Proceedings of the Canadian Society for Mechanical Engineering International Congress 2023
CSME Congress 2023
May 28-31, 2023, Sherbrooke, QC, Canada

Characterizing Driving Behavior and Link to Fuel Consumption for University Campus Shuttle Minibuses

Yang Liu *, Hamidreza Abediasl , Amir Ansari , Mahdi Shahbakhti
Department of Mechanical Engineering, University of Alberta, Edmonton, Canada
*lyang10@ualberta.ca

Abstract**—**This paper focuses on the effect of aggressive driving behavior on fuel consumption of a vehicle. Different from the traditional statistical analysis method, this paper adopts the frequency domain analysis method to analyze driving aggressiveness and apply a quantitative driving aggressiveness evaluation metric. At the same time, the fuel consumption impact caused by the driving aggressiveness under different driving situations is analyzed. The results are demonstrated for two university shuttle bus. Fuel consumption rate of each vehicle is determined by using available on-board diagnostics (OBD) data including intake air mass flow rate of engine and air/fuel equivalence ratio. The experimental results show that the degree of influence of driving aggressiveness on fuel consumption is not the same in different driving situations. The higher the speed of the driving situation, the greater the difference in fuel consumption caused by driving aggressiveness.

Keywords-driving behavior; frequency domain analysis; driving aggressiveness; driving situations; fuel consumption

I. INTRODUCTION

The University of Alberta manages more than 170 vehicles, hereafter referred to as the University of Alberta fleet vehicles. These fleet vehicles consume about 205,000 liters of fuel per year, and the CO2 greenhouse gas (GHG) emissions released by these vehicles is about 564,000 kg each year. To this end, a research project under the University of Alberta Energy Management and Sustainable Operations (EMSO) aims to achieve the goals of operational fuel and cost reduction, GHG emission reduction, and campus air quality improvement. This paper focuses on characterizing the excess fuel consumption caused by a driver's driving behavior.

The research on driver's driving behavior (also referred to as driving behavior for short) usually consists of three parts as shown in Fig.1 [1, 2, 3]. Within the scope of driver self-influence, research directions can be divided into three categories, that is, (i) finding the optimal driving cycle, (ii) driving behavior prediction, and (iii) driving characteristic classification [4, 5, 6]. The research on driving characteristic classification can be further divided into three categories, including driver emotion impact analysis, unsafe behavior analysis, and driving aggressiveness analysis [7, 8, 9]. This study focuses on the driving aggressiveness (DA) in driving behavior. Traditional DA research methods focus on the statistics of acceleration [10]. However, statistical methods will ignore the detailed performance of driving behavior. These detailed behaviors are needed for the analysis of DA in different driving situations. Therefore, this paper adopts the frequency domain analysis method to analyze the DA under different driving situations [11]. In Fig. 1, the focus area of this paper is highlighted in green blocks.

A clustering algorithm, K-means algorithm [12], is applied to classify driving situations. According to the characteristics of the experimental vehicles, multiple driving situations ranging from low-speed crawling driving situations to high-speed urban driving situations are divided to study DA in detail. To analyze the difference in fuel consumption (FC) at the scale of the whole driving cycle, a model that uses mass flow rate of intake air (MAF) to estimate fuel consumption rate (FCR) is used [13].

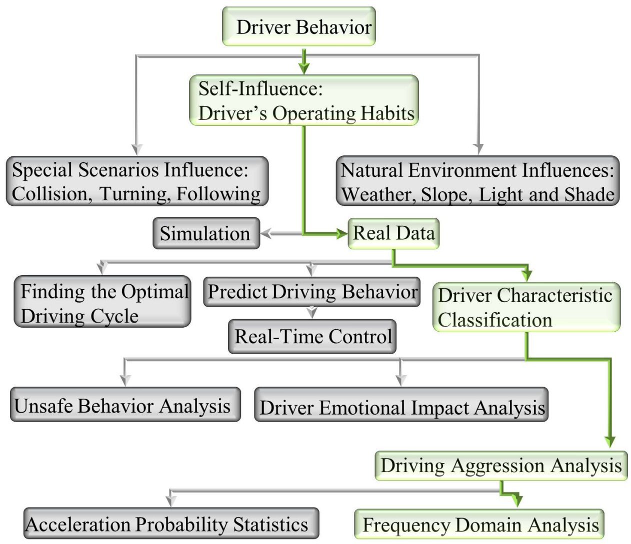


Fig.1 Classification of driving behavior analysis. Green blocks show the focus area of this study.

The main contributions from this paper include: (i) real word data collection and identification of different driving situations of two campus minibuses, (ii) analysis of the DA differences under different driving situations, and (iii) investigation of the differences in FC caused by differences in DA under different driving situations.

This paper is organized as follows: Section 2 explains the vehicle experimental setup. The method of frequency domain analysis, driving situations classification algorithm, and models to estimate FC are described in Section 3. Section 4 specifically shows the driver's DA and FC differences in different driving situations and finally Ssection 5 presents the summary and conclusions.

II. VEHICLE EXPERIMENTAL SETUP

A. Vehicle Selection

In this paper, two minibuses from the University of Alberta fleet vehicles are selected for study. Two minibus vehicles are driven by two drivers respectively. The driving path of the minibus is fixed. This allows to properly reflect the difference in vehicles' performance (such as FC) caused by the difference in the driving behavior of the drivers. The two minibuses transfer students between the main campus and Campus Saint-Jean (CSJ). By using two similar minibuses (Table 1) to analyze driving behavior, the difference caused by vehicles is reduced.

As shown in Fig. 2, both minibuses travel between the University of Alberta (UA) main campus and CSJ campus along the same route. Mark 1 in Fig. 2 is the UA main campus and Mark 2 in Fig. 2 is the 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.

B. Data Collection

In this study, Freematics One+ on-board diagnostics (OBD) data loggers, shown in Fig.3 (c), were used to collect minibus vehicle data. The vehicle OBD interface is shown in red box in Fig. 3 (a) and the zoomed picture is shown in Fig. 3 (b). In this study, vehicle data for a total of 17 days was collected from the two minibuses leading to 83,685 collected data.

C. Microtrip Database

The traditional driver behavior analysis is in the driving cycle scale, which omits the driver behavior difference in different driving situations. Microtrip is a small-scale "driving cycle", which is obtained by slicing driving cycles [14]. A Microtrip is defined as a trip between two consecutive time when the vehicle speed is zero. Microtrip database for the collected data is shown in Fig. 4. According to Microtrip's maximum and average vehicle speed, different driving situations can be divided, such as creeping situations, urban low-speed driving situations, and urban high-speed driving situations, highway driving situations [15]. By dividing the Microtrip database to obtain driving situations, the differences in driving behavior are analyzed from a microscopic perspective.

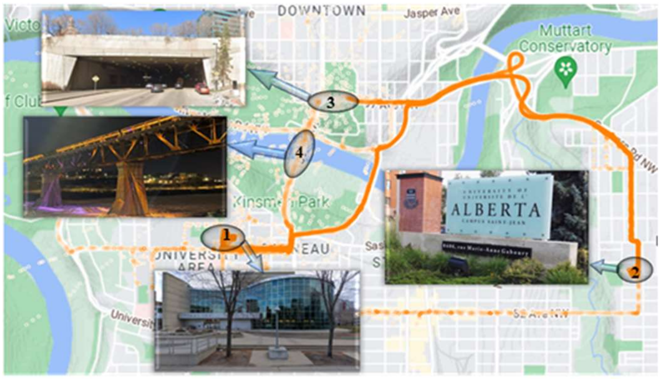


Fig.2 The GPS route map for minibuses

TABLE 1 INFORMATION OF UA MINIBUSES

Vehicle	Minibus A	Minibus B
Photo		
Makers & Model	Ford E450	Ford E450
Year	2018	2020
Rated Power	325 hp	325 hp
Engine Size	7.3 L V8	7.3 L V8

III. METHOD FOR ANALYZING DRIVING BEHAVIOR

A. Determine the Driving Situations

Each Microtrip has a certain value of average vehicle speed and maximum vehicle speed which are the most important driving characteristic parameters [16], as shown in Fig. 4. Each point in Fig. 4 is a Microtrip. By using the data in Fig. 4, a machine learning model is developed to form the sub-Microtrip database. Each sub-Microtrip database presents one of driving situations from creeping situations to high-speed driving situations. In this study, the K-means algorithm is used to determine the sub-Microtrip database.

The K-means algorithm is a typical unsupervised learning algorithm, which is mainly used to automatically classify similar samples into one category. In the clustering 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 paper is the Euclidean Distance method [17] because Euclidean Distance method is simple to calculate and it can speed up the algorithm.

For using K-means algorithm, firstly, the sample set X in Equation (1) and the number of clusters k should be determined. Each element in X is called an object.

$$X = \{X_1, X_2, X_3, \dots X_n\} \tag{1}$$

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 (2). Initialized k cluster centers are usually

Fig.3 The minibus cabin showing the OBD port and OBD data logger for data collection

composed of k points which are randomly selected from *X*.

$$C = \{C_1, C_2, C_3, \dots C_k\}, 1 < k \leq n \tag{2}$$

Equation (3) is used to calculate the Euclidean distance (*dis*) from each object to each cluster center.

$$dis(X_i, C_j) = \sqrt{\sum_{t=1}^{n} \left(X_{it} - C_{jt}\right)^2} \tag{3}$$

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 (4).

$$S = \{S_1, S_2, S_3, \dots S_K\} \tag{4}$$

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 new cluster center (C_{inew}) is as follows.

$$C_{inew} = \frac{\sum_{X_i \in S_i} X_i}{|S_i|} \tag{5}$$

After obtaining the new cluster center, an interactive process is followed to calculate 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 paper uses cross-validation to select the optimal *k* according to the loss function [18].

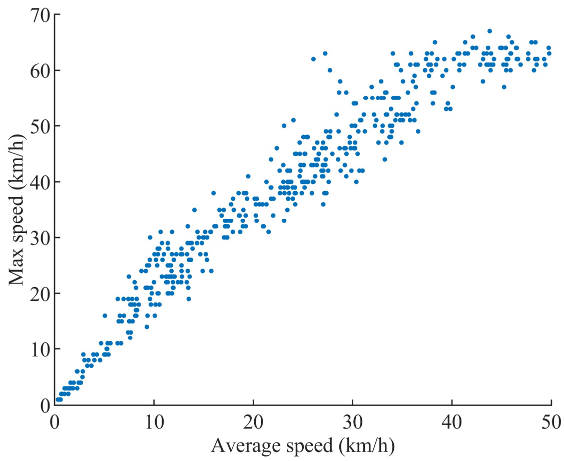


Fig.4 Microtrip database for UA minibuses

$$J = \sum_{i=1}^{n} min_k \|X_i - C_k\|^2 \tag{6}$$

The loss function (*J*) will eventually have an elbow point. According to the elbow point, the best *k* value is selected.

By using K-means algorithm, the sub-Microtrip database is determined and each cluster of the Microtrip is a driving situation. The result is shown in Fig. 5 (a). In Fig. 5 (a), each point represents a Microtrip, and points with the same color mean they are in the same driving situation. In Microtrip Database part, shown in Fig. 5 (b), each figure 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.

B. *Frequency Domain Analysis*

Microtrip is a signal sequence with a limited time. Since OBD receives signals at a frequency of 1Hz, this 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 Equation can be seen in Equation (7)

$$X(k) = \sum_{n=0}^{N-1} x(n) e^{-j\frac{2\pi}{N}}, \; k = 0, \dots, N-1 \tag{7}$$

Using Fast Fourier Transform (FFT) can speed up the transformation time of DFT and optimize the algorithm.

Based on DFT and Parseval theorem [19], as shown in Equation (8), it is known that the energy of the time-domain signal sequence and frequency-domain signal sequence is conserved.

$$\sum_{n=0}^{N-1} |x(n)|^2 = \frac{1}{N} \sum_{n=0}^{N-1} |X(k)|^2 \tag{8}$$

The content $\frac{1}{N}|X(k)|^2$ on the right side of Equation (8) is Periodogram [20]. 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 exactly the fluctuation of driving speed, that is, the embodiment of DA. 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. Also, DA is not altered by the speed mean. 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. This makes DA less sensitive. 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 limit of this high and low frequency is 1Hz [21]. Thus, the numerical expression of driver aggressiveness can be shown by Equation (9).

$$DA = \frac{HF}{LF+HF} \quad (9)$$

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 process of DA on the Microtrip database is shown in Fig. 5 for Frequency Domain Analysis.

As shown in Fig. 5 (c), each Microtrip under different driving situations was analyzed in the frequency domain. Using Equation (9) to calculate the DA value of each Microtrip, the DA of different drivers in different driving situations can be obtained.

C. Fuel Consumption Estimation

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

$$FCR(t) = \frac{MAF(t)}{\lambda(t)\times AFR_{stoich}} \quad (10)$$

By using Equation (10), 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 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.

FC is an important parameter that reflects the vehicle's driving economy. Vehicle fuel consumption is obtained by adding up the times of FCR and time interval:

$$FC = \sum FCR(t) \times \Delta t \quad (11)$$

The data recorded by OBD data loggers is used to obtain the relatively accurate FC of the experimental driving cycle using Equations (10) and (11).

IV. Driving Behavior for Minimus Drivers

According to the methods in Section 3.A and Section 3.B, the driving behavior of the two minibus drivers at the University of Alberta is analyzed.

Fig. 6 shows driving aggressiveness of drivers A and B in different driving situations. In Fig. 6, the x-axis represents the magnitude of the DA value. The most aggressive driving occurs

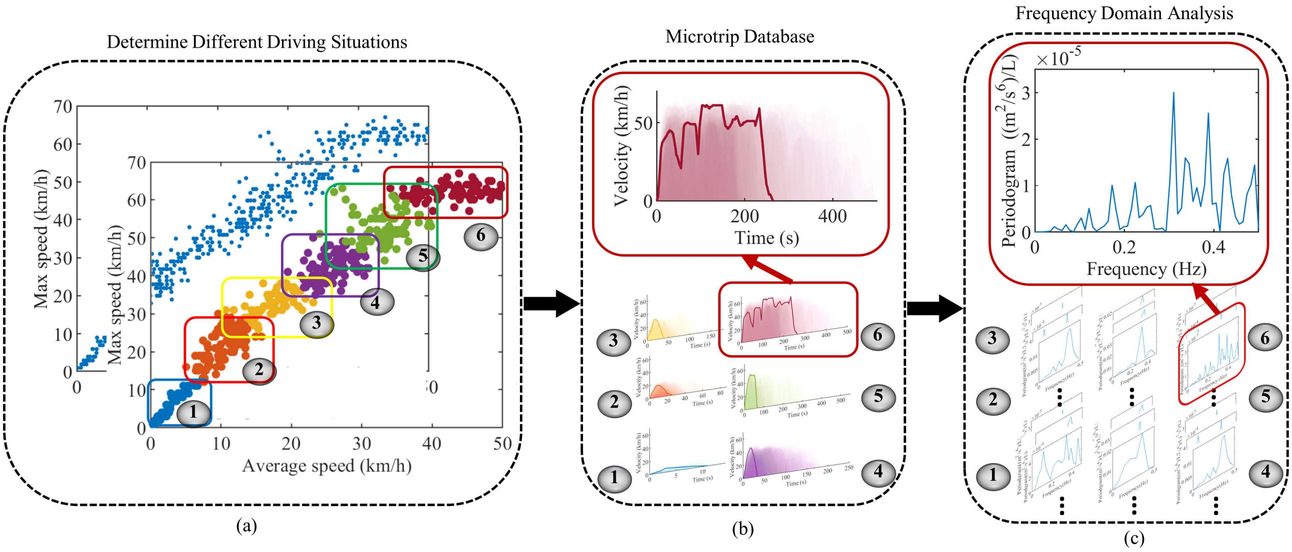


Fig.5 Driving situation division and frequency domain analysis flow

when the DA is close to 1, and the smoother the driving occurs when the DA is close to 0. Each sub-database on the y-axis represents different driving situations. Situation 1 with the blue color represents low-speed creeping driving situation; Situation 2 with the orange color represents creeping driving situation; Situation 3 with the yellow color represents high-speed creeping driving situation; Situation 4 with the purple color represents low-speed urban driving situation; Situation 5 with the green color represents urban driving situation; Situation 6 with the red color represents high-speed urban driving situation.

For each driving situation, each point represents the DA value of the Microtrip under that driving situation. The area plot represents the probability density of Microtrip's DA distribution under each situation. The boxplot in Fig. 6 shows the upper bound, upper quartile, median, lower quartile, and lower bound of the data. It can be seen from Fig. 6 that driver B is more aggressive than driver A in all situations. Such a difference in DA is reflected in fuel consumption, and driver B will consume more fuel. As shown in Fig. 7, the lower bound value of fuel consumed by driver B is greater than the upper bound value of fuel consumed by driver A.The average FC of driver A is 7502 g and the average FC for driver B is 16304 g. The FC of driver B is more than twice that of driver A.

Fig. 8 shows the difference in fuel consumption of drivers A and B under different driving situations. It shows that the difference in fuel consumption caused by the driver's DA is different. In situation 2, the average DA value for the drivers A and B are 0.31 and 0.76 (0.45 difference), and the difference in average FC value is 196 g. However, in the situation 6, the average DA value between the drivers 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 situations, the higher the DA difference, the larger the FC difference; thus, large DA leads to large FC. At the same time, the greater the average speed of the driving situation, the greater the excess FC caused by a same DA. Moreover, Fig. 9 shows that an aggressive driver drives faster than a conservative driver in all driving scenarios. It means that the proposed definition, DA, is consistent with the common sense.

V. Summary and Conclusions

This paper uses the frequency domain analysis method to determine the driving behavior particularly to quantize driver aggressiveness. The method is illustrated for two drivers of the University of Alberta shuttle minibuses. Using the clustering algorithm, different driving situations are divided into six groups, and the driving behaviors of different driving situations are analyzed separately.

Data samples from seventeen 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 is more aggressive in all driving situations. Such a difference in DA value is reflected in the fuel consumption, where the average FC in a driving cycle scale of driver B is about 5000 g more than driver A. When driving situations at higher driving speeds, it usually has higher DA. Moreover, even if the driver's DA values are close under different driving situations (e.g., Driver B's average DA is 0.86 and 0.89 in situation 5 and situation 6 respec-

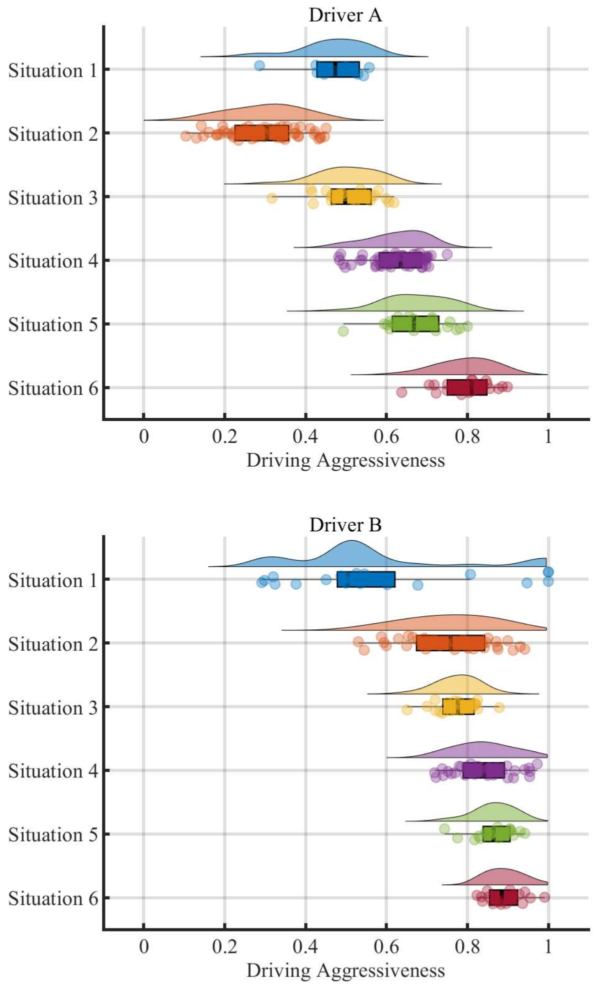


Fig. 6 Driving aggressiveness of drivers A and B in different driving situations

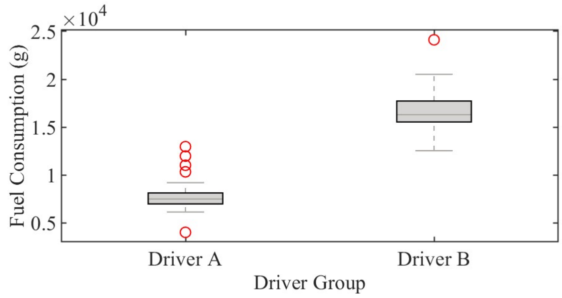
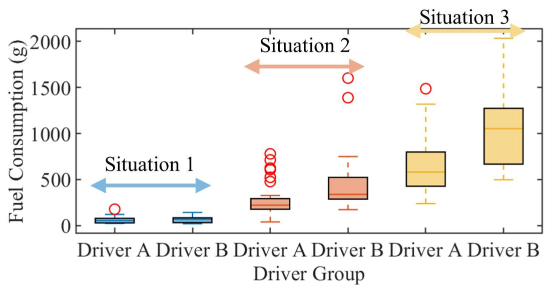


Fig. 7 Differences in fuel consumption by drivers at the driving cycle scale



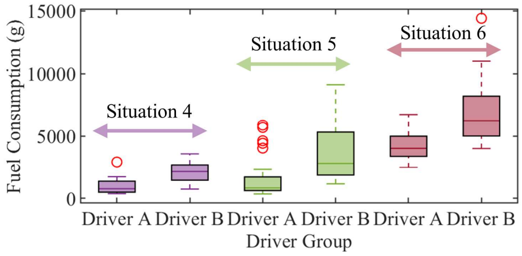


Fig.8 Differences in fuel consumption by drivers at the driving situation scale

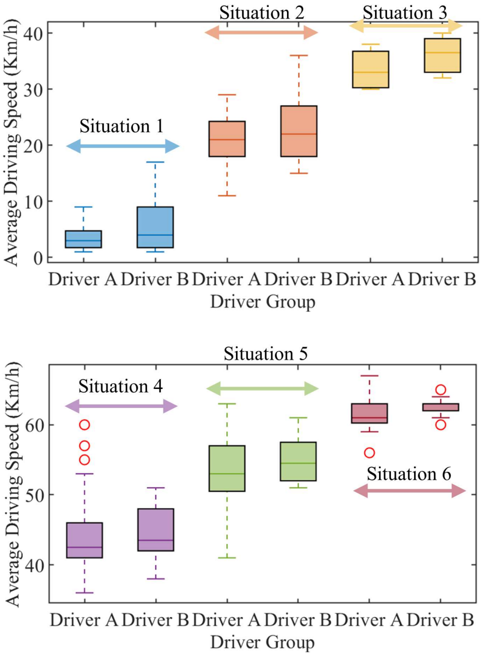


Fig.9 Differences in average driving speed by drivers at the driving situation scale

-tively), the average FC is difference (3427 g difference for Driver B between situation 5 and situation 6).

In future research, we will analyze the differences in driver behavior of other vehicles in the University of Alberta fleet, and create a guideline to assist the fleet drivers to avoid excessively high DA.

ACKNOWLEDGMENTS

The authors would like to thank Jim Laverty, the managers of Transportation Services at the University of Alberta for his support in vehicle testing. The authors also thank Energy Management and Sustainable Operations (EMSO) office of the University of Alberta, especially Michael Versteege for supporting this study. Finally, the authors would like to thank the minibus drivers of the University of Alberta for their cooperation in the study.

REFERENCES

[1] Ma, X., Ma, Z., Zhu, X., Cao, J., & Yu, F. "Driver behavior classification under cut-in scenarios using support vector machine based on naturalistic driving data. " SAE Technical Paper, No. 2019-01-0136. 2019.

[2] Yang, J., Yan, X., Xue, Q., & Li, X. "How do signs and warning influence driving behaviors at flashing light grade crossings?." Journal of Transportation Safety & Security vol.14.5,pp. 843-872, 2022.

[3] Faria, M. V., Baptista, P. C., Farias, T. L., & Pereira, J. "Assessing the impacts of driving environment on driving behavior patterns." Transportation vol. 47.3,2020, pp. 1311-1337.

[4] Linda, Ondrej, and Milos Manic. "Improving vehicle fleet fuel economy via learning fuel-efficient driving behaviors." 2012 5th International Conference on Human System Interactions. IEEE, 2012.

[5] Toledo, Tomer, Haris N. Koutsopoulos, and Moshe Ben-Akiva. "Estimation of an integrated driving behavior model." Transportation Research Part C: Emerging Technologies vol. 17.4, 2009, pp. 365-380.

[6] Xu, W., Wang, J., Fu, T., Gong, H., & Sobhani, A. "Aggressive driving behavior prediction considering driver's intention based on multivariate-temporal feature data." Accident Analysis & Prevention vol. 164, 2022, pp. 106477.

[7] Seo, H., Shin, J., Kim, K. H., Lim, C., & Bae, J. "Driving Risk Assessment Using Non-Negative Matrix Factorization With Driving Behavior Records." IEEE Transactions on Intelligent Transportation Systems vol. 23.11, 2022, pp. 20398-20412.

[8] Karamali, F., Akbari, H., Saberi, H. R., Dehdashti, A., Ziloochi, M. H., Behzadi, M., & Kashani, M. M. "Dangerous driving behaviors among professional drivers of Kashan." International Archives of Health Sciences vol. 7.4, 2020, pp. 215.

[9] Li, J., Zhou, Y., Ge, Y., & Qu, W. "Sensation seeking predicts risky driving behavior: the mediating role of difficulties in emotion regulation." Risk analysis, 2022.

[10] Bingham, Chris, Chris Walsh, and Steve Carroll. "Impact of driving characteristics on electric vehicle energy consumption and range." IET Intelligent Transport Systems vol. 6.1, 2012, pp. 29-35.

[11] Liu, Zifan, Andrej Ivanco, and Zoran 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. 5.2, 2016, pp. 249-258.

[12] Fotouhi, Abbas, and M. J. S. I. Montazeri-Gh. "Tehran driving cycle development using the k-means clustering method." Scientia Iranica vol. 20.2, 2013, pp. 286-293.

[13] Moradi, Ehsan, and Luis Miranda-Moreno. "Vehicular fuel consumption estimation using real-world measures through cascaded machine learning modeling." Transportation Research Part D: Transport and Environment vol. 88, 2020, pp. 102576.

[14] Sundarkumar, G. G., BV, S. B., Munigety, C. R., & Arora, A. S. "A time series clustering based approach for construction of real-world drive cycles." Transportation Research Part D: Transport and Environment 97, 2021, pp. 102896.

[15] P. Couch and Jon Leonard. "Characterization of drayage truck duty cycles at the Port of Long Beach and Port of Los Angeles." TIAX LLC,2011.

[16] Liu, Yang, Amir Ansari, and Mahdi Shahbakhti. "Identification of the Driving Cycle for University Fleet Vehicles." Canadian Society of Mechanical Engineers (CSME) 2022 International Congress, Jun, 2022.

[17] Singh, Archana, Avantika Yadav, and Ajay Rana. "K-means with Three different Distance Metrics." International Journal of Computer Applications vol. 67.10, 2013.

[18] Bhaskaran, V. Murali. "Improving the efficiency of image clustering using modified non euclidean distance measures in data mining." International Journal of Computers Communications & Control vol. 9.1, 2014, pp. 56-61.

[19] Hughes, Harold K. "The physical meaning of parseval's theorem." American Journal of Physics vol. 33.2, 1965, pp. 99-101.

[20] Yang, K., Al Haddad, C., Yannis, G., & Antoniou, C. "Classification and Evaluation of Driving Behavior Safety Levels: A Driving Simulation Study." IEEE Open Journal of Intelligent Transportation Systems vol. 3, 2022, pp. 111-125.

[21] Liu, Zifan, Andrej Ivanco, and Zoran Filipi. "Quantification of drive cycle's rapid speed fluctuations using Fourier analysis." SAE International Journal of Alternative Powertrains vol. 4.1, 2015, pp. 170-177.