Driving Cycle and Driver Behavior Analysis for University of Alberta Fleet Vehicles

by

Yang Liu

A thesis submitted in partial fulfillment of the requirements for the degree of

Master of Science

Department of Mechanical Engineering
University of Alberta

Abstract

This thesis as a part of the University of Alberta Energy Management and Sustainable Operations (EMSO) program aims to design an intelligent management tool for the University of Alberta fleet vehicles to achieve the fuel consumption reduction, and driver behavior improvement. To achieve this goal, developing the fleet vehicles' driving cycles is essential. The driving cycle is affected by factors such as vehicle application, driving area, etc. This thesis divided the fleet vehicles into four categories according to the vehicle applications and developed distinctive driving cycles for each category.

This study used Freematics one + OBD data loggers to collect real-world driving data of vehicles. The collected data are divided into Microtrips, which are trips between two consecutive times when the vehicle speed is zero. Multiple driving scenarios are generated by using an unsupervised clustering algorithm to classify Microtrips. Driving cycles for different vehicle categories are generated by combining data from different driving scenarios.

The results show that different vehicle categories have different driving cycle characteristics. Those vehicles running in the campus area (such as utility and trade, shuttle minibus, and University of Alberta police services category vehicles) have a low average velocity which are in the range of 17.5 km/h to 24.6 km/h. In contrast, the average operating velocity for highway-running vehicles, like casual rental category vehicles, is 51.9 km/h. The University of Alberta police services category vehicles have the largest ratio of idle time in the driving cycle which is 41.1 %, but other category vehicles' idle time ratio to the driving cycle are in the range of 17.0 %

to 27.4 %.

Driver behavior, especially driver aggressiveness, directly affects a vehicle's fuel consumption (FC). Two shuttle minibuses and one Ford Escape Plug-In Hybrid Electric Vehicle (PHEV) with three fixed drivers were selected to do the test. The driving route was fixed and the vehicle model between shuttle minibuses was identical. The collected data was then used to develop and assess driving aggressiveness (DA). Different from the traditional statistical analysis method, this thesis adopts the frequency domain analysis method to analyze DA and apply a quantitative DA evaluation metric. According to the frequency of occurrence of driving data in different driving situations, revised average fuel consumption (RAFC) was used to analyze the effect of DA on FC.

The results show that DA can have adverse impact on certain driving scenarios. The bigger the DA value the greater the average fuel consumption, but the RAFC value may be different. When the DA value is close to 1, the driver drives more aggressively and causes more fuel consumption. On the contrary, when DA is close to 0, the driver drives smoothly and consumes less fuel. To reduce driver aggressiveness and achieve economic driving with low fuel consumption, future research may focus on driver training programs for key driving scenarios including urban driving scenario and highway driving scenario.

Preface

This thesis is the original work done by Yang Liu including writing the thesis. The research in this thesis is part of a collaborative research project led by Professor Mahdi Shahbakhti at the University of Alberta. The data analysis in chapter 3 and chapter 4 and conclusion in chapter 5 are Yang Liu's original work. Yang Liu was involved in the design of experiments for collecting minibus data for driver behaviour analysis. The vehicular data collection and experimental setup were done by Hamidreza Abediasl and Amir Ansari.

Part of Chapter 3 of this thesis has been published as Y. Liu, A. Ansari, M. Shahbakhti, " Identification of the Driving Cycle for University Fleet Vehicles ", Canadian Society of Mechanical Engineers (CSME) 2022 International Congress, 6 pages, Jun. 5-8, 2022, Edmonton, AB, Canada. 102.

Part of Chapter 4 of this thesis has been published as Y. Liu, H. Abediasl, A. Ansari, M. Shahbakhti, " Characterizing Driving Behavior and Link to Fuel Consumption for University Campus Shuttle Minibuses", Canadian Society of Mechanical Engineers (CSME) 2023 International Congress, 6 pages, May 28-31, 2023, Sherbrooke, QC, Canada.

Yang Liu was responsible for the data analysis as well as the as well as the manuscript composition. Hamidreza Abediasl and Amir Ansari were responsible for the data collection the data collection and contributed to manuscript edits. Mahdi Shahbakhti was the supervisory author and was involved with concept formation and manuscript composition.

Acknowledgements

First of all, I would like to thank my supervisor Mahdi Shahbakht for his guidance. It was he who took me across the border of scientific research. Secondly I want to thank Hamidreza Abediasl and Amir Ansari. Although we are from different countries, we can still work happily together. I also want to thank my parents for supporting me to get to where I am today. Finally, I would like to thank my girlfriend Zoya Xu for making me feel so happy living in Canada.

Table of Contents

List of Tables

List of Figures

List of Symbols

Variables

$\%Acc$ Time spent on acceleration divided by the total driving time

$\%Cruise$ Time spent on cruise divided by the total driving time

$\%Dec$ Time spent on deceleration divided by the total driving time

$\%Idle$ Time spent on idling divided by the total driving time

α_i Principal component

Δt Time interval (t)

λ_i Eigenvalues

Σ Covariance matrix

$h(n)$ Digitizer impulse response sequence

n Signal sequence length

$x(n)$ Original input signal sequence

$y(n)$ Output sequence after FRR filtering

A Matrix of feature vector

$A\%$ Matrix of proportional feature vector

Acc_{avg} Average acceleration of a driving cycle (m/s^2)

C Cluster centers set

C_{inew} New cluster center

Dec_{avg} Average deceleration of a driving cycle (m/s^2)

S Cluster set

$Stop/km$ Number of vehicle stops per kilometer

t_m Missing time data (t)

t_{m-1} Former time data of missing data (t)

V_{avg}	Average velocity of a driving cycle (Km/h)
Ve_{avg}	Average velocity of a driving cycle except idle (Km/h)
X	Sample set
x_{iavg}	Average target parameter for all OBD recorded data
x_i	Average target parameter for a candidate drive cycle
v_m	Missing velocity data (Km/h)

Abbreviations

AFC Average fuel consumption.

AFR Actual air/fuel ratio.

CSJ Campus Saint-Jean.

DA Driving aggressiveness.

DDS Dynamometer driving schedules.

DFT Discrete Fourier Transform.

EMSO Energy management and sustainable operations.

EPA Environment protection agency.

EV Electric vehicle.

FC Fuel consumption.

FCR Fuel consumption rate.

FRR Forward-reverse filtering.

GHG Greenhouse gas.

GPS Global positioning system.

HF High-frequency.

HHT Hilbert-Huang transform.

IMF Intrinsic mode function.

LED Light-emitting diode.

LF Low- frequency.

MAF Mass air flow.

NYCC New York city driving cycle.

OBD On-board diagnostics.

OPDA Occurrence probability of driver aggressiveness.

PCA Principal component analysis.

PHEV Plug-in hybrid electric vehicle.

PKE Positive kinetic energy.

RAFC Revised-average-fuel consumption.

RMS Route mean square.

S3VM Semi-supervised support vector machine.

SCE Safety critical event.

SI Spark Ignition.

UA University of Alberta.

UAPS University of Alberta protective services.

WL Window length.

Chapter 1

Introduction

1.1 Motivation

Faced with rising greenhouse gas (GHG) emissions, Canada has an ambitious plans to achieve a GHG reduction target of 40% below 2005 levels by 2030 and net zero emissions by 2050, to improve the environment [1]. The Energy Management and Sustainable Operations (EMSO) program at University of Alberta aims to achieve 66K tonnes fewer GHG emissions/year [2]. The EMSO program has continued and consists of sixteen sub-projects, including fleet management, light-emitting diode (LED) equipment renewal, solar energy utilization, and ventilation and heating equipment renewal [2]. Due to the diversity and complexity of its content, EMSO projects are completed by different departments and teams.

Our team was responsible for intelligent fleet management at the University of Alberta (hereby called UAlberta fleet vehicles). The overall goal of our team included UAlberta representative driving cycle development categorized for vehicle applications, eco-driving analysis of the vehicle collected data based on driving behaviors, reducing the UAlberta fleet operational cost by minimizing fuel/energy consumption, and developing an emission inventory,

Figure 1.1 shows the key research aspects of the effort for creating an intelligent management tool for the University of Alberta fleet vehicles. The work in this thesis mainly centres on two parts: i) driving cycle identification for each UAlberta fleet

application. These driving cycles can be used for assessing fuel consumption and emission from the university vehicles, and also creating a renewal plan for university fleet vehicles to select optimum vehicle type for each vehicle application, ii) driver behaviour analysis which illustrates the extent of extra fuel consumption by aggressive driving behavior.

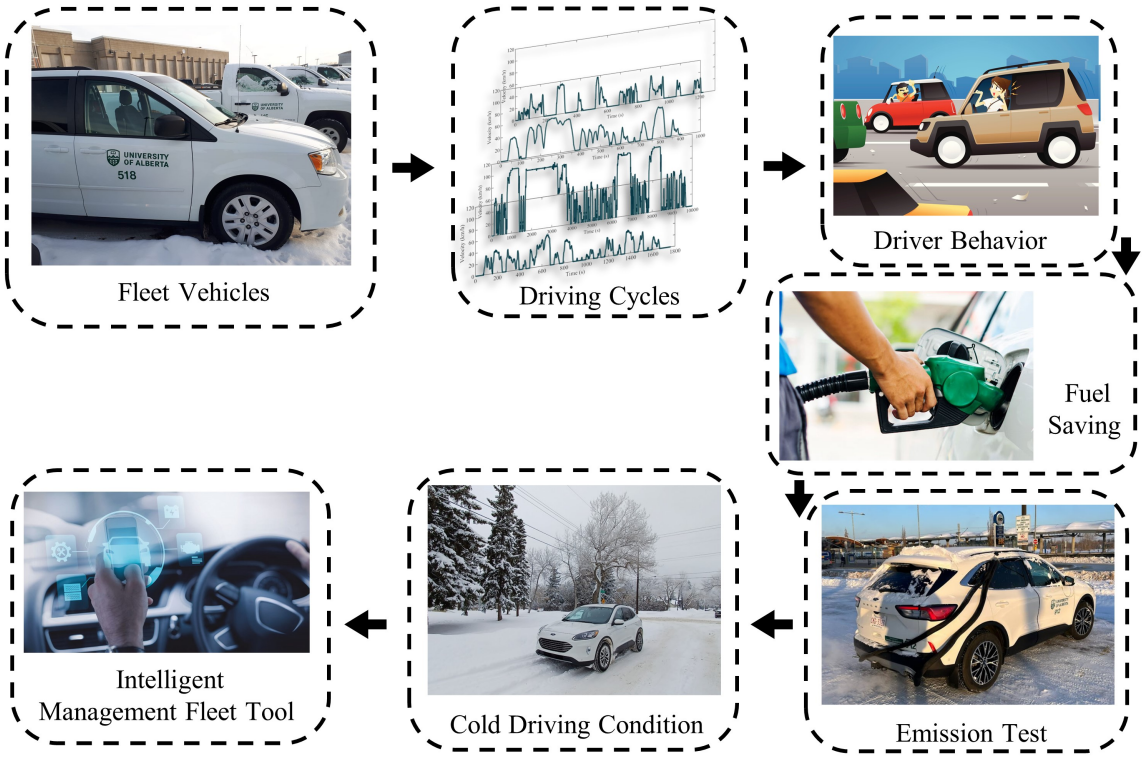


Figure 1.1: Research aspects of developing an intelligent fleet management system for UAlberta fleet vehicles as a part of the energy management and sustainable operations (ESMO) program at the university

1.2 Background

The existing UAlberta fleet vehicles include 172 vehicles with five different vehicle types, as listed in Table 1.1. UAlberta fleet vehicles include 150 trucks and vans, and those trucks and vans have engine sizes ranging from 3 liters to 7.6 liters. Using Table 1.1 data, the fleet estimated annual fuel consumption is over 205,000 liters which leads to about 564,291 kg CO2 greenhouse gas (GHG) emissions based on

Canada Clean Fuel Standard [3]. The estimated operational cost of the UAlberta fleet vehicles is about $530K annually.

The estimated fuel consumption of the UAlberta fleet is shown in Figure 1.2. This is calculated based on the UAlberta fleet data in 2019-2020 fiscal year to determine an average mileage for each vehicle type of the fleet. The data in Figure 1.2 shows an estimated fuel consumption of 205,388 liters for the whole fleet. 165 of the fleet vehicles run with gasoline fuels and the remaining 7 vehicles (1 bus, 6 trucks) run with diesel fuel. Considering Edmonton's average fuel price of $1.1/liter for gasoline and $1.15/liter for diesel during 2019-2020 fiscal year [4, 5], the estimated fuel cost for the UAlberta fleet will be about $226,000. The estimated fleet fuel cost of $226,000 shows about 42% of the fleet operational cost is spent on fuel. The remaining operational costs include i) engine oil, brake fluid, automatic transmission fluid, oil filters, etc., ii) UAlberta labour cost for repairs and maintenance, iii) outside university contract work (e.g., wheel alignment), iv) replacement parts (e.g., wipers, bulbs), v) decals, and vi) overhead (admin time). By optimizing and minimizing the size of the fleet, the operational and maintenance cost of the fleet should decrease.

The vehicle types are varied, so that the speed profile, time duration, and fuel consumption are different. That is the reason why driving cycles for different vehicle applications should be analyzed. Additionally, different driving styles, especially aggressive driving, can cause excessive fuel consumption. Analyzing drivers' aggressive behavior can help the fleet manager to develop training program for drivers and also assign fleet vehicles to drivers by consider driving aggressiveness.

1.3 Driving Cycle

A vehicle driving cycle is the vehicle speed profile versus time for a certain type of vehicle in a specified environment. Drive cycles are developed by different countries and organizations to evaluate vehicle performance in various ways, such as fuel consumption, electric vehicle (EV) autonomy, and tailpipe emissions. Another use for

Table 1.1: University of Alberta fleet vehicle types, fuel consumption rate, and average annual travel distances

Vehicle Type	**Number of Vehicles**	**Fuel Consumption [L/100 km]** [6]	**Average Travel [km/year]**
Sedans	12	6.0	8830
SUVs	6	11.8	10930
Trucks	73	14.3	9244
Vans	79	18.4	4249
Buses	2	47.0	7472

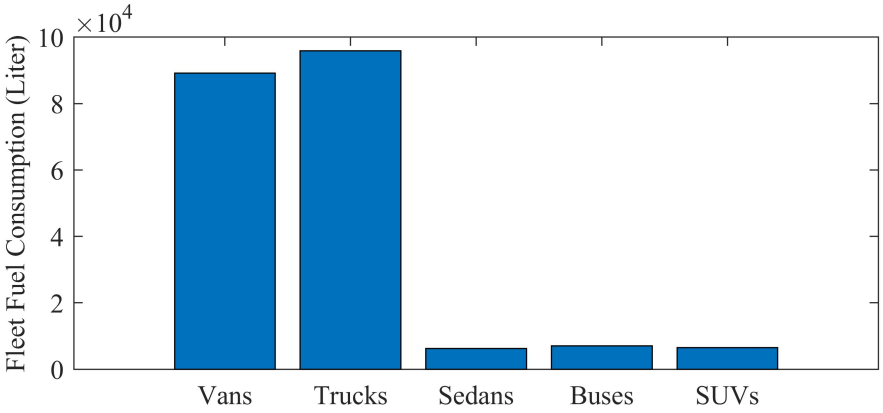


Figure 1.2: Estimated annual total fuel consumption for each vehicle type for the UAlberta fleet vehicles using the data from Table 1.1

the driving cycle is vehicle simulation. Driving cycles are used in propulsion system simulations to predict the performance of internal combustion engines, transmissions, electric drive systems, batteries, fuel cell systems, and similar components [7–9].

The relationship between the driving cycle and fuel consumption/emissions is that the driving cycle can significantly affect how much fuel a vehicle consumes and how much emission the vehicle produces. This is because different driving cycles can lead to different patterns of acceleration, deceleration, and idling, which can affect the efficiency of the engine and the amount of pollutants emitted [10]. For example, a driving cycle that involves frequent stops and starts with a lot of idling time can

increase fuel consumption and emissions, as the engine is required to work harder to accelerate from a standstill and run longer to keep the vehicle running while idling. On the other hand, a driving cycle that involves more steady-state driving at a constant speed can be more fuel-efficient and produce less emissions

A driving cycle can also be used to estimate driving style. The number of vehicles' starts and stops in a driving cycle reflects at the driver's driving style. A driver with more start-stop actions will cause more emissions [11]. The driving range of electric vehicles is obtained based on driving cycles. In addition, driving cycles can also be used to estimate the charge of state of EVs [12]. The flowchart in Figure 1.3 summarizes different applications of driving cycle.

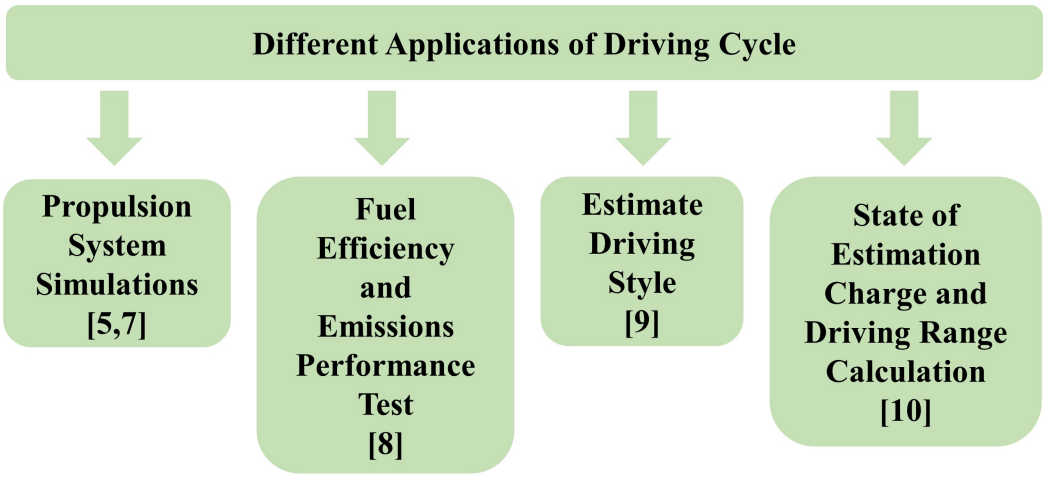


Figure 1.3: Different applications of driving cycles

Different driving cycles can be used to test the fuel efficiency and emissions performance of vehicles under different driving conditions. Driving cycles can also be used to compare the performance of different vehicles. There are three most commonly used driving cycle categories, shown in Figure 1.4 and Figure 1.5, and they include (i) United States Environmental Protection Agency (EPA) Vehicle Chassis Dynamometer Driving Schedules (DDS), (ii) Economic Commission for Europe Dynamometer

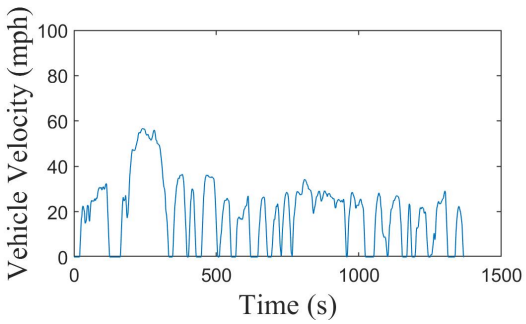
Operating Cycles, and (iii) Driving schedules specified in Japanese Technical Standards [13].

Figure 1.4 shows the DDS driving cycle category [13]. In Figure 1.4, Figure 1.4a is the EPA urban dynamometer driving cycle, Figure 1.4b is the New York city driving cycle (NYCC), Figure 1.4c is the US06 which is a high acceleration aggressive driving cycle, Figure 1.4d is the LA-92 driving cycle for Class 3 Heavy-Duty vehicles, Figure 1.4e is the federal test procedure motorcycle 1-B driving cycle, and Figure 1.4f is the federal test procedure driving cycle. DDS driving cycle category mainly consists of driving cycles for urban and highway driving scenarios. According to the applications of the vehicle and traffic conditions, a variety of driving cycles in urban and highway driving scenarios have been developed. For example, LA-92 is the driving cycle developed for heavy duty vehicles and NYCC is the driving cycle that features low-speed stop-and-go traffic conditions.

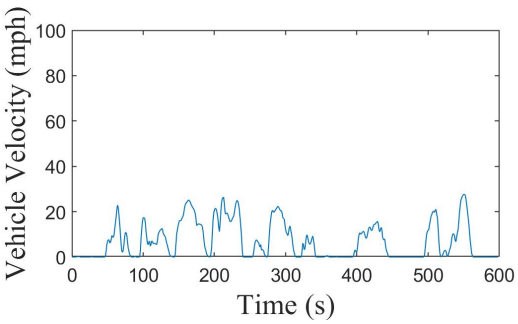
Figure 1.5 shows the Economic Commission for Europe Dynamometer Operating Cycles, and Driving schedules specified in Japanese Technical Standards driving cycle category [13]. In Figure 1.5, Figure 1.5a is the part 1 UN/ECE elementary urban driving cycle, Figure 1.5b is the part 2 UN/ECE elementary urban driving cycle, Figure 1.5c is the UN/ECE urban driving cycle for low-powered vehicles, Figure 1.5d is the Japanese 10 mode driving cycle, and Figure 1.5e is the Japanese 15 mode driving cycle.

None of above driving cycles, including US EPA driving cycles, EU driving cycles, and Japan driving cycles, can represent UAlberta fleet vehicles because UAlberta fleet vehicles are mainly running in campus area where the vehicle speed limit is 30 km/h. In addition, campus area is more crowded, so vehicles need to have frequent stop-starts. Based on the vehicle operating characteristics, this thesis aims to generate driving cycles for UAlberta fleet vehicles.

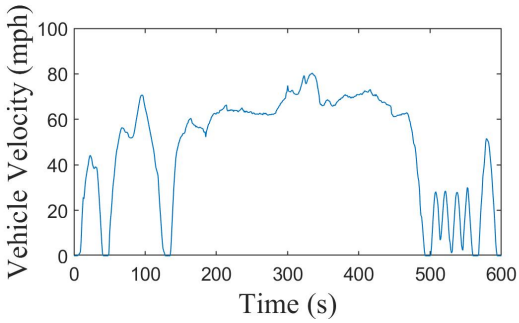
There are substantial research conducted to develop location-specific driving cycles. Uditha et al. [14] developed representative and economical driving cycles for a given



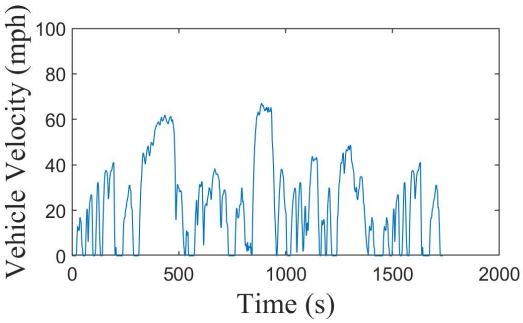
(a) EPA urban dynamometer driving cycl



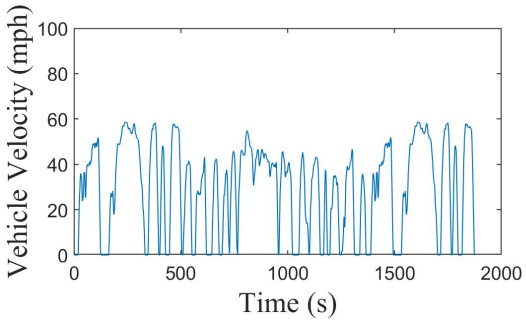
(b) New York city driving cycle



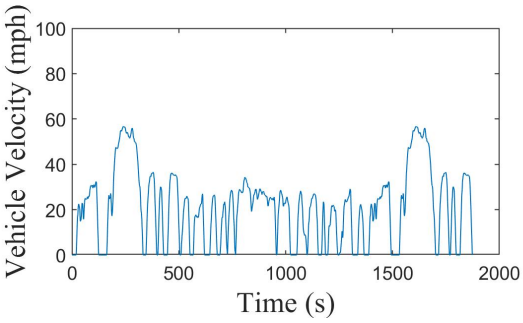
(c) US06 driving cycle



(d) LA-92 driving cycle



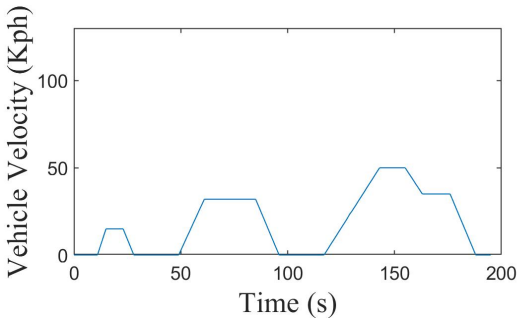
(e) Federal test procedure 1-B driving cycle



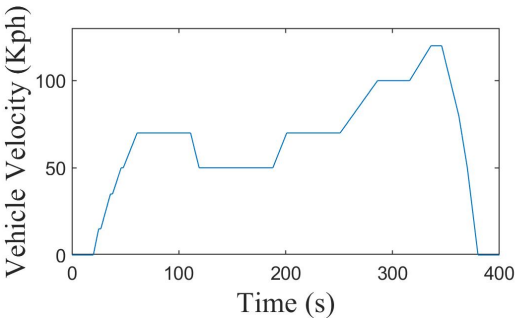
(f) Federal test procedure driving cycle

Figure 1.4: The US environmental protection agency driving cycles for emission standards

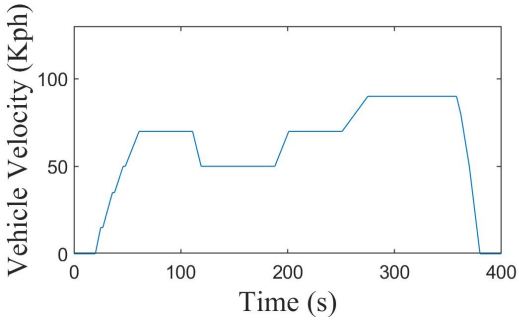
geographic location. Uditha pointed out that the driving cycle can be considered one of the main models for estimating emission inventories. It can be used for various purposes, such as setting emission standards, for traffic management purposes, or to determine travel times. Uditha also pointed out that when the driving cycle development is in its infancy, the basic parameters commonly used in many countries can be used for cycle evaluation. Commonly used parameters are: average speed, average running speed, average acceleration, average deceleration, percent idling, percent ac-



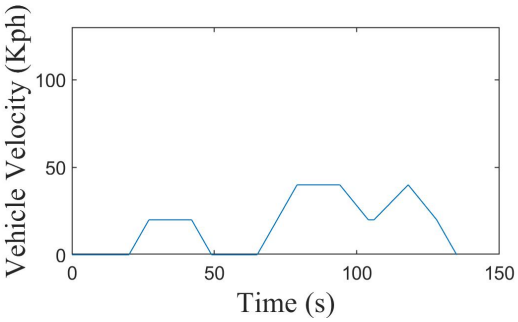
(a) Part 1 UN/ECE elementary urban driving cycle



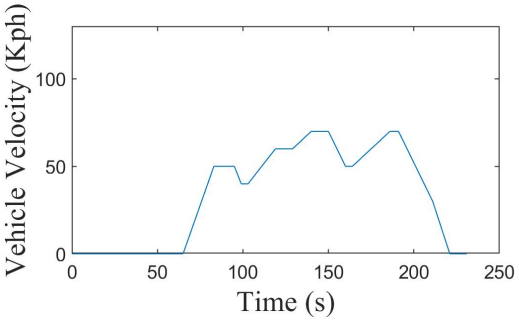
(b) Part 2 UN/ECE elementary urban driving cycle



(c) UN/ECE urban driving cycle for low-powered vehicles



(d) Japanese 10 mode driving cycle



(e) Japanese 15 mode driving cycle

Figure 1.5: EU and Japan driving cycles

celeration, percent deceleration, percent cruise, route mean square acceleration and positive kinetic energy (PKE).

Arvind et al. [15] expanded the range of parameters for evaluating driving cycles. They used Average Velocity, Maximum Velocity, Average Acceleration, Maximum acceleration, Time share of acceleration, Average Deceleration, Maximum Deceleration, Time share of Deceleration, and Time share of Idle parameters to assess the gener-

ated driving cycle. The generated driving cycle needs a method to judge whether it can represent driving behavior or not. Route mean square (RMS) algorithm is commonly used to validate the generated driving cycles. The common parameters used in driving cycles' RMS include average speed, average running speed, average acceleration, average deceleration, percent idling, percent acceleration, percent deceleration, percent cruise, and so on [16]. However, those parameters not as equally important. Therefor, to design weight for each parameter to use the RMS algorithm is necessary [16].

Driving cycle analysis can be simplified by classifying experimental data to generate driving scenarios. Jean-Baptiste Gallo divided the Microtrip, which is a time series of speed, into four driving scenarios with a certain speed standard. Those driving scenarios including i) idle driving scenario: vehicle speed is equal to zero, ii) creeping driving scenario: Microtrip maximum speed in the range from 0 to 8.2 mph, iii) Low-Speed Transient driving scenario: Microtrip maximum speed in the range from 8.2 to 22 mph, and High-Speed Transient driving scenario: Microtrip maximum speed in the range from 22 to 52 mph. The desired driving cycle can be obtained by selecting representative Microtrips in different driving scenarios and combining the representative Microtrips according to certain rules [17].

The Tehran driving cycle was developed by Fotouhi et al. [18], by using Mcrotrips and applying the k-means method to cluster the Microtrip database. Fotouhi used average vehicle speed and idling time as variables for evaluating the suitability of driving cycles. Driving scenarios are identified by clustering Microtrips. Four driving scenarios including congested, urban, out-urban and highway driving were divided to study the Tehran driving cycle. The Tehran driving cycle was characterized by a distance of more than 14.4 km, a duration of 1533 seconds, an average speed of 33.84 km/h and an idling time of 15.3%.

Chugh et al. [19] used Microtrips and applied the velocity and acceleration frequency matrices formed by each sequence to develop the Delhi driving cycle. Chugh

evaluated the effectiveness of the developed driving cycle in terms of average speed, percent idling time, percent acceleration/deceleration and percent cruise time. By clustering Microtrip, four driving scenarios of crowded, semi-urban, urban and non-urban driving are obtained. When generating driving cycles, Chugh also took into account the characteristics of the class of traffic conditions. The Delhi driving cycle was characterized by a mileage of 7.8 km, a duration of 1565 seconds, an average speed of 25.5 km/h and an idling time of 29.9%.

In studies for generating driving cycles, Tamsanya et al. [20] designed an driving cycle generation algorithm based on average velocity. In Tamsanya's research, driving cycles are generated by randomly combining Microtrips.

Shi et al. [21] constructed a driving cycle using a Markov process from stochastic process theory. Shi used the transition matrix of vehicle data to determine the statistical characteristics. Modal events were randomly selected from the data and combined until the desired cycle length was reached. Shi's research serves as a guiding theory that can be applied in the development of different kinds of driving cycles according to the specific experimental situation.

Justin et al. [22] developed a robust, data-driven approach to Markov chains. This method can capture the real behavior in the driving cycle without deconstructing the raw velocity-time series. Justin used the number of velocity states, drive cycle length, and Markovian repetitions to evaluate the effectiveness of generating drive cycles. Justin's research found that the accuracy of candidate drive cycles is largely dependent on the number of Markovian repetitions.

The study by Huiming et al. [23] found that different vehicles in the same region have different driving cycle characteristics. Based on the measured data, the usage and driving characteristics of electric vehicles in Beijing were investigated. The driving characteristics of the EV were compared with existing standard cycles. The results show that the actual data differ significantly from the standard driving conditions in many respects. The development of the drive cycle, therefore, takes into

account the differences in the test vehicles.

Shortcomings : There are two drawbacks for the existing driving cycles research. The first one is the lack of weight design for parameters. In the driving cycle appraising algorithm, if all parameters are calculated with the same weight, the influence of insignificant parameters will be amplified. Another shortcoming is the inefficient calculation of driving scenarios and driving cycles. This thesis aims to address these shortcomings (see detial in Section 3.3.2 and Section 3.6).

1.4 Driver Behavior

Driver behavior refers to the actions and habits of a driver while operating a vehicle. It includes a range of factors such as driving style, attitude, skill level, and experience. Driver aggression is a subset of driver behavior that refers to aggressive or reckless driving practices [24]. This thesis focus on driver aggressive analysis.

Driver aggression can have a significant impact on fuel consumption. Aggressive driving, such as rapid acceleration, sudden braking, and excessive speeding, can cause a vehicle's fuel conversion efficiency to decrease by up to 33% on highways and 5% in the city, according to the US Department of Energy [25]. This is because aggressive driving requires more energy for heavy acceleration, resulting in increased fuel consumption. Rapid acceleration requires the engine to work harder to produce power, which leading to more fuel consumption. Sudden braking wastes the kinetic energy that the car has gained from accelerating, which then needs to be regained through more fuel use. In addition, excessive speeding increases wind resistance, which also requires more energy and fuel to maintain the vehicle's speed.

Fangfang et al. [26] found through real-world data of 38 drivers in Chengdu, China, that traffic control signals are the influencing factors that affect vehicle fuel consumption and emissions. Idling due to parking at traffic intersections is a major contributor to fuel consumption losses. An aggressive driver consumes more fuel when the vehicle starts and brakes, which greatly exacerbates fuel consumption. At

the same time, Fangfang et al. also designed a comprehensive emission model to obtain the real-time emission and fuel consumption of vehicles.

Javier et al. [27] designed an intelligent traffic system, through data mining and a neural network system, based on the analysis of the driver's characteristics, the relationship between the driver's style and the fuel consumption of the vehicle. In Javier's experiment, the real-time fuel consumption of the vehicle was derived from two parameters; Fuel Flow and Vehicle Speed.

Wenshuo et al. [28] used a semi-supervised method, a semi-supervised support vector machine (S3VM), to classify driver aggressiveness. In Wenshuo's experiment, different from the traditional driving scenario coordinate system which uses vehicle average speed and maximum speed, the vehicle speed and throttle opening were used as the coordinate system to mark the label, and also the S3VM algorithm was used to obtain the classification of the driver's driving state.

Tzirakis et al. [29] tested driving aggressiveness differences among drivers of twelve vehicles in a closed experimental field. The test site used by Tzirakis et al is characterized by many uphill, downhill, and humps. This caused the driver to make more accelerations and decelerations. In the experiment, the fuel consumption was measured by a portable vehicle fuel consumption analyzer directly. Experimental results showed that compared with smooth driving, aggressive driving leads to increased fuel consumption, ranging from 78.5% to 137.3% for gasoline vehicles and 116.3% to 128.3% for diesel vehicles .

Deepak Hari et al. [30] installed a retrofit driver behavior improvement device on 15 vehicles in the UK and tested it in a real-world environment. Fuel savings of 7.6% were achieved through improved driving behavior by reducing harsh acceleration and earlier shifting to higher gears. Through simulation experiments on urban roads and high-speed roads, Emilia et al. [31] found that aggressive driving behavior greatly affects the fuel consumption and emissions of vehicles. Especially on urban roads, aggressive driving increases fuel consumption by 30% and pollutant emissions by

40%.

Glareh et al. [32] utilized driving cycles to represent differences in driving aggression. Experimental results showed that medium truck drivers are more aggressive than heavy truck drivers. This aggressiveness resulted in medium-duty trucks emitting 60% more harmful substances than heavy-duty trucks.

Eunsol et al. [33] used the in-vehicle data capture device to record the factors that may affect the driver's aggressiveness. A binary logistic regression model was established with influencing factors as independent variables and driving aggressiveness as independent variables. The main factors affecting driving aggressiveness were found through mathematical research through the logistic regression model. Experimental resulted show that mileage is an important factor affecting driver aggressiveness. In addition, the driver's mental health is also a key factor affecting driving performance.

Shortcomings : The existing research's method on driver aggressiveness analysis are mainly focus on statistical analysis method. Moreover, the driving scenarios selection range is not comprehensive, focusing on small scale driving scenario such as uphill, downhill, and humps. Based on those features, this thesis applies an algorithm to quantify driver aggressiveness and use unsupervised machine learning method to generate a large scale driving scenarios. Detail information can be found in Section 3.4 and Section 4.3.1.

1.5 Thesis Objectives and Scope

Driving cycles can provide a standardized way to test a vehicle's fuel efficiency and emissions. A consistent drive cycle allows for a fair test comparison of vehicles. Test results can be used to design and optimize vehicle components. The commonly used driving cycles are mainly under generated urban and highway driving scenarios. Inadequate work has been done on the development of driving cycles on university campus vehicles. Different from urban and rural area vehicles, campus vehicles have many unique characteristics, such as lower average speed and more frequent starts

and stops. The University of Alberta is located in Edmonton where has longer winters with low ambient temperatures. The snow in winter will make vehicles travel more slowly. Therefore, it is necessary to analyze the driving cycle of the UAlberta fleet vehicles.

There are more than 170 vehicles on the UA campus, it is important to develop a campus driving cycle to analyze vehicle fuel consumption and emissions. In this thesis, according to the characteristics of the fleet vehicles of the University of Alberta, the fleet vehicles are divided into four categories, and driving cycles are developed for each category.

Driver behavior, especially aggressive driver behavior, directly influences vehicle fuel consumption. However, the existing research mainly provides statistic criteria for judging driver aggressiveness in the driving cycle scale. Large scale driver aggressiveness analysis will loss detail information so it is difficult to formulate accurate training programs for drivers. This thesis will start from the frequency domain and provide a metric for different driving scenarios of driver s to measure driver aggressiveness. At the same time, this thesis will analyze the difference in driver aggressiveness and fuel consumption caused by this difference in different driving scenarios.

The main outcomes/benefits of this thesis are

i) A driving cycle generation method based on average speed is designed. This method is applied and tested for extensive vehicle data sets.

ii) Extraction of a drive cycle for each of four categories of university fleet vehicles. These driving cycles can be used for fuel consumption and emission analysis of the fleet vehicles as well as deciding for the university vehicles renewal plan to choose optimum vehicle type for each category of applications.

iii) A numerical driving aggressiveness (DA) evaluation method based on frequency domain analysis is designed. When the DA value is closer to 0, the driving is more smooth, and when the DA value is closer to 1, the driving is more

aggressive. This method can be used to compare university vehicle drivers and identify driving situations that lead to aggressive driving, thus, driver training programs, e.g. by design of driving simulation platforms, can be designed to train university drivers.

iv) Proposed a revised average fuel consumption (RAFC) metric to better reflect the impact of DA on vehicle fuel consumption. This method considers the impact of the occurrence probability of DA in different driving scenarios.

1.6 Thesis Outline

Figure 1.6 shows the structure of this thesis. This thesis consists of five parts including introduction, experimental setup, driving cycle analysis, driver behavior analysis, and conclusion.

In the experimental setup chapter, detailed information of the vehicle test setup will be introduced, including the vehicle type, vehicle instrumentation, engine parameters, Global Positioning System (GPS) map of the driving routes, and data collection. In addition, according to the characteristics of the collected data, this chapter designs a set of data preprocessing methods to enhance the usability of the data. In chapter 3, major driving cycle conceptions including Microtrip, Kinematic Fragment, and Assessment Metrics are introduced first. Next, the algorithms used in this chapter, including the clustering algorithm for classifying driving scenarios and the Principal Component Analysis (PCA) algorithm for evaluating parameter weight, will be explained. Finally, the Microtrip and sub-Microtrip databases for each vehicle class, the driving cycle generation algorithm, and the final results are discussed in detail. Chapter 4 introduces driver behavior conceptions including driver behavior, and driver aggressiveness. Next, the algorithms used in this chapter, including frequency domain algorithms for numerically expressing driver aggressiveness and algorithms for estimating fuel consumption, will be discussed. Finally, the driver aggressiveness

and fuel consumption differences of different drivers under different driving scenarios are obtained. The last chapter concluded this thesis by listing the main findings from this thesis.

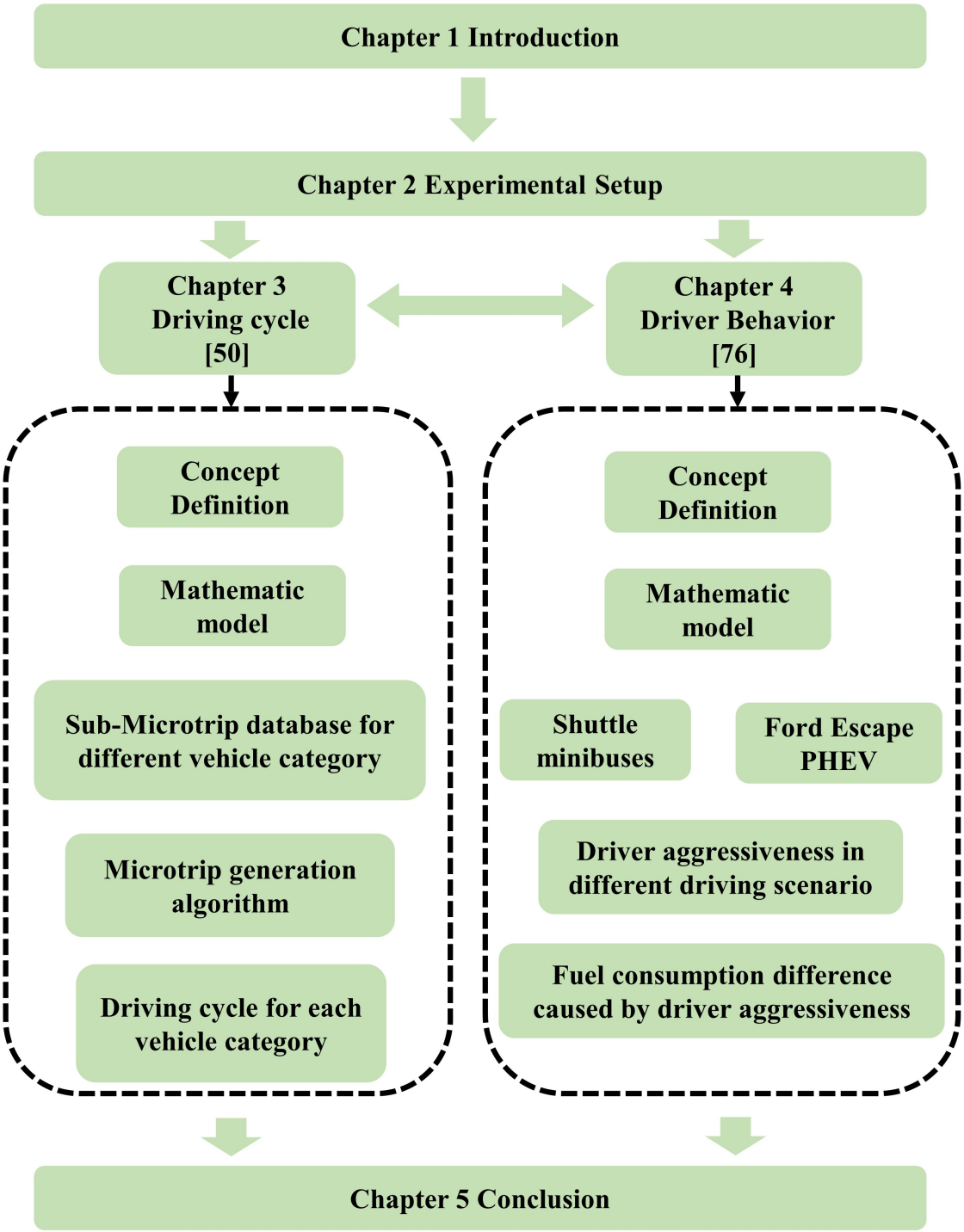


Figure 1.6: Thesis flow including chapters and resulting publications

Chapter 2

Experimental Setup

2.1 Introduction

Currently, UAlberta fleet vehicles include 172 vehicles, most of which are trucks or vans. According to the application of each vehicle, the fleet vehicles in this research were divided into 4 categories, namely: Utility or Trades; Casual Rentals; Shuttle Minibus; University of Alberta Protective Services (UAPS). Within each category of vehicles, several vehicles are selected to represent all other vehicles in the same category. Vehicle characteristics and instruments installed to collect data are the focus of this chapter. In addition, this chapter also introduces the main running routes of different categories of vehicles and their characteristics.

2.2 Data Collection

Today's vehicles all have on-board diagnostics (OBD) capability for vehicle self-diagnostic and reporting. To this end, extensive vehicle operational data is available and can be recorded through OBD port in vehicles. The OBD data can be read by a standardized digital communication port in vehicles.

In this study, Freematics One+ OBD data loggers, shown in Figure 2.1, were used to collect vehicles' data. The Freematics ONE+ is an Arduino programmable vehicle telemetry prototyping platform in the form of an OBD dongle that plugs into the vehicle's OBD port and works as a standalone smart device that can communicate

with the vehicle's Engine Control Unit (ECU). Freematics One+ OBD data logger has a dual-core Central Processing Unit (CPU) that can process data quickly in real time, and store the processed data in the internal flash memory or microSD card. Freematics One+ OBD data logger can also use cellular network or WiFi to transmit data to Freematics Hub. Freematics Hub is a cross-platform open-source telemetry and IoT data server software. It receives real-time data from multiple Freematics ONE+ OBD data loggers and provides a simple REST API for third party applications to access real-time and historical data. Freematics Hub is efficient and has low system resource requirements. It requires no application to maintain a connection with the telematics device, while being deployable on any types of computes. Figure 1 shows the workflow of the Freematics One+ OBD data logger [34, 35].

During use, Freematics One+ OBD data loggers were inserted into vehicle OBD interface to read the vehicle running data. At the same time, the location of the vehicle will also be transmitted to Freematics One+ OBD data logger in real time through GPS signals. Freematics data loggers are equipped with a GNSS model M8030, and the GPS accuracy of horizontal position is 2.0 m. All data uses a dedicated telemetry transmission protocol to transmit high-throughput and low-latency vehicle data to the Freematics Hub with minimal data overhead. In the end, the historical data from Freematics Hub were stored into the database for free call by participants. Table 2.1 shows the OBD parameters that were used to generate driving cycle and analyze driver behavior.

This thesis focuses on the differences in driving cycles between different vehicle applications and differences in driver behavior among different drivers. Given, the test data are mainly in the same season, this thesis does not focus on studying the seasonal influence on vehicle driving cycle and driver behavior.

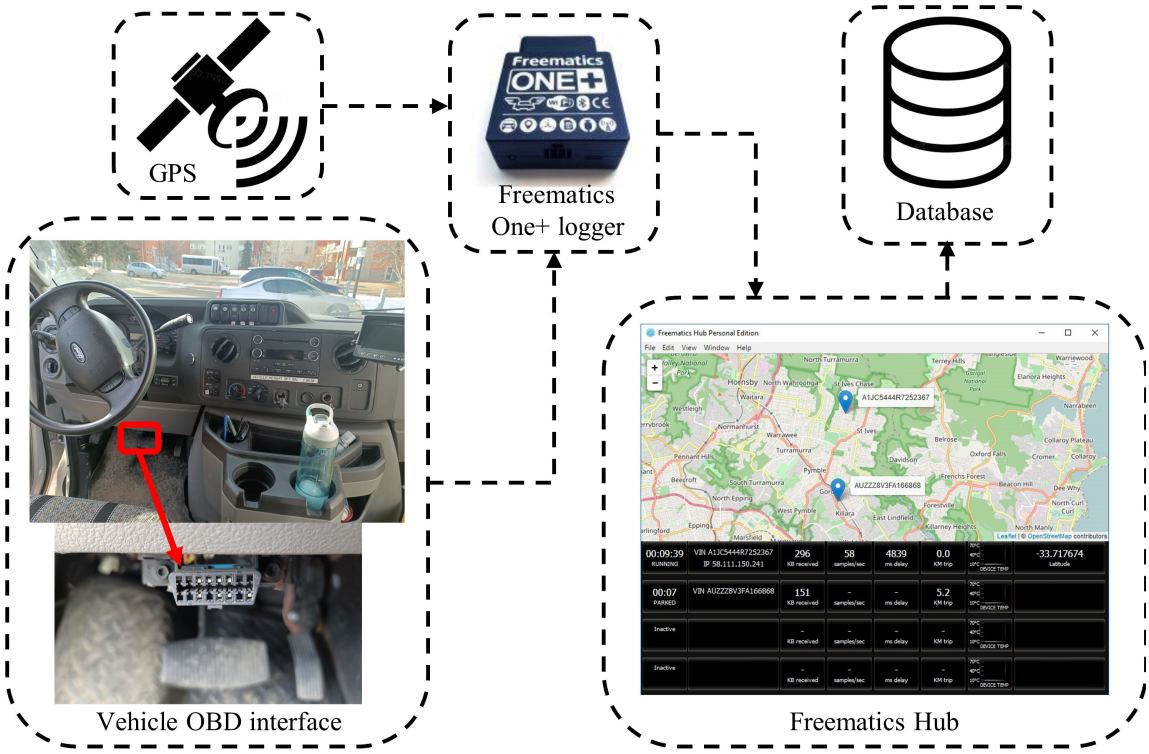


Figure 2.1: The workflow of Freematics One+ OBD data logger

Table 2.1: Parameters collected by Freematics One+ to generate driving cycle and analyze driver behavior

Parameter Name	**Unit**
Vehicle ID	-
GPS Time stamp	s
GPS latitude	°
GPS longitude	°
Vehicle speed	Km/h
Engine speed	RPM
Throttle position	°
Pedal position	°
Intake air mass flow rate (MAF)	g/s
Short and long term fuel trims	-
Air-fuel equivalence ratio (λ)	-

2.3 Vehicle Selection

Because there are more than 170 vehicles owned by UAlberta fleet vehicles, the cost of installing experimental equipment on all vehicles to collect data is too large. Moreover, many vehicles have similar or even identical models. At the same time, the geographical areas that different vehicles use also follow certain rules. The Utility or Trade and UAPS vehicles are running in the campus area. In contract, Casual Rental vehicles are mainly running on highways. The following content will introduce the operating characteristics, selected representative vehicle models, and specific vehicle driving trajectories according to the categories of vehicles.

2.3.1 Utility or Trades

The University of Alberta 's utility vehicles are mainly running near the university campus. The types of these vehicles in Utility or Trades category are mainly trucks and vans. In this experiment, six vehicles are selected to represent the University of Alberta trucks and vans category. The data for this study is collected for four months of operation from these six vehicles. In total, 716,794 data points were obtained. The specifications of vehicles are shown in Table 2.2 to Table 2.7.

Table 2.2: Specification of the tested vehicle A in utility and trades category

Vehicle	Dodge RAM 2500 (0429)
Model year	2011
Engine size	5.7 Liters
Fuel Type	Gasoline
Rated power	383 hp @ 5,600 rpm
Rated torque	400 lb-ft @ 4,000 rpm
Data collection duration	Aug. 17, 2021 to Nov. 24, 2021

Table 2.3: Specification of the tested vehicle B in utility and trades category

Vehicle	Ford E250 (0441)
Model year	2012
Engine size	4.6 Liters
Fuel Type	Gasoline (flex fuel)
Rated power	225 hp @ 4,800 rpm
Rated torque	286 lb-ft @ 3,500 rpm
Data collection duration	Aug. 30, 2021 to Dec. 16, 2021

Table 2.4: Specification of the tested vehicle C in utility and trades category

Vehicle	Ford E250 (0443)
Model year	2012
Engine size	4.6 Liters
Fuel Type	Gasoline (flex fuel)
Rated power	225 hp @ 4,800 rpm
Rated torque	286 lb-ft @ 3,500 rpm
Data collection duration	Aug. 30, 2021 to Dec. 7, 2021

Table 2.5: Specification of the tested vehicle D in utility and trades category

Vehicle	Ford F250 (0451)
Model year	2012
Engine size	6.2 Liters
Fuel Type	Gasoline (flex fuel)
Rated power	385 hp @ 5,500 rpm
Rated torque	405 lb-ft @ 4,500 rpm
Data collection duration	Aug. 19, 2021 to Dec. 24, 2021

Table 2.6: Specification of the tested vehicle E in utility and trades category

Vehicle	Chevrolet Express (0484)
Model year	2014
Engine size	5.3 Liters
Fuel Type	Gasoline (flex fuel)
Rated power	310 hp @ 5,200 rpm
Rated torque	334 lb-ft @ 4,500 rpm
Data collection duration	Sep. 3, 2021 to Jan. 20, 2021

Table 2.7: Specification of the tested vehicle F in utility and trades category

Vehicle	Ford Transit (0505)
Model year	2016
Engine size	3.7 Liters
Fuel Type	Gasoline (flex fuel)
Rated power	275 hp @ 6,000 rpm
Rated torque	260 lb-ft @ 4,000 rpm
Data collection duration	Aug. 30, 2021 to Dec. 13, 2021

Table 2.2 to Table 2.7 show that the utility trucks used by University of Alberta are all mostly light-duty trucks with engine size less than 6 liter. Such a vehicle model is in line with the characteristics of campus work. In campus, Large vehicles require a lot of steering space and are not suitable for campus work. At the same time, the vehicles working on the campus start and stop frequently with substantial idling time. The larger the vehicles, the more fuel consumption.

The running trajectory of the University of Alberta trucks and vans category can be obtained by GPS data from OBD logger. As shown in Figure 2.2, most of the vehicle trajectories are within the University of Alberta campus, which is framed by

a black dotted lines.

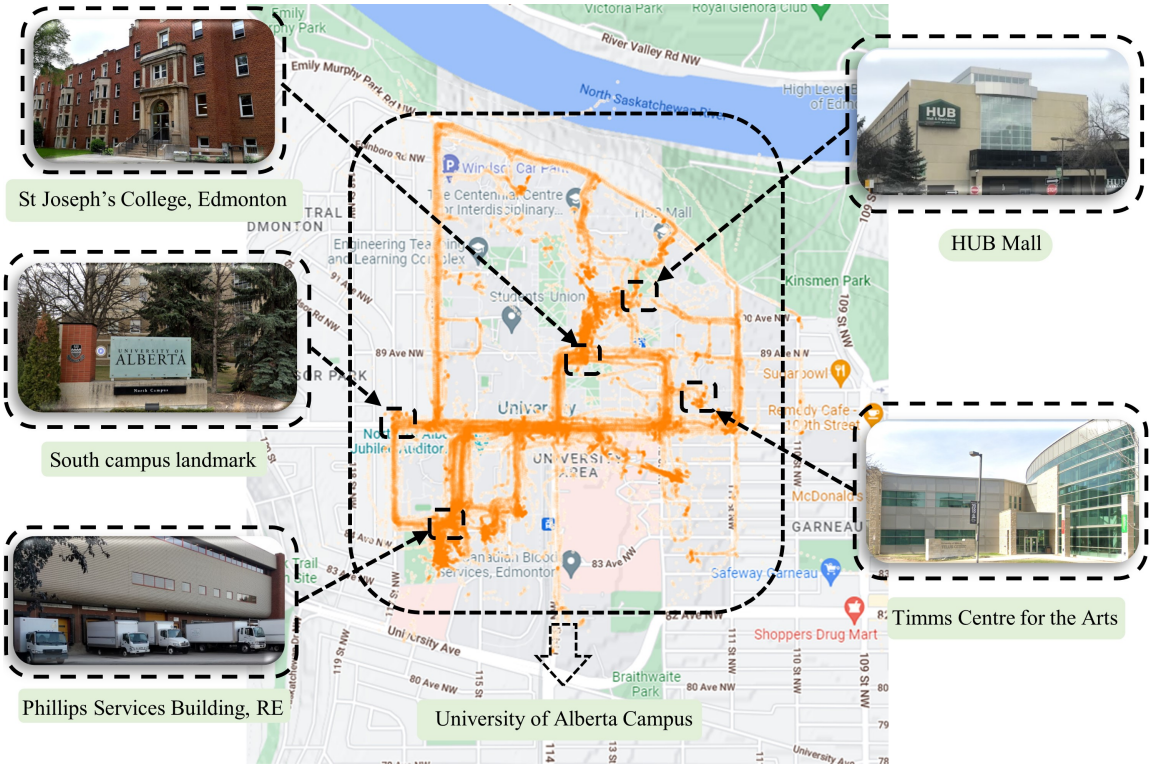


Figure 2.2: The GPS route map for the tested utility or trade vehicles; Data from Aug. 17, 2021 to Jan. 20, 2022

At the same time, in Figure 2.2, the landmark buildings near University of Alberta campus are also marked, they are St Joseph's College, South campus landmark, Phillips services building, HUB mall, and Timms Center for the arts respectively. Readers can find satellite images of nearby roads on Google Maps based on landmarks.

2.3.2 Casual Rentals

According to work management and demands of UAlberta, not all vehicles will be put into service. For the casual vehicles, University of Alberta fleet vehicles' manager will rent them out. Vehicle models in this category is similar to the vehicle model in Utility or Trades category. But their running trajectories, driving scenarios, and the number of starts and stops are quite different. In this experiment, a total of 4 vehicles were selected to represent the casual rental category. The specific information of the

selected vehicles is list in the Table 2.8 to Table 2.11.

Table 2.8: Specification of the tested vehicle A in casual rentals category

Vehicle	Ford F250 (0486)
Model year	2015
Engine size	6.2 Liters
Fuel Type	Gasoline (flex fuel)
Rated power	385 hp @ 5,500 rpm
Rated torque	405 lb-ft @ 4,500 rpm
Data collection duration	May 16, 2022

Table 2.9: Specification of the tested vehicle B in casual rentals category

Vehicle	Dodge Grand Caravan (0518)
Model year	2016
Engine size	3.6 Liters
Fuel Type	Gasoline (flex fuel)
Rated power	283 hp @ 6,400 rpm
Rated torque	260 lb-ft @ 4,400 rpm
Data collection duration	May 12, 2022 to Sep. 8, 2022

Through Table 2.8 to Table 2.11, it can be found that the selected vehicles are close to the vehicles in utility and trades category, all of which are light-duty trucks.

The trajectory map of Casual Rentals category vehicles is shown in Figure 2.3. It can be seen from Figure 2.3 that vehicles of Casual Rentals category mostly run on highways. Casual rental vehicles consumed more fuel than those urban driving vehicles as high speed driving cause high fuel consumption.

Table 2.10: Specification of the tested vehicle C in casual rentals category

Vehicle	Dodge Grand Caravan (0488)
Model year	2015
Engine size	3.6 Liters
Fuel Type	Gasoline
Rated power	283 hp @ 6,400 rpm
Rated torque	260 lb-ft @ 4,400 rpm
Data collection duration	May 12, 2022 to Sep. 14, 2022

Table 2.11: Specification of the tested vehicle D in casual rentals category

Vehicle	Ford F350 (0587)
Model year	2021
Engine size	6.2 Liters
Fuel Type	Gasoline (flex fuel)
Rated power	385 hp @ 5,750 rpm
Rated torque	430 lb-ft @ 3,800 rpm
Data collection duration	May 12, 2022 to Sep. 13, 2022

2.3.3 Shuttle Bus

Except the main campus located by the North Saskatchewan River, University of Alberta (UA) also has a campus named Campus Saint-Jean (CSJ) to provide students with French courses. UA has two minibuses to shuttle students between the two campuses. The passenger capacity of the minibus is about 30 passengers, so the vehicles are larger than noraml vans and consume more fuel. At the same time, the trajectory of the shuttle minibuses is fixed. This kind of trajectory is ideal for driver performance analysis as vehicles with the same driving trajectory will reduce the difference on vehicle fuel consumption caused by driven distance. Under such a

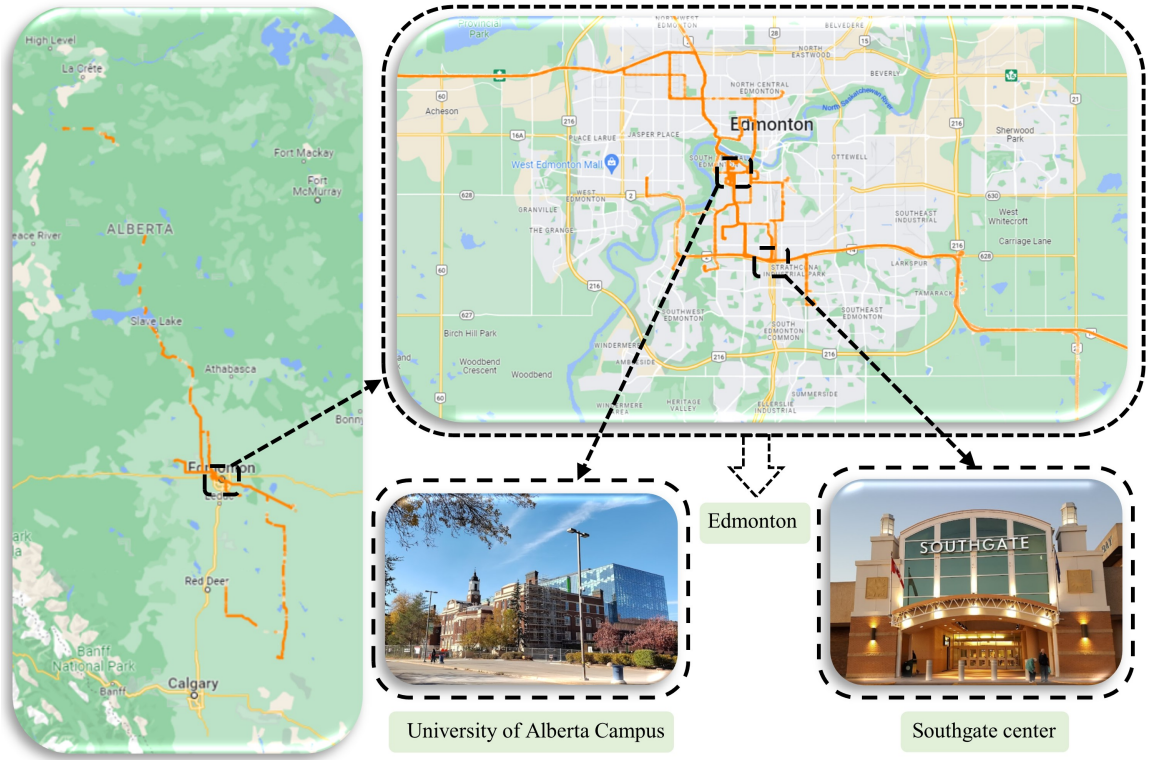


Figure 2.3: The GPS route map for the tested casual rental vehicles; Data from May. 12, 2022 to Sep. 14, 2022

control variable study, the fuel consumption of the vehicle is mainly caused by the difference in the driving style of the personnel. This part will be discussed in detail in Chapter 4. The specification of the two vehicles is shown in Table 2.12 and Table 2.13.

Table 2.12: Specification of the tested vehicle A in shuttle minibus category

Vehicle	FORD E-450 (0438)
Model year	2018
Engine size	7.3 Liters
Fuel Type	Gasoline (flex fuel)
Rated power	325 hp @ 3,800 rpm
Rated torque	420 lb-ft @ 3,000 rpm
Data collection duration	Apr. 4, 2022 to Apr. 21, 2022

Figure 2.4 shows the driving track of the minibus. At the same time, the University

Table 2.13: Specification of the tested vehicle B in shuttle minibus category

Vehicle	FORD E-450 (0416)
Model year	2018
Engine size	7.3 Liters
Fuel Type	Gasoline (flex fuel)
Rated power	325 hp @ 3,800 rpm
Rated torque	420 lb-ft @ 3,000 rpm
Data collection duration	Apr. 4, 2022 to Apr. 21, 2022

of Alberta north campus and CSJ are also marked with black dotted lines in the figure. The speed limit in the campus area in the picture is 30 km/h, and the speed limit in urban areas is 50 km/h. The distance from UA campus to CSJ is 6.1 Km, and the distance from CSJ to UA is 6.3 km.

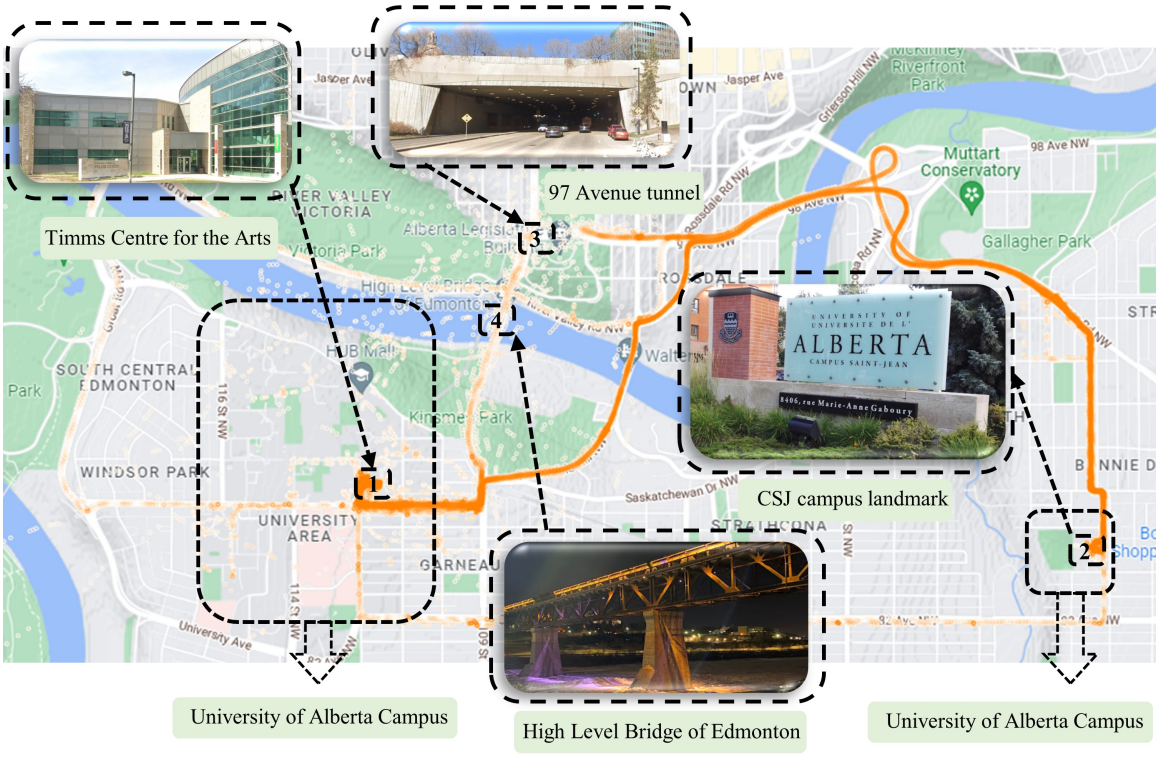


Figure 2.4: The GPS route map for the tested shuttle minibuses; Data from Apr. 4, 2022 to Apr. 21, 2022

As shown in Figure 2.4, both minibuses travel between the UA main campus and CSJ campus along the same route. Mark 1 in Figure 2.4 is the Timms Centre for the Arts in UA main campus, which is also the departure station of shuttle minibus, and Mark 2 in Figure 2.4 is the CSJ campus landmark in Saint Jean campus. The reason why the lines of the GPS route map at the Marks 3 and 4 locations are not clear is because the route at Mark 3 will pass through a tunnel, and the route at Mark 4 will pass through an iron bridge, High Level Bridge of Edmonton. The tunnel and iron bridge made the receiving and sending of GPS signals unstable.

2.3.4 University of Alberta Protective Service

The UAlberta fleet vehicles has a total of 6 sport utility vehicles (SUV), most of which are used for security patrols. UAPS vehicles patrol within the main campus of UA. However, the specific patrol trajectory cannot be publiced due to security concerns. In this experiment, a total of two vehicles were selected to represent the category of emergency vehicles. The specification of the vehicle are shown in Table 2.14 and Table 2.15.

Table 2.14: Specification of the tested vehicle A in UAPS category

Vehicle	Ford Explorer Police (0093)
Model year	2017
Engine size	3.7 Liters
Fuel Type	Gasoline (flex fuel)
Rated power	300 hp @ 6,250 rpm
Rated torque	280 lb-ft @ 4,500 rpm
Data collection duration	Jun. 23, 2022 to Feb. 2, 2023

As shown in Table 2.14 and Table 2.15, the models of the vehicles are all the same. Due to the limitation of safety factors, this thesis mainly studies the driving cycle of

Table 2.15: Specification of the tested vehicle B in UAPS category

Vehicle	Ford Explorer Police (0095)
Model year	2017
Engine size	3.7 Liters
Fuel Type	Gasoline (flex fuel)
Rated power	300 hp @ 6,250 rpm
Rated torque	280 lb-ft @ 4,500 rpm
Data collection duration	Jun. 21, 2022 to Oct. 9, 2022

UAPS.

2.3.5 Ford Escape Plug-In Hybrid Electric Vehicle

In the EMSO project, sedan vehicles are also classified in the casual rental category. One Sedan vehicles, Ford Escape Plug-In Hybrid Electric Vehicle (PHEV), is selected for the cold climate experiment. In the cold climate test, one driver did one normal driving and one aggressive driving through a fixed driving route. Thus, the data from this experiment can also be used for driver behavior analysis. The information, trajectory, and database of the selected sedan vehicle are shown below.

Vehicle Information

The specific information of the vehicle is shown in Table 2.16.

Test Rote

Figure 2.5 shows the select test driving route. The reason to select this route is that this route covers a broad range of driving conditions which is close to shuttle minibus driving conditions. The route includes campus area (maximum speed of 30 km/h), residential area (maximum speed of 40 km/h), and urban area (60 – 100 km/h).

Table 2.16: Specification of Ford Escape Plug-In Hybrid Electric Vehicle

Vehicle Make/ Model	Ford Escape PHEV
Model Year	2021
Vehicle Body Style	Compact SUV
Fuel Type	Gasoline / Battery
Engine Size	2.5 L
Engine Power	221 hp (@ 6,250 rpm)
Battery Capacity	14.4 kWh

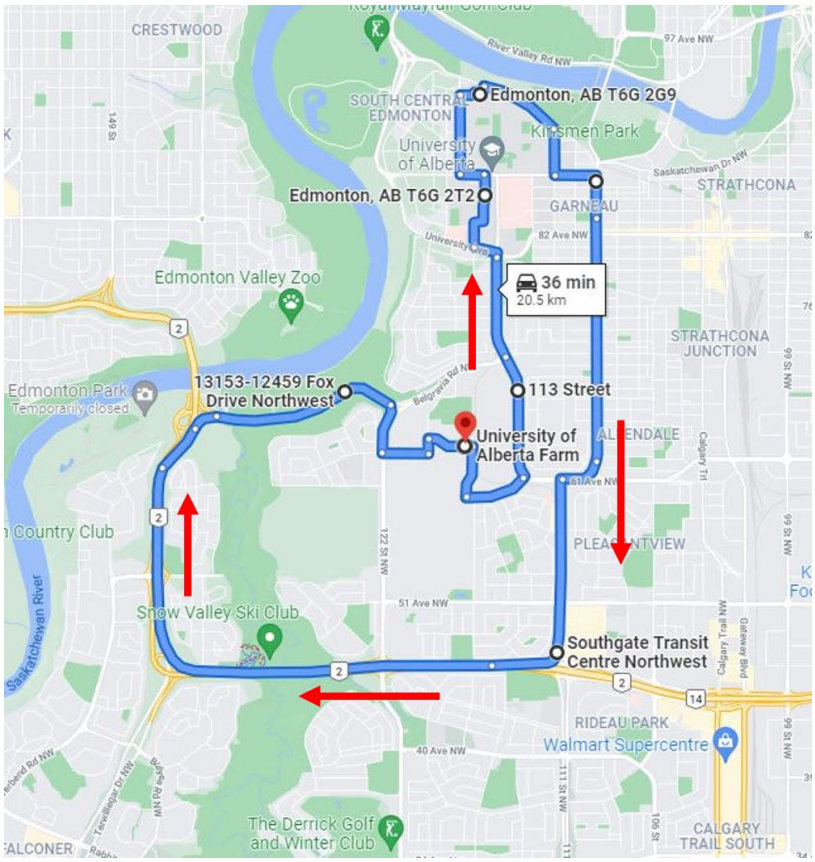


Figure 2.5: The GPS route map for Ford Escape PHEV; Use data from Mar. 26, 2022 to Mar. 27, 2022

Fuel Measurement

Because the vehicle model is different, the amount of data that the OBD data logger can read is also different. Freematics one + can't read the data like MAF and

AFR_{stoich} from Ford Escape. Therefore, an ultrasonic fuel flow meter by Sentronics (FlowSonic LF Low-Flow Sensor) was installed on Fordd Escape PHEV to measure instantaneous fuel consumption. Figure 2.6 shows the fuel flow meter installation in the fuel path of the tested vehicles. FlowSonic LF Low-Flow Sensor has many advantages including (i) low-volume fuel flow measurement (e.g., idling condition of a small fleet vehicle), (ii) different fuels measurement (e.g., gasoline, diesel), (iii) high measurement accuracy, and (iv) small and lightweight to install easily on any engine. Therefore, the lowSonic LF Low-Flow Sensor was installed on Ford Escape PHEV for the experiment done in this study.

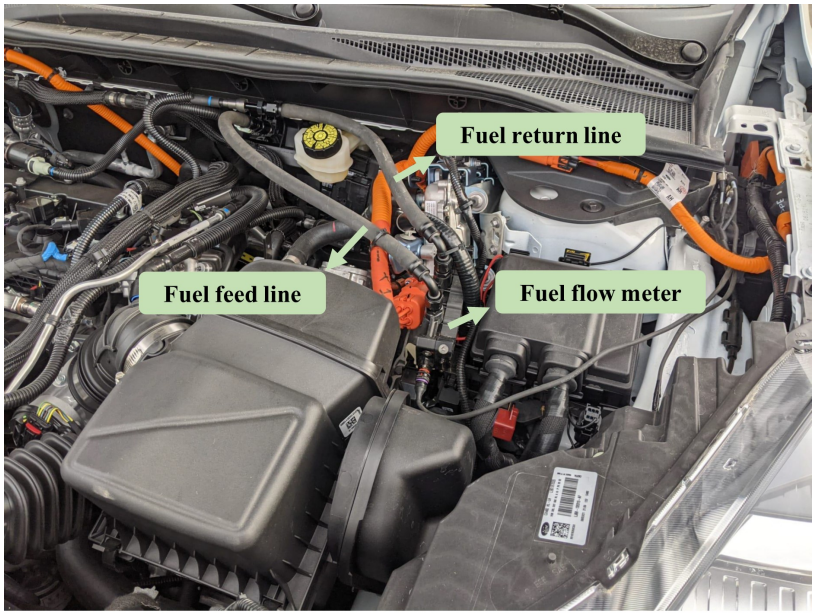


Figure 2.6: Ultrasonic fuel flow meter for Ford Escape PHEV

CAN Data Collection

Figure 2.7 shows the data collection process by using CANedge2 data logger. By using CANedge2 data logger, OBD data is collected through CAN bus. Fuel flow measurement data is collected by CSS Electronics CANedge2 CAN bus.

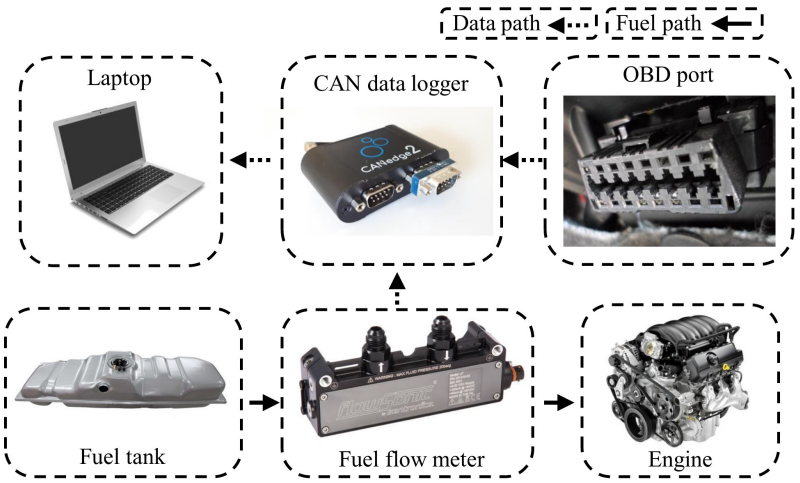


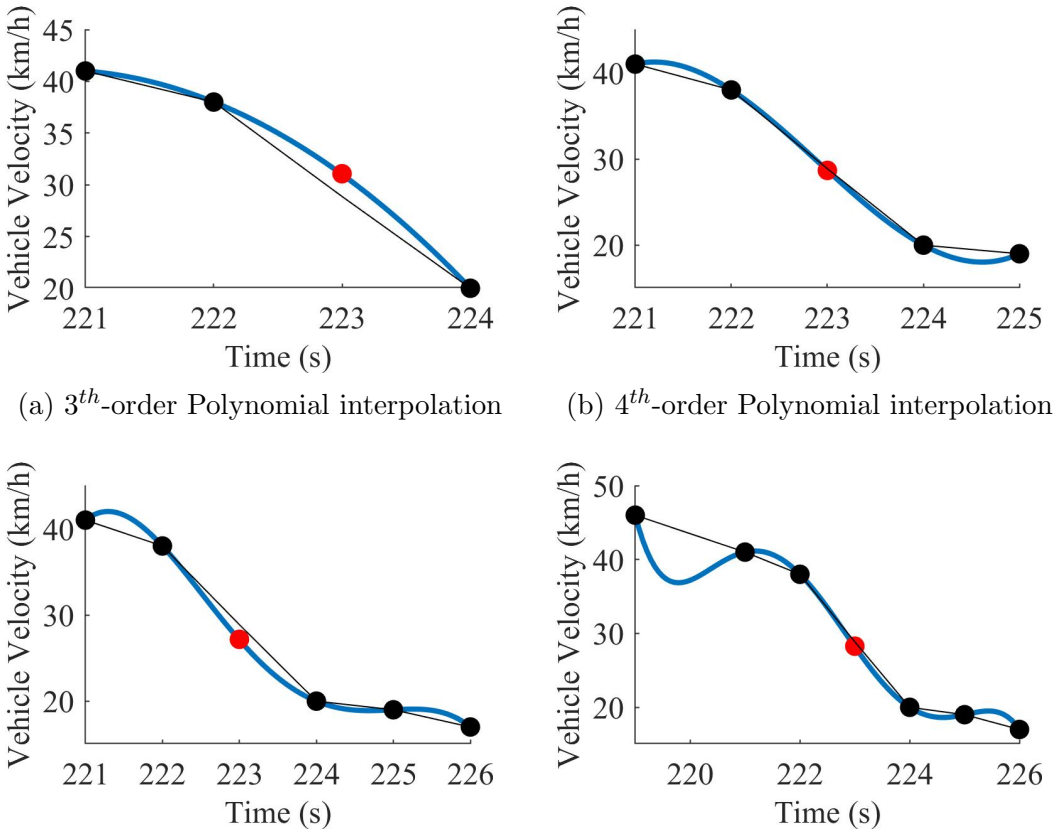
Figure 2.7: Schematics of the data collection process Ford Escape PHEV

2.4 Data Pre-process

The data obtained by OBD loggers cannot be used directly, due to missing data, and unstable sampling rate at some time stamps. In this section, according to the Freematics+ OBD characteristics used in this experiment, a data pre-processing method is designed to avoid data errors.

2.4.1 Data Interpolation

Data collection by Freematics+ OBD sometimes have missing data points. The missing data is filled by using a polynomial interpolation algorithm. Polynomial interpolation is the procedure of fitting a polynomial of degree n to a set of $n+1$ data points. In this thesis, according to the Runge's phenomenon [36] the 4^{th}-order Polynomial interpolation algorithm is used. If the order is too low, the interpolation will lead to underfitting, and if the order is too high, it will lead to large oscillations between data points [36]. Runge's phenomenon also happens in our test data which is shown in Figure 2.8. In Figure 2.8, the blue line is interpolating polynomial, the black dots are raw data points, the red dots are interpolated points, the black lines are linear links among data. Figure 2.8 shows that for the sample presented data, the 4^{th}-order polynomial performs better than other orders of polynominals.



(a) 3^{th}-order Polynomial interpolation (b) 4^{th}-order Polynomial interpolation

(c) 5^{th}-order Polynomial interpolation (d) 6^{th}-order Polynomial interpolation

Figure 2.8: Effect of polynomial order on interpolation accuracy; Use 2022.8.9 shuttle minibus data

In this thesis, since the frequency of collection is fixed, the time point t_m of missing data can be obtained by Equation (2.1).

$$t_m = t_{m-1} + \Delta t \tag{2.1}$$

t_m is the missing time data, t_{m-1} is the former time data of missing data, and Δt is the time interval. In this thesis, the sampling frequency of the data is 1Hz after down-sampling. Therefor, the Δt is 1 second.

The 4^{th} order interpolation equation, when used, requires 4 known data points to infer the unknown data points. In this thesis, the first two time data points t_{m-1} and t_{m-2} , and the last two time data points of the missing data points t_{m+1} and t_{m+2} is used as known time parameter points. The velocity data parameter is denoted by

y_{m-1}, y_{m-2}, y_{m+1}, and y_{m+2} . The matrix form of the 4^{th} order interpolation equation is shown as Equation (2.2)

$$\begin{bmatrix} 1 & t_{m-2} & t_{m-2}^2 & t_{m-2}^3 \\ 1 & t_{m-1} & t_{m-1}^2 & t_{m-1}^3 \\ 1 & t_{m+1} & t_{m+1}^2 & t_{m+1}^3 \\ 1 & t_{m+2} & t_{m+2}^2 & t_{m+2}^3 \end{bmatrix} \begin{Bmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \end{Bmatrix} = \begin{Bmatrix} y_{m-2} \\ y_{m-1} \\ y_{m+1} \\ y_{m+2} \end{Bmatrix} \tag{2.2}$$

Bring in the data points to solve the matrix to get the value of the a vector. According to Equation (2.3), the missing data points are determined.

$$v_m = \ a_0 + a_1 t_m + a_2 t_m^2 + a_3 t_m^3 \tag{2.3}$$

Here is an example in Figure 2.9, taking the shuttle minibus data on April 9, 2022, the original data and interpolated data are shown.

2.4.2 Data Re-sampling

The sampling frequency of the OBD data logger used in this thesis is not stable at 1 hz. This thesis adapts a down sampling method that extracts sample values in the window length (WL) to create a new signal. The number of new signal samples decreases as the WL value increases. The specific down-sampling equation is shown in Equation (2.4).

$$y(n) = \ x\left(nWL\right) \tag{2.4}$$

In Equation (2.4), x(n) is the original signal, y(n) is the downsampled signal

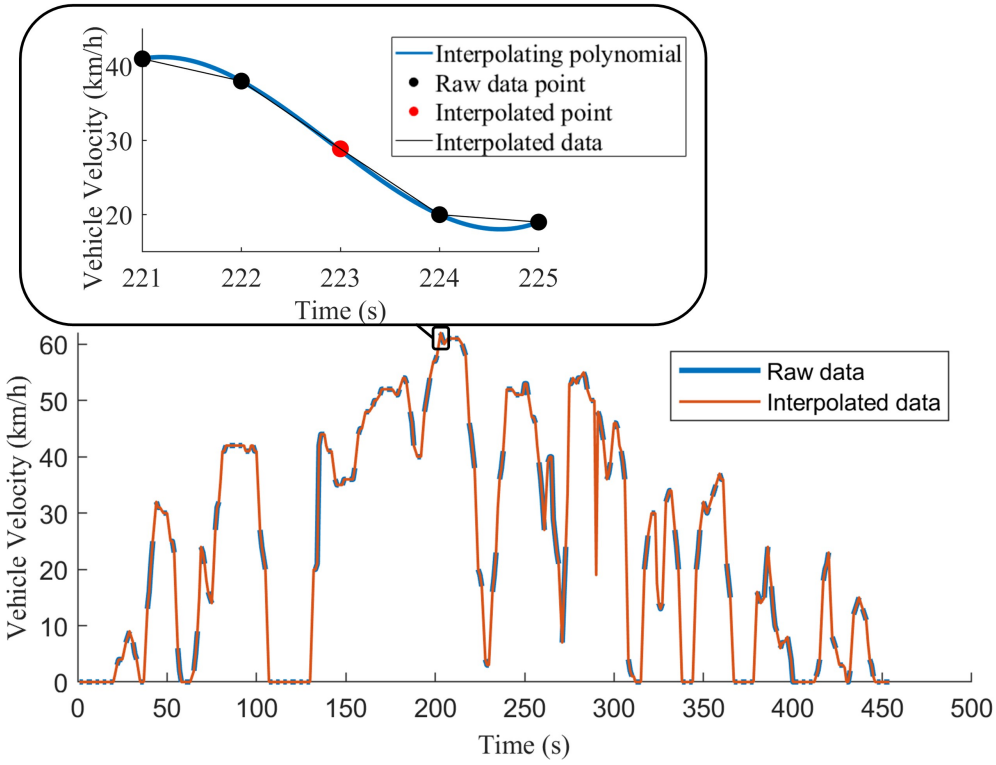


Figure 2.9: The comparison between the interpolated data and the original data based on the shuttle minibus data on April 9, 2022

Chapter 3

Driving Cycle Establish for Different Vehicle Categories

3.1 Introduction

In this chapter, some concepts related to the driving cycle will be introduced first. Secondly, the algorithm used in generating the driving cycle is introduced. The database for different vehicle categories and the algorithm for generating driving cycles are discussed in detail in Section 3.4 and Section 3.5. Finally, driving cycles for each vehicle category from UAlberta fleet vehicles are given.

3.2 Concept Definition

3.2.1 Microtrips and Microtrip Databases

A general method of creating driving cycle is breaking the recorded natural driving cycle data into pieces and combining those pieces with a certain algorithm later [37]. Microtrips are those pieces. So, when getting the recorded natural driving cycle data, one need is cut the data at the time stamps when vehicle velocity is zero [38]. A Microtrip is a trip between two consecutive times when the vehicle speed is zero. By using data from the Utility or Trades category, a sample set of Microtrip is shown in Figure 3.1. In Figure 3.1, each velocity-time series surrounded by a rectangle is a Microtrip.

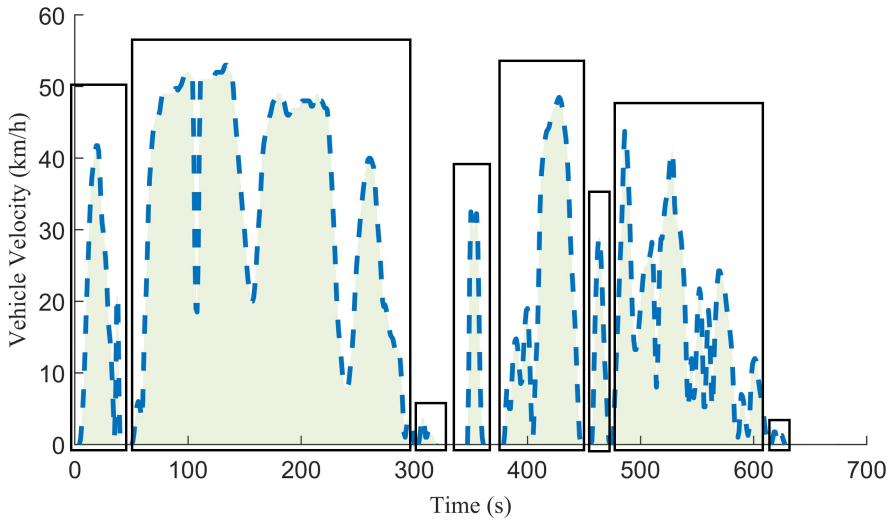
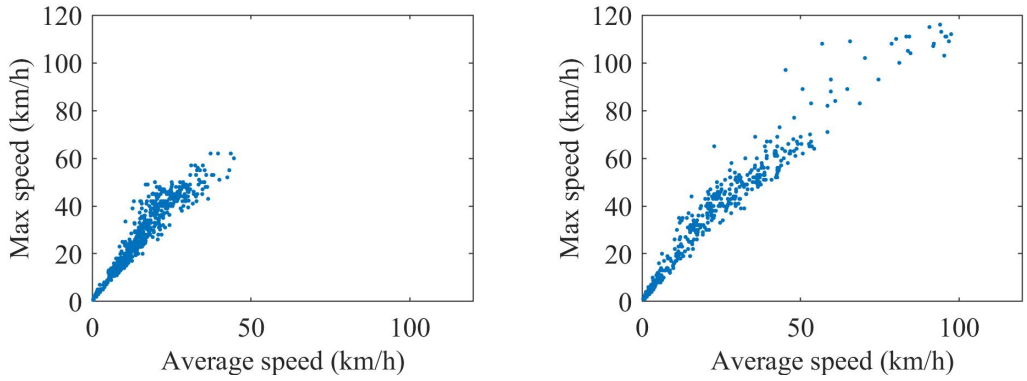


Figure 3.1: Depiction of Microtrips extraction for 2021/10/1 vehicle velocity data from a vehicle from utility or trades category.

A collection of Microtrip forms a Microtrip database. Figure 3.2 shows the Microtrip database for all vehicle categories. Microtrip database for Utility or Trades category is shown in Figure 3.2a, casual rental category Microtrip database is shown in Figure 3.2b, Figure 3.2c shows the Microtrip database for shuttle minibus, and UAPS' Microtrip database is shown in Figure 3.2d.

The reason for showing Microtrip database with the average and maximum speed is that the driving scenarios of each vehicle category can be obtained by combining these two parameters with the clustering algorithm [18]. The detail information of clustering algorithm is shown in Section 3.2.2, and the driving scenarios are shown in Section 3.4.

It can be seen from the Figure 3.2 that vehicles working near the University of Alberta campus have speeds of Microtrips within 80 km/h. The maximum speed of the casual rental vehicle's Microstrip has reached 120 km/h, reflecting the characteristics of rental vehicles running on expressways. In addition, the Microtrip's average speed of casual rental vehicles is mainly below 60 km/h, which shows that casual



(a) Utility or trade vehicles' Microtrip database

(b) Casual rental vehicles' Microtrip database

(c) Shuttle minibuses' Microtrip database

(d) UAPS' Microtrip database

Figure 3.2: Microtrip database for different vehicle categories for all OBD recorded data from Aug. 17, 2021 to Oct. 9, 2022

rental vehicles mainly travel in the residential and urban areas. The data points in Figure 3.2d are more scattered than other vehicles driving on the University of Alberta campus, which means that the acceleration of UAPS will be greater than other vehicles. On the contrary, the Microtrip data points of the utility or trade vehicles are more concentrated, indicating that the utility or trade vehicles run more smoothly and the acceleration will be relatively small.

3.2.2 Kinematic Fragment

A Kinematic fragment is a meticulous division of driving segments compared to Microtrips [39]. It contains acceleration, deceleration, cruise and idle driving modes.

The criteria to generate Kinematic fragments are explained in Table 3.1. As it shows, the driving segments with a certain range of velocity and acceleration are used to determine the driving state. This research considers the idle duration and does not consider ultra-short idling (less than 5 seconds) as an idle state. An example of kinematic fragments is illustrated in Figure 3.3.

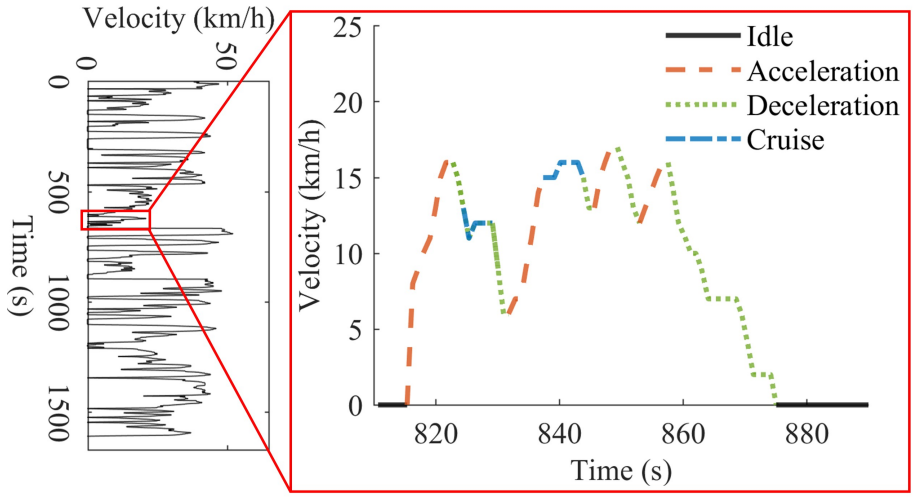


Figure 3.3: An example of kinematic fragments by a using utility vehicles' data from 2021/9/30 test

Table 3.1: Principles used for kinematic fragments

Kinematic State	**v (km/h)**	**a (m/s^2)**	**t (s)**	**Cruise State with $t < 5$ s**
Idle	$= 0$	–	> 0	–
Cruise	> 0	≥ -0.15 & ≤ 0.15	> 5	–
Acceleration	> 0	> 0.15	> 0	$V(k+1) > V(k)$
Deceleration	> 0	< -0.15	> 0	$V(k+1) < V(k)$

As it shows, each Microtrip consists of multiple Kinematic fragments. Those Kinematic fragments will be used to calculate the assessment metrics which are used to

judge whether a generated driving is acceptable. The assessment metrics are shown in Section 3.2.3.

3.2.3 Assessment Metrics

Generating a driving cycle is to find the driving cycle that best represents the recorded driving data. Statistical metrics are used to determine whether the generated driving cycle can represent all recorded driving cycles [16]. Some of the common metrics include the average velocity of a driving cycle (V_{avg}), the average velocity of a driving cycle except idle (Ve_{avg}), average acceleration of a driving cycle (Acc_{avg}), average deceleration of a driving cycle (Dec_{avg}), time spent on idling divided by the total time ($\%Idle$), time spent on cruise divided by the total time ($\%Cruise$), time spent on acceleration divided by the total time ($\%Acc$), time spent on deceleration divided by the total time ($\%Dec$), and the number of vehicle stops per kilometer ($Stop/km$). These metrics are also called target parameters. The values of each target parameters are the median of the data.

Table 3.2 shows the utility or trade vehicles' target parameters and their values. Similarly, Table 3.3 shows the casual rental vehicles' target parameters and their values. The values of target parameters for shuttle minibuses and UAPS vehicles are shown in Table 3.4 and Table 3.5 respectively.

Table 3.2: Utility or trade vehicles' assessment metrics values; Data from Aug. 17, 2021 to Jan. 20, 2022

Target Parameters	V_{avg} **(km/h)**	Ve_{avg} **(km/h)**	Acc_{avg} **(m/s^2)**	Dec_{avg} **(m/s^2)**	$Stop/km$
Value	14.6	20.1	0.36	-0.36	3.0
Target Parameters	$\%Acc$	$\%Dec$	$\%Cruise$	$\%Idle$	
Value	27.8	28.1	16.7	27.4	

Table 3.3: Casual rental vehicles' assessment metrics values; Data from May. 12, 2022 to Sep. 14, 2022

Target Parameters	V_{avg} **(km/h)**	Ve_{avg} **(km/h)**	Acc_{avg} **(m/s^2)**	Dec_{avg} **(m/s^2)**	$Stop/km$
Value	58.3	70.7	0.46	-0.47	0.4
Target Parameters	$\%Acc$	$\%Dec$	$\%Cruise$	$\%Idle$	
Value	14.9	14.6	52.9	17.6	

Table 3.4: Shuttle minibuses' assessment metrics values; Data from Apr. 4, 2022 to Apr. 21, 2022

Target Parameters	V_{avg} **(km/h)**	Ve_{avg} **(km/h)**	Acc_{avg} **(m/s^2)**	Dec_{avg} **(m/s^2)**	$Stop/km$
Value	25.0	33.2	0.44	-0.45	1.1
Target Parameters	$\%Acc$	$\%Dec$	$\%Cruise$	$\%Idle$	
Value	24.8	24.5	25.9	24.7	

Table 3.5: UAPS' driving assessment metrics values; Data from Jun. 21, 2022 to Oct. 9, 2022

Target Parameters	V_{avg} **(km/h)**	Ve_{avg} **(km/h)**	Acc_{avg} **(m/s^2)**	Dec_{avg} **(m/s^2)**	$Stop/km$
Value	13.7	23.2	0.76	-0.73	4.0
Target Parameters	$\%Acc$	$\%Dec$	$\%Cruise$	$\%Idle$	
Value	21.2	21.2	16.4	41.1	

By investigating Table 3.2 to Table 3.5, it can be found that casual rental vehicles have the highest average velocity. Moreover, UAPS vehicles have the highest acceler-

ation and deceleration, and utility or trade vehicles have the lowest acceleration and deceleration, which confirms that the conclusion in Section 3.2.1 about the acceleration value is correct. Furthermore, Table 3.3 gives more details about the driving state. The cruise driving state of the casual rental vehicle is relatively long that takes up more than half of the driving time. Table 3.5 shows that the UASP category vehicles have the most idle driving state. The driving style of casual rental vehicles is mainly long-term high-speed driving. The driving status of other vehicles will be more evenly distributed based on the $\%Acc$, $\%Dec$, and $\%Cruise$ values.

3.2.4 Driving Scenarios

Driving scenarios are situations and conditions that drivers may encounter while operating a vehicle on a road [40]. These scenarios can vary widely depending on factors such as the type of road, the time of day, the weather conditions, and the specific geographical location. Generally, driving scenarios can be divided into three broad categories for different studies which are shown in Figure 3.4. Three categories are (i) Environmentally Influenced Driving Scenarios, (ii) Driving Speed Influenced Driving Scenarios, and (iii) Driving Area Influenced Driving Scenarios.

Environmentally Influenced Driving Scenarios usually contain parts like night driving scenario, adverse weather conditions driving scenario, and so on [40–42]. The night driving scenario involves driving in low light conditions, which can make it more difficult to see and react to potential hazards on the road. Drivers in this scenario should make sure their headlights are on and be extra vigilant for pedestrians, bicyclists, and other hazards that may be difficult to see. The adverse weather conditions driving scenario includes driving in the rain, snow, ice, fog, or other hazardous weather conditions that can make driving more challenging. Drivers in this scenario should slow down and allow extra time for braking and steering and should avoid sudden movements or sharp turns.

Driving Area Influenced Driving Scenarios usually contain parts like lead sudden

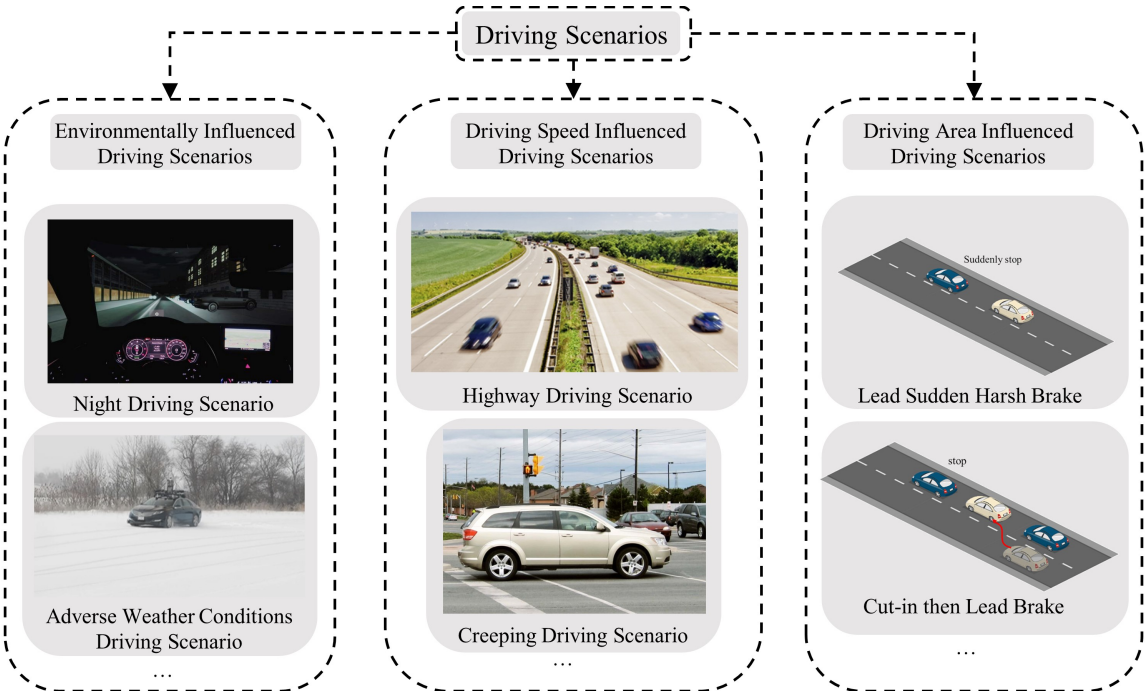


Figure 3.4: Classification for type of analysis for driving scenarios.

harsh brake driving scenario, cut-in then lead brake, and so on [43]. The research on this category mainly focuses on driving safety, especially those conditions which relate to driving conflicts. A conflict is defined as a driving situation that results in a driver crashing or evasive maneuvers [44]. When drivers in scenarios that are required a rapid, severe, evasive maneuver to avoid a crash, the driving task is referred to as Safety Critical Event (SCE). When the drivers are required an evasive maneuver occurring at less magnitude than a SCE, the driving task is given the label of Safety Relevant Event (SRE). Based on SCE and SRE criteria, to analyze French in-depth accident database VOIESUR [45], European naturalistic driving study UDRIVE and Field Operational Test MOOVE [46], several safety-critical scenarios, like lead sudden harsh brake driving scenario, cut-in then lead brake are determined.

Driving Speed Influenced Driving Scenarios usually contain parts like highway driving scenario, city driving scenario, creeping driving scenario, and so on. The highway driving scenario involves driving on a freeway or interstate, where drivers typically travel at higher speeds and must be aware of merging traffic, changing lanes, and nav-

igating exits. Drivers in this scenario need to be alert, maintain a safe distance from other vehicles, and follow posted speed limits [47]. City driving scenario involves navigating through urban environments, including dealing with traffic congestion, stop-and-go driving, and navigating through intersections and one-way streets. Drivers in this scenario need to be extra vigilant, especially when pedestrians or bicyclists are present and should be prepared for sudden stops or changes in traffic flow [48]. Creeping driving scenario is a situation where a driver moves the vehicle forward at a very slow speed, typically a few kilometers per hour or less, in order to maintain control of the vehicle and navigate through a congested area. This scenario is often encountered in slow-moving traffic, parking lots, or other situations where the driver needs to be extra cautious and maintain a high level of control over the vehicle [49].

In this research, Driving Speed Influenced Driving Scenarios is the key part. An unsupervised machine algorithm is used to generate different driving scenarios, including creeping driving scenario, low-speed city driving scenario, high-speed city driving scenario, etc, which will be discussed in Section 3.4.

3.3 Machine Learning Algorithm

3.3.1 K-means

Each Microtrip has a certain value of average vehicle speed and maximum vehicle speed which are the most important driving characteristic parameters, as shown in Figure 3.2 [50]. Each point in Figure 3.2 is a Microtrip. By using the data in Figure 3.2, a machine learning model is developed to form the sub-Microtrip database for different vehicle categories. Each sub-Microtrip database presents one of the driving scenarios from creeping driving scenarios to high-speed driving scenarios. In this study, the K-means algorithm is used to determine the sub-Microtrip database.

The K-means algorithm is a common unsupervised learning algorithm, which is mainly used to automatically classify similar samples into one category. In the cluster-

ing algorithm, samples are divided into different categories according to the similarity among samples. For different similarity calculation methods, different clustering results will be obtained. The similarity calculation method used in this thesis is the Euclidean Distance method [51], because the Euclidean Distance method is simple to calculate and it can speed up the algorithm.

For using the K-means algorithm, firstly, the sample set X in Equation (3.1) and the number of clusters k should be determined. Each element in X is called an object.

$$X = \{X_1, X_2, X_3, \ldots X_n\} \tag{3.1}$$

The goal of the K-means algorithm is to gather n objects into the specified k clusters according to the similarity between objects. For K-means, it is needed to initialize k cluster centers as listed in Equation (3.2). Initialized k cluster centers are usually composed of k points which are randomly selected from X.

$$C = \{C_1, C_2, C_3, \ldots C_k\}, \ 1 < k \leq n \tag{3.2}$$

Equation (3.3) is used to calculate the Euclidean distance (dis) from each object to each cluster center.

$$dis\,(X_i, C_j) = \sqrt{\sum_{t=1}^{n} (X_{it} - \ C_{jt})^2} \tag{3.3}$$

Next, the distance from each object to each cluster center is compared, and the object is assigned to the cluster with the nearest cluster center. Therefore, k clusters S is determined in Equation (3.4).

$$S = \{S_1, S_2, S_3, \ldots S_K\} \tag{3.4}$$

The K-means clustering algorithm uses the center to define the characteristics of the cluster. The new cluster center of the cluster is the mean value of all objects in the cluster in each dimension. The equation to calculate the new cluster center (C_{inew}) is as follows.

$$C_{inew} = \frac{\sum_{X_i \in S_i} X_i}{|S_i|} \tag{3.5}$$

After obtaining the new cluster center, an interactive process is followed to calculate the distance from each object to each cluster center to determine the new cluster center. Until the value of the cluster center does not change, the classification of the data is completed.

For the selection of k value, this thesis uses cross-validation to select the optimal k according to the loss Equation (3.6) [52].

$$Loss = \sum_{i=1}^{n} min_k \left\| X_i - C_k \right\|^2 \tag{3.6}$$

The loss Equation (3.6) will eventually have an elbow point. According to the elbow point, the best k value is selected. However, sometimes the elbow point is not easy to see directly. At this time, $\Delta Loss$ is needed to find the elbow point. The equation of $\Delta Loss$ is shown blow

$$\Delta Loss(i) = Loss(i+1) - Loss(i) \tag{3.7}$$

When the elbow point of $\Delta Loss$ is found, assuming that the found $\Delta Loss$ value is $\Delta Loss(i)$, then the value of $Loss(i+1)$ is the elbow point of the loss function data.

By using the K-means algorithm, the sub-Microtrip database is determined and each cluster of the Microtrip is a driving situation. Sub-Microtrip will be detialed in Section 3.4.

3.3.2 Principal Component Analysis

PCA is an unsupervised machine learning algorithm and a technique for exploring high-dimensional data structures. It is mainly used for dimensionality reduction of data. Through dimensionality reduction, features that are easier found, and the processing speed of valuable information from samples can be accelerated.

The logic of PCA is that the data is transformed from the original coordinate system to the new coordinate system, and the choice of the new coordinate system is determined by the data itself. The first new coordinate axis selects the direction with the largest variance in the original data, and the second new coordinate axis selects the direction that is orthogonal to the first coordinate axis and has the largest variance. This process is repeated for the number of features in the original data. Most of the variance is contained in the first few new axes. Therefore, to ignore the remaining coordinate axes can perform dimensionality reduction on the data [53].

In PCA, the distance between the coordinate axes of the new coordinate system and the original data has a distance relationship. There is a link between the original data and the new axes through the variance. The variance is the degree of dispersion of a random variable, that is, the measure of the distance between a group of numbers and its average value. The variance is also the expected value of the square of the deviation between a random variable and its population mean or sample mean [54]. Through reverse thinking, the insignificant coordinate axis can be regarded as the insignificance of the parameter, that is, the weight of the parameter is designed through the primary and secondary of the coordinate axis of the new coordinate system. The specific algorithm of this weight designer is as follows.

When using PCA, the parameters are not distinguished as independent variables and dependent variables; thus, all parameters are treated equally. The parameters are arranged into an array X, and the covariance matrix "Σ" of different parameter arrays are calculated. The analysis is then simplified by examining arrays in new

spaces through mathematical transformations, which is shown in Equation (3.8).

$$\begin{cases} Y_1 &= \alpha_1^T X = a_{11}x_1 + \ldots + a_{1p}x_p \\ Y_2 &= \alpha_2^T X = a_{21}x_1 + \ldots + a_{2p}x_p \\ \ldots & \\ Y_p &= \alpha_p^T X = a_{p1}x_1 + \ldots + a_{pp}x_p \end{cases} \tag{3.8}$$

$$cov\left(Y_i, Y_k\right) = \alpha_i^T \sum \alpha_k = 0 \tag{3.9}$$

$$\left\|\alpha_i\right\|^2 = \alpha_i^T \alpha_i = 1 \tag{3.10}$$

Where, a_{ij} represents the value of row "i" and column "j" in the covariance matrix "Σ" . The principal component (PC) is α_i in the Equation (3.8), which is also the eigenvector e_i of the covariance matrix "Σ". Equation (3.9) and Equation (3.10) show that the vectors of PC in the new coordinate system are pairwise orthogonal, which explains that each coordinate axis is independent of the other in the new space. Moreover, it is the orthogonality of those vectors that greatly simplifies the research difficulty of the problem as orthogonality indicates that the subsequent design of the weight allocator can ensure that the weights of the parameters will not overlap. α are used to get the matrix of feature vector A as Equation (3.11) shows.

$$A = \begin{bmatrix} \alpha_1^T \\ \alpha_2^T \\ \ldots \\ \alpha_p^T \end{bmatrix} = \begin{bmatrix} a_{11} & \ldots & a_{1p} \\ a_{21} & \ldots & a_{2p} \\ \ldots & \ldots & \ldots \\ a_{p1} & \ldots & a_{pp} \end{bmatrix} \tag{3.11}$$

The eigenvalues λ_i is obtained by calculating the variance of Y_i from Equation (3.12) to Equation (3.14). The larger the eigenvalue λ_i, the more important the coordinate

axis of the corresponding new coordinate system is, that is, the data set is mainly close to the new coordinate axis which has larger λ_i.

$$Y_i = \alpha_i^T X \tag{3.12}$$

$$var\,(Y_i) = \ max\,\|\alpha_i\| = \alpha^T \sum \alpha \ \left(with\ \alpha_i^T \sum \alpha_j = 0\right) \tag{3.13}$$

$$var\,(Y_i) = \alpha_i^T \sum \alpha_i = \lambda_i \tag{3.14}$$

Next, a single λ_i is divide by the $\Sigma\lambda_i$ to get $\lambda_i\%$ and the elements of the covariance matrix are divided by the sum of the elements in the same column to get $a_{(i,j)}\%$. The equations are shown in the following. Finally, Equation (3.17) is obtained that yields the matrix of the proportional feature vector.

$$\lambda_i\% = \frac{\lambda_i}{\sum \lambda_i} = \left(\frac{\lambda_i}{\lambda_1 + \ldots + \lambda_i + \ldots \lambda_p}\right) \times 100\% \tag{3.15}$$

$$a_{i,j}\% = \ \frac{a_{i,j}}{\sum\limits_{i=1}^{p} a_{i,j}} \times \lambda_i\% \tag{3.16}$$

$$A\% = \begin{bmatrix} a_{11}\% & \ldots & a_{1p}\% \\ a_{21}\% & \ldots & a_{2p}\% \\ \ldots\% & \ldots & \ldots \\ a_{p1} & \ldots & a_{pp}\% \end{bmatrix} \tag{3.17}$$

By summing of each row of the matrix of proportional feature vector $A\%$, the weight of each parameter is obtained as shown in Equation (3.18). By summing of

each column of the matrix of proportional feature vector $A\%$, the $\lambda_i\%$ of each PC is obtained. In order to avoid confusion of equation parameters, subsequent $\lambda_i\%$ is represented by $\lambda_j\%$.

$$W_i = \sum_{j=1}^{p} a_{i,j} \tag{3.18}$$

$$\lambda_j\% = \sum_{i=1}^{p} a_{i,j} \tag{3.19}$$

Figure 3.5 is the matrix of the proportional feature vector for vehicle target parameters. In this thesis, the target parameters of all driving cycles are calculated and combined into a target parameter matrix. Each row represents the target parameters, and each column represents the eigenvalues. Inputting the data matrix into the previously designed weight allocator yields Figure 3.5.

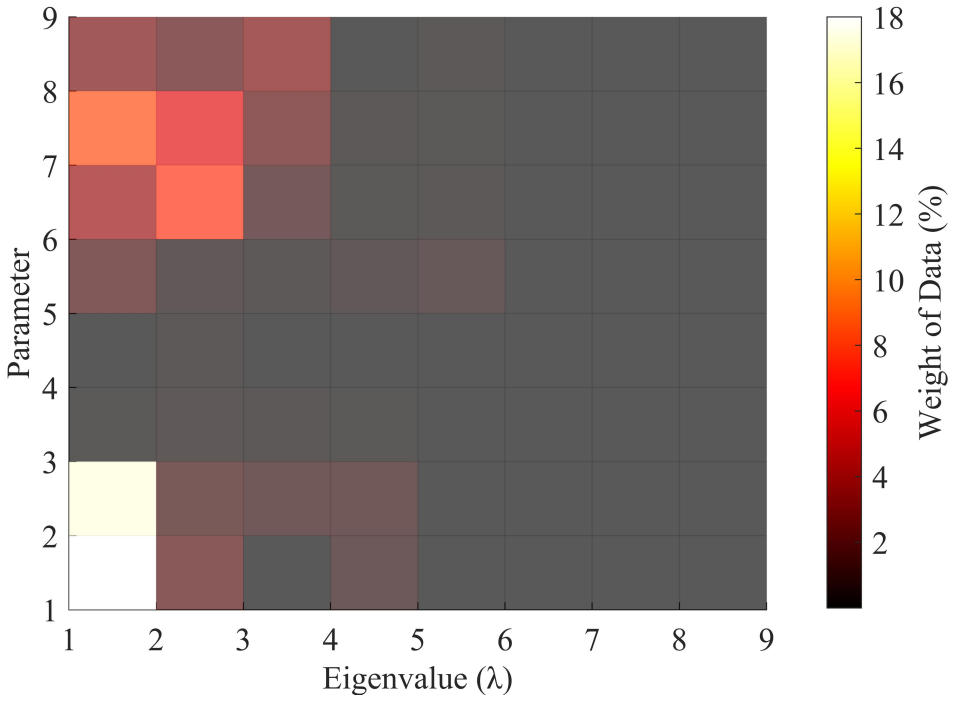


Figure 3.5: The matrix of the proportional feature vector based on the OBD recording data from Aug. 17, 2021 to Oct. 9, 2022 for all vehicle category

In Figure 3.5, parameter 1 to 9 are V_{avg},Ve_{avg}, Acc_{avg}, Dec_{avg}, $\%Idle$, $\%Cruise$, $\%Acc$, $\%Dec$, and $Stop/km$, respectively. Color bar shows the importance of each data, which is the weight for each data. Add the data of each row of Figure 3.5 can get the weight of each parameter, and add the data in each column of Figure 3.5 can get the eigenvalues for each PC.

3.3.3 Revised Root Mean Square Algorithm

The Root Mean Square (RMS) is an algorithm used to determine the deviation between the result data column and the target data column. RMS is one of four evaluation functions commonly used in machine learning [55]. The root means square error, also known as the standard error, is the square root of the ratio of the square of the deviation between the observed value and the true value to the number of observations. The root mean square error is used to measure the deviation between the observed value and the true value. The standard error is very sensitive to the large or small error in a set of measurements, and the standard error can be used as a criterion for assessing the precision of a measurement process. RMS is calculated as follows:

$$RMS = \sqrt{\sum_{n} \left(\frac{x_i - x_{i.avg}}{x_{i.avg}} \right)^2} \tag{3.20}$$

$$x_i = average\ target\ parameter\ for\ a\ candidate\ drive\ cycle$$

$$x_{iavg} = average\ target\ parameter\ for\ all\ OBD\ recorded\ data$$

However, the RMS algorithm does not take into account the weight for all parameters. If the RMS algorithm is directly used as the evaluation function of the driving cycle, there will be an effect of parameter error accumulation. Therefore, in

this research, the Principal component analysis algorithm is used to design weight for each assessment metric parameters. the detailed information can be got from Section 3.3.2. The weight of each parameter is shown in Figure 3.6.

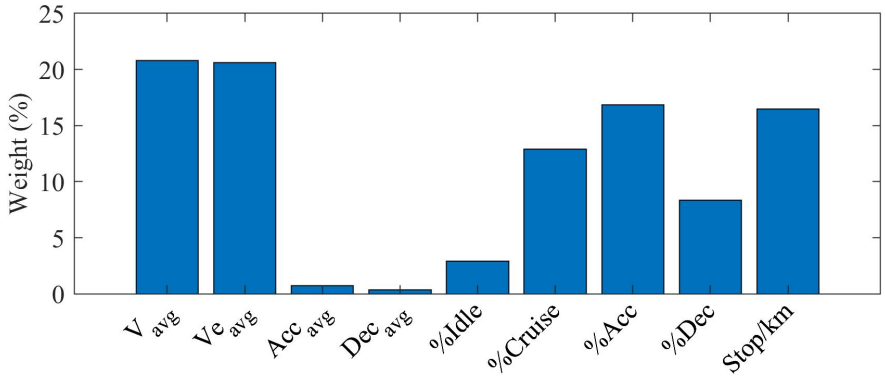


Figure 3.6: Weights for target parameters based on the OBD recording data from Aug. 17, 2021 to Oct. 9, 2022 for all vehicle category

Through Section 3.3.3 to Section 3.3.2, the revised RMS ($RMS_{revised}$) is determined.

$$RMS_{revised} = \sqrt{\sum_{n} \left(\frac{x_i - x_{i.avg}}{x_{i.avg}} \right)^2} \times W_i \tag{3.21}$$

As shown in Equation (3.21), each parameter is assigned a weight when it is used. Thus, during use, the problem of over-weighting due to parameter correlation is resolved.

3.4 Driving Scenarios for Different Vehicle Category

The generation of the driving cycle for the UAlberta fleet vehicles in this research faced the problem of an excessive number of Microtrips. Therefore, it is important to simplify the Microtrip database. To this end, at first, Microtrips are divided into several scenarios, then simplify each scenario, and finally get a simplified database,

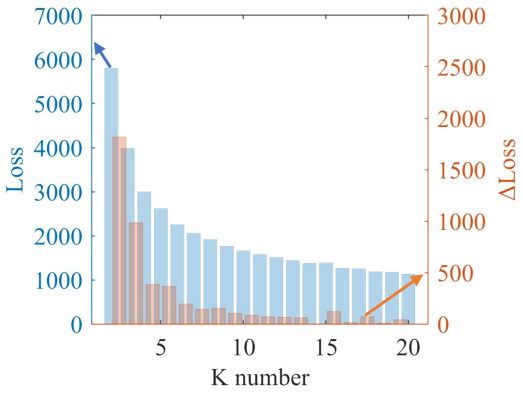
which is also called sub-Microtrip database. Through this categorization and simplifying method, the database can be simplified quickly.

Average speed (V_{avg}) and maximum speed (V_{max}) are important parameters to reflect the driving state. For example, when the V_{avg} and V_{max} are close to 0, vehicles are usually in idle driving scenarios, but when the V_{avg} and V_{max} are high, vehicles are usually in a cruising state. Therefore, using the V_{avg} and V_{max} to distinguish the Microtrips can let driving scenarios be determined.

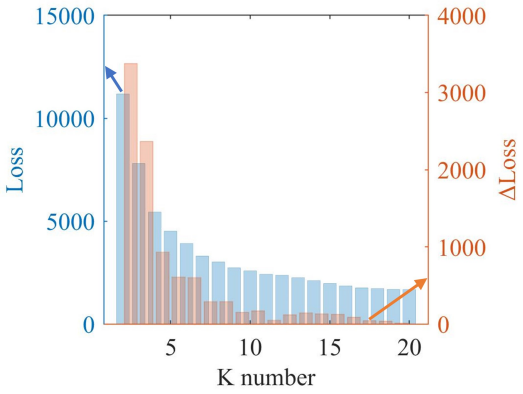
The traditional driving scenarios dividing method is based on the speed range evenly divided method. For example, the driving scenarios are divided into steps of $5km/h$ (0 - 5 km/h is the first driving scenario, 5 - 10 km/h is the secondary driving scenario, and so on) [20]. However, such a driving scenarios determination method has two drawbacks. The first drawback is that the division steps of different vehicle categories are different. Vehicles travelling on campus should have smaller steps, while vehicles travelling on highways should have larger steps. Another problem is how to determine the value of the step size. As shown in the Figure 3.2, it is difficult to determine how to divide the data. Therefore, a driving scenarios determination algorithm based on an unsupervised machine learning method is applied in this thesis.

As mentioned before in Section 3.2.4, clustering a Microtrip database is a kind of method to get driving scenarios. As introduced in the previous, this thesis uses the speed-based driving scenario division method, because the experimental data in this thesis are real-world vehicle driving data. The Microtrip database of different vehicle categories can be classified using the unsupervised learning algorithm K-means. The number of clusters can be obtained from the Equation (3.7). The result is shown below.

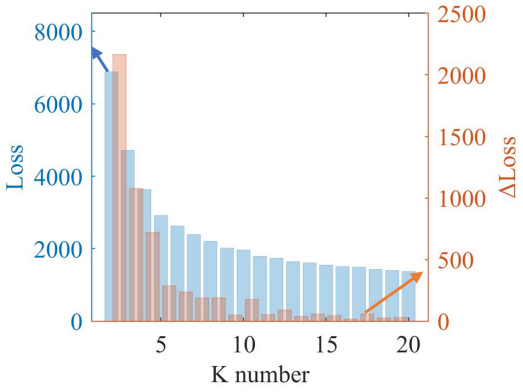
Figure 3.7a is the result for utility or trade vehicle category, Figure 3.7b is the result for casual rental vehicle category, Figure 3.7c is the result for shuttle minibuses category, and Figure 3.7d is the result for UAPS category. From the Figure 3.7, the blue histogram is the *Loss* value, which is difficult to find the elbow point. Therefore,



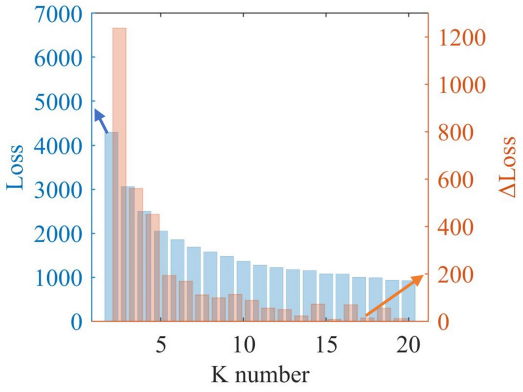
(a) *Loss* and $\Delta loss$ for utility or trade vehicles



(b) *Loss* and $\Delta loss$ for casual rental vehicles



(c) *Loss* and $\Delta loss$ for shuttle minibuses



(d) *Loss* and $\Delta loss$ for UAPS

Figure 3.7: *Loss* and $\Delta loss$ for different vehicle category for all OBD recorded data from Aug. 17, 2021 to Oct. 9, 2022

use the $\Delta Loss$ value obtained by the Equation (3.7) to find the elbow point. It can be seen from the Figure 3.7 that the utility or trade vehicle category and casual rental vehicle category both need five cluster centers, but the shuttle minibuses category and UAPS category need six cluster centers.

The selection for the final number of clusters in this study is confirmed with the physical interpretation of vehicle velocity data and purpose of this study to understand different driving situations by university vehicles as well understanding driver aggressiveness in different scenarios. The number of driving scenarios is linked to the type of driving area. For instant, low-speed creeping driving scenario links to the

crossroad area, campus driving scenario with speeds below 30 km/h links to university area, etc.

Another factor that affects the number of driving scenarios is speed. Take utility or trades vehicle data as an example, high-speed resident driving scenario's average driving speed is 30.32 km/h, but low-speed resident driving scenario's average driving speed is 18.61 km/h. Driving speed is related to driving aggressiveness. More aggressive driving usually has higher driving speeds. Dividing Micortrips with large differences in driving speed into different driving scenarios can make the comparison of driving aggressiveness more reliable.

In addition, the choice of the number of driving scenarios is also affected by the Microtrip duration. Take utility or trades vehicle data as an example, the Microtrip duration range of Creeping driving scenario is mainly within 15 s, the low-speed campus driving scenario Microtrip duration is mainly within 60 s, the high-speed campus driving scenario Microtrip duration is mainly within 100 s, etc. The duration of Microtrip directly affects energy consumption. Such a division can ensure that the impact of driver aggression on fuel consumption is sufficiently accurate. Otherwise, the variance of data statistics in the same driving area will be large, which means that the correlation of the data is weak.

The Microtrip database divided into driving scenarios is shown in Figure 3.8 to Figure 3.11.

After getting driving scenarios, $RMS_{revised}$ and PCA from Section 3.3.3 and Section 3.3.2 are applied to get the five most representative Microtrips in each driving scenario interval to simplify the Microtrip database. The simplified Microtrip database with driving scenarios classification is called sub-Microtrip database. The sub-Microtrip database for different vehicle categories is shown from Figure 3.12 to Figure 3.15. In each driving scenario category, the solid lines are the best representative Microtrips. The remaining Microtrips are represented by the area plot, and the darker the color, the more frequent the driving situation in the range appears.

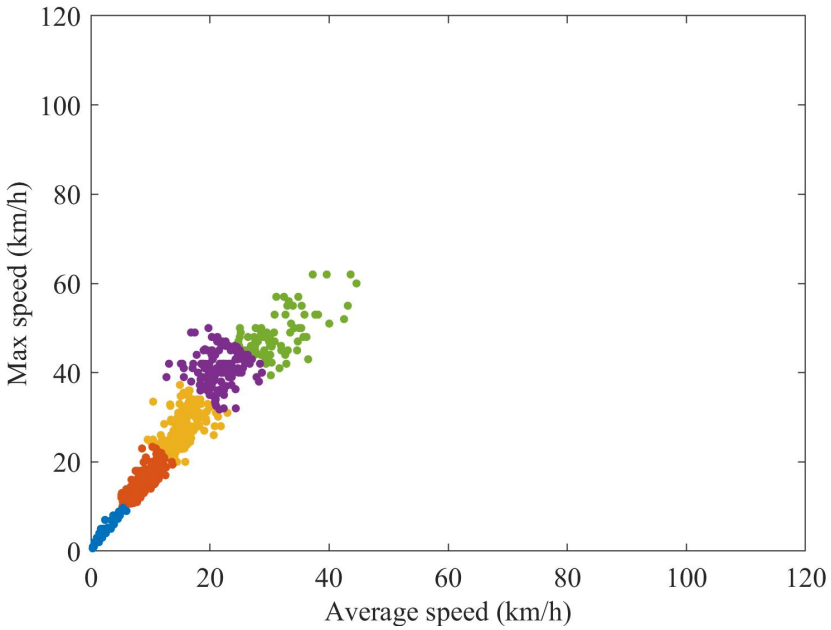


Figure 3.8: Driving scenarios for utility or trade vehicles; Data from Aug. 17, 2021 to Jan. 20, 2022

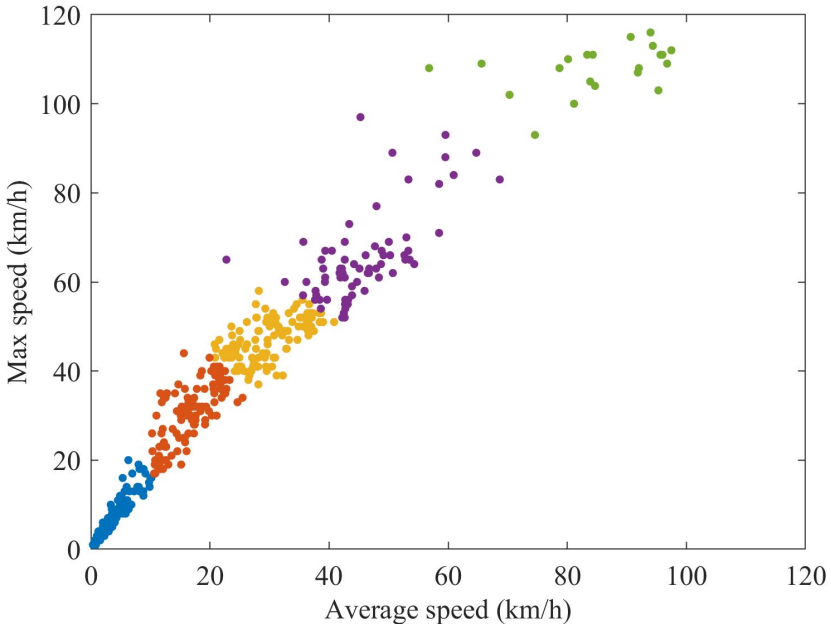


Figure 3.9: Driving scenarios for casual rental vehicles; Data from May. 12, 2022 to Sep. 14, 2022

In Figure 3.12, Figure 3.12a is the creeping driving scenario, Figure 3.12b is the low-speed campus driving scenario, Figure 3.12c is the high-speed campus driving

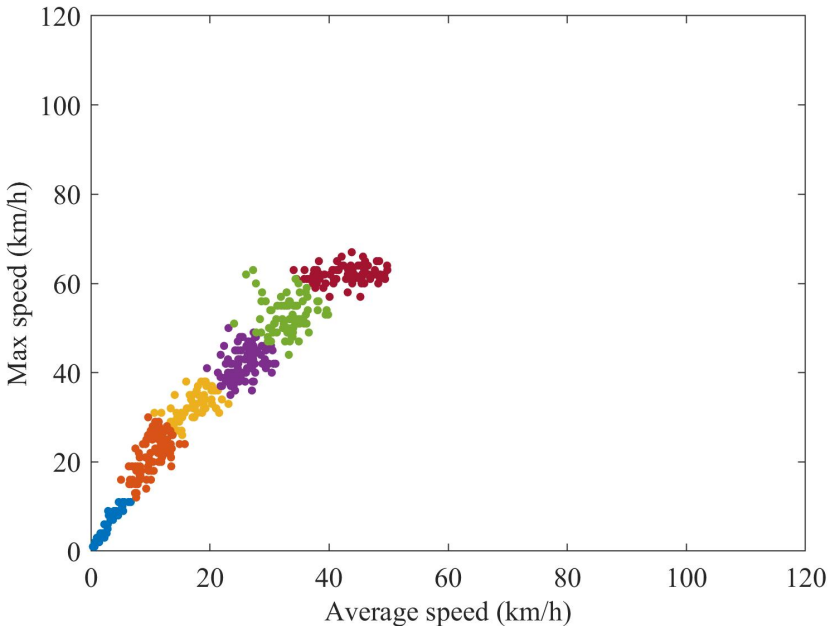


Figure 3.10: Driving scenarios for shuttle minibuses; Data from Apr. 4, 2022 to Apr. 21, 2022

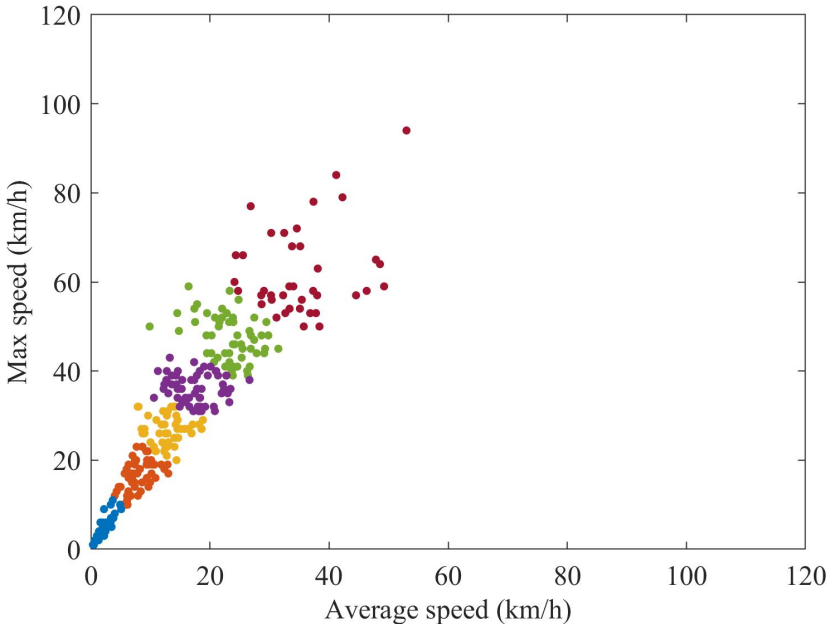
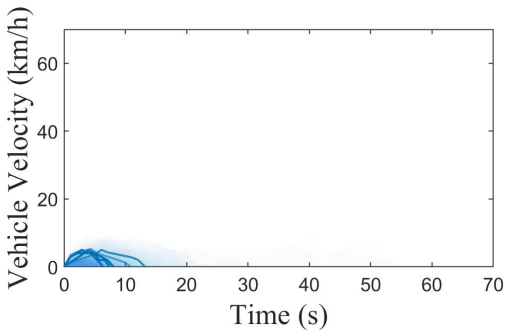


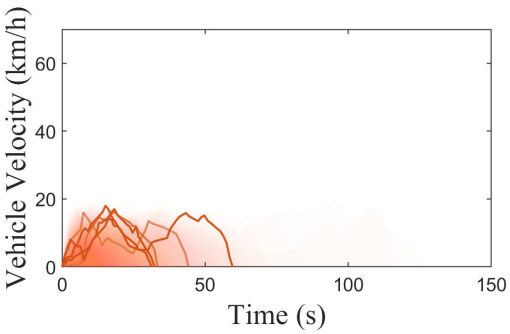
Figure 3.11: Driving scenarios for UAPS; Data from Jun. 21, 2022 to Oct. 9, 2022

scenario, Figure 3.12d is the low-speed resident driving scenario, and Figure 3.12e is the high-speed resident driving scenario.

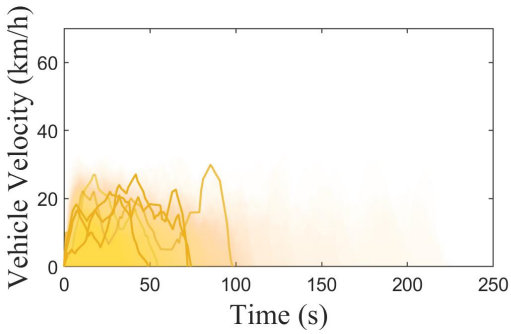
For creeping driving scenarios, the time interval of Microtrip is less than one



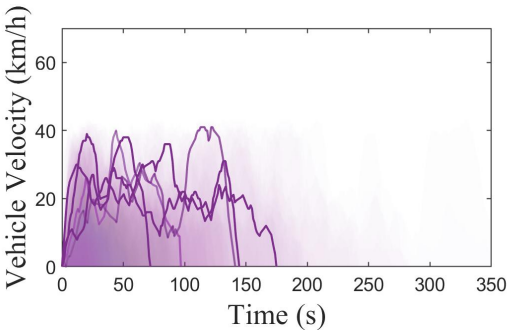
(a) Creeping driving scenario



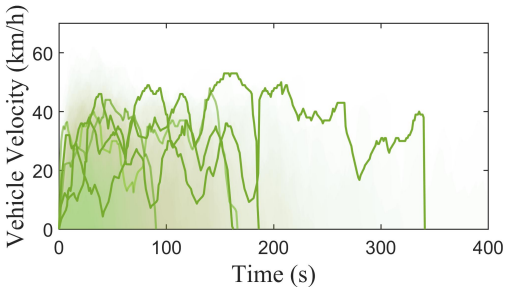
(b) Low-speed campus driving scenario



(c) High-speed campus driving scenario



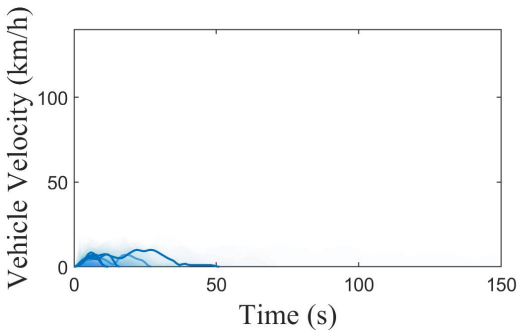
(d) Low-speed resident driving scenario



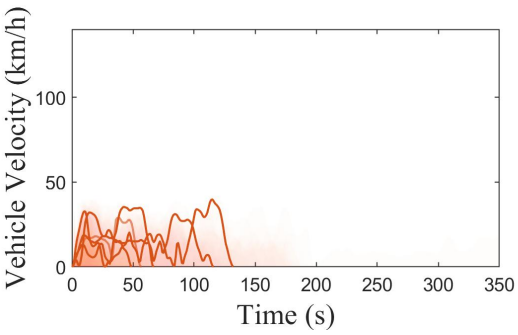
(e) High-speed resident driving scenario

Figure 3.12: Sub-Microtrip database for utility or trade vehicle category ; Data from Aug. 17, 2021 to Jan. 20, 2022

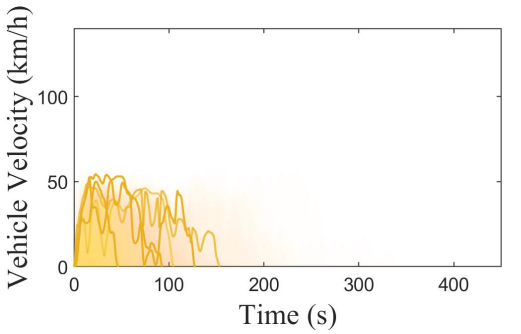
minute, and the maximum driving speed does not exceed 10 km/h. It belongs to the characteristics of driving in crowded areas. The campus driving scenario is a driving scenario where the vehicle is running in an area with people on the campus. It can be seen that the driving speed of the vehicle is within the campus speed limit of 30 km/h. Both low-speed and high-speed resident driving scenarios have longer time intervals. The greater the average speed of Microtrips, the longer the time interval of Microtrips.



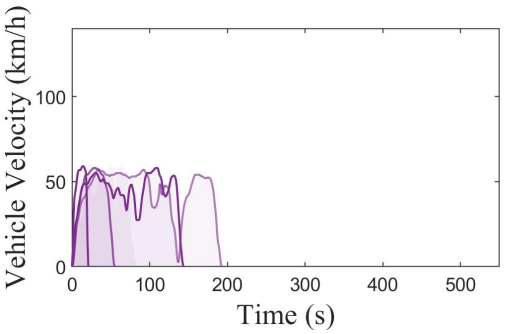
(a) Creeping driving scenario



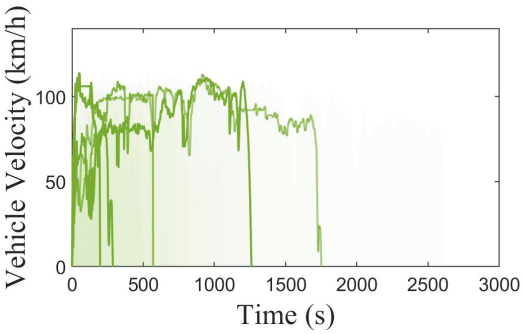
(b) Low-speed resident driving scenario



(c) High-speed resident driving scenario



(d) Urban driving scenarios

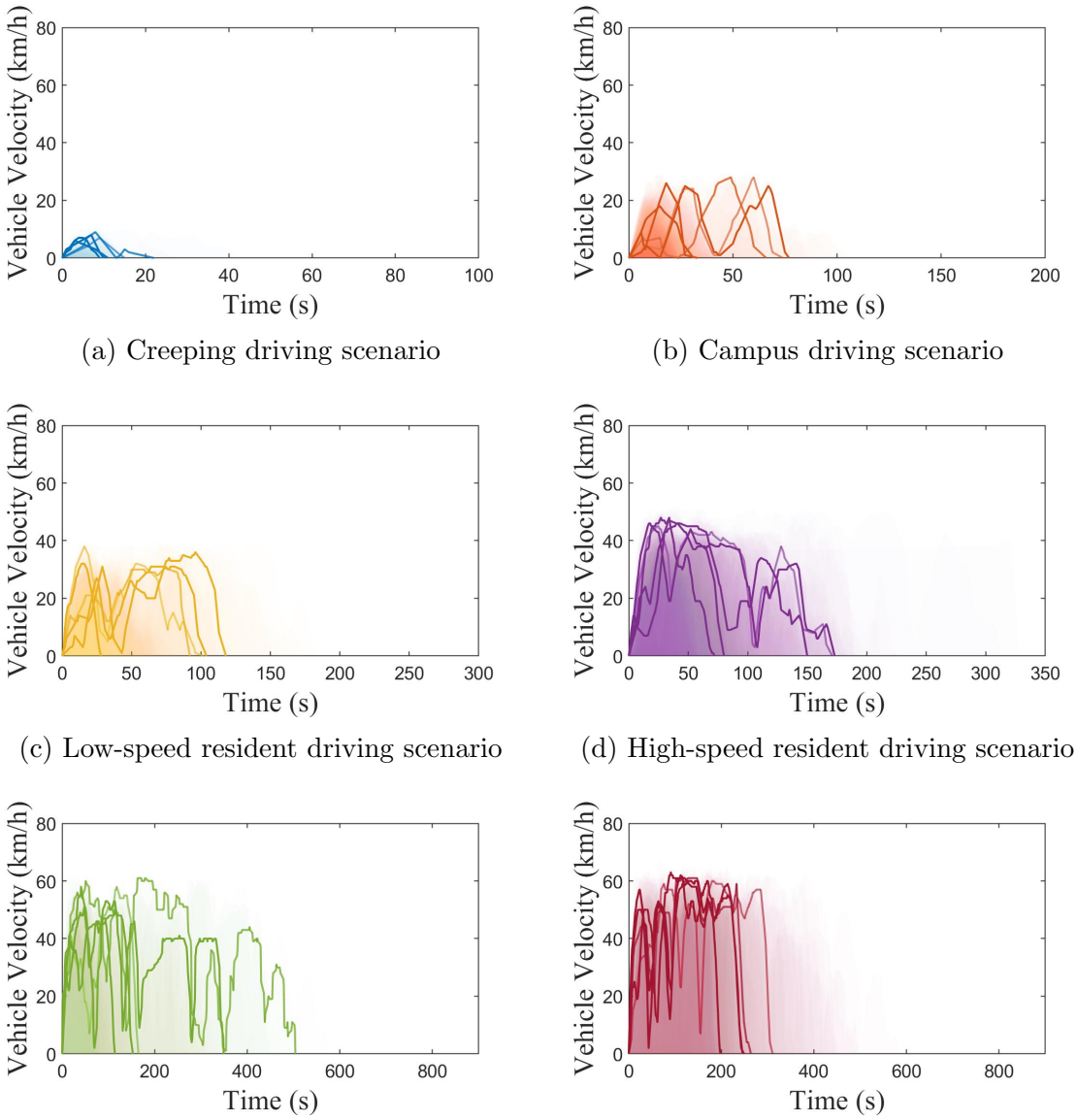


(e) Highway driving scenario

Figure 3.13: Sub-Microtrip database for casual rental vehicle category; Data from May. 12, 2022 to Sep. 14, 2022

In Figure 3.13, Figure 3.13a is the creeping driving scenario, Figure 3.13b is the low-speed resident driving scenario, Figure 3.13c is the high-speed resident driving scenario, Figure 3.13d is the urban driving scenario, and Figure 3.13e is the highway urban driving scenario. The driving scenarios for casual rental vehicles have higher speeds and longer time intervals than utility and trade vehicles. Only casual rental vehicles have highway driving scenarios. Therefore, the assessment metric for casual

rental vehicles have the highest average speed and the lowest number of $Stops/m$, which are 58.3 km/h and 0.4 respectively.



(a) Creeping driving scenario

(b) Campus driving scenario

(c) Low-speed resident driving scenario

(d) High-speed resident driving scenario

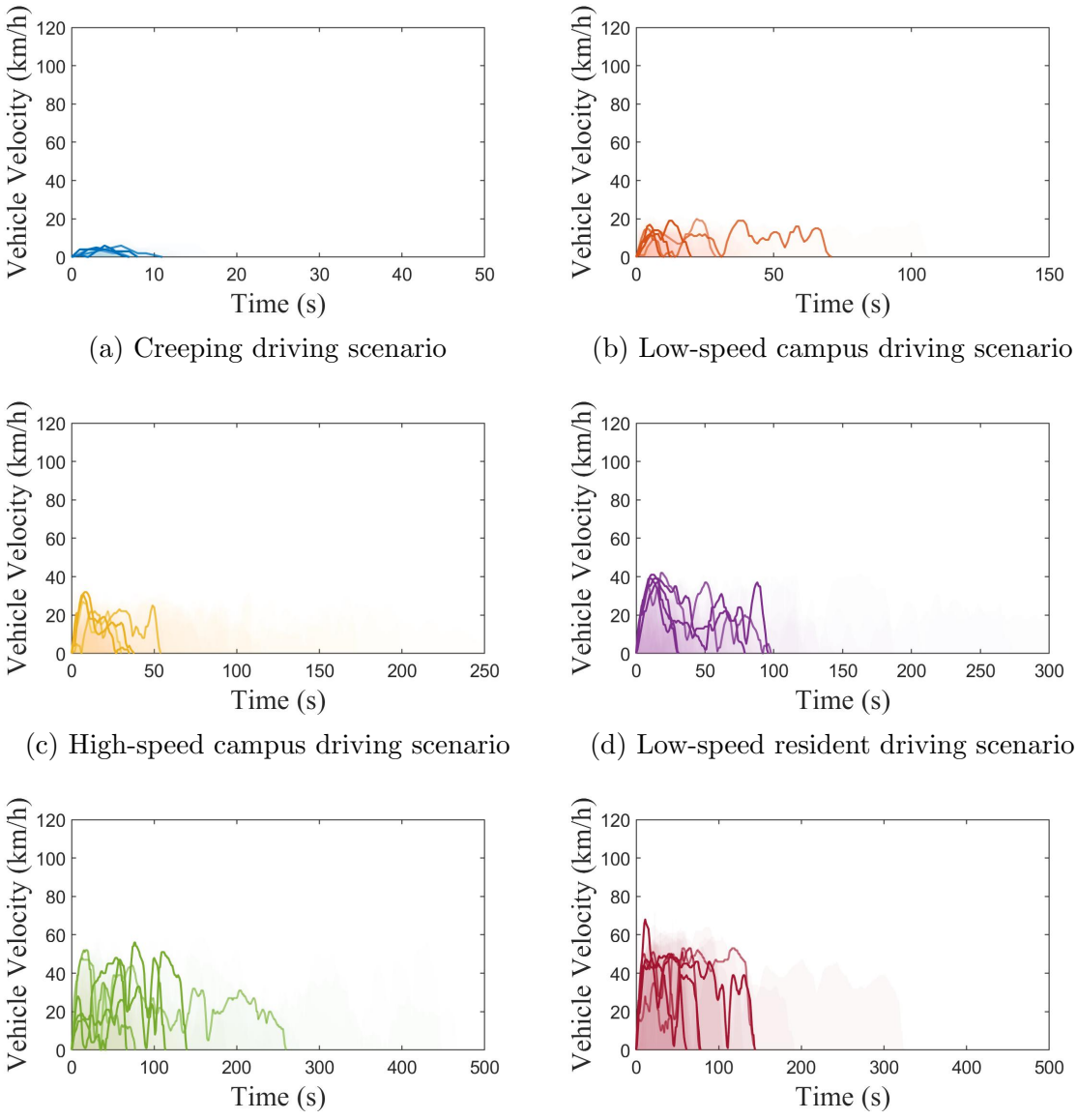
(e) Low-speed urban driving scenario

(f) High-speed urban driving scenario

Figure 3.14: Sub-Microtrip database for shuttle minibus vehicle category; Data from Apr. 4, 2022 to Apr. 21, 2022

In Figure 3.14, Figure 3.14a is the creeping driving scenario, Figure 3.14b is the campus driving scenario, Figure 3.14c is the low-speed resident driving scenario, Figure 3.14d is the high-speed resident driving scenario, Figure 3.14e is the low-speed urban driving scenario, Figure 3.14f is the high-speed urban driving scenario. Com-

pared with casual rental vehicles, shuttle minibuses have a higher speed in both low-speed resident driving scenario and high-speed resident driving scenario. Even in the same driving scenario, different vehicle models and different driving routes can lead to different driving characteristics. This also illustrates the necessity of building driving cycles for different vehicle categories.



(a) Creeping driving scenario

(b) Low-speed campus driving scenario

(c) High-speed campus driving scenario

(d) Low-speed resident driving scenario

(e) High-speed resident driving scenario

(f) Urban driving scenario

Figure 3.15: Sub-Microtrip database for UAPS vehicle category; Data from Jun. 21, 2022 to Oct. 9, 2022

In Figure 3.15, Figure 3.15a is the creeping driving scenario, Figure 3.15b is the

low-speed campus driving scenario, Figure 3.15c is the high-speed campus driving scenario, Figure 3.15d is the low-speed resident driving scenario, Figure 3.15e is the high-speed resident driving scenario, Figure 3.15f is the urban driving scenario. UAPS have the most campus driving scenarios which resulting in a higher number of starts and stops. Morover, because of the speed limit in the campus area, UAPS has the lowest average speed assessment metric value.

3.5 Microtrip Generation Algorithm

Based on the generation method of the sub-Microtrip database from section Section 3.4, a driving cycle is constructed according to the algorithm illustrated in Figure 3.16. Different from the conventional completely random driving cycle construction method, which randomly combines Microtrips and then judges the performance to determine whether the generated driving cycle is available. The algorithm in Figure 3.16 uses the average speed of Microtrips to accelerate the process.

In the process of combining the Microtrips into a driving cycle, because the Microtrip does not include the idle driving state, the driving cycle consisting of Microtrips should not include the idle driving state. Here, VR_{avg} is used, which is the average speed of a driving cycle consisting of Microtrips without any inserted idle driving states. Because the idle driving state is abandoned during the combination, the total expected driving cycle time duration is also reduced, so $Time_{Desired}$ is used to represent the expected driving cycle time without the idle driving state. $Time_{DC}$ in the algorithm is the time duration of the Microtrip combined driving cycle.

As Figure 3.16 shows, firstly, a Microtrip is randomly selected and its VR_{avg} is calculated. If VR_{avg} is equal to Ve_{avg}, then the algorithm randomly selects a Microtrip from the Microtrip database. If VR_{avg} is larger than Ve_{avg}, then the algorithm randomly selects the next driving state from the Microtrip database for which the speed range is higher than the previously selected Microtrip and randomly selects a new Microtrip to connect the previous Microtrip, otherwise, the algorithm randomly

selects the next driving state interval which is lower than the previously selected one and keep other steps unchanged. The process is then continued until the time is greater than the $Time_{Desired}$. Next, the $RMS_{revised}$ of the generated driving cycle is calculated, if $RMS_{revised}$ is greater than 1, the generated driving cycle is regarded as an invalid driving cycle; thus, the driving cycle is deleted and the process is restarted. Based on the above process, the driving cycle except for the idle driving state is obtained.

For the allocation of idle times, the start and the end idle times of the driving cycle are obtained by calculating the average ratio of the start and end idle times to the total time from all recorded data. The driving cycle's idle time between Microtrips is the total idle time minus the start and end idle time divided by the number of Microtrips in the driving cycle. The driving cycle time is the average value of all OBD recorded driving cycle times.

3.6 Driving Cycles for Each Vehicle Category

In this section, driving cycles, values of target parameters for driving cycles, and $RMS_{revised}$ value for different vehicle categories are shown below. Figure 3.17 shows the driving cycle for the utility or trade vehicle category. Table 3.6 shows the driving cycles' target parameter values for the utility or trade vehicle category. As mentioned in Section 3.4, the duration of the driving cycle is the average value of all OBD recorded driving cycle duration time. Therefore, the cycle time for utility and trades vehicles is 1761 seconds. From all OBD recorded driving cycles, the idle driving state time accounted for 33.7% of the total time. Thus, for utility and trade vehicles, the duration of idle driving state time is 594.6 s. The idle time of the driving cycle at the beginning and the end of the total driving cycle time are 28.5 s and 39.2 s, respectively. $RMS_{revised}$ is 0.27, which means the proposed driving cycle in Figure 3.17 can well represent the driving characteristics of the tested fleet vehicles in utility or trade vehicles category.

Table 3.6: The driving cycles' target parameter values for utility or trade vehicle category

Target Parameters	V_{avg} **(km/h)**	Ve_{avg} **(km/h)**	Acc_{avg} **(m/s^2)**	Dec_{avg} **(m/s^2)**	$Stop/km$
Value	17.5	21.2	0.34	-0.34	3.1
Target Parameters	$\%Acc$	$\%Dec$	$\%Cruise$	$\%Idle$	$RMS_{revised}$
Value	25.9	23.5	23.2	27.4	0.27

Figure 3.18 depicts the driving cycle for the casual rental vehicle category. Table 3.7 shows the driving cycles' target parameter values for the casual rental vehicle category. The cycle time for casual rental vehicles is 9136 seconds. From all OBD recorded driving cycles, the idle driving state time accounted for 17.6% of the total time. Thus, the duration of idle driving state time is 1607 s. The idle time of the driving cycle at the beginning and the end of the total driving cycle time are 132 s and 343 s, respectively. It can be seen that the driving cycle for the casual rental vehicle category has the longest driving duration, and it's the only vehicle category which has the highway driving scenario.

Table 3.7: The driving cycles' target parameter values for casual rental vehicle category

Target Parameters	V_{avg} **(km/h)**	Ve_{avg} **(km/h)**	Acc_{avg} **(m/s^2)**	Dec_{avg} **(m/s^2)**	$Stop/km$
Value	51.9	70.1	0.51	-0.51	0.3
Target Parameters	$\%Acc$	$\%Dec$	$\%Cruise$	$\%Idle$	$RMS_{revised}$
Value	14.7	15.0	49.3	17.0	0.87

Figure 3.19 shows the driving cycle for the shuttle minibus category. Table 3.8

shows the driving cycles' characteristics for shuttle minibus category. The cycle time for shuttle minibus is 980 seconds. From all OBD recorded driving cycles, the idle driving state time accounted for 27.4% of the total time. Thus, in this study, the duration of idle driving state time is 242 s. The idle time of the driving cycle at the beginning and the end of the total driving cycle time are 35 s and 13 s, respectively. Compared with utility and trades vehicles, the shuttle minibus' driving cycle has fewer cruise fragments and more acceleration and deceleration. It means that minibus drivers are more aggressive than utility and trade vehicles driver. To relieve aggressive driving for minibus drivers can save fuel and produce less emissions.

Table 3.8: The driving cycles' target parameter values for shuttle minibus category

Target Parameters	V_{avg} **(km/h)**	Ve_{avg} **(km/h)**	Acc_{avg} **(m/s^2)**	Dec_{avg} **(m/s^2)**	$Stop/km$
Value	24.6	29.4	0.41	-0.41	1.7
Target Parameters	$\%Acc$	$\%Dec$	$\%Cruise$	$\%Idle$	$RMS_{revised}$
Value	29.1	31.1	16.0	23.8	0.22

Figure 3.20 shows the driving cycle for the UAPS vehicle category. Table 3.9 shows the driving cycles' characteristics for UAPS vehicle category. The cycle time for UAPS vehicles is 1221 seconds. From all OBD recorded driving cycles, the idle driving state time accounted for 41.1% of the total time. Thus, in this study, the duration of idle driving state time is 502 s. The idle time of the driving cycle at the beginning and the end of the total driving cycle time are 20 s and 15 s, respectively. UAPS vehicles have the most idle fragments which means that UAPS vehicles are more about waiting than patrolling. This is also very understandable, that the campus area has a large flow of people, and it is inconvenient to patrol for a long time.

Table 3.9: The driving cycles' target parameter values for UAPS vehicle category

Target Parameters	V_{avg} **(km/h)**	Ve_{avg} **(km/h)**	Acc_{avg} **(m/s^2)**	Dec_{avg} **(m/s^2)**	$Stop/km$
Value	18.3	20.2	0.66	-0.65	5.1
Target Parameters	$\%Acc$	$\%Dec$	$\%Cruise$	$\%Idle$	$RMS_{revised}$
Value	22.4	23.0	13.6	41.0	0.18

3.7 Conclusion

This chapter elaborates on the algorithms and results for the driving cycles for the UAlberta fleet vehicles. A method based on unsupervised machine learning is used to generate driving scenarios, and a PCA algorithm is used to design the weights of different parameters. Based on the operating characteristics of the vehicles, four driving cycles were developed for the UAlberta fleet vehicles. They are utility and trade vehicles driving cycle, casual rental vehicles driving cycle, shuttle minibuses driving cycle, and UAPS driving cycle. The driving cycle can be used to test the fuel efficiency and emissions of vehicles.

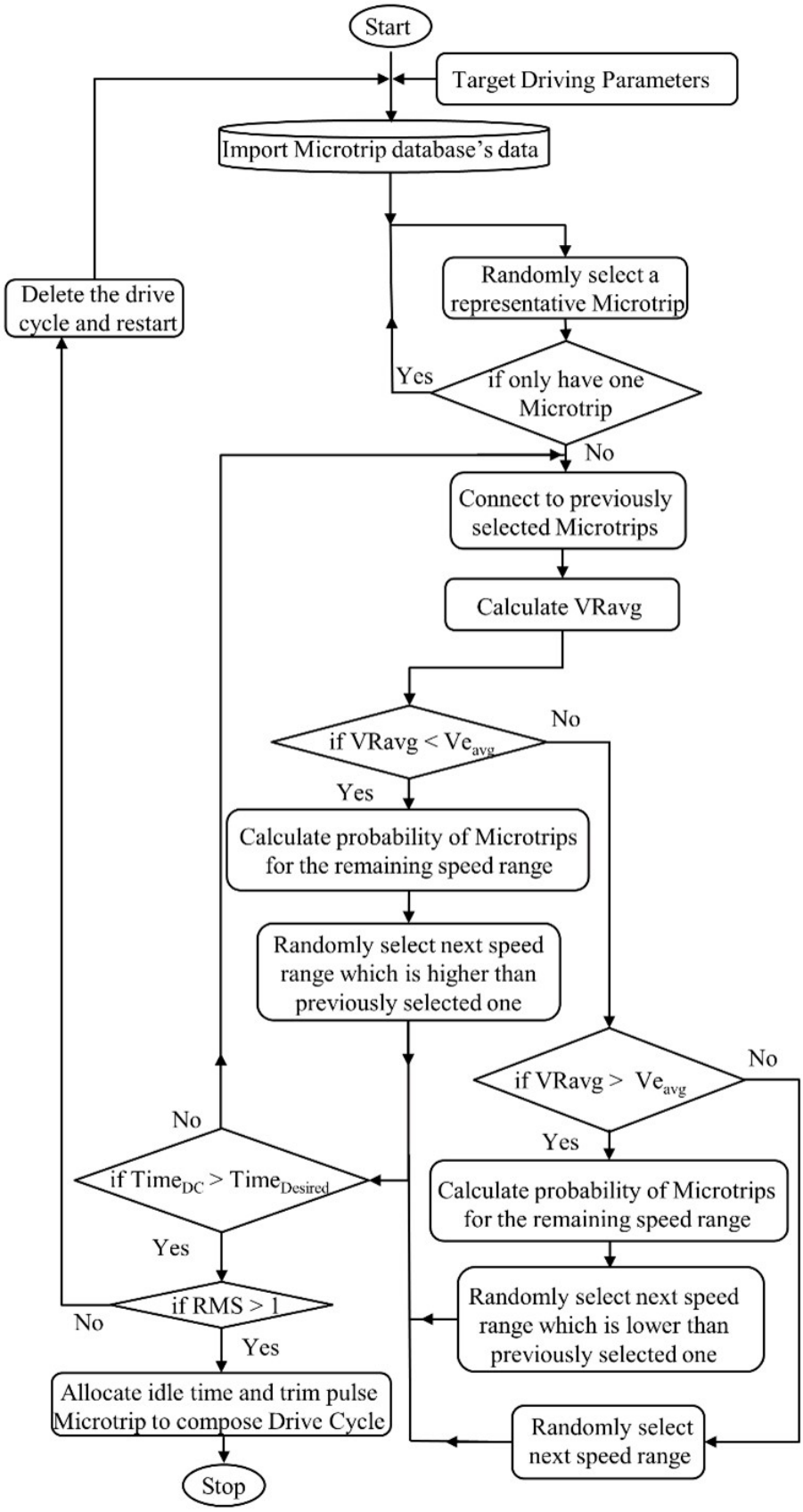


Figure 3.16: Flowchart of the designed algorithm to construct the driving cycle

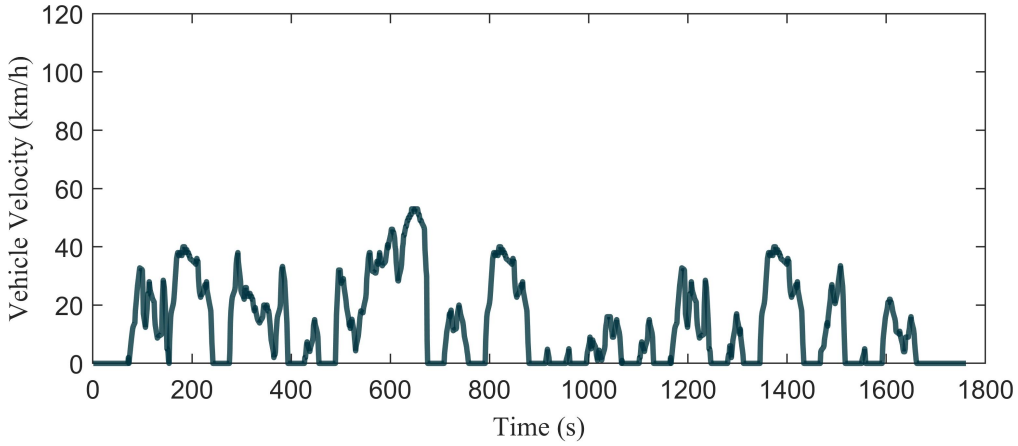


Figure 3.17: The driving cycle for utility or trade vehicles

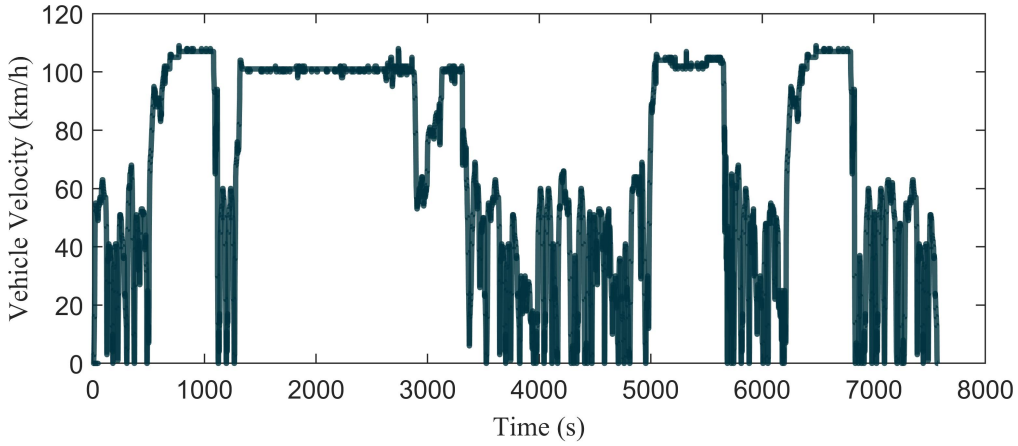


Figure 3.18: The driving cycle for casual rental vehicles

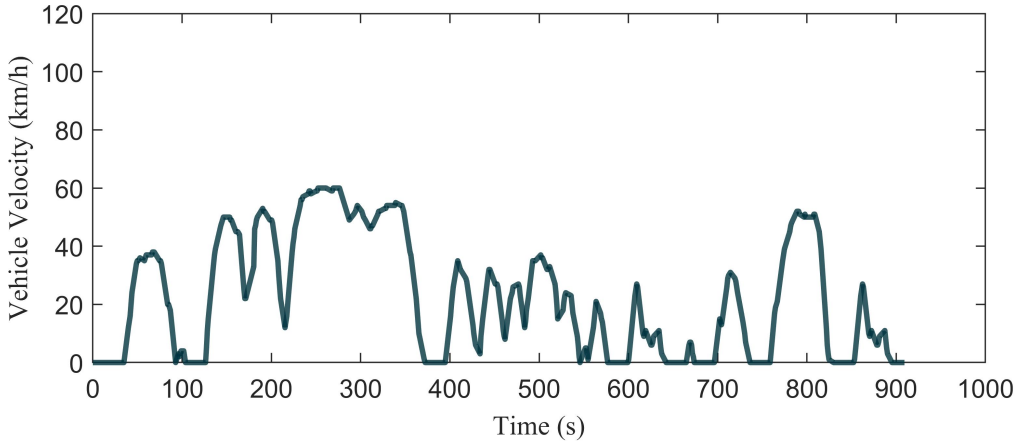


Figure 3.19: The driving cycle for shuttle minibuses

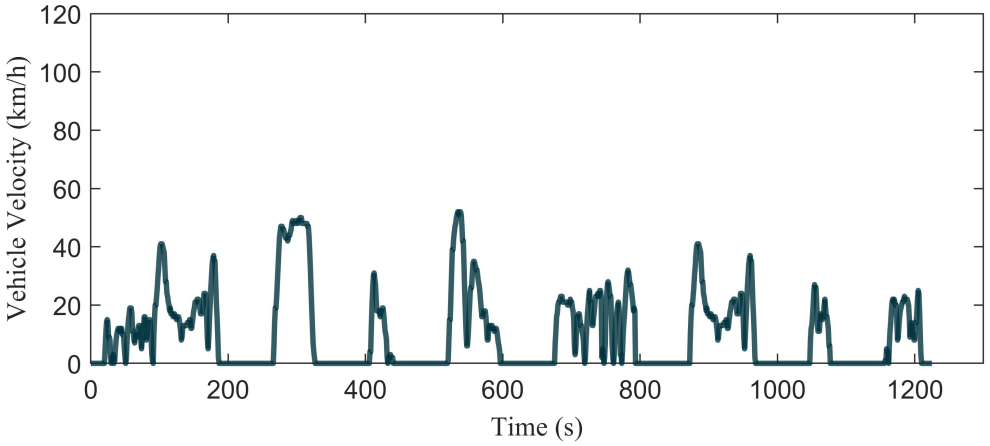


Figure 3.20: The driving cycle for UAPS

Chapter 4

Driver Behavior Assessment and Fuel Consumption

4.1 Introduction

In this chapter, the numerical measurement of driver aggressiveness and the impact of driving aggression on fuel consumption will be detailed. Two types of vehicles, shuttle minibus and sedan vehicles, were used for testing. For shuttle minibuses, two drivers were selected for testing on a fixed route, and for sedan vehicle, one driver are selected to do one normal driving and one aggressive driving. In order to analyze the impact of driver aggressiveness on fuel consumption, driver aggressiveness will be assessed based on different driving scenarios.

4.2 Concept Definition

4.2.1 Driver Behavior

Driver's driving behavior (also referred to as driver behavior for short) refers to the actions, decisions, and attitudes of drivers when driving a vehicle on a road [56]. It covers a wide range of behaviors, including how drivers obey traffic rules, deal with distractions, react to unexpected situations, manage emotions and interact with other road users [57–61]. At the same time, drivers behave differently in different areas and environments [62, 63]. Driver behavior can have a major impact on road safety Unsafe

driving behaviors such as speeding, aggressive driving, and distracted driving increase the risk of accidents and injuries.

The research on driver behavior usually includes three parts as shown in Figure 4.1. Different driving scenarios affect driver behavior in various ways, as drivers may need to adapt different behavior and decisions to the specific situations they encounter on the road.

Some examples of driving scenarios that may affect driver behavior include: (i) Traffic congestion: In heavy traffic, drivers may become more irritable, frustrated, and aggressive, leading to behaviors such as tailgating, weaving in and out of lanes, and honking [64]. (ii) Country Roads: Driving on narrow, winding roads with limited visibility may require more concentration and defensive driving skills, as well as slower speeds [65]. (iii) Urban areas: On busy city streets, drivers may need to navigate intersections, pedestrian crossings, and cycle lanes while dealing with heavy traffic and unpredictable road users [66]. Overall, different driving scenarios affect driver behavior in complex ways, depending on factors such as risk level, stress level, and familiarity with the environment. Understanding these factors can help drivers anticipate potential behavioral challenges and make timely adjustments to ensure safer, more efficient driving.

Some examples of how natural environmental conditions can affect driver behavior include: (i) Weather: Rain can reduce visibility, increase stopping distances, and create slippery roads that make it harder for the driver to maintain control of the vehicle; ice and snow can make roads wet, which can lead to skids and accidents; fog can reduce visibility and create dangerous driving conditions, especially at high speeds [67–69]. (ii) Slope, also known as a ramp, can have a significant impact on driver behavior as they affect vehicle speed, acceleration and handling. When going uphill, the vehicle may need to accelerate to maintain a steady speed, which affects fuel efficiency and engine performance. On a downhill slope, the vehicle may accelerate more easily, which increases the risk of speeding and loss of control [70].

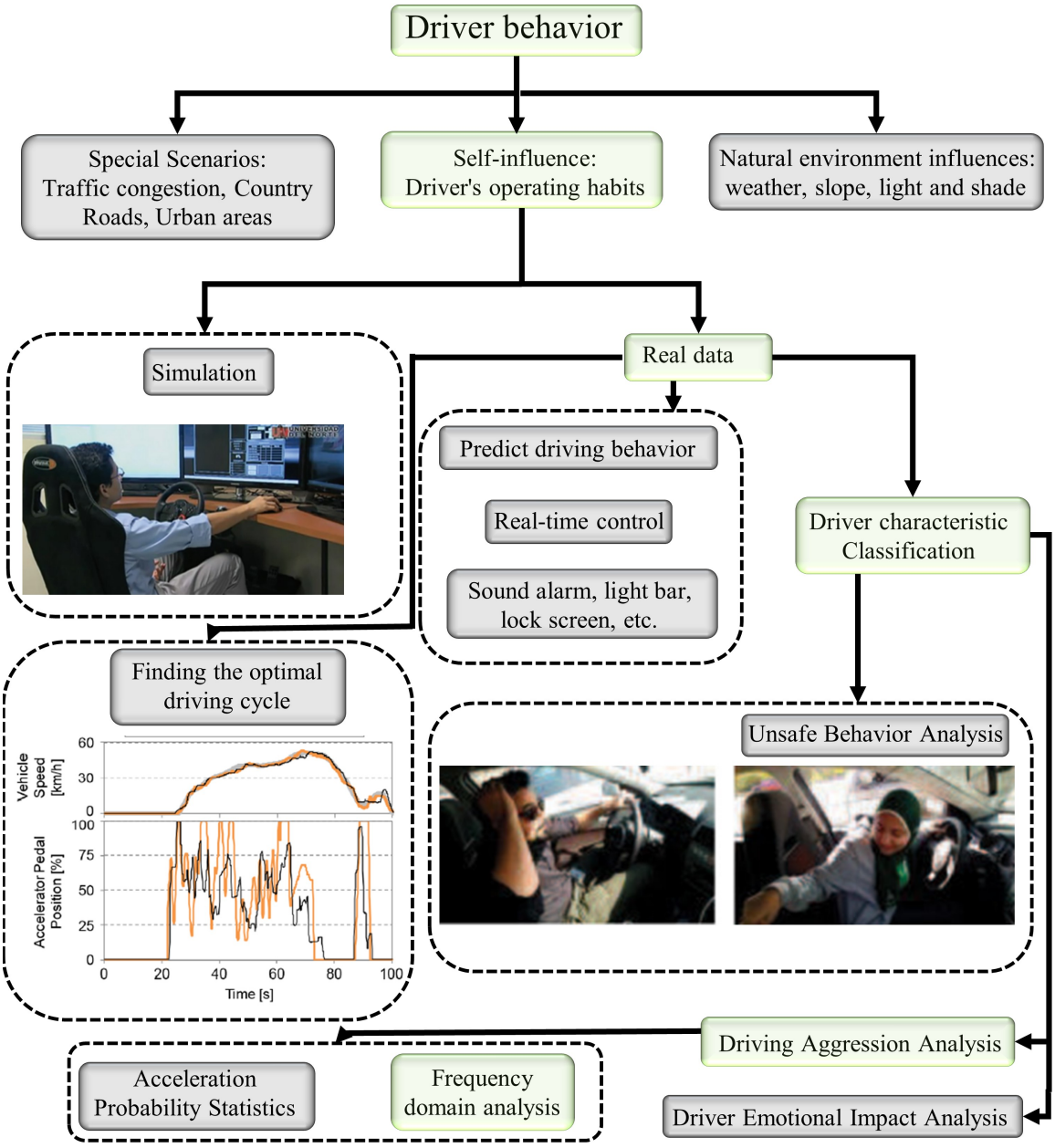


Figure 4.1: Classification of driving behavior analysis. Green blocks show the focus area of this study.

(iii) Light and shadow can create contrast in the visual environment, making objects, signs or other road users more difficult to see [71]. Natural environmental conditions can significantly affect driver behavior, requiring drivers to adjust their behavior and decisions to ensure safe and efficient driving. Knowing these conditions and taking proper precautions can help drivers avoid accidents and reach their destinations safely.

In terms of experimental methods, driver behavior research can be classified as simulation experiments and on-road driving tests. Simulation experiments are generally carried out on the driving simulator, as shown in the simulation module in Figure 4.1. Within the scope of driver self-influence, research directions can be divided into three categories, that is, (i) finding the optimal driving cycle: This kind of research is mainly aimed at vehicles with fixed driving routes and road conditions [72], (ii) driving behavior prediction: This type of research is mainly used for real-time control of the driver to ensure that the driver is in a safe driving state. Real-time control can also be used to guide drivers to drive economically. Common forms of feedback include sound alarms, light bars, and screen lock [73], and (iii) driving characteristic classification.

The research on driving characteristic classification can be further divided into three categories, including (i) driver emotion impact analysis: By understanding how different emotions affect driver behavior, traffic safety organizations and law enforcement agencies can develop targeted interventions and education programs to promote safe and responsible driving [74], (ii) unsafe behavior analysis: Unsafe behavior analysis is the process of identifying and studying the behaviors and actions that contribute to unsafe driving practices. By understanding these unsafe behaviors, traffic safety organizations and law enforcement agencies can develop targeted interventions and education programs to promote safe and responsible driving [75], and (iii) driving aggressiveness analysis [**liu2023identification**]. In Figure 4.1, the focus area of this thesis is highlighted by green blocks.

4.2.2 Driver Aggressiveness

This study focuses on driving aggressiveness (DA). DA refers to a pattern of driving behavior characterized by hostility, impatience, and a tendency to take risks while driving [76]. The performance in the data is that compared with the normal driving mode, the aggressive driving speed is faster, and the acceleration and deceleration

are more frequent and the range is larger. Aggressive driving increases the risk of accidents and has serious consequences for aggressive drivers and other road users.

Traditional DA research methods focus on the statistics of acceleration [77]. However, statistical methods will ignore the detailed performance of driving behavior. For instance, consider the shuttle minibus category data as an example. There are two minibuses in the shuttle minibus category. The two minibuses are driven by two fixed drivers, called Driver A and Driver B. The driving route of the minibus is also fixed. If one counts the acceleration data of driver A and B and makes a violin figure as shown in Figure 4.2; it can be seen from the Figure 4.2 that the acceleration values of Driver A and B are all concentrated around 0. However, it is difficult to see from the Figure 4.2 that which vehicle has more driving acceleration. This will mislead the reader that Driver A and B have the same driving aggressiveness. However, the driving fuel consumption of Driver A and B is quite different. It can be seen from the Figure 4.3 that the lowest fuel consumption value of driver B is greater than the highest fuel consumption value of driver A. If one analyzes DA from a driving scenarios perspective such as creeping driving scenarios and so on, the error of DA calculated by traditional statistical method is even bigger.

Frequency domain analysis transforms the data into the frequency domain. In the frequency domain, the characteristics of the data are displayed sequentially from low frequency to high frequency. The information of the signal is not lost during conversion [78]. Therefore, this research adopts the frequency domain analysis method to analyze the DA under different driving situations.

4.3 Mathematic model

4.3.1 Frequency Domain Analysis

In this study, the driver's DA and the fuel consumption difference from different driving scenarios was analyzed. As introduced in Section 3.4, this thesis uses Microtrip

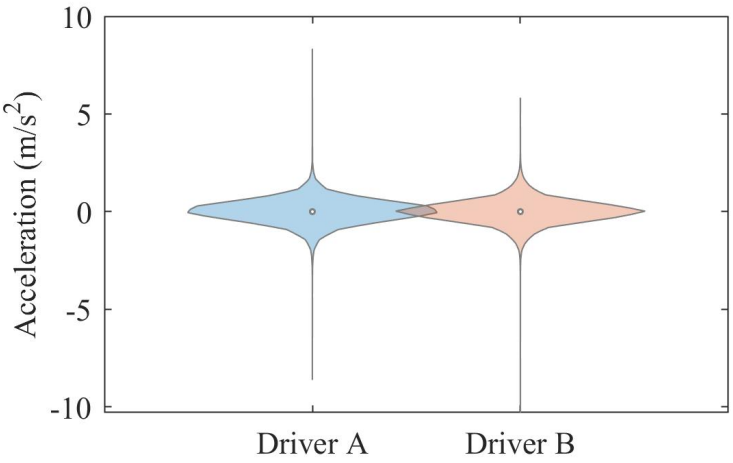


Figure 4.2: The acceleration violin statistics of the drivers of the shuttle minibuses (Use data of shuttle minibus from Apr. 4, 2022 to Apr. 21, 2022)

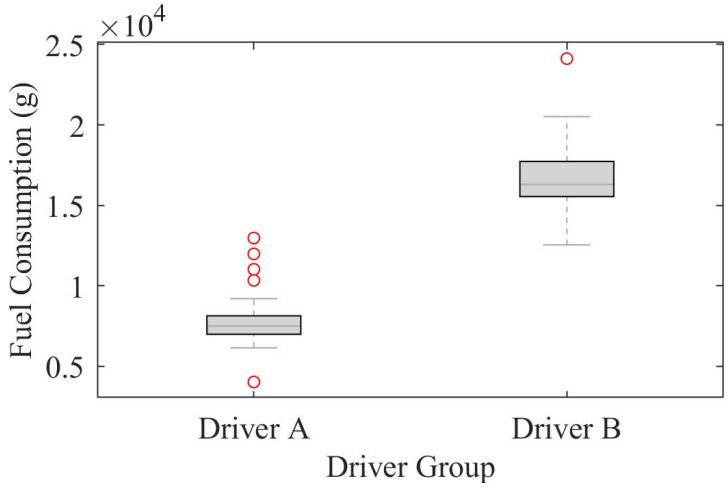


Figure 4.3: Differences in fuel consumption by shuttle minibus drivers (Use data of shuttle minibus from Apr. 4, 2022 to Apr. 21, 2022)

to identify driving scenarios. Microtrip is a signal sequence with a limited time and a certain frequency (1Hz). Microtrip signal sequence can be regarded as a discretized version of continuous time-driving velocity. Discrete Fourier Transform (DFT) is used to analyze Microtrip signal in the frequency domain. DFT maps length-N signals into a set of N discrete frequency components. The DFT representation is shown in Equation (4.1)

$$X(k) = \sum_{n=0}^{N-1} x(n) e^{-j\frac{2\pi}{N}},\ k = 0, \ldots, N-1 \tag{4.1}$$

In Equation (4.1), $x(n)$ is the discrete signals in the time domain, $X(k)$ is the discrete signals in the frequence domain, and N is the original signal length

Based on DFT and Parseval theorem, as shown in Equation (4.2), the energy of the time-domain signal sequence and frequency-domain signal sequence is conserved [79].

$$\sum_{n=0}^{N-1} |x(n)|^2 = \frac{1}{N} \sum_{n=0}^{N-1} |X(k)|^2 \tag{4.2}$$

The content $\frac{1}{N} |X(k)|^2$ on the right side of Equation (4.2) is periodogram [80]. The value under the periodogram area is exactly the variance of the signal in the time domain. The variance of the time-domain signal is the fluctuation of driving speed. Therefore, the DA can be reflected by the $\frac{1}{N} |X(k)|^2$ value of the frequency domain signal, and this description is quantitative.

When performing frequency domain analysis, it is first necessary to subtract the average velocity of the Microtrip signal, because the zero-mean signal occupies too much energy in the frequency domain. At the same time, the zero-mean signal cannot reflect the driver's driving performance as zero-mean signal is the driving state that vehicle is idling. The analysis of DA in different speed ranges can be analyzed in different driving situations.

Moreover, to analyze the Microtrip signal, it is necessary to obtain the second-order derivative to obtain the jerk trace. Because if the velocity signal is directly subjected to frequency analysis, the low-speed component signal accounts for too much energy in the frequency domain, which desensitize the DA model [78]. The low-frequency (LF) components of the velocity signal in the frequency domain are usually caused by the driving environment, while the high-frequency (HF) speed components are

caused by the driving behavior, and the cutoff frequency between LF and HF is the frequency at which high frequency accounts for 80% of the total energy. [81].

Thus, the numerical expression of driver aggressiveness can be shown by Equation (4.3).

$$\text{DA}= \frac{\text{HF}}{\text{LF+HF}} \tag{4.3}$$

Where, HF represents high-frequency energy, and LF represents low-frequency energy. Thus, the numerical range of DA is between 0 and 1, and the closer it is to 0, the smoother the driver is driving, and the closer it is to 1, the more aggressively the driver is driving. The flow chart of DA research process on the Microtrip database is shown in Figure 4.4.

Figure 4.4 shows the flow of frequency domain analysis. Firstly, as shown in Figure 4.4(a), the Microtrip database is divided into sub-Microtrip databases with different driving scenarios using a clustering algorithm (see Section 3.4 for details). Figure 4.4(b) shows the shuttle minibus sub-Microtrip database. Each figure in Figure 4.4(b) is composed of multiple Microtrip lines, and it is determined as sub-Microtrip database (one sub-Microtrip is zoomed for demonstration). In each sub-Microtrip database, one Microtrip line is bolded for demonstration. As shown in Figure 4.4(c), each Microtrip under different driving scenarios was analyzed in the frequency domain. Using Equation (4.2) to calculate the DA value of each Microtrip, the DA of different drivers in different driving situations can be obtained. Figure 4.4(d) shows the specific DA calculation process of Microtrip circled by the red box in Figure 4.4(c). One can add periodogram value with frequencies lower than 0.1 HZ to calculate LF, and add periodogram value with frequencies higher than 0.1 HZ to calculate HF. By appling the Equation (4.3), the DA of Microtrip is determined.

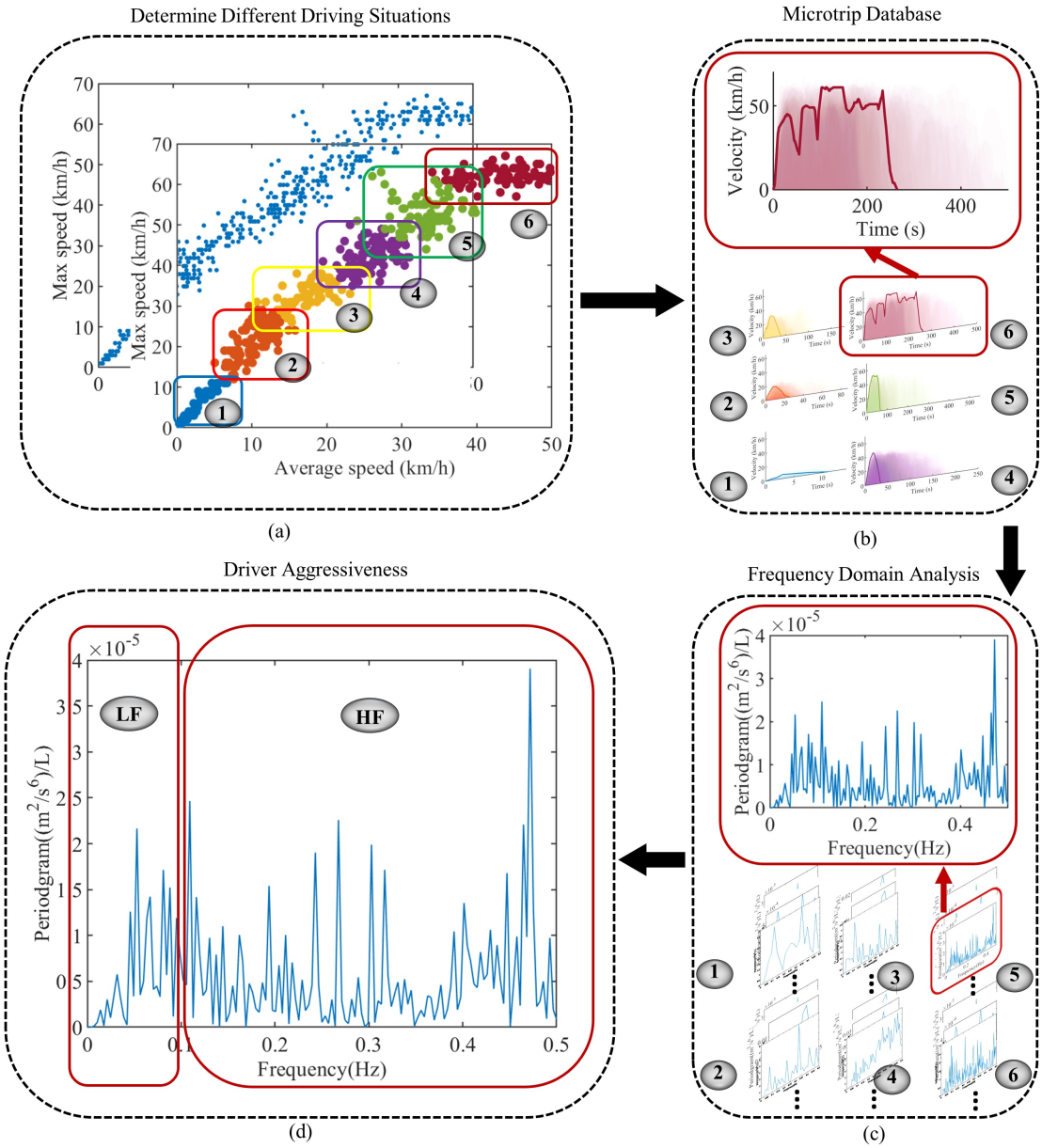


Figure 4.4: The flow chart of frequency domain analysis on driver aggressiveness in different driving scenarios (Use data of shuttle minibus from Apr. 4, 2022 to Apr. 21, 2022)

4.3.2 Fuel Consumption Rate

The OBD data loggers don't provide fuel consumption rate (FCR) data directly. Therefore, engine data including Intake Mass Air flow (MAF), the air-to-fuel ratio at the stoichiometric level (AFR_{stoich}), and λ which is the ratio of the actual air/fuel ratio (AFR) to AFR_{stoich} are used to estimate vehicle fuel consumption rate [82].

$$\mathrm{FCR\,(t)} = \frac{\mathrm{MAF\,(t)}}{\lambda(t) \times \mathrm{AFR_{stoich}}} \tag{4.4}$$

In this study, the short and long-term fuel trims from OBD data are used to calculate the air-fuel equivalence ratio accurately for the tested vehicle with spark ignition (SI) engine [83].

$$\lambda\,(\mathrm{t}) = \frac{1}{\left(1 + \frac{Short-term\ fuel\ trim}{100}\right) \times \left(1 + \frac{Long-term\ fuel\ trim}{100}\right)} \tag{4.5}$$

By using Equation (4.4), FCR data for a driving cycle is obtained. To properly assess the driving behavior and its link to FCR, all driving data from each driver is used to create Microtrips that from a driving cycle. Vehicle data from each driver is divided into small pieces to generate Microtrips. At the same time FCR is also divided into pieces following the same segmentation method as Microtrip. Therefore, each Microtrip contains information including, velocity, time, and actual FCR data.

Total fuel consumption (FC) is calculated by

$$FC = \sum FCR(t) \cdot \Delta t \tag{4.6}$$

The data recorded by OBD data loggers are used to obtain the relatively accurate FC of the experimental driving cycle using Equation (4.4) and Equation (4.6).

Figure 4.5 uses the shuttle minibus OBD record data on April 5, 2022 to show the instantaneous fuel consumption and vehicle speed.

4.4 Driver Aggressiveness in Different Driving Scenarios

4.4.1 Shuttle Minibuses' Driving Aggressiveness

According to the methods in Section 2.3.3, Section 3.4 and Section 4.3, the driving behavior of the University of Alberta two shuttle minibuses drivers are analyzed.

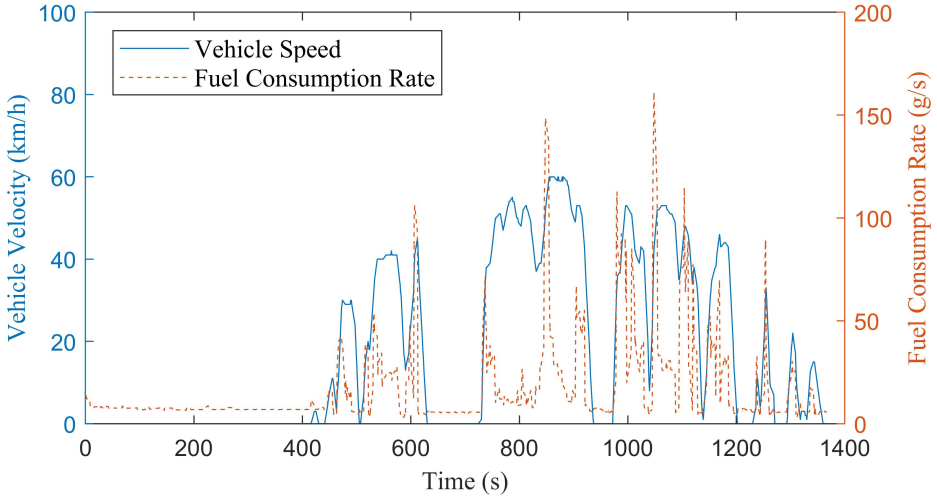


Figure 4.5: Time series of vehicle speed and recorded fuel consumption (Data source: April 5, 2022 shuttle minibus; ID: 0438)

Figure 4.6 shows the driving aggressiveness of drivers A and B in different driving scenarios. In Figure 4.6, the x-axis represents the magnitude of the DA value. The most aggressive driving occurs when the DA is close to 1, and the smoother the driving occurs when the DA is close to 0. Each sub-Microtrip database on the y-axis represents different driving situations.

For each driving scenario, each point in Figure 4.6 represents the DA value of the Microtrip under that driving scenario. The area plot represents the probability density of Microtrip's DA distribution under each scenario. The boxplot in Figure 4.6 shows the upper bound, upper quartile, median, lower quartile, and lower bound of the data.

As shown in Figure 4.6, Driver B is more aggressive than Driver A in all scenarios since average DA values of Driver B are higher by 7% to 45% compared those by the Driver A. Figure 4.6 shows that for shuttle minibus drivers, the driver's DA gradually increases with the increase of driving speed. The DA value range of Driver A in each driving scenario is relatively concentrated, and the difference between the upper and lower quartiles of DA ranged from 0.09 to 0.12. This also shows that for Driver A, the aggressive change in driving is mainly affected by the driving speed. When the

driving speed is higher, the driver usually has to step on the pedal harder to brake to a safe speed. For the Driver B, the difference between the upper and lower quartiles of DA ranged from 0.07 to 0.16 in different driving scenarios. Moreover, the DA value of Driver B under the creeping driving scenario and campus driving scenario has a large change span. The DA upper and lower bounds value ranged from 0.29 to 0.80 and 0.53 to 0.94 respectively. Therefore, in order to improve the driving performance of Driver B, it is first necessary to improve his tendency of driving aggression at low speed.

4.4.2 Ford Escape PHEV Vehicle's Driving Aggressiveness

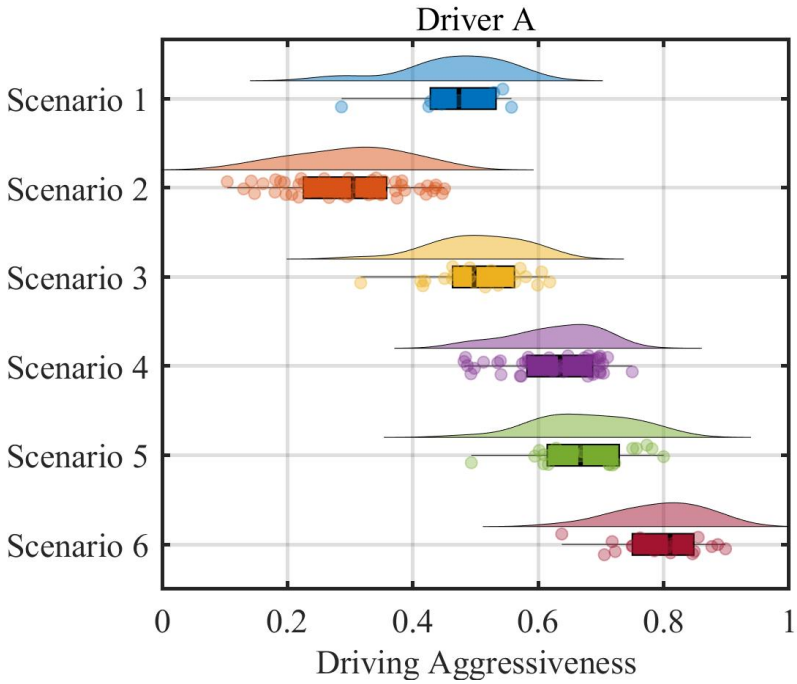
Result in Section 4.4.1 were from the analysis for two different drivers. This section aims to analyze driving of a same driver but for different days with different driving state of vehicle driving.

Driving scenarios for Ford Escape PHEV

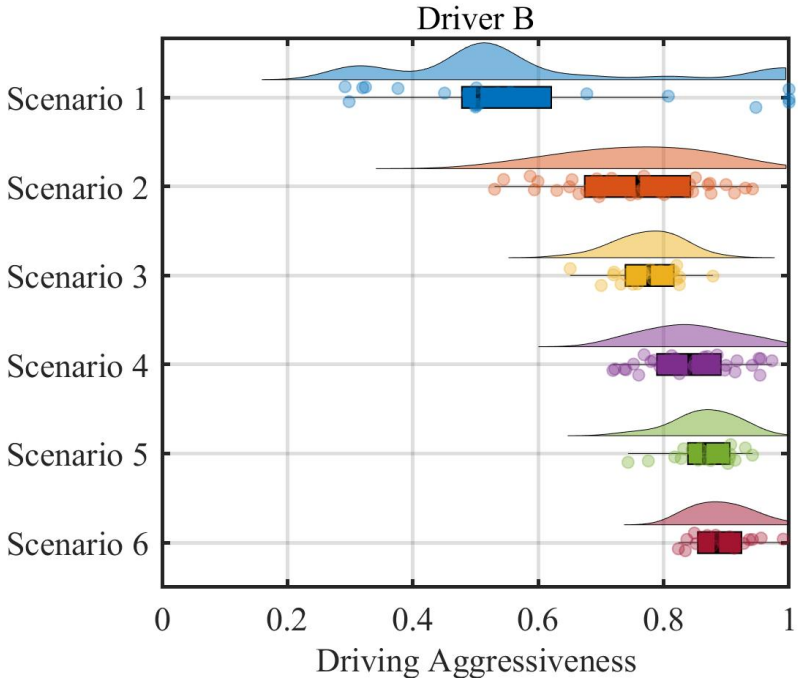
The Microtrip database and the driving scenarios of the Ford Escape PHEV can be obtained by using the algorithm in Section 3.3. The result is shown in Figure 4.7a and Figure 4.7b. The reason of why choice max speed and average speed as label is explained in Section 3.2.1. In Figure 4.7c, $Loss$ and $\Delta Loss$ is defined in Section 3.4. It can be seen from the Figure 4.7a and Figure 4.7b that the driving scenarios of the test vehicle are divided into five categories. They are creeping driving scenario, low-speed residential driving scenario, high-speed residential driving scenario, low-speed urban driving scenario, and high-speed urban driving scenario. Compared with shuttle minibus Microtrip database as shown in Figure 3.10, the driving scenarios of both Sedan vehicle and shuttle minibuses are very close.

Driving aggressive for Ford Escape PHEV

Figure 4.8 shows the driver aggressiveness for the Ford Escape PHEV test vehicle. Because the driver did cold climate emission test, one normal driving state test and

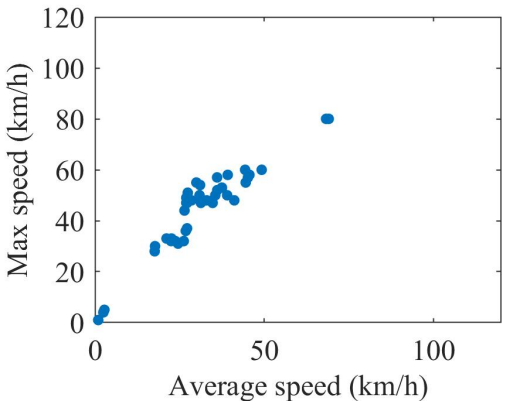


(a) Driving aggressiveness for drivers A

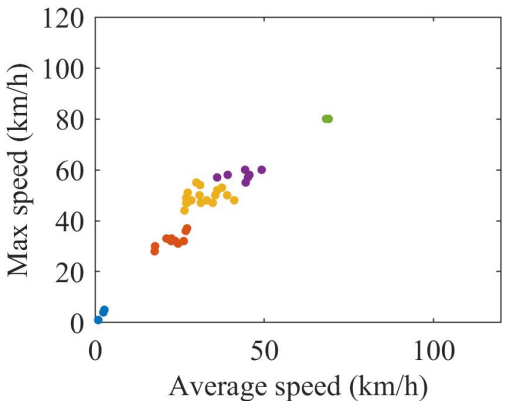


(b) Driving aggressiveness for drivers B

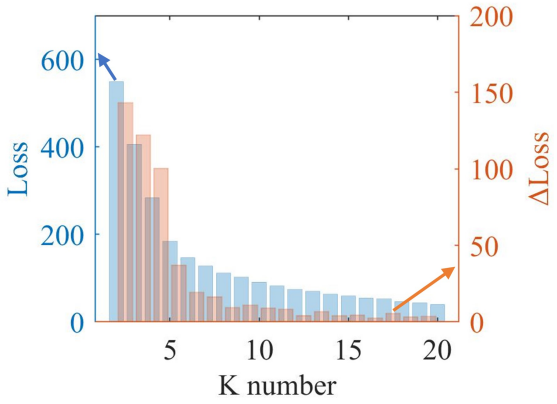
Figure 4.6: Driving aggressiveness of shuttle minibus drivers A and B in different driving scenarios (Scenario 1-6 are defined in Figure 3.10; Use data of shuttle minibus from Apr. 4, 2022 to Apr. 21, 2022)



(a) Ford Escape PHEV's Microtrip database



(b) Driving scenarios for Ford Escape PHEV



(c) *Loss* and $\Delta loss$ for Ford Escape PHEV
(*Loss* and $\Delta Loss$ is defined in Section 3.4)

Figure 4.7: Microtrip database and sub-Microtrip database for Ford Escape PHEV (Use data of Ford Escape PHEV from Mar. 26, 2022 to Mar. 27, 2022)

one aggressive driving state test had been done in two days with the same driving route. By observing the distribution of DA in different scenarios, it can be determined which day of driving is more aggressive. It can be seen from Figure 4.8 that in 2022/03/27, the driver had creeping driving conditions, but in 2022/03/26, the driver did not have creeping driving conditions. Therefore, Figure 4.8a indicates the driver driving in a normal driving state and Figure 4.8b indicates the driver driving in an aggressive driving state.

From Figure 4.8, one can find that regardless of the driving state, the driver's DA distribution is relatively compact in each driving scenarios. For both normal

driving state and aggressive driving state, the biggest difference between the upper and lower quartiles of DA is 0.02, which indicate that driving aggressiveness is mainly determined by the driver behavior. In addition, as the driving speed increases, the value of DA becomes more concentrated. In low-speed residential driving scenario, the biggest difference between the upper and lower bounds of DA are 0.016 and 0.025 respectively, and in low-speed urban driving scenario, the biggest difference between the upper and lower bounds of DA are 0.003 and 0.005 respectively. The scattered DA values at low speeds indicate that driving aggressiveness is more affected by the driving environment in low speeds driving state compared with high speed driving state. Therefore, when training drivers, focus on high-speed driving scenarios first.

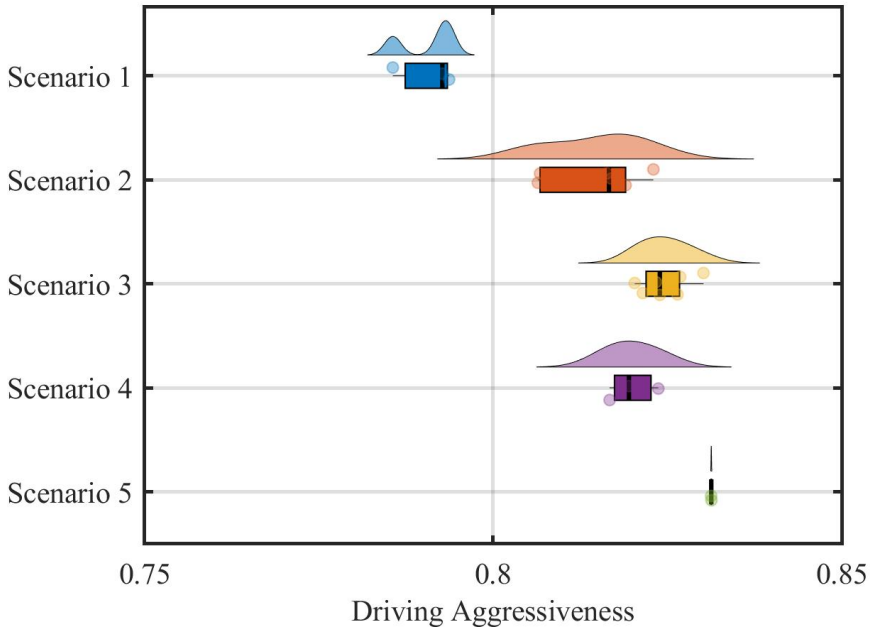
4.5 Fuel Consumption Difference Caused by Driver Aggressiveness

In this section, the effect of DA on vehicle fuel consumption is studied. Because the vehicle model and the driving routes used in this thesis for driver behavior studies are consistent, the fuel consumption difference are mainly caused by DA behavior. By comparing fuel consumption in different driving scenarios, the impact of DA on fuel consumption can be determined. A more precise analysis should take into account the distribution of Microtrip in different driving scenarios. Therefor, a revised average fuel consumption (RAFC) metric is proposed to better reflect the impact of DA on fuel consumption.

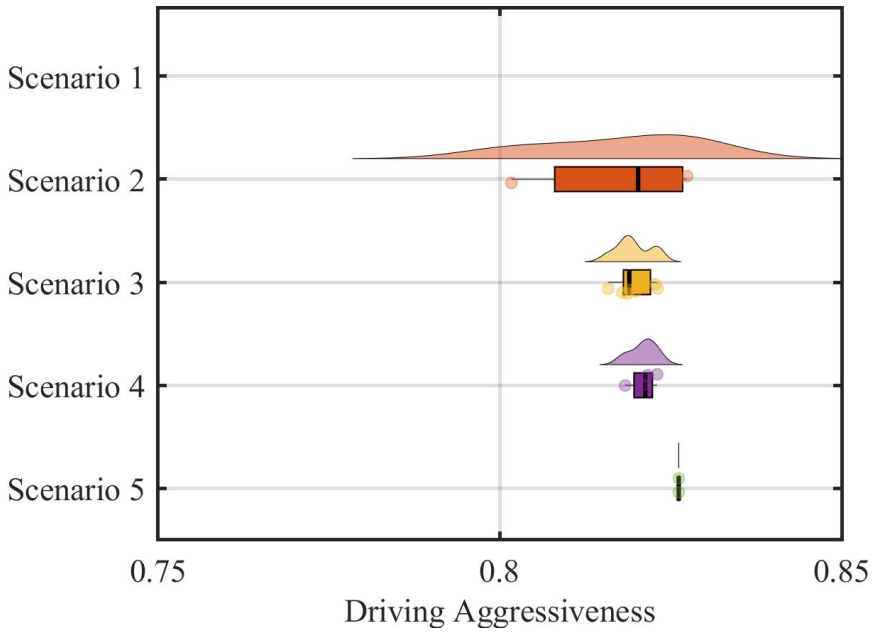
4.5.1 Shuttle Minibuses' Fuel Consumption Caused by Driving Aggressiveness

Figure 4.9 shows the correlation coefficient value for shuttle minibus driver A in different driving scenarios. The result shows that DA and fuel consumption are highly correlated in different driving scenarios.

Figure 4.10 shows the difference in fuel consumption (FC) of drivers A and B under



(a) Driving aggressiveness for normal driving state



(b) Driving aggressiveness for aggressive driving state

Figure 4.8: Driving aggressiveness of Ford Escape PHEV in different driving scenarios (Scenario 1-5 are defined in Figure 4.7b; Use data of Ford Escape PHEV from Mar. 26, 2022 to Mar. 27, 2022)

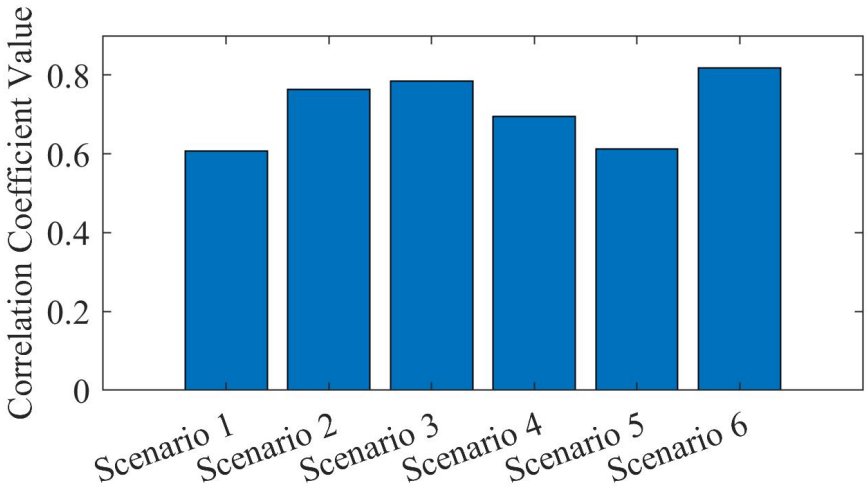


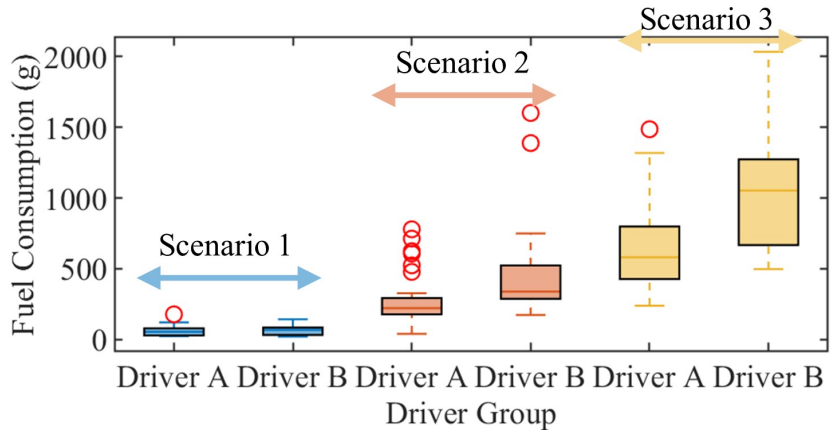
Figure 4.9: The correlation coefficient value for the link between DA and vehicle fuel consumption for the shuttle minibus driver A in different driving scenarios

different driving scenarios. In scenario 2, the average DA value for the driver A and B are 0.31 and 0.76 (0.45 difference), and the difference in average FC value is 196 g. However, in scenario 6, the average DA value between the driver A and B is 0.82 and 0.89 (0.07 difference), but the difference in average FC value is 2904 g. This shows that in the same scenarios, the higher the DA difference, the larger the FC difference.

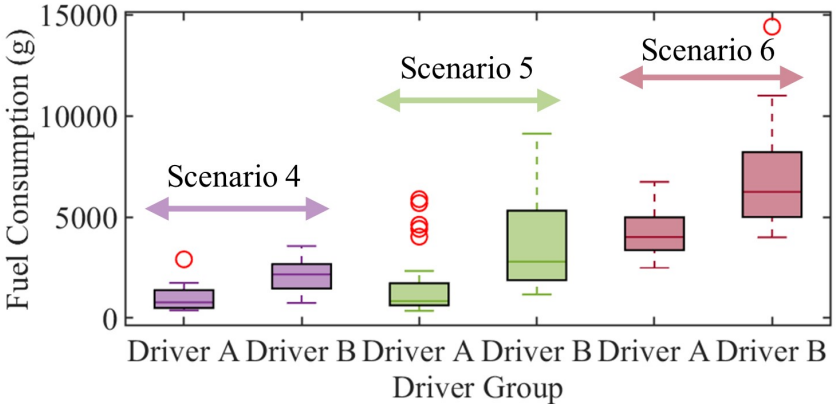
However, Figure 4.10 cannot properly show the impact of the occurrence probability of DA in different driving scenarios on fuel consumption. The occurrence probability is the value that data number in one driving scenario divide by total data number. In Figure 4.6, the density of data points in different driving scenarios is different. Therefore, Figure 4.11 is used to show the probability of DA occurrence in different driving scenarios. For minibus drivers A and B, the DA is more likely to occur in driving scenarios 2 and 4.

Based on the occurrence probability of DA (OPDA) in different driving scenarios, a revised average-fuel consumption (RAFC) formulation is proposed to better describe the impact of DA on fuel consumption of vehicles. The RAFC equation is:

$$RAFC = OPDA \cdot AFC \tag{4.7}$$



(a) Fuel consumption for driver A



(b) Fuel consumption for driver B

Figure 4.10: Fuel consumption for shuttle minibus two drivers in different driving scenarios (Scenario 1-6 are defined in Figure 3.10; Use data of shuttle minibus from Apr. 4, 2022 to Apr. 21, 2022)

Where AFC is the average fuel consumption in driving scenarios. The calculated value is shown in Figure 4.12. For Driver B, one can find that Driver B consume more fuel in scenario 5 than scenario 4 (the average fuel consumption is 2202 g and 2716 g respectively) in Figure 4.10b. The results can mislead people into thinking that driving scenario 5 is more important. But if the OPDA is considered, it can be seen from Figure 4.12 that the RAFC of Driver B is 526 g in driving scenario 4 and only 478 g in driving scenario 5. Therefore, in order to improve driver behavior, driving scenario 4 should be considered more than driving scenario 5.

From Figure 4.12, it can be clearly seen that for minibus drivers, the extra fuel

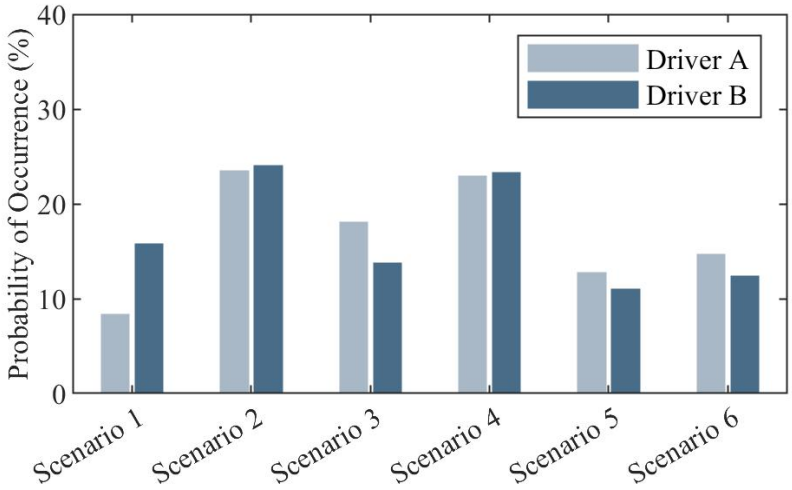


Figure 4.11: The occurrence probability of driving aggressiveness in different driving scenarios for shuttle minibuses two drivers (Scenario 1-6 are defined in Figure 3.10; Use data of shuttle minibus from Apr. 4, 2022 to Apr. 21, 2022)

consumption caused by DA mainly occurs in high-speed driving scenarios. Also, Figure 4.12 shows that extra fuel consumption mainly occurs in driving scenario 4, high-speed resident driving scenario, which has 377 g RAFC difference. Scenario 5 and 6, low-speed and high speed urban driving scenario, are also the primary driving scenarios that result in additional fuel consumption, which have 208 g and 297 g RAFC difference respectively. The value of RAFC points out key scenarios for improving driver's driving behavior.

4.5.2 Ford Escape PHEV Vehicle' Fuel Consumption Caused by Driving Aggressiveness

Through Figure 4.8 in Section 4.5.1, it was found that the driver performing aggressive driving does not mean that the driver is in an aggressive driving situation in all driving scenarios. Morover, it can be seen from Figure 4.8 that the experimental data is limited in driving scenarios 1 and 5, so when analyzing fuel consumption, only driving scenarios 2, 3 and 4 are included. The test result shows that the driver's total fuel consumption is 1649 g under aggressive driving, and the total fuel consumption under normal driving is 1201 g. However, the driver with aggressive driving only

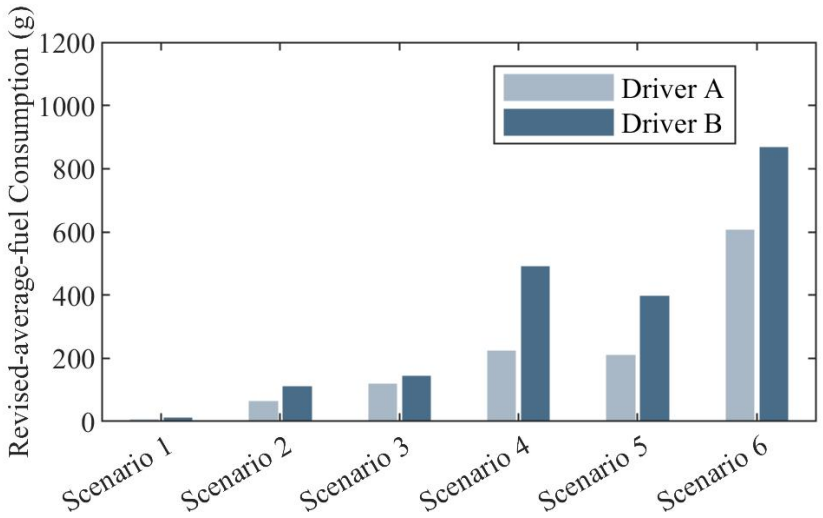


Figure 4.12: Differences in revised average fuel consumption in different driving scenarios for shuttle minibuses two drivers (scenario 1-6 are defined in Figure 3.10; Use data of shuttle minibus from Apr. 4, 2022 to Apr. 21, 2022)

drives more aggressively in low-speed residential driving scenarios. This phenomenon can also be verified by Figure 4.13. In Figure 4.13, driver with aggressive driving only has more fuel consumption in scenario 2. The OPDA and RAFC values for Ford Escape are shown in Figure 4.14 and Figure 4.15, respectively.

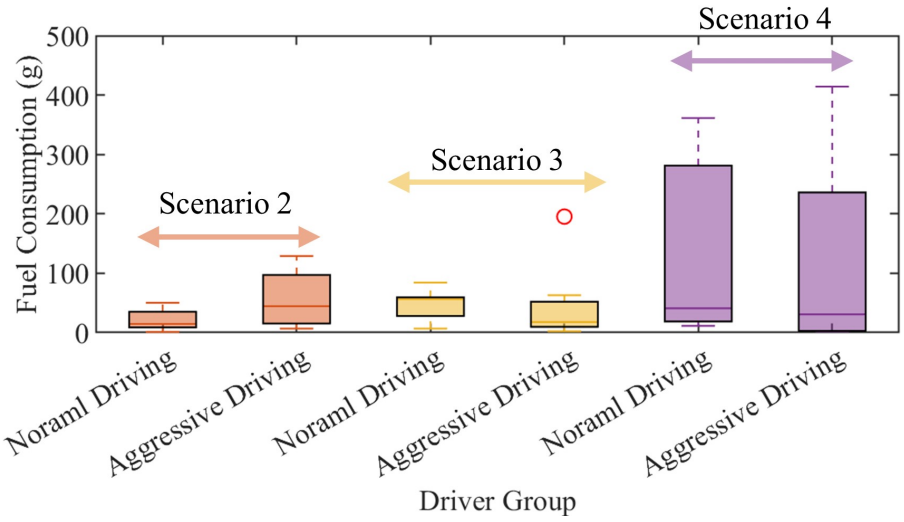


Figure 4.13: Fuel consumption for Ford Escape PHEV driver with different driving behaviors in different driving scenarios (scenario 2-4 are defined in Figure 4.7b; Use data of Ford Escape PHEV from Mar. 26, 2022 to Mar. 27, 2022)

Different from Figure 4.13, the RAFC values for aggressive driving are always

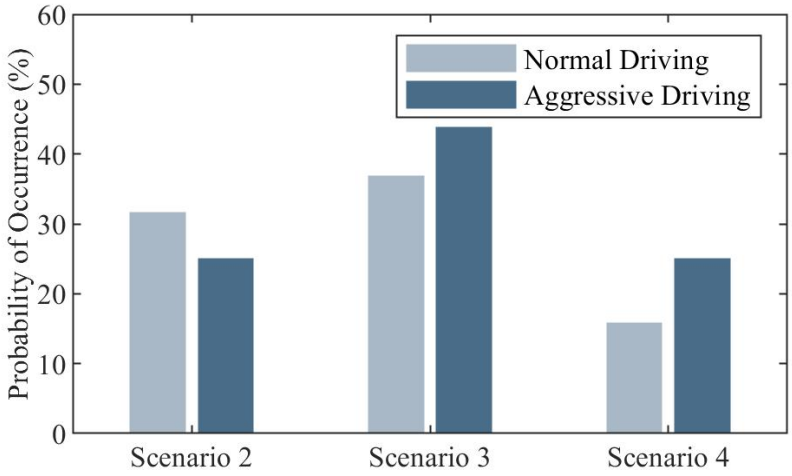


Figure 4.14: The occurrence probability of driving aggressiveness in different driving scenarios for Ford Escape PHEV driver (Scenario 2-4 are defined in Figure 4.7b; Use data of Ford Escape PHEV from Mar. 26, 2022 to Mar. 27, 2022)

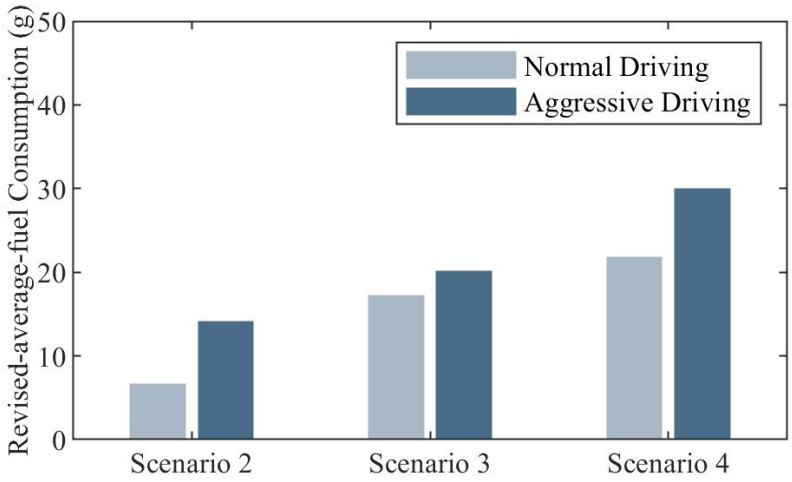


Figure 4.15: Differences in revised average fuel consumption (RAFC) in different driving scenarios for Ford Escape PHEV (Scenario 2-4 are defined in Figure 4.7b; Use data of Ford Escape PHEV from Mar. 26, 2022 to Mar. 27, 2022)

higher than normal driving RAFC values. This shows that RAFC can better reflect the impact of DA on fuel consumption than fuel consumption value. The maximum RAFC difference is occured in scenarios 4, which the aggressive driving has 8.11 g extra RAFC than of normal driving. It shows that for Ford Escape PHEV, it is still necessary to pay attention to high-speed driving situations, so as to improve the

driver's driving behavior.

Such results also illustrate the importance of studying driver aggressiveness at the driving scenarios scale. Through the study of driver aggressiveness in different driving scenarios. It can better guide drivers to improve their driving styles and thus realize feul efficient which leads to low tailpipe CO_2 emission. In addition, because the UAlberta fleet vehicles have a variety models of vehicles, energy saving and emission reduction can be achieved through vehicle allocation matching drivers aggressiveness. For example, drivers who are more aggressive in low-speed driving scenarios can be assigned to HEVs that run electric motors at low vehicle speed; thus, minimizing tailpipe emissions.

4.6 Conclusion

In this chapter, the frequency domain analysis method was used to quantize driver aggressiveness. Three drivers from the University of Alberta shuttle minibuses and Ford Escape PHEV Sedan vehicle were analyzed. Using the clustering algorithm, different driving scenarios were divided into six and five groups for shuttle minibuses and Ford Escape PHEV respectively. The driving behaviors and fuel consumption in different driving scenarios were analyzed separately.

Data samples from 17 days of testing two minibuses were used to illustrate DA for two drivers, named Driver A and Driver B. It was found that driver B's driving behavior was more aggressive in all driving situations. Such a difference in DA value was reflected in the fuel consumption, where the average fuel consumption in a driving cycle scale of driver B was about 5000 g more than that of driver A. When driving scenarios at higher driving speeds, it usually has a higher DA. Moreover, considering the occurrence probability of driving aggressiveness in different driving scenarios is different. This thesis used RAFC plots to better illustrated how much extra fuel consumption was caused by driving aggression. The results of RAFC also illustrate which driving scenarios are key parts. For Ford Escape PHVE, the testing result

indicated that driver aggressiveness was concentrated in a few scenarios.

In future research, the differences in driver behavior of other vehicles in the UAlberta fleet can be analyzed using the method from this thesis. Next, a guideline can be developed to assist the fleet drivers to avoid excessively high DA to minimizing fuel consumption and GHG emission.

Chapter 5

Conclusion and Future Work

5.1 Thesis Contributions

This thesis focus on developing driving cycles of the UAlberta fleet vehicles and analyzing the driving aggressiveness for different drivers. Research on driver behavior provides a tool for how the University of Alberta can assess driver performance and subsequently plan driver training programs or consider driver aggressiveness in assigning vehicles. As a result, the energy saving and emission reduction goals of the EMSO project can be achieved.

The main contributions of this thesis are:

i Development of driving cycle for each application of university vheicles: According to the vehicle characteristics of the UAlberta fleet vehicles, the fleet is divided into several categories, which are the utility and trade category, casual rental category, shuttle minibus category, and UAPS category. In this thesis, unique driving cycles were developed for each vehicle category.

ii Establishment of a method characterise driver aggressiveness (DA): Using the frequency domain analysis method to analyze the DA in different driving scenarios.

iii Linking the DA with fuel consumption: The impact of DA on driving fuel consumption under different driving scenarios had been analyzed. This thesis

also proposes to revised average fuel consumption (RAFC) to better reflect the impact of DA with fuel consumption. Use RAFC to identify driving scenarios that prioritize improving driver behavior.

The main innovations of the thesis are:

i A parameter weight assignment method is designed using the PCA algorithm. Using this weight assignment method can eliminate the duplication between parameters.

ii Using an unsupervised clustering algorithm, driving scenarios for different vehicle categories were developed by clustering the labels of the average and maximum speeds of Microtrips.

iii By using the Microtrip database and the average velocity of a driving cycle except idle (Ve_{avg}) parameters, an algorithm for the development of driving cycles was designed. The R values between the generated driving cycles and the collected data are all lower than 0.5 which indicates that the algorithm is robust.

iv An algorithm was designed to reflect driver aggressiveness based on discrete fourier transform. When the calculation result is closer to 0, it means that the driving is more stable, and on the contrary, when the result is closer to 1, it means that the driving is more aggressive.

v Using engine data (i.e., MAF, AFR_{stoich}, and λ), vehicle fuel consumption rate was estimated and then linked to driver behavior. Next, driver aggressiveness along with fuel consumption in different driving scenarios were analyzed. The difference in driver aggressiveness of different drivers and the difference in fuel consumption caused by driver aggressiveness was identified.

5.2 Conclusions

Here are the main conclusions from this study:

- Research on the driving cycle has shown that the driving cycle varies widely for different vehicle applications. The driving cycle for casual rental vehicles has the longest duration, 9136 seconds, but all other driving cycles' duration are less than 2000 seconds.

- The characteristic of different driving cycles (e.g., average speed, number of starts, stops per kilometer, and the idle ratio) are also quite different. For example, casual rental category vehicles have the highest average running velocity (V_{avg}) which is 58.3 km/h. While other vehicles' V_{avg} ranged from 13.7 km/h to 25.0 km/h. The results underscore the necessity of developing driving cycles for different vehicles.

- DA can be better analyzed by using the frequency domain analysis method. Traditional statistical driving aggression analysis methods miss detail information. For different drivers of the shuttle minibus, although the distribution of acceleration is similar, the fuel consumption is very different. The average fuel consumption of Driver B is 16304 g, but Driver A's average fuel consumption is only 12969 g. Frequency domain analysis method can numerically express driving aggressiveness. The most aggressive driving occurs when the result is close to 1, and the smoother the driving occurs when the result is close to 0.

- Research on driver behavior showed that the driver's aggressiveness state is different in different driving situations. For shuttle minibus, Driver B is more aggressive than Driver A in all scenarios since average DA values of Driver B are higher by 7% to 45% compared those by the Driver A. For Ford Escape PHEV, aggressive driving state doesn't have creeping driving scenarios. To relieve driver aggressiveness, research should focus on key driving scenarios.

- This thesis use RAFC values to better demonstrate the effect of DA on fuel consumption. For minibus drivers, the extra fuel consumption caused by DA mainly occurs in high-speed driving scenarios. The biggest RAFC difference happened in high-speed resident driving scenario, which is 377 g.

5.3 Future Work

Building upon the findings from this thesis, future work may include the following research avenues:

1. Design an optimization framework to utilize driving cycle for each university vehicle category to determine optimum i) fleet composition for university vehicles, ii) assign vehicle type for each driving assignment at the university.

2. Integrating driving cycles for estimating tailpipe emission for university vehicles.

3. Adding road grade information to developed driving cycles to improve their usage for estimating vehicular fuel consumption and emission.

4. Analyzing driver aggressiveness for university drivers for various vehicles (e.g., university utility and trade vehicles) and creating a training program and a monitoring system to improve driving behavior of university drivers.

5. Investigation into the change of the drive cycle by the change of season, i.e., winter vs summer since there is more idling time and lower average vehicle speed in winter driving vs summer driving.

6. Considering the effect of vehicle slip in measured vehicle velocity for calculation of drive cycles, particularly for winter driving data. This will require vehicle instrumentation to measure wheel speed vs other methods of measuring vehicle speed (e.g., using high-accuracy GPS measurements, or using high-precision inertial measurement units, IMUs). In addition, this information can be used

along with the drive cycle of vehicles to determine potential regenerative energy available for winter vs summer driving; thus, allowing a more accurate estimation/prediction of energy consumption for electrified vehicles during winter driving.

Bibliography

[1] *2030 emissions reduction plan: Clean air, strong economy*, https://www.canada.ca/en/services/environment/weather/climatechange/climate-plan/climate-plan-overview/emissions-reduction-2030.html, Accessed May 15, 2023.

[2] *Green technology*, https://www.ualberta.ca/facilities-operations/projects-initiatives/energy-climate-action/green-technology/index.html, Accessed May 15, 2023.

[3] *Canada gazette, part i, volume 154, number 51: Clean fuel regulations*, https://gazette.gc.ca/rp-pr/p1/2020/2020-12-19/html/reg2-eng.html, Accessed May 15, 2023.

[4] *Edmonton gasoline prices, litre*, https://www.globalpetrolprices.com/Canada/Edmonton/gasoline_prices/, Accessed May 15, 2023.

[5] *Alberta diesel prices, litre*, https://www.globalpetrolprices.com/Canada/Alberta/diesel_prices/, Accessed May 15, 2023.

[6] *The official u.s. government source for fuel economey information*, https://www.fueleconomy.gov/, Accessed May 21, 2023.

[7] K. Brundell-Freij and E. Ericsson, "Influence of street characteristics, driver category and car performance on urban driving patterns," *Transportation Research Part D: Transport and Environment*, vol. 10, no. 3, pp. 213–229, 2005.

[8] E. Ericsson, "Variability in urban driving patterns," *Transportation Research Part D: Transport and Environment*, vol. 5, no. 5, pp. 337–354, 2000.

[9] E. Ericsson, "Independent driving pattern factors and their influence on fuel-use and exhaust emission factors," *Transportation Research Part D: Transport and Environment*, vol. 6, no. 5, pp. 325–345, 2001.

[10] L. F. Quirama, M. Giraldo, J. I. Huertas, and M. Jaller, "Driving cycles that reproduce driving patterns, energy consumptions and tailpipe emissions," *Transportation Research Part D: Transport and Environment*, vol. 82, p. 102 294, 2020.

[11] K. M. Sentoff, L. Aultman-Hall, and B. A. Holmén, "Implications of driving style and road grade for accurate vehicle activity data and emissions estimates," *Transportation Research Part D: Transport and Environment*, vol. 35, pp. 175–188, 2015.

[12] X. Zhao, Q. Yu, J. Ma, Y. Wu, M. Yu, and Y. Ye, "Development of a representative ev urban driving cycle based on a k-means and svm hybrid clustering algorithm," *Journal of Advanced Transportation*, vol. 2018, pp. 1–18, 2018.

[13] *Dynamometer drive schedules*, https://www.epa.gov/vehicle-and-fuel-emissions-testing/dynamometer-drive-schedules, Accessed May 15, 2023.

[14] U. Galgamuwa, L. Perera, S. Bandara, *et al.*, "Developing a general methodology for driving cycle construction: Comparison of various established driving cycles in the world to propose a general approach," *Journal of Transportation Technologies*, vol. 5, no. 04, p. 191, 2015.

[15] A. Satish, A. Sabu, J. X. Saldanha, and A. Nagesh, "Development of drive cycle using fleet data for two-wheelers in indian market," SAE Technical Paper, Tech. Rep., 2020.

[16] P Couch and J. Leonard, "Characterization of drayage truck duty cycles at the port of long beach and port of los angeles," *TIAX LLC*, 2011.

[17] J.-B. Gallo, "Development of representative regional delivery drive cycles for heavy-duty truck tractors," *SAE International Journal of Commercial Vehicles*, vol. 7, no. 2014-01-9024, pp. 337–347, 2014.

[18] A. Fotouhi and M. Montazeri-Gh, "Tehran driving cycle development using the k-means clustering method," *Scientia Iranica*, vol. 20, no. 2, pp. 286–293, 2013.

[19] S. Chugh *et al.*, "Development of delhi driving cycle: A tool for realistic assessment of exhaust emissions from passenger cars in delhi," SAE Technical Paper, Tech. Rep., 2012.

[20] S. Tamsanya, S. Chungpaibulpatana, and B. Limmeechokchai, "Development of a driving cycle for the measurement of fuel consumption and exhaust emissions of automobiles in bangkok during peak periods," *International Journal of Automotive Technology*, vol. 10, pp. 251–264, 2009.

[21] S. Shi *et al.*, "Research on markov property analysis of driving cycle," in *2013 IEEE Vehicle Power and Propulsion Conference (VPPC)*, IEEE, 2013, pp. 1–5.

[22] J. D. Bishop, C. J. Axon, and M. D. McCulloch, "A robust, data-driven methodology for real-world driving cycle development," *Transportation Research Part D: Transport and Environment*, vol. 17, no. 5, pp. 389–397, 2012.

[23] H. Gong, Y. Zou, Q. Yang, J. Fan, F. Sun, and D. Goehlich, "Generation of a driving cycle for battery electric vehicles: A case study of beijing," *Energy*, vol. 150, pp. 901–912, 2018.

[24] A. Bouhoute, R. Oucheikh, K. Boubouh, and I. Berrada, "Advanced driving behavior analytics for an improved safety assessment and driver fingerprinting," *IEEE Transactions on Intelligent Transportation Systems*, vol. 20, no. 6, pp. 2171–2184, 2018.

[25] *Driving more efficiently*, https://www.energy.gov/energysaver/driving-more-efficiently, Accessed February 15, 2023.

[26] F. Zheng, J. Li, H. Van Zuylen, and C. Lu, "Influence of driver characteristics on emissions and fuel consumption," *Transportation Research Procedia*, vol. 27, pp. 624–631, 2017.

[27] J. E. Meseguer, C. K. Toh, C. T. Calafate, J. C. Cano, and P. Manzoni, "Drivingstyles: A mobile platform for driving styles and fuel consumption characterization," *Journal of Communications and networks*, vol. 19, no. 2, pp. 162–168, 2017.

[28] W. Wang, J. Xi, A. Chong, and L. Li, "Driving style classification using a semisupervised support vector machine," *IEEE Transactions on Human-Machine Systems*, vol. 47, no. 5, pp. 650–660, 2017.

[29] E Tzirakis, F Zannikos, and S Stournas, "Impact of driving style on fuel consumption and exhaust emissions: Defensive and aggressive driving style," in *Proceedings of the 10th international conference on environmental science and technology*, Global Network for Environmental Science and Technology (Global-NEST), 2007, pp. 1497–1504.

[30] D. Hari, C. J. Brace, C. Vagg, J. Poxon, and L. Ash, "Analysis of a driver behaviour improvement tool to reduce fuel consumption," in *2012 International Conference on Connected Vehicles and Expo (ICCVE)*, IEEE, 2012, pp. 208–213.

[31] E. M. Szumska and R. Jurecki, "The effect of aggressive driving on vehicle parameters," *Energies*, vol. 13, no. 24, p. 6675, 2020.

[32] G. Amirjamshidi and M. J. Roorda, "Development of simulated driving cycles for light, medium, and heavy duty trucks: Case of the toronto waterfront area," *Transportation research part D: transport and environment*, vol. 34, pp. 255–266, 2015.

[33] E. Cho, S. Park, and C. Oh, "Analysis of intrinsic factors leading to aggressive driving behavior to derive safety policy implications for bus drivers," *Transportation research record*, vol. 2676, no. 11, pp. 664–675, 2022.

[34] G. Halfacree, *Freematics' one+ aims to set your odb-ii port free with wi-fi, bluetooth, cellular data logging*, https://www.hackster.io/news/freematics-one-aims-to-set-your-odb-ii-port-free-with-wi-fi-bluetooth-cellular-data-logging-d4cd2aa78a73, Accessed March 15, 2023.

[35] *Freematics hub*, https://freematics.com/hub/, Accessed March 15, 2023.

[36] D. KAREN, A. SAMER, W. LINDSEY, and L. YONG, *Piecewise interpolation: Cubic spline interpolation*, https://engcourses-uofa.ca/books/numericalanalysis/piecewise-interpolation/cubic-spline-interpolation/, Accessed March 15, 2023.

[37] J. Lin and D. A. Niemeier, "An exploratory analysis comparing a stochastic driving cycle to california's regulatory cycle," *Atmospheric Environment*, vol. 36, no. 38, pp. 5759–5770, 2002.

[38] Z. Xiao, Z. Dui-Jia, and S. Jun-Min, "A synthesis of methodologies and practices for developing driving cycles," *Energy procedia*, vol. 16, pp. 1868–1873, 2012.

[39] J. Zhang, Z. Wang, P. Liu, and Z. Zhang, "Energy consumption analysis and prediction of electric vehicles based on real-world driving data," *Applied Energy*, vol. 275, p. 115 408, 2020.

[40] G. Yang, X. Song, C. Huang, Z. Deng, J. Shi, and B. Zhou, "Drivingstereo: A large-scale dataset for stereo matching in autonomous driving scenarios," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2019, pp. 899–908.

[41] S. M. Easa, M. J. Reed, F. Russo, E. Dabbour, A. Mehmood, and K. Curtis, "Effect of increasing road light luminance on night driving performance of older adults," *International Journal of Civil and Environmental Engineering*, vol. 4, no. 8, pp. 201–208, 2010.

[42] K. Yoneda, N. Suganuma, R. Yanase, and M. Aldibaja, "Automated driving recognition technologies for adverse weather conditions," *IATSS research*, vol. 43, no. 4, pp. 253–262, 2019.

[43] D. He, *Understanding and Supporting Anticipatory Driving in Automated Vehicles*. University of Toronto (Canada), 2020.

[44] L. GUYONVARCH *et al.*, *Data driven scenarios for ad/adas validation*, 2020.

[45] *Vehicle occupant infrastructure road user safety study Voiesur*, https://anr.fr/Project-ANR-11-VPTT-0007, Accessed March 4, 2023.

[46] N. van Nes, *European naturalistic driving and riding for infrastructure & vehicle safety and environment*, https://cordis.europa.eu/project/id/314050/reporting, Accessed March 4, 2023.

[47] J. Li, B. Dai, X. Li, X. Xu, and D. Liu, "A dynamic bayesian network for vehicle maneuver prediction in highway driving scenarios: Framework and verification," *Electronics*, vol. 8, no. 1, p. 40, 2019.

[48] L. C. Silva, J. J. Eckert, M. A. Lourenco, F. L. Silva, F. C. Corrêa, and F. G. Dedini, "Electric vehicle battery-ultracapacitor hybrid energy storage system and drivetrain optimization for a real-world urban driving scenario," *Journal of the Brazilian Society of Mechanical Sciences and Engineering*, vol. 43, no. 5, p. 259, 2021.

[49] M. Rodrigues, G. Gest, A. McGordon, and J. Marco, "Adaptive behaviour selection for autonomous vehicle through naturalistic speed planning," in *2017 IEEE 20th International Conference on Intelligent Transportation Systems (ITSC)*, IEEE, 2017, pp. 1–7.

[50] Y. Liu, A. Ansari, and M. Shahbakhti, "Identification of the driving cycle for university fleet vehicles," 2022.

[51] A. Singh, A. Yadav, and A. Rana, "K-means with three different distance metrics," *International Journal of Computer Applications*, vol. 67, no. 10, 2013.

[52] V. M. Bhaskaran *et al.*, "Improving the efficiency of image clustering using modified non euclidean distance measures in data mining," *International Journal of Computers Communications & Control*, vol. 9, no. 1, pp. 56–61, 2014.

[53] Chutai, *Machine learning-principle and advantages and disadvantages of principal component analysis (pca) algorithm*, https://www.cnblogs.com/lsm-boke/p/11760224.html, Accessed May 30, 2022.

[54] *Variance*, https://zh.wikipedia.org/zh-cn/%E6%96%B9%E5%B7%AE, Accessed May 31, 2022.

[55] Zhen, *Four evaluation functions in machine learning*, https://juejin.cn/post/6844903599298838541, Accessed April 14, 2022.

[56] J. A. Michon, "A critical view of driver behavior models: What do we know, what should we do?" *Human behavior and traffic safety*, pp. 485–524, 1985.

[57] Y Alver, M. Demirel, and M. Mutlu, "Interaction between socio-demographic characteristics: Traffic rule violations and traffic crash history for young drivers," *Accident Analysis & Prevention*, vol. 72, pp. 95–104, 2014.

[58] K. Young, M. Regan, and M. Hammer, "Driver distraction: A review of the literature," *Distracted driving*, vol. 2007, pp. 379–405, 2007.

[59] N. Schaap, B. Van Arem, and R. Van der Horst, "Drivers' behavioural reactions to unexpected events," in *Proceedings of 10th TRAIL Congress-TRAIL in Perspective*, 2008.

[60] B. Scott-Parker, "Emotions, behaviour, and the adolescent driver: A literature review," *Transportation research part F: traffic psychology and behaviour*, vol. 50, pp. 1–37, 2017.

[61] P. M. Salmon, M. G. Lenne, G. H. Walker, N. A. Stanton, and A. Filtness, "Exploring schema-driven differences in situation awareness between road users: An on-road study of driver, cyclist and motorcyclist situation awareness," *Ergonomics*, vol. 57, no. 2, pp. 191–209, 2014.

[62] X. Qin, N. Zhang, W. Zhang, and M. Meitner, "How does tunnel interior color environment influence driving behavior? quantitative analysis and assessment experiment," *Tunnelling and Underground Space Technology*, vol. 98, p. 103 320, 2020.

[63] J. Michaels *et al.*, "Driving simulator scenarios and measures to faithfully evaluate risky driving behavior: A comparative study of different driver age groups," *PloS one*, vol. 12, no. 10, e0185909, 2017.

[64] D. de Waard, C. Dijksterhuis, and K. A. Brookhuis, "Merging into heavy motorway traffic by young and elderly drivers," *Accident Analysis & Prevention*, vol. 41, no. 3, pp. 588–597, 2009.

[65] A. Sovtic, D. Adelberger, and M. Wang, "Probabilistic prediction of driving behavior on country roads," in *2022 European Control Conference (ECC)*, IEEE, 2022, pp. 783–789.

[66] R. Zheng, K. Nakano, H. Ishiko, K. Hagita, M. Kihira, and T. Yokozeki, "Eye-gaze tracking analysis of driver behavior while interacting with navigation systems in an urban area," *IEEE Transactions on Human-Machine Systems*, vol. 46, no. 4, pp. 546–556, 2015.

[67] S. H. Hamdar, L. Qin, and A. Talebpour, "Weather and road geometry impact on longitudinal driving behavior: Exploratory analysis using an empirically supported acceleration modeling framework," *Transportation research part C: emerging technologies*, vol. 67, pp. 193–213, 2016.

[68] D. Bosanquet *et al.*, "Driving on ice: Impaired driving skills in current methamphetamine users," *Psychopharmacology*, vol. 225, pp. 161–172, 2013.

[69] X. Li, X. Yan, and S. Wong, "Effects of fog, driver experience and gender on driving behavior on s-curved road segments," *Accident Analysis & Prevention*, vol. 77, pp. 91–104, 2015.

[70] A. Kondyli and L. Elefteriadou, "Modeling driver behavior at freeway–ramp merges," *Transportation research record*, vol. 2249, no. 1, pp. 29–37, 2011.

[71] A. Mukhtar, L. Xia, and T. B. Tang, "Vehicle detection techniques for collision avoidance systems: A review," *IEEE transactions on intelligent transportation systems*, vol. 16, no. 5, pp. 2318–2338, 2015.

[72] Z. Song, J. Hou, S. Xu, M. Ouyang, and J. Li, "The influence of driving cycle characteristics on the integrated optimization of hybrid energy storage system for electric city buses," *Energy*, vol. 135, pp. 91–100, 2017.

[73] X. Geng, H. Liang, B. Yu, P. Zhao, L. He, and R. Huang, "A scenario-adaptive driving behavior prediction approach to urban autonomous driving," *Applied Sciences*, vol. 7, no. 4, p. 426, 2017.

[74] M. Weber, J. Giacomin, A. Malizia, L. Skrypchuk, V. Gkatzidou, and A. Mouzakitis, "Investigation of the dependency of the drivers' emotional experience on different road types and driving conditions," *Transportation research part F: traffic psychology and behaviour*, vol. 65, pp. 107–120, 2019.

[75] X. Zhou, C. Huang, Y. Zhou, *et al.*, "Unsafe behavior analysis and risk measurement of traffic accidents in mountainous highway tunnel," *Mathematical Problems in Engineering*, vol. 2022, 2022.

[76] E. Berdoulat, D. Vavassori, and M. T. M. Sastre, "Driving anger, emotional and instrumental aggressiveness, and impulsiveness in the prediction of aggressive and transgressive driving," *Accident Analysis & Prevention*, vol. 50, pp. 758–767, 2013.

[77] C. Bingham, C. Walsh, and S. Carroll, "Impact of driving characteristics on electric vehicle energy consumption and range," *IET Intelligent Transport Systems*, vol. 6, no. 1, pp. 29–35, 2012.

[78] Z. Liu, A. Ivanco, and Z. S. Filipi, "Impacts of real-world driving and driver aggressiveness on fuel consumption of 48v mild hybrid vehicle," *SAE International Journal of Alternative Powertrains*, vol. 5, no. 2, pp. 249–258, 2016.

[79] N. Bohr, "Atomic theory and the description of nature," *American Journal of Physics*, vol. 30, no. 9, pp. 658–660, 1962.

[80] K. Yang, C. Al Haddad, G. Yannis, and C. Antoniou, "Classification and evaluation of driving behavior safety levels: A driving simulation study," *IEEE Open Journal of Intelligent Transportation Systems*, vol. 3, pp. 111–125, 2022.

[81] Z. Liu, A. Ivanco, and Z. Filipi, "Quantification of drive cycle's rapid speed fluctuations using fourier analysis," *SAE International Journal of Alternative Powertrains*, vol. 4, no. 1, pp. 170–177, 2015.

[82] E. Moradi and L. Miranda-Moreno, "Vehicular fuel consumption estimation using real-world measures through cascaded machine learning modeling," *Transportation Research Part D: Transport and Environment*, vol. 88, p. 102 576, 2020.

[83] A. Ansari, H. Abediasl, P. R. Patel, V. Hosseini, C. R. Koch, and M. Shahbakhti, "Estimating instantaneous fuel consumption of vehicles by using machine learning and real-time on-board diagnostics (obd) data," 2022.

Appendix A: MSc. Publications

A.1 Refereed Conference Papers in Proceedings

1. **Y. Liu**, H. Abediasl, A. Ansari, M. Shahbakhti, " Characterizing Driving Behavior and Link to Fuel Consumption for University Campus Shuttle Minibuses ", Canadian Society of Mechanical Engineers (CSME) 2023 International Congress, 6 pages, May 28-31, 2023, Sherbrooke, QC, Canada.

2. **Y. Liu**, A. Ansari, M. Shahbakhti, " Identification of the Driving Cycle for University Fleet Vehicles ", Canadian Society of Mechanical Engineers (CSME) 2022 International Congress, 6 pages, Jun. 5-8, 2022, Edmonton, AB, Canada.

Appendix B: Thesis Files

B.1 Program and Data File Summary

The following files were created in this thesis.

B.2 Chapter 1

Table B.1: Chapter 1 figure files

File name	File Description
thesis scheme.pptx, thesis scheme.jpg	Figure 1.1
EMSO Project_Overview.pdf	Figure 1.2
dcc.pptx, dc category.jpg	Figure 1.3
epa 1.fig, epa 1.jpg	Figure 1.4a
epa 2.fig, epa 2.jpg	Figure 1.4b
epa 3.fig, epa 3.jpg	Figure 1.4c
epa 4.fig, epa 4.jpg	Figure 1.4d
epa 5.fig, epa 5.jpg	Figure 1.4e
epa 6.fig, epa 6.jpg	Figure 1.4f
euja 1.fig, euja 1.jpg	Figure 1.5a
euja 2.fig, euja 2.jpg	Figure 1.5b
euja 3.fig, euja 3.jpg	Figure 1.5c
euja 4.fig, euja 4.jpg	Figure 1.5d
euja 5.fig, euja 5.jpg	Figure 1.5e
tf.pptx, thesis flow.jpg	Figure 1.6

Table B.2: Chapter 1 Matlab script files

File name	File Description
ch_1_estimate_FC.m (Used to plot estimated total fuel consumption for each vehicle type for the UAlberta fleet vehicles)	Figure 1.2
epa_drivingCycle.m (Used to plot the US environmental protection agency driving cycles and EU & Japan driving cycles)	Figure 1.4a - Figure 1.4e

Table B.3: Chapter 1 data files

File name	File Description
drive cycle.xlsx (Contain all US environmental protection agency driving cycles and EU and Japan driving cycles' data)	Figure 1.4a - Figure 1.4e

B.3 Chapter 2

Table B.4: Chapter 2 figure files

File name	File Description
gps1.pptx, gps1.jpg	Figure 2.2
gps2.pptx, gps2.jpg	Figure 2.3
gps3.pptx, gps3.jpg	Figure 2.4
OBD.pptx, OBD.jpg	Figure 2.1
Test.pdf, test route.jpg	Figure 2.5
FF.pptx, fuel flow.jpg	Figure 2.6
collect.pptx, Schematics of the data collection process.jpg	Figure 2.7
inter.pptx, interpolation data.jpg	Figure 2.8
interpolate.pptx	Figure 2.9

Table B.5: Chapter 2 Matlab script files

File name	File Description
Data_Interpolation.m (Used to plot the interpolation figure and plot the polynomial order on interpolation accuracy figure)	Figure 2.8 - Figure 2.9
inter_data.m (Used to plot the enlarged part of the interpolation figure)	Figure 2.9

Table B.6: Chapter 2 data files

File name	**File Description**
Unit429.xlsx - Unit505.xlsx (Include all utility or trade vehicles' OBD record data)	Figure 2.2
Unit486.xlsx - Unit587.xlsx (Include all casual rental vehicles' OBD record data)	Figure 2.3
Unit438.xlsx - Unit416.xlsx (Include all shuttle minibuses' OBD record data)	Figure 2.4

B.4 Chapter 3

Table B.7: Chapter 3 figure files

File name	File Description
microtrip.fig, microtrip.jpg	Figure 3.1
ua truck database.fig, ua truck database.jpg	Figure 3.2a
rental database.fig, rental database.jpg	Figure 3.2b
minibus database.fig, minibus database.jpg	Figure 3.2c
police database.fig, police database.jpg	Figure 3.2d
kf.pptx, kinematic fragment.jpg	Figure 3.3
driving scenarios.pptx, driving scenarios.jpg	Figure 3.4
pca hot.fig, pca hot.jpg	Figure 3.5
pca weight.fig, pca weight.jpg	Figure 3.6
loss ua truck.fig, loss ua truck.jpg	Figure 3.7a
loss rental.fig, loss rental.jpg	Figure 3.7b
loss minibus.fig, loss minibus.jpg	Figure 3.7c
loss police.fig, loss police.jpg	Figure 3.7d
scenario ua.fig, scenario ua.jpg	Figure 3.8
scenario rental.fig, scenario rental.jpg	Figure 3.9
scenario minibus.fig, scenario minibus.jpg	Figure 3.10
scenario police.fig, scenario police.jpg	Figure 3.11

Table B.8: Chapter 3 figure files

File name	File Description
sub ua 1.fig, sub ua 1.jpg	Figure 3.12a
sub ua 2.fig, sub ua 2.jpg	Figure 3.12b
sub ua 3.fig, sub ua 3.jpg	Figure 3.12c
sub ua 4.fig, sub ua 4.jpg	Figure 3.12d
sub ua 5.fig, sub ua 5.jpg	Figure 3.12e
srental 1.fig, sub rental 1.jpg	Figure 3.13a
srental 2.fig, sub rental 2.jpg	Figure 3.13b
srental 3.fig, sub rental 3.jpg	Figure 3.13c
srental 4.fig, sub rental 4.jpg	Figure 3.13d
srental 5.fig, sub rental 5.jpg	Figure 3.13e
sub mini 1.fig, sub mini 1.jpg	Figure 3.14a
sub mini 2.fig, sub mini 2.jpg	Figure 3.14b
sub mini 3.fig, sub mini 3.jpg	Figure 3.14c
sub mini 4.fig, sub mini 4.jpg	Figure 3.14d
sub mini 5.fig, sub mini 5.jpg	Figure 3.14e
sub mini 6.fig, sub mini 6.jpg	Figure 3.14f
spolice 1.fig, sub police 1.jpg	Figure 3.15a
spolice 2.fig, sub police 2.jpg	Figure 3.15b
spolice 3.fig, sub police 3.jpg	Figure 3.15c
spolice 4.fig, sub police 4.jpg	Figure 3.15d
spolice 5.fig, sub police 5.jpg	Figure 3.15e
spolice 6.fig, sub police 6.jpg	Figure 3.15f
DC.pptx, DC flow chart.jpg	Figure 3.16
dc ua.fig, dc ua.jpg	Figure 3.17
dc rental.fig, dc rental.jpg	Figure 3.18
dc minibus.fig, dc minibus.jpg	Figure 3.19
dc uaps.fig, dc uaps.jpg	Figure 3.20

Table B.9: Chapter 3 Matlab script files

File name	**File Description**
Microtrip_determine.m (Used to plot Microtrip example figure)	Figure 3.1
get_microtrip.m (Used to plot Microtrip database for different vehicle categories)	Figure 3.2a - Figure 3.2d
LookForAccDecCruise.m (Used to plot kinematic fragment example figure)	Figure 3.3
parameter_for_total.m (Used to plot the matrix of the proportional feature vector and weights for target parameters figures)	Figure 3.5 - Figure 3.6
judge_k_number.m (Used to plot $Loss$ and $\Delta loss$ for different vehicle category)	Figure 3.7a - Figure 3.7d
draw_figure.m (Used to plot driving scenarios for different vehicle category)	Figure 3.8 - Figure 3.11
microtrip_database.m (Used to plot sub_microtrip database for different vehicle category)	Figure 3.12a - Figure 3.15f
driving_cycle.m (Used to plot driving cycle for different vehicle category)	Figure 3.17 - Figure 3.20
table1.m (Used to get driving fragment)	Figure 3.8 - Figure 3.20
table2.m (Used to generate database)	Figure 3.8 - Figure 3.20

Table B.10: Chapter 3 driving cycle files

File name	File Description
driving_cycle.xlsx (Contain all generate driving cycles for different vehicle categories)	Figure 3.17 - Figure 3.20

Table B.11: Chapter 3 data files

File name	File Description
Unit429.xlsx - Unit505.xlsx (Include all utility or trade vehicles' OBD record data)	Figure 3.2a & Figure 3.7a & Figure 3.12a - Figure 3.12e
Unit486.xlsx - Unit587.xlsx (Include all casual rental vehicles' OBD record data)	Figure 3.2b & Figure 3.7b & Figure 3.13a - Figure 3.13e
Unit438.xlsx - Unit416.xlsx (Include all shuttle minibuses' OBD record data)	Figure 3.2c & Figure 3.7c & Figure 3.14a - Figure 3.14f
Unit93.xlsx - Unit95.xlsx (Include all UAPS' OBD record data)	Figure 3.2d & Figure 3.7d & Figure 3.15a - Figure 3.15f

B.5 Chapter 4

Table B.12: Chapter 4 figure files

File name	File Description
DB.pptx, driver behavior.jpg	Figure 4.1
acc compare.fig, acc compare.jpg	Figure 4.2
fuel compare.fig, fuel compare.jpg	Figure 4.3
DF.pptx, a flow fig.jpg	Figure 4.4
af.fig, actual fuel.jpg	Figure 4.5
test database.fig, test database.jpg	Figure 4.7a
test subdatabase.fig, test subdatabase.jpg	Figure 4.7b
test loss.fig, test loss.jpg	Figure 4.7c
Da agg.fig, Da agg.jpg	Figure 4.6a
Db agg.fig, Db agg.jpg	Figure 4.6b
test a.fig, test a.jpg	Figure 4.8a
test b.fig, test b.jpg	Figure 4.8b
correlation coefficient shuttle.fig, correlation coefficient shuttle.jpg	Figure 4.9
fba.fig, fuel boxlpot a.jpg	Figure 4.10a
fbb.fig, fuel boxlpot b.jpg	Figure 4.10b
pm.fig, prpbability for minibus.jpg	Figure 4.11
Rm.fig, RAFC minibus.jpg	Figure 4.12
test fuel.fig, test fuel.jpg	Figure 4.13
pt.fig, prpbability for test.jpg	Figure 4.14
Rt.fig, RAFC test.jpg	Figure 4.15

Table B.13: Chapter 4 Matlab script files

File name	File Description
acc_compare.m (Used to plot acceleration compare figure)	Figure 4.2
compare_UArental.m (Used to plot fuel consumption compare figure)	Figure 4.3
rainplot.m (Used to plot driver aggressiveness figures)	Figure 4.6a - Figure 4.8b
box_togother.m (Used to plot fuel consumption figures for minibuses)	Figure 4.6a - Figure 4.8b
cc_shuttle.m (Used to plot correlation coefficient value for shuttle minibus driver A)	Figure 4.9
RAFC_mini.m (Used to plot revised average fuel consumption and occurrence probability of driving aggressiveness figures for minibuses)	Figure 4.11 - Figure 4.12
Ford_fuel.m (Used to plot fuel consumption figures for Ford Escape PHEV)	Figure 4.13
RAFC.m (Used to plot revised average fuel consumption and occurrence probability of driving aggressiveness figures for Ford Escape PHEV)	Figure 4.14 - Figure 4.15
get_driver_agg.m (Used to calculate driver aggressive)	Figure 4.6a - Figure 4.8b

Table B.14: Chapter 4 data files

File name	File Description
normal.xlsx - aggressive.xlsx (Include all Ford Escape PHEV's OBD record data)	Figure 4.7a - Figure 4.7c & Figure 4.8a - Figure 4.8b & Figure 4.13 & Figure 4.11 - Figure 4.12
minibus_record_data.xlsx (Include all shuttle minibuses' OBD record data)	Figure 4.1 & Figure 4.2 & Figure 4.3 & Figure 4.6a - Figure 4.6b & Figure 4.10a - Figure 4.10b & Figure 4.14 - Figure 4.15