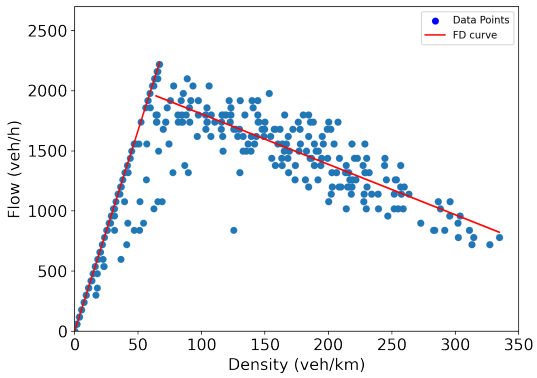
(a) 0% of ACC vehicle



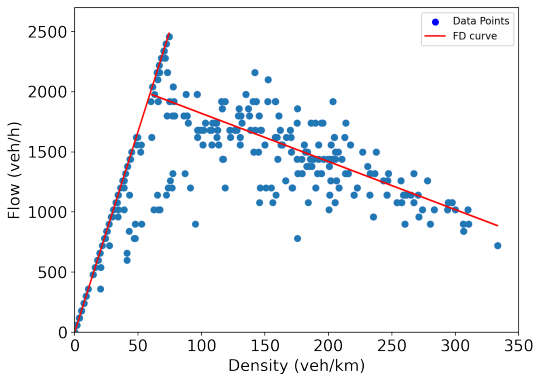
(b) 20% of ACC vehicle



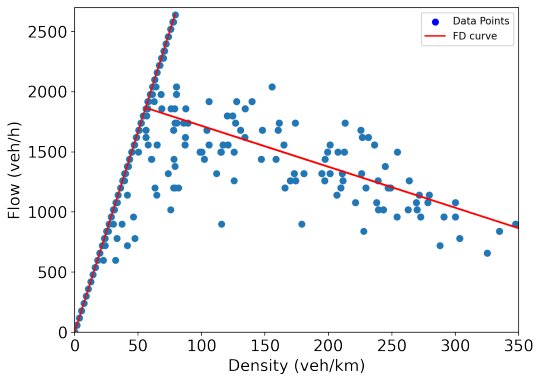
(c) 40% of ACC vehicle



(d) 60% of ACC vehicle



(e) 80% of ACC vehicle



(f) 100% of ACC vehicle

Figure 6.2: Fundamental diagram of ACC vehicle heterogeneity simulation

Table 6.2: Maximum capacity and capacity drop of ACC vehicle heterogeneity simulation

ACC penetration rate	Maximum capacity	Capacity Drop
0%	2160	224.71
20%	2220	285.47
40%	2280	364.06
60%	2220	274.20
80%	2460	537.72
100%	2640	851.99

The merging rate of the onramp is shown in Figure 6.3. The analysis of the merging rate and queue discharging rate primarily focuses on the merging lane. This focus is rooted in the recognition that the merging lane is highly sensitive to changes in the AV penetration rate. The loop detector records data at 5-minute intervals, and as such, the x-axis unit represents

increments of 5 minutes. Notably, in the initial 50 minutes, there is a significant disparity in merging rates between pure ACC traffic and traffic with other ACC penetration rates, with ACC traffic demonstrating notably higher merging rates. However, as time progresses, these differences gradually diminish. Subsequently, after 100 minutes, pure ACC vehicles scenario exhibit the highest merging rates. In conjunction with the speed contour illustrated in Figure 6.1, the congestion starts to spill back at approximately 3000s coinciding with the onset of the merging rate decline. The subsequent rise in the merging rate aligns with the resolution of congestion. The merging rate result shows that heightened ACC penetration yields a substantially greater merging rate during non-congested periods, converging to similar capacities during congestion.

Meanwhile, with regard to the queue discharging rate, the 100 percent penetration rate scenario consistently surpasses others. This finding indicates that ACC traffic maintains a higher queue discharging rate during both congested and uncongested conditions. Although the trend of merging capacity is not clear, the queue discharging rate increases with the rise of ACC vehicle penetration rates.

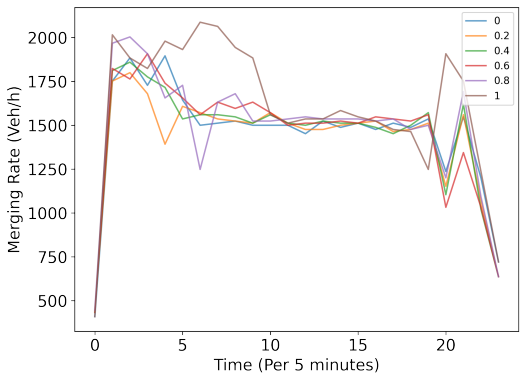


Figure 6.3: Dynamic merging rate in heterogeneous same-level AVs scenario

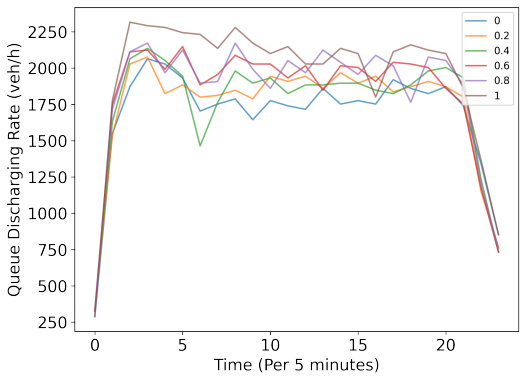


Figure 6.4: Queue discharge rate in heterogeneous same-level AVs scenario

The outcomes of congestion patterns align with the anticipated results of the corresponding parameters. The main difference between ACC vehicles and human-driving vehicles is the desired time gap and maximum acceleration. The average acceleration of ACC vehicles is smaller than that of human-driving vehicles, then the low-speed area will spill further upstream. The smaller desired time gap offers the ACC vehicles a higher queue discharging rate.

6.2. Heterogeneity towards different levels of AVs

The statistics of Simulation 2 are presented in Table 6.3. In comparison to the all-human driving vehicle scenarios, the 2030 and 2035 scenarios exhibit higher average time losses. The augmentation of the AV penetration rate leads to a reduction in traffic efficiency. The average time loss has risen to 275.79 seconds and the average speed reduced to 21.18m/s from 22.06m/s.

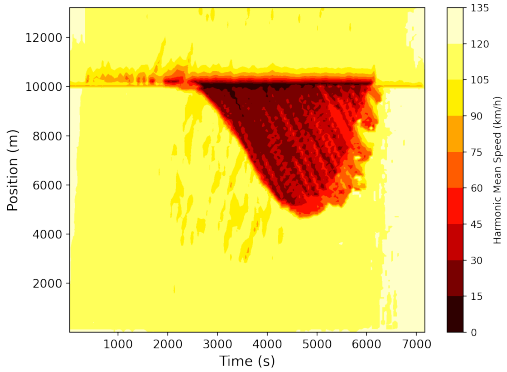
Regarding the 2040 scenario, noteworthy observations emerge when comparing it to the 2030 and 2035 scenarios. Specifically, in the 2040 scenario, there is a discernible reduction in the average time loss, accompanied by an increase in average speed. These trends signify traffic improvements attributed to the increased presence of ACC vehicles, even in the face of heightened traffic heterogeneity. Nonetheless, it is crucial to acknowledge that the negative effects of heterogeneity persist. In this context, traffic efficiency in the 2040 scenario, although improved compared to the preceding scenarios, still falls short of the efficiency achieved in the absence of AVs.

Table 6.3: Statistics of different level AVs heterogeneity simulation

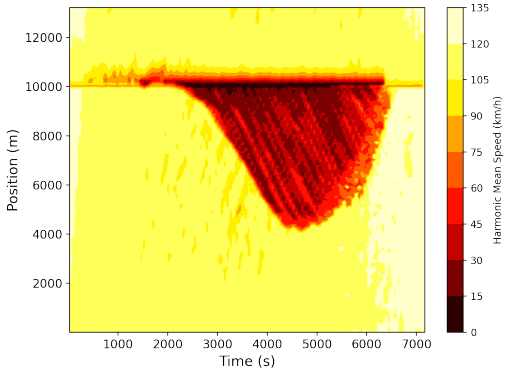
Vehicle share			Simulation Result	
HDV	ACC	HAV	Average Time Loss (s)	Average Speed (m/s)
100%	0%	0%	191.19	23.82
82%	15%	3%	249.76	22.06
74%	20%	6%	275.79	21.18
50%	38%	12%	243.13	21.93

The velocity profile is depicted in Figure 6.5. In contrast to the 2035 scenario, the adoption rate of ACC has risen from 15% to 20%, while the HAV penetration rate has increased from 3% to 6%. This upsurge in the heterogeneous rate has led to the emergence of prolonged wide-moving jams.

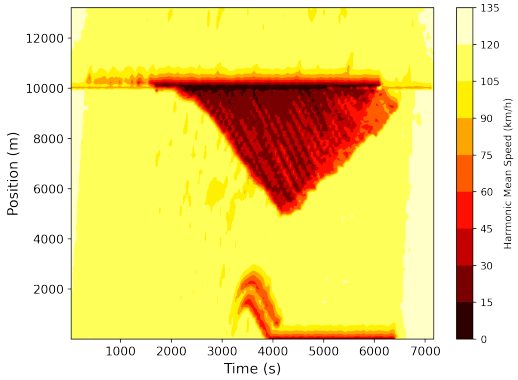
Furthermore, in the 2035 scenario, congestion has manifested in locations other than the bottleneck. This occurrence can be attributed to human errors. Within the SUMO, the model integrates parameters such as sigmaleader, sigmagap, and sigmaerror, representing the magnitude of estimation error for the leading vehicle's speed, the magnitude of estimation error for the gap between the vehicle and the leading vehicle, and the magnitude of driving error, respectively. In our research, AVs, encompassing both ACC-equipped vehicles and HAVs, are assumed to possess error parameters set to zero, implying that AVs operate without errors, aligning with common expectations. Conversely, human drivers are assigned default values for error parameters, representing the imperfect driving behavior inherent to humans. Notably,



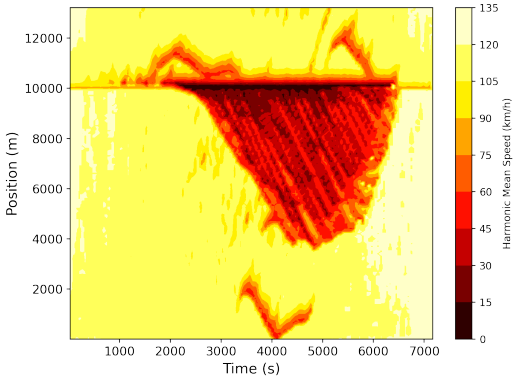
(a) No AV scenario



(b) 2030 Year scenario



(c) 2035 Year scenario



(d) 2040 Year scenario

Figure 6.5: Speed contour of 2030 and 2035 scenario

when the error parameters for Human-Driven Vehicles are set to zero, the congestion dissipates. As illustrated in Figure 6.6 and Figure 6.7, the absence of human error eradicates congestion at the beginning of the road and downstream of the bottleneck.

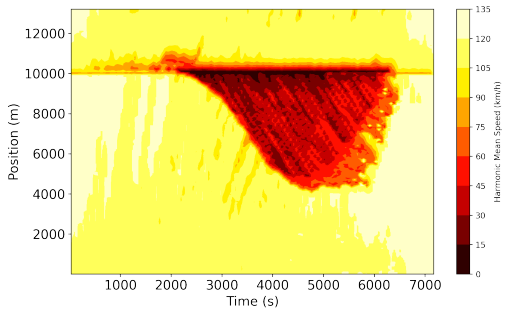


Figure 6.6: 2035 year scenario without human error

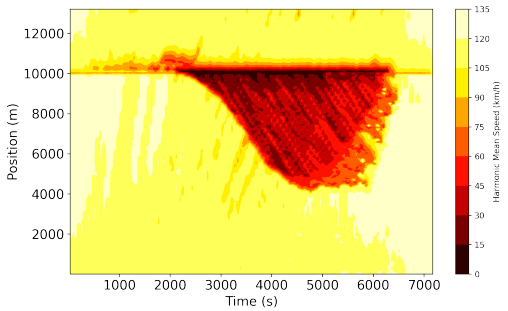
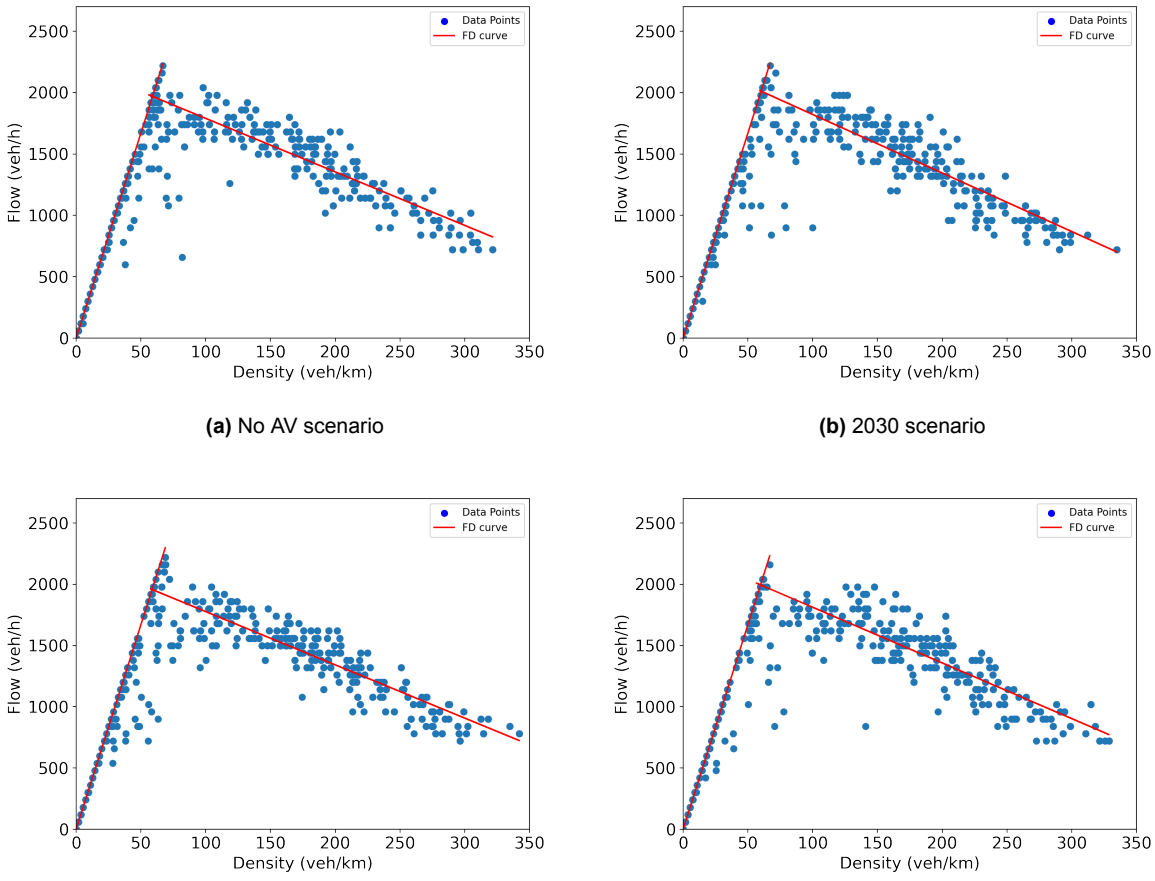


Figure 6.7: 2040 year scenario without human error

The presence of congestion due to human error underscores the instability of traffic condi-

tions in the 2035 and 2040 scenarios. This phenomenon is not observed in the 2030 scenario, this indicates that the traffic condition could be the critical point, and a higher penetration of different levels of AV will result in an unstable traffic condition. This hints at a broader issue with traffic stability in high HAV penetration rate scenarios. In the context of the 2040 scenario, it is evident that congestion propagates further upstream. Furthermore, congestion is observed to emanate from non-bottleneck positions. These observations indicate that while traffic efficiency shows improvement with the increasing presence of heterogeneous AV traffic, traffic stability continues to decline compared to the 2035 scenario.

Figure 6.8 illustrates the fundamental diagram depicting various penetration rates of mixed levels of AVs.



(a) No AV scenario

(b) 2030 scenario

(c) 2035 scenario

(d) 2040 scenario

Figure 6.8: Fundamental diagram of 2030 and 2035 scenario

The capacity and capacity drops are shown in Table 6.4. The maximum capacity initially increases, followed by a decline observed between the 2035 and 2040 scenarios. A similar trend is also evident for the capacity drop, with an initial increase succeeded by a subsequent decrease. This phenomenon can be attributed to the negative impact of heterogeneity on ca-

pacity, while the increase in the share of vehicles equipped with ACC significantly contributes to the rise in the maximum capacity.

Table 6.4: Maximum capacity and capacity drop of heterogeneity towards different levels of AVs

Scenario	Maximum capacity	Capacity Drop
No AV	2160	224.71
2030	2220	240.99
2035	2220	307.02
2040	2160	196.88

The merging rate and queue discharge rate are displayed in Figure 6.9 and Figure 6.10, respectively. In terms of the merging capacity, the trends of 2030, 2035 and 2040 are similar. However, there are still some differences. The merging rate at 5 minutes and 20 minutes is lower. Due to the fact that the merging rate is from the second loop detector. This could be because the vehicles move to the left side lane before the loop detector. That could also be the reason that The queue discharging rate of 2035 is much smaller before the congestion. Compared to the full human-driving vehicle scenario, the 2030 and 2035 scenarios have a relatively smaller merging capacity and queue discharging rate. In the context of the 2040 scenario, the situation exhibits improvement. The results align with the fundamental diagram.

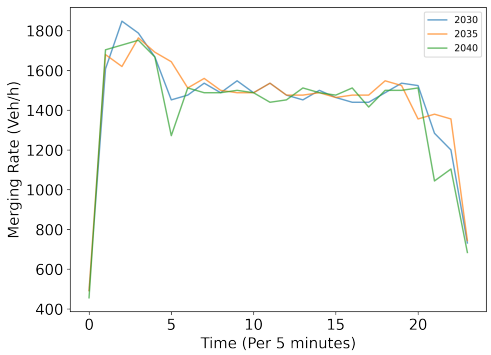


Figure 6.9: Dynamic merging rate in heterogeneous different-level AVs scenario

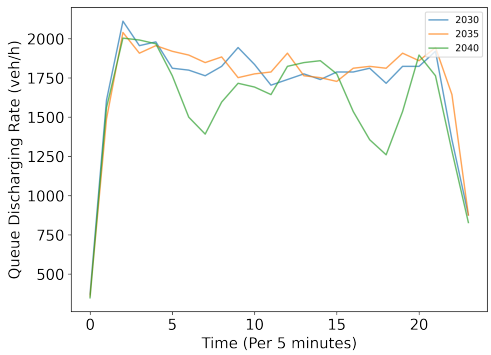


Figure 6.10: Queue discharging rate in heterogeneous different-level AVs scenario

7

Conclusion and disscussion

In this chapter, the primary findings and conclusions of the study will be presented. The specific details of these conclusions are introduced by answering the research questions. Subsequently, the limitations encountered in the course of this research will be explored, and the chapter will conclude by providing recommendations for future development.

7.1. Answer to the research questions

The main findings of this research were obtained after analyzing the data and simulation results. The analysis revealed that heterogeneous behavior not only exists across different automation levels but also within the same level. It was found that ACC vehicles have different time gap distributions when compared under similar conditions. Regarding the simulation results, it was observed that the introduction of multi-level autonomous vehicles into the traffic system may have a negative impact on both traffic efficiency and stability. The main result of the simulation highlights the impact of AV heterogeneity on traffic flow.

Does empirical data exhibit evidence of autonomous vehicle heterogeneity?

The first research question could be answered with the result of the data analysis part. Within the same automation level, the ACC vehicles manifest heterogeneity. The heterogeneity of the ACC vehicle could be seen in the TTC, desired time gap, and acceleration. As for the TTC, the vehicles that are analyzed have pretty different CDFs. And the start point: the minimum TTC starts differently. Some ACC vehicles have a 2.5 seconds minimum TTC and some vehicles have a larger minimum TTC which is larger than 3 seconds. Then the time gap distribution of the different vehicles under different conditions are examined. The distribution

of AVs' time gap is quite different under different acceleration conditions or speed conditions. Then the acceleration and the deceleration profile of the different ACC vehicles are different. Some vehicles are more likely to produce sharp acceleration or sharp deceleration than other vehicles. And the maximum acceleration/deceleration are different. In conclusion, Some ACC vehicles are more likely to take more risky behavior than others. The heterogeneity of the AVs manifests in the TTC, desired time gap, and acceleration.

The heterogeneity also shows in the parameters that are calibrated. As shown in Figure 4.4, the shape of the ACC vehicle desired gap distribution is nail shape and the largest number of data concentrate on about 1.1s. However, the range of the desired time gap distribution is from 0.6 seconds to 1.6 seconds. the range is pretty wide. As for the acceleration, the maximum acceleration of the ACC vehicle is mostly below $3m/s^2$ but ranges from $0.8m/s^2$ to $3m/s^2$. As for the HAV, the desired time gap is less concentrated and the range is still wide. There is also a wide range of acceleration distribution.

What methodology is most suitable for modeling the inherent heterogeneity in autonomous vehicles?

In this research, a method utilizing car-following model parameter distributions is employed to model heterogeneous AV behavior. For each AV automation level, data calibration is performed on the car-following model, resulting in calibrated parameters. These calibrated parameters are subsequently utilized to identify the best-fitting distributions, particularly in relation to the desired time gap and maximum acceleration. These distributions effectively capture the heterogeneity observed in AVs' desired time gaps and acceleration profiles.

Regarding differences in AV automation levels, our approach involves the simultaneous presence of ACC vehicles, HDVs, and HAVs. For each automation level, the percentage of vehicles anticipated in future years are taken into account for their presence in the simulation.

What are the traffic flow impacts of autonomous vehicles?

Within the context of AVs at the same automation level, there is a noteworthy observation: as the penetration rate of ACC vehicles increases, traffic efficiency initially decreases before subsequently increasing. This phenomenon can be attributed to the relatively shorter desired time gap and reduced acceleration of ACC-equipped vehicles in comparison to Human-Driven Vehicles. Congestion generated by the traffic becomes less frequent but harder to resolve as the ACC penetration rate increases. It's worth noting that the capacity drop becomes more pronounced with the increasing ACC penetration rate. Furthermore, heightened ACC penetration results in a significantly greater merging rate during non-congested periods, eventually reaching similar capacities during congestion. In addition, the introduction of multi-level AVs into the traffic system may have adverse effects on both traffic efficiency and stability. Traffic

efficiency is expected to decrease in the near future. The emergence of unusual congestion due to human error suggests that traffic becomes less stable when different levels of AVs are introduced. While the year 2040 shows improvements compared to a decade ago, it does not perform as well as the no-AV condition. The ultimate conclusion drawn from this analysis underscores the necessity of validating AV performance comprehensively before embarking on large-scale implementation endeavors.

7.2. Limitation

The limitations of this research primarily pertain to two aspects: data availability and modeling methods. The first limitation relates to data availability. The primary data sources of this research include the open ACC dataset, the processed Waymo dataset, and the processed Lyft 5 dataset, which are relatively limited in scope. Specifically, the ACC dataset comprises approximately 200 trajectories, which is considerably smaller when compared to the over 8000 trajectories available for HAV datasets. This disparity in data size poses constraints on our ability to capture the heterogeneity of AVs. To address this limitation, increasing the dataset size, possibly by subdividing single trajectories or gathering more extensive data, could lead to a more comprehensive representation of AV behavior. Another issue arising from data availability pertains to the traffic context. In this research, due to the absence of detailed traffic context provided by the database, we differentiate between urban and non-urban environments. Additionally, for HAVs, an unsupervised machine learning method is employed. The research could potentially be enhanced when databases with more comprehensive information become available, allowing for the use of more specific data as input.

Concerning our modeling method, both car-following and lane-changing models are employed to simulate AV behavior. However, using the car-following model alone introduces simplifications that may not accurately represent AV behavior. Although the IDM has been utilized successfully by other researchers to model both HAVs and AVs, there remain critical opinions about its applicability. While the need for better modeling methods is acknowledged, the car-following model currently serves as a reasonable approach for simulating longitudinal AV behavior, given the absence of superior alternatives.

Another limitation arises from the default parameters used for lane-changing models, which are uniform for ACC vehicles, HAVs, and HDVs in this research. These parameters may not align with real-world scenarios. Additionally, the lane-changing parameters for HAVs are based on a report without empirical calibration, limiting their accuracy. Utilizing lane-changing parameters derived from empirical data could enhance the credibility of models. Furthermore, this research may be limited by the isolation of lane-changing and car-following models. De-

veloping a more integrated, aggregate model could yield more persuasive results.

It's important to note that once universal standards for AVs are established, addressing the impact of AV heterogeneity could become a more precise endeavor. Nevertheless, our research sheds light on potential traffic flow issues stemming from AV implementation, which may contribute to the establishment of such standards.

7.3. Concluding remarks

In conclusion, the simulation results presented in Sections 6.1 and 6.2 shed light on the complex dynamics of AV heterogeneity and its effects on traffic flow. It is evident that as the penetration of ACC and HAV increases in the near future, there will be challenges and negative impacts on traffic flow performance.

As we move forward into this transitional phase, it is essential to address these challenges proactively in order to navigate through this not-so-brief period smoothly. This will require a combination of technological advancements, policy interventions, and collaboration between stakeholders. The key elements that help alleviate the shock of heterogeneous AV traffic are as follows:

- Harmonized AV Technologies: It is crucial to promote the standardization and harmonization of AV technologies. This can help reduce the variability in AV behavior on the road and facilitate smoother interactions with conventional vehicles.
- Regulatory Framework: Developing a flexible and adaptive regulatory framework that can accommodate evolving AV technologies is vital. This framework should encourage innovation while ensuring safety and efficiency.
- Data Sharing and Collaboration: Encouraging data sharing among AV manufacturers and relevant authorities can lead to a better understanding of traffic patterns and improve overall traffic management strategies.

In essence, the challenges posed by the heterogeneity of autonomous vehicles can be mitigated through a combination of technological innovations, well-thought-out policies, and cooperation between stakeholders. As we prepare for the future of the traffic system, it is important to recognize that the impacts of AVs on traffic flow are not set in stone but can be influenced by our collective efforts. With a thoughtful and proactive approach, we can navigate this transition period and ultimately harness the full potential of autonomous vehicles to enhance the efficiency and safety of our traffic systems.

References

[1] Ankit Anil Chaudhari et al. “Calibrating Wiedemann-99 model parameters to trajectory data of mixed vehicular traffic”. In: *Transportation research record* 2676.1 (2022), pp. 718–735.

[2] WS Atkins. “Research on the impacts of connected and autonomous vehicles (CAVs) on traffic flow”. In: *Stage 2: Traffic Modelling and Analysis Technical Report* (2016).

[3] Mohamed Berrazouane et al. “Analysis and Initial Observations on Varying Penetration Rates of Automated Vehicles in Mixed Traffic Flow utilizing SUMO”. In: *2019 IEEE International Conference on Connected Vehicles and Expo (ICCVE)*. 2019, pp. 1–7. DOI: 10.1109/ICCVE45908.2019.8965065.

[4] Andrew L Berthaume et al. “Variations in driver behavior: An analysis of car-following behavior heterogeneity as a function of road type and traffic condition”. In: *Transportation research record* 2672.37 (2018), pp. 31–44.

[5] Jane Bierstedt et al. “Effects of next-generation vehicles on travel demand and highway capacity”. In: *FP Think Working Group* 8 (2014), pp. 10–1.

[6] SC Calvert, WJ Schakel, and JWC Van Lint. “Will automated vehicles negatively impact traffic flow?” In: *Journal of advanced transportation* 2017 (2017).

[7] Joint Research Centre. *Open ACC Database*. url:http://data.europa.eu/89h/9702c950-c80f-4d2f-982f-44d06ea0009f. European Commission, 2020.

[8] Jakob Erdmann. “SUMO’s lane-changing model”. In: *Modeling Mobility with Open Data: 2nd SUMO Conference 2014 Berlin, Germany, May 15-16, 2014*. Springer. 2015, pp. 105–123.

[9] Xuan Fang et al. “Effects of automated vehicle models at the mixed traffic situation on a motorway scenario”. In: *Energies* 15.6 (2022), p. 2008.

[10] Xin-Yue Guo, Geng Zhang, and Ai-Fang Jia. “Study on mixed traffic of autonomous vehicles and human-driven vehicles with different cyber interaction approaches”. In: *Vehicular Communications* 39 (2023), p. 100550. ISSN: 2214-2096. DOI: https://doi.org/10.1016/j.vehcom.2022.100550. URL: https://www.sciencedirect.com/science/article/pii/S2214209622000973.

[11] John Houston et al. “One thousand and one hours: Self-driving motion prediction dataset”. In: *Conference on Robot Learning*. PMLR. 2021, pp. 409–418.

[12] Xiangwang Hu et al. “Processing, assessing, and enhancing the Waymo autonomous vehicle open dataset for driving behavior research”. In: *Transportation Research Part C: Emerging Technologies* 134 (2022), p. 103490. ISSN: 0968-090X. DOI: https://doi.org/10.1016/j.trc.2021.103490. URL: https://www.sciencedirect.com/science/article/pii/S0968090X21004769.

[13] Sae International. “Taxonomy and definitions for terms related to driving automation systems for on-road motor vehicles”. In: *SAE international* 4970.724 (2018), pp. 1–5.

[14] Ozgenur Kavas-Torris et al. *The Effects of Varying Penetration Rates of L4-L5 Autonomous Vehicles on Fuel Efficiency and Mobility of Traffic Networks*. 2023. arXiv: 2306.01177 [eess.SY].

[15] Arne Kesting and Martin Treiber. “Calibrating Car-Following Models by Using Trajectory Data”. In: *Transportation Research Record: Journal of the Transportation Research Board* 2088.1 (Jan. 2008), pp. 148–156. DOI: 10.3141/2088-16. URL: https://doi.org/10.3141%2F2088-16.

[16] Arne Kesting, Martin Treiber, and Dirk Helbing. “Enhanced intelligent driver model to access the impact of driving strategies on traffic capacity”. In: *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 368.1928 (Oct. 2010), pp. 4585–4605. DOI: 10.1098/rsta.2010.0084. URL: https://doi.org/10.1098%5C%2Frsta.2010.0084.

[17] Jiwon Kim and Hani S. Mahmassani. “Correlated Parameters in Driving Behavior Models: Car-Following Example and Implications for Traffic Microsimulation”. In: *Transportation Research Record* 2249.1 (2011), pp. 62–77. DOI: 10.3141/2249-09. eprint: https://doi.org/10.3141/2249-09. URL: https://doi.org/10.3141/2249-09.

[18] Guopeng Li et al. *Large Car-following Data Based on Lyft level-5 Open Dataset: Following Autonomous Vehicles vs. Human-driven Vehicles*. 2023. arXiv: 2305.18921 [eess.SY].

[19] Pengfei Liu and Wei (David) Fan. “Exploring the impact of connected and autonomous vehicles on freeway capacity using a revised Intelligent Driver Model”. In: *Transportation Planning and Technology* 43.3 (2020), pp. 279–292. DOI: 10.1080/03081060.2020.1735746. eprint: https://doi.org/10.1080/03081060.2020.1735746. URL: https://doi.org/10.1080/03081060.2020.1735746.

[20] Qiong Lu et al. “The impact of autonomous vehicles on urban traffic network capacity: an experimental analysis by microscopic traffic simulation”. In: *Transportation Letters* 12.8 (2020), pp. 540–549. DOI: 10.1080/19427867.2019.1662561. eprint: https://doi.org/10.1080/19427867.2019.1662561. URL: https://doi.org/10.1080/19427867.2019.1662561.

[21] Michail Makridis et al. “OpenACC. An open database of car-following experiments to study the properties of commercial ACC systems”. In: *Transportation Research Part C: Emerging Technologies* 125 (2021), p. 103047. ISSN: 0968-090X. DOI: https://doi.org/10.1016/j.trc.2021.103047. URL: https://www.sciencedirect.com/science/article/pii/S0968090X21000772.

[22] Michail A Makridis et al. “Characterising driver heterogeneity within stochastic traffic simulation”. In: *Transportmetrica B: Transport Dynamics* (2022), pp. 1–19.

[23] Vicente Milanés and Steven E. Shladover. “Modeling cooperative and autonomous adaptive cruise control dynamic responses using experimental data”. In: *Transportation Research Part C: Emerging Technologies* 48 (2014), pp. 285–300. ISSN: 0968-090X. DOI: https://doi.org/10.1016/j.trc.2014.09.001. URL: https://www.sciencedirect.com/science/article/pii/S0968090X14002447.

[24] Evangelos Mintsis et al. “Modelling, simulation and assessment of vehicle automations and automated vehicles’ driver behaviour in mixed traffic”. In: *TransAID Deliverable* 3.1 (2018).

[25] Christopher Nowakowski et al. “Cooperative Adaptive Cruise Control: Driver Acceptance of Following Gap Settings Less than One Second”. In: *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* 54.24 (2010), pp. 2033–2037. DOI: 10.1177/154193121005402403. eprint: https://doi.org/10.1177/154193121005402403. URL: https://doi.org/10.1177/154193121005402403.

[26] Saskia Ossen and Serge P Hoogendoorn. “Heterogeneity in car-following behavior: Theory and empirics”. In: *Transportation research part C: emerging technologies* 19.2 (2011), pp. 182–195.

[27] Saskia Ossen, Serge P. Hoogendoorn, and Ben G. H. Gorte. “Interdriver Differences in Car-Following: A Vehicle Trajectory–Based Study”. In: *Transportation Research Record* 1965.1 (2006), pp. 121–129. DOI: 10.1177/0361198106196500113. eprint: https://doi.org/10.1177/0361198106196500113. URL: https://doi.org/10.1177/0361198106196500113.

[28] Mitra Pourabdollah et al. “Calibration and evaluation of car following models using real-world driving data”. In: *2017 IEEE 20th International Conference on Intelligent Transportation Systems (ITSC)*. 2017, pp. 1–6. DOI: 10.1109/ITSC.2017.8317836.

[29] Vincenzo Punzo, Biagio Ciuffo, and Marcello Montanino. “Can Results of car-following Model Calibration Based on Trajectory Data be Trusted?” In: *Transportation Research Record* 2315.1 (2012), pp. 11–24. DOI: 10.3141/2315-02. eprint: https://doi.org/10.3141/2315-02. URL: https://doi.org/10.3141/2315-02.

[30] Vincenzo Punzo, Zuduo Zheng, and Marcello Montanino. “About calibration of car-following dynamics of automated and human-driven vehicles: Methodology, guidelines and codes”. In: *Transportation Research Part C: Emerging Technologies* 128 (2021), p. 103165. ISSN: 0968-090X. DOI: https://doi.org/10.1016/j.trc.2021.103165. URL: https://www.sciencedirect.com/science/article/pii/S0968090X21001832.

[31] Jaehyun So et al. “Automated emergency vehicle control strategy based on automated driving controls”. In: *Journal of Advanced Transportation* 2020 (2020), pp. 1–11.

[32] Julian Staiger and Simeon Calvert. “Empirical Analysis of Longitudinal and Lateral Vehicle Dynamics of Automated Vehicles”. In: Jan. 2021.

[33] Peter Sukennik. *Micro-Simulation Guide for Automated Vehicles—Final*. 2020.

[34] Pei Sun et al. “Scalability in Perception for Autonomous Driving: Waymo Open Dataset”. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. June 2020.

[35] Zhanbo Sun et al. “Modeling Car-Following Heterogeneities by Considering Leader–Follower Compositions and Driving Style Differences”. In: *Transportation Research Record* 2675.11 (2021), pp. 851–864. DOI: 10.1177/03611981211020006. eprint: https://doi.org/10.1177/03611981211020006. URL: https://doi.org/10.1177/03611981211020006.

[36] Malin Sundbom, Paolo Falcone, and Jonas Sjöberg. “Online driver behavior classification using probabilistic ARX models”. In: *16th International IEEE Conference on Intelligent Transportation Systems (ITSC 2013)*. 2013, pp. 1107–1112. DOI: 10.1109/ITSC.2013.6728380.

[37] Martin Treiber and Arne Kesting. “Traffic flow dynamics”. In: *Traffic Flow Dynamics: Data, Models and Simulation, Springer-Verlag Berlin Heidelberg* (2013), pp. 983–1000.

[38] Jahnavi Yarlagadda and Digvijay S Pawar. “Heterogeneity in the driver behavior: an exploratory study using real-time driving data”. In: *Journal of advanced transportation* 2022 (2022).

[39] Haiyang Yu et al. “Automated vehicle-involved traffic flow studies: A survey of assumptions, models, speculations, and perspectives”. In: *Transportation Research Part C: Emerging Technologies* 127 (2021), p. 103101. ISSN: 0968-090X. DOI: https://doi.org/10.1016/j.trc.2021.103101. URL: https://www.sciencedirect.com/science/article/pii/S0968090X21001224.

[40] K. Yuan et al. “A geometric Brownian motion car-following model: towards a better understanding of capacity drop”. In: *Transportmetrica B: Transport Dynamics* 7.1 (2019), pp. 915–927. DOI: 10.1080/21680566.2018.1518169. eprint: https://doi.org/10.1080/21680566.2018.1518169. URL: https://doi.org/10.1080/21680566.2018.1518169.

[41] Fangfang Zheng et al. “Analyzing the impact of automated vehicles on uncertainty and stability of the mixed traffic flow”. In: *Transportation Research Part C: Emerging Technologies* 112 (2020), pp. 203–219. ISSN: 0968-090X. DOI: https://doi.org/10.1016/j.trc.2020.01.017. URL: https://www.sciencedirect.com/science/article/pii/S0968090X19308095.

[42] Meixin Zhu et al. “Modeling car-following behavior on urban expressways in Shanghai: A naturalistic driving study”. In: *Transportation Research Part C: Emerging Technologies* 93 (2018), pp. 425–445. ISSN: 0968-090X. DOI: https://doi.org/10.1016/j.trc.2018.06.009. URL: https://www.sciencedirect.com/science/article/pii/S0968090X18308635.