Proceedings of the Canadian Society for Mechanical Engineering International Congress 2022
CSME Congress 2022
June 5-8, 2022, Edmonton, AB, Canada

Identification of the Driving Cycle for University Fleet Vehicles

Yang Liu[1*], Amir Ansari[1], Mahdi Shahbakhti[1]
[1]Department of Mechanical Engineering, University of Alberta, Edmonton, Canada
*lyang10@ualberta.ca

Abstract**—**A driving cycle represents the operating conditions of a vehicle as a function of vehicle speed and time. It is used for assessment of vehicle energy consumption, tailpipe emissions, and driving behavior. A driving cycle depends on a vehicle application, geographical regions, and driving zones (e.g, urban vs. highway). This paper collects driving data from the University of Alberta fleet vehicles and develops a driving cycle for fleet vehicles. The driving cycle is generated based on the Microtrip combination method. Extraction of the driving cycle is based on using the Principal Component Analysis method, and developing an algorithm for calculating the weights of different parameters using statistical analysis to avoid the excessive weighting caused by the similarity of the physical meaning of the parameters. The process of combining Microtrips into driving cycles is simplified through database simplification and is accelerated by designing an algorithm for selecting Microtrips. The results are presented for five selected university vehicles with same application. The identified driving cycle consists of 24.5% Acceleration, 23.9% Deceleration, 3.9% Cruising, 47.7% Idling. The resulting driving cycle serves as a utility for the development of an intelligent fleet management system.

Keywords-Driving cycle; Microtrips; Principal Component Analysis; Fleet vehicles

I. INTRODUCTION

The University of Alberta manages over 170 fleet vehicles (hereafter referred to as the UAlberta fleet vehicles). That consume approximately 205,000 liters of fuel annually and thus produces 564,291 kg of CO_2 greenhouse gas (GHG) per year according to the Canada Clean Fuel Standard. The goal is to reduce the operation cost and GHG emission from the UAlberta fleet vehicles via an intelligent management and selection of fleet vehicles. An important need is to determine a representative driving cycle of the UAlberta fleet vehicles. This paper aim to address this need.

A driving cycle is a time series that describes the changes in the vehicle speed within a certain period of time. Driving cycles can be used to (i) evaluate vehicle performance, especially energy consumption and tailpipe emissions, (ii) calibration of vehicle control and energy management strategies, (iii) identify the optimum powertrain framework and vehicle type for certain applications and fleet renewal.

Some of the existing standard driving cycles for light-duty vehicles include Federal Test Procedure cycles FTP 75 and FTP 72, Supplemental Federal Test Procedure US06, Highway Fuel Economy Test (HFET), World harmonized Light-duty vehicles Test Procedure (WLTC), and China Light-Duty Vehicle Test Cycle (CLTC). These standard driving cycles are derived from broad driving condition that are inappropriate to apply to UAlberta fleet vehicles. To this end, this paper aims to collect extensive vehicle data and design an appropriate driving cycle development algorithm for identifying the UAlberta fleet vehicles' driving cycle.

Common methods to develop driving cycles include the Microtrip method [1], clustering method [2], and Markov chain method [3]. In this paper, the Microtrip method is used to construct the driving cycle since it is less complex to develop compared to the Markov method and it is more robust than the clustering methods. Microtrip is a fragment between two idle driving states in a driving cycle [4]. Selection of settings to identify Microtrip is important, because Microtrips are the main building blocks of a driving cycle. Abas et al. [5] and Sundarkumar et al. [6] proposed defining a Microtrip based on vehicle speed and engine speed, respectively. In order to verify whether the driving cycle generated by the Microtrip method is acceptable, assessing driving parameter metrics that reflect the driving characteristics is essential. Xiao et al. [7] and Chugh et al. [8] indicated commonly used parameters for evaluating the performance of generated driving cycles. After obtaining the key parameters, it is essential to quantitatively determine the performance of the generated driving cycle through an algorithm. Gallo et al. [9] proposed the root mean square (RMS) algorithm to quantitatively assess the performance of driving cycles. This algorithm is applied in this study.

The main contributions from this paper include: (i) vehicle instrumentation and collection of extensive vehicle data via a cellular network and on board diagnosis (OBD) data loggers, (ii) developing a new method to form a weight distribution for different parameters of a driving cycle, by mathematically expanding the principal component analysis (PCA) algorithm, (iii) creating an efficient method to identify a driving cycle by simplifying the Microtrip database, and (iv) identifying a driving cycle for the university fleet vehicles.

This paper is organized as follows: Section 2 explains the vehicle experimental setup. The Driving Cycle Algorithm design is described in Section 3. Section 4 showcases the simplification of the database and the final driving cycle is developed in Section 5. Finally, Section 6 presents the summary and conclusions.

II. VEHICLE EXPERIMENT SETUP

A. Data Collection

In this study, Freematics One+ data loggers were used to record data from vehicle. As Fig. 1 shows, each vehicle is equipped with one OBD data logger that collects vehicle and engine information such as vehicle speed and engine speed. Using the cellular option, OBD data loggers are able to send real time data to our designed data server at UAlberta. In the server there are multiple programs which receive and process the data into csv file and subsequently store them in the database. The data for this study is collected for four months of operation from five vehicles. In total, 209,813 data points were obtained. The specification of vehicles are shown in Table 1.

B. Processing and Analysis of Data

1) Data Resampling: The data obtained by OBD loggers sometimes cannot be used directly, due to noise and missing data at some time stamps. The collected data in this study was record by 2 Hz sampling frequency. To this end, zero-phase digital filter was used and the data was resampled to 1 Hz.

2) Determine Microtrips: The conventional method of constructing driving cycles is by "slicing" recorded driving data and combining them according to certain algorithms. Each segment "sliced" from the recorded driving data is called a Microtrip. A properly constructed Microtrip is the cornerstone of a driving cycle. It is necessary to ensure that the information flow of a single Microtrip is as complete and independent as possible, so that when forming a driving cycle, the information flow is not discontinuous. A stop is defined when both vehicle speed and engine speed are zero. To this end, this paper selects the trip between two complete stops of a vehicle as the Microtrip.

3) Kinematic Fragments: A Kinematic fragment [10] is a meticulous division of driving segments compared to Microtrips. It contains acceleration, deceleration, cruise and idle driving modes. The criteria to generate Kinematic fragments are explained in Table 2. As it shows, the driving segments with a certain range of velocity and acceleration are used to determine the driving state. This paper 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 Fig. 2. As it shows, each Microtrip consists of multiple Kinematic fragments.

4) 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 [11]. Some of common metrics include the average velocity of a driving cycle (V_{avg}), 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

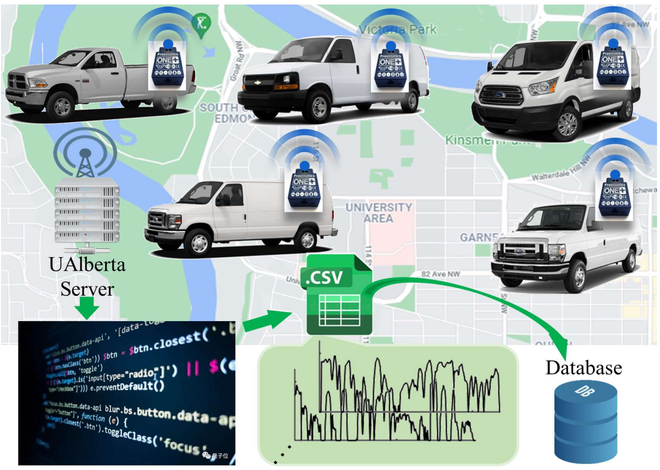


Fig. 1. Designed Framework for Collecting Data from UAlberta Fleet Vehicles

TABLE 1 SPECIFICATION OF THE TESTED VEHICLES IN THE UALBERTA FLEET

Unit No.	Makes	Model	Year	Rated Power (hp)	Engine Size (Liters)
V0429	DODGE	RAM 2500	2011	350	5.7
V0441	FORD	E-250	2012	259	4.6
V0443	FORD	E-250	2012	259	4.6
V0484	CHEVROLET	EXPRESS	2014	341	4.8
V0505	FORD	TRANSIT	2016	310	3.7

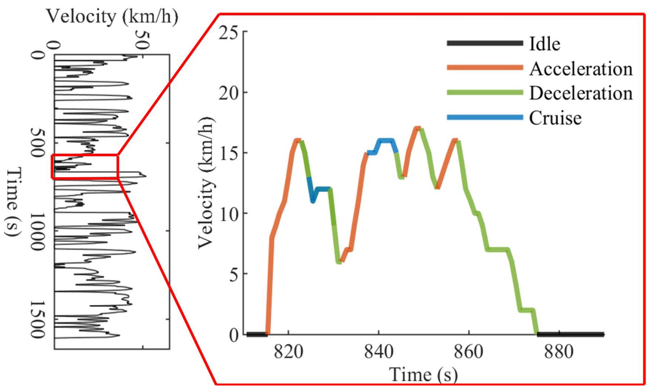


Fig. 2. An Example of Kinematic Fragments in This Study

TABLE 2 PRINCIPLES UESD FOR KINEMATIC FRAGMENTS

Kinematic State	v (km/h)	a (m/s^2)	t (s)	Cruise State with t<5s
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)

stops per kilometer (Stop/km). All driving data recorded by OBD are classified according to Table 2 and mentioned metrics. The results are shown in Table 3. It can be noted that the average speed is significantly lower than that of existing standard driving cycles, and there is less cruising driving.

III. Algorithm to Assign Weights of Target Parameters

Identification of the driving cycle of the fleet can be understood as finding or creating a driving cycle that can represent the characteristics of all recorded data as much as possible. It is a kind of data mining, so an algorithm is needed to test this representativeness.

The mean square error (RMS) algorithm can be used to assess how close the target parameters are compared to each proposed driving cycle. Because RMS can amplify the deviation error between the sample and the characteristic data, the smaller the RMS, the better the sample driving cycle represents the overall driving characteristics. RMS is determined by:

$$RMS = \sqrt{\sum_n \left(\frac{x_i - x_{i.avg}}{x_{i.avg}}\right)^2} \tag{1}$$

x_i = average target parameter for a candidate drive cycle

$x_{i.avg}$ = average target parameter for all OBD recorded data

The RMS does not take into account the weights among different parameters, while the importance of different target parameter metrics is different. Directly using RMS will make the important parameter characteristics insufficiently matched and the secondary parameters occupy too much share of the results. How to assign appropriate weights among target parameters is important.

PCA is usually used as a dimensionality reduction algorithm by using statistical methods to get the correlation among parameters. When using PCA, the parameters are not distinguished as independent variables and dependent variables; thus, all parameters are treated equally. We arranged the parameters into an array X and calculated the covariance matrix Σ of different parameter arrays. The analysis is then simplified by examining arrays in new spaces through mathematical transformations, which is shown in Equation (2).

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

$$cov(Y_i, Y_k) = \alpha_i^T \sum \alpha_k = 0 \tag{3}$$

$$\|\alpha_i\|^2 = \alpha_i^T \alpha_i = 1 \tag{4}$$

TABLE 3 UALBERTA FLEET VEHICLES DRIVING CHARACTERISTICS

Target Parameters	V_{avg} (km/h)	Ve_{avg} (km/h)	Acc_{avg} (m/s^2)	Dec_{avg} (m/s^2)	Stop/km
Value	14.7	21.1	1.9	-2.6	0.5
Target Parameters	**%Acc**	**%Dec**	**%Cruise**	**%Idle**	
Value	27.0	32.8	6.4	33.7	

In Equation (2), a_{ij} represents the value of row "i" and column "j" in the covariance matrix Σ. α_i in the equation is the principal component (PC), which is also the eigenvector $\boldsymbol{e_i}$ of the covariance matrix Σ. Equations (3) and (4) show that the PC vectors in the new space are pairwise orthogonal. Moreover, it is the orthogonality of this vector that greatly simplifies the research difficulty of the problem. We used α to get the matrix of feature vector A as Equation (5):

$$A = \begin{bmatrix} \alpha_1^T \\ \alpha_2^T \\ \cdots \\ \alpha_p^T \end{bmatrix} = \begin{bmatrix} a_{11} & \dots & a_{1p} \\ a_{21} & \dots & a_{2p} \\ \cdots & \cdots & \cdots \\ a_{p1} & \dots & a_{pp} \end{bmatrix} \tag{5}$$

The eigenvector λ_i is obtained by calculating the variance of Y_i from Equations (6) to (8).

$$Y_i = \alpha_i^T X \tag{6}$$

$$var(Y_i) = max\|\alpha_i\| = 1\alpha^T \sum \alpha \quad (with\ \alpha_i^T \sum \alpha_j = 0) \tag{7}$$

$$var(Y_i) = \alpha_i^T \sum \alpha_i = \lambda_i \tag{8}$$

Next, we divide a single λ_i by the $\sum \lambda_i$ 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 (11) is obtained that yields the matrix of proportional feature vector.

$$\lambda_i\% = \frac{\lambda_i}{\sum \lambda_i} = \left(\frac{\lambda_i}{\lambda_1 + \cdots + \lambda_i + \cdots \lambda_p}\right) \times 100\% \tag{9}$$

$$a_{i,j}\% = \frac{a_{i,j}}{\sum_{i=1}^{p} a_{i,j}} \times \lambda_i\% \tag{10}$$

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

By summing of each row of the matrix of proportional feature vector $A\%$, the weight of each parameter is obtained.

$$W_i = \sum_{j=1}^{p} a_{i,j} \tag{12}$$

Different from the traditional PCA algorithm, the proportional covariance matrix can be used to obtain eigenvalues λ_i for dimensionality reduction in the new space and to obtain the weight W_i of each parameter. For the proportional coefficient matrix, the sum of each column is λ and the sum of each row is the weight of each parameter. Fig. 3 shows the weights for the target parameters.

Based on the above Equations (1) to (12), the revised RMS

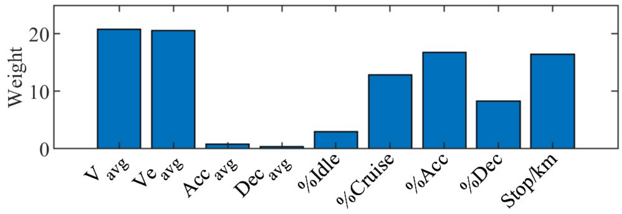


Fig. 3 Selected Weights for the Target Parameters

is determined by:

$$RMS_{revised} = \sqrt{\sum_n \left(\frac{x_i - x_{i.avg}}{x_{i.avg}}\right)^2 \times W_i} \qquad (13)$$

IV. MICROTRIP DATABASES

The construction of the driving cycle for the UAlberta fleet in this paper faced the problem of excessive number of Microtrips. Therefore, it is important to simplify the Microtrip database through an algorithm to have an efficient operation. To this end, at first Microtrips are divided into several categories, and then simplify each category, and finally get a simplified database. Through this method of categorizing and simplifying, the database can be simplified quickly, and the problem of excessive simplification in the simplification process can be avoided.

Average speed is an important parameter to reflect the driving state. For example, when the average vehicle speed is close to 0, a vehicle is usually at an idle state, and when the average speed is high, the vehicle is usually in a cruising state. Most of the others are in the aggression driving state, which is mainly composed of acceleration and deceleration kinematic fragments. Therefore, using the average speed to distinguish the Microtrips can let it be classified according to the driving state without being too complicated.

The traditional method of dividing the driving state by the average speed is based on evenly divided speed range method. However, the driving situation of the UAlberta fleet vehicles is very different from that of urban vehicles. This division method may lead to uneven distribution of Microtrips in different driving state intervals. In this project, if using an evenly divided method the result is shown in Table 4. The percentage of Microtrips in each speed division is defined as %No in Table 4 and Table 5. The %No with an average speed in the range of 0 to 5 km/h is 9% more than that in the range of 25 km/h to remaining. Such division method will lead to that some databases have a large number of samples, and those database simplification will miss many data information.

To solve this problem, a driving state division method based on number of Microtrip is proposed by Equation (14), where the evenly divided Microtrip number (EDMN) method aims to minimize the error between the %No and total Microtrip divided by the number of intervals. In order to have enough samples for each driving state intervals, in this study, it is stipulated that the number of Microtrips in each driving state interval cannot be less than 300. This leads to 6 driving state intervals listed in Table 5. The %No in each interval is not exactly equal, because the number of Microtrips cannot be divided by N without leaving a remainder, which is also why EDMN is to find the minimum error.

$$EDMN = min\left(\sum\left(\%No - \frac{100}{N}\right)\right) \qquad (14)$$

N = number of driving state intervals

After getting a division of the driving state, RMS and PCA from section 4 are applied to get the five most representative Microtrips in each driving status interval to simplify the Microtrip database. Fig. 4 shows the Microtrip database. In each

TABLE 4 EVEN DISTRIBUTION OF SPEED FOR FLEET DATA

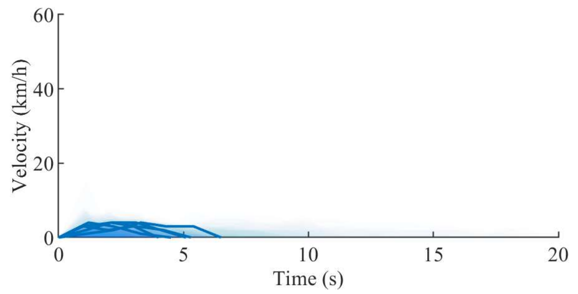
Average Speed Range (km/h)	0~5	6~10	11~15	16~20	21~25	>25
% No	21.8	15.1	18.8	18.2	13.9	12.0

TABLE 5 EVEN DISTRIBUTION OF NUMBER OF MICROTRIP FOR FLEET DATA

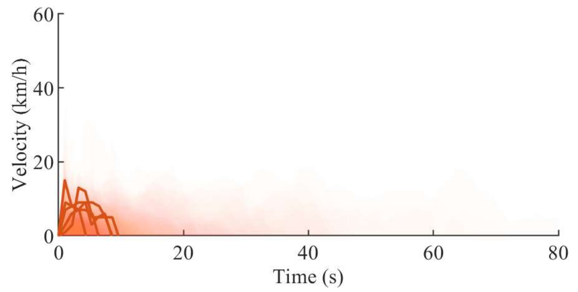
Average Speed Range (km/h)	0~3	4~8	9~13	14~17	18~23	>23
% No	15.8	14.4	18.2	16.2	18.8	16.7

driving state categories, the solid lines are the best representative Microtrips. The remaining Microtrips are represented by the area chart, and the darker the color, the more frequent the driving situation in the range appears.

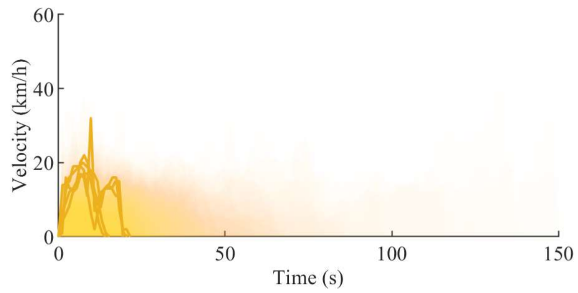
Fig. 4a shows that most of the Microtrip fragments' speed in this driving state interval is closed to 0 and Fig. 4f shows that most of the Microtrip driving state intervals have a large number of kinematic fragments in high-speed cruising state compared with other situations. This also shows that the theory that the driving state can be roughly divided according to the average



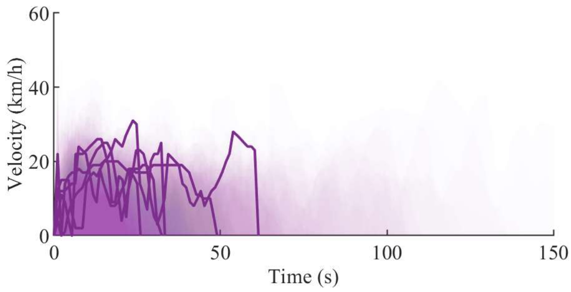
(a) 0~3 km/h



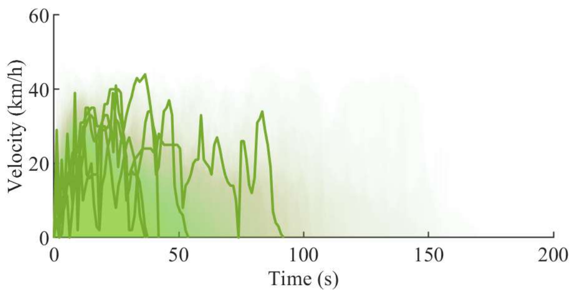
(b) 3~8 km/h



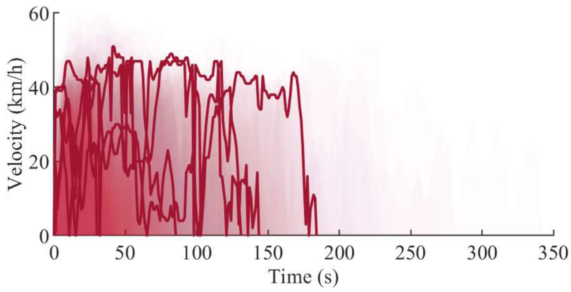
(c) 8~13 km/h



(d) 13~17 km/h



(e) 17~23 km/h



(f) 23 km/h ~all remaining

Fig. 4. Microtrip Databases for Each Driving Category

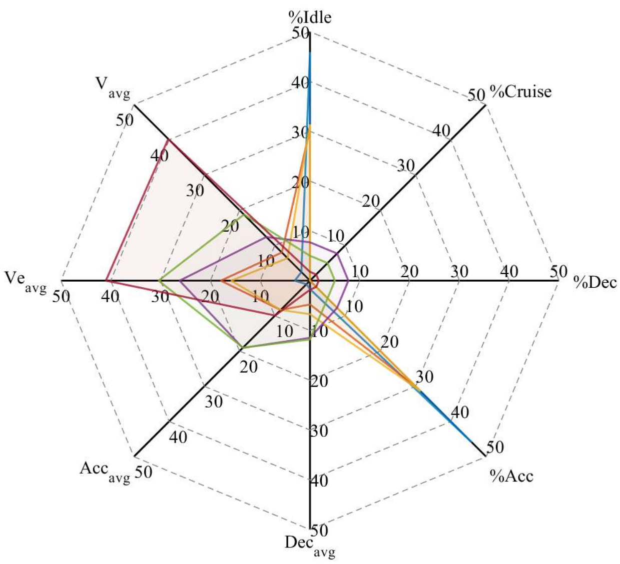


Fig. 5 Weights of Traget Parameters for Microtrip Database

speed is correct. Fig. 5 shows the weights of target parameters in Microtrip database. Each color represents a Microtrip database category, and the Microtrip driving state category represented by each color is consistent with that in Fig. 4.

V. CONSTRUCTION OF DRIVE CYCLE

Based on the Microtrip database from section iv, a driving cycle is constructed according to the algorithm illustrated in Fig. 6. Different from the conventional completely random driving cycle construction method [1], which randomly combines Microtrips and then judge the performance to determine whether it is available. The algorithm in Fig. 6 uses the average speed of Microtrips to accelerate the process of generating the driving cycle.

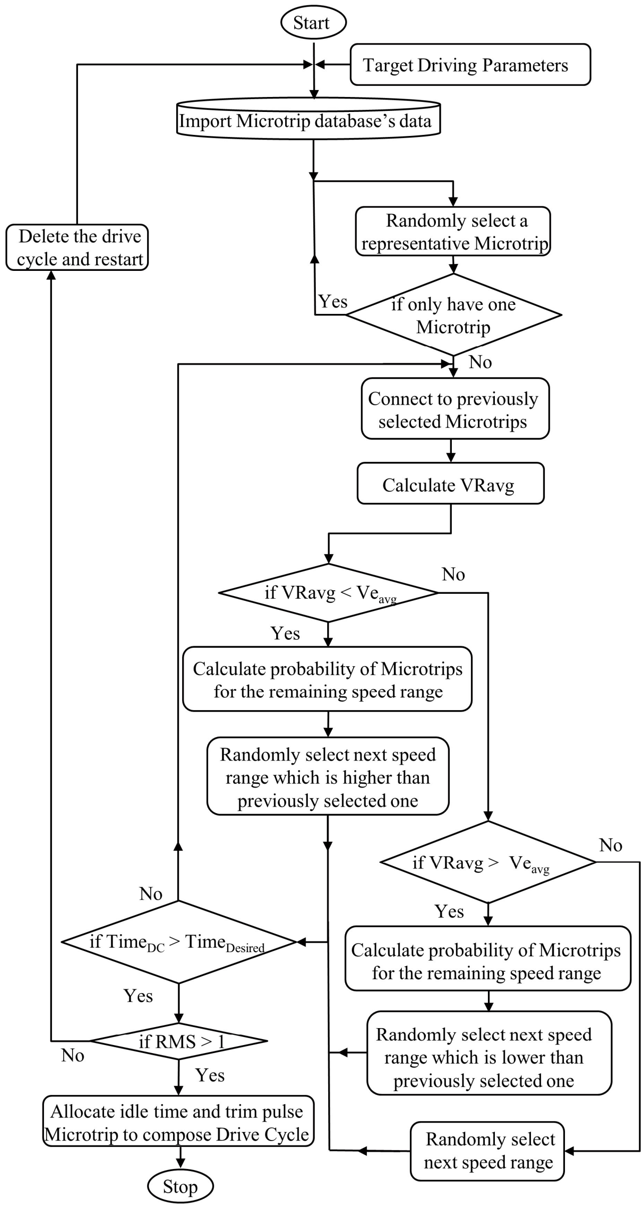


Fig. 6 Flowchart of the Designed Algorithm to Construct the Driving Cycle

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 idle driving state. $Time_{DC}$ in the algorithm is the time duration of the Microtrip combined driving cycle.

As Fig. 6 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 of the generated driving cycle is calculated, if RMS 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 time 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 time. In this study, the cycle time is 1761 seconds. From all OBD recorded driving cycles, the idle driving state time accounted for 33.7% of the total time. Thus, in this study, 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. The final driving cycle is shown in Fig. 7. RMS is 0.27, which means the proposed driving cycle in Fig. 7 can well represent the driving characteristics of the tested fleet vehicles.

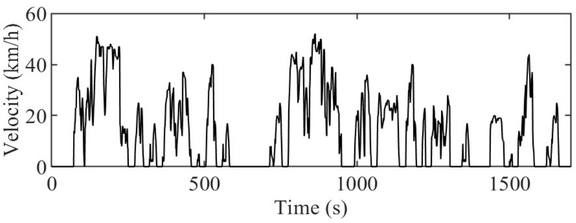


Fig. 7 Final Driving Cycle for the Tested UAlberta Fleet Vehicles

VI. SUMMARY AND CONCLUSIONS

In this study, a driving cycle was determined by collecting and analysis of extensive data from the UAlberta fleet vehicles. A root mean square algorithm with a weight correction and a Microtrip combination algorithm were developed and applied on the designed database. The design of optimum weights in the RMS algorithm improved the accuracy of the algorithm. In addition, the simplification conducted on the database improved the efficiency of the designed algorithm. The final driving cycle has an average velocity of 12.6 km/h, the maximum velocity of 52 km/h, with a total duration of 1761 seconds.

Future work includes collecting data from fleet vehicles with different applications and derive a specific driving cycle for each vehicle application, utilizing the algorithm from this study.

ACKNOWLEDGMENT

The authors would like to thank Jim Laverty, the manager and the staff of Transportation Services at University of Alberta for their support in vehicle testing. The authors also thank Energy Management and Sustainable Operations (EMSO) of the University of Alberta, especially Michael Versteege and Shannon Leblanc for supporting this study.

REFERENCES

[1] M. André, R. Joumard, A. Hickman and D. Hassel, "Actual car use and operating conditions as emission parameters: derived urban driving cycles", Science of The Total Environment, vol. 146-147, pp. 225-233, 1994.

[2] Fotouhi A, Montazeri-Gh M, "Tehran driving cycle development using the k-means clustering method". Scientia Iranica, 2013, 20(2): 286-293.

[3] Q. Shi, Y. Zheng, R. Wang and Y. Li, "The study of a new method of driving cycles construction", Procedia Engineering, vol. 16, pp. 79-87, 2011.

[4] G. Tzeng and J. Chen, "Developing A Taipei motorcycle driving cycle for emissions and fuel economy", Transportation Research Part D: Transport and Environment, vol. 3, no. 1, pp. 19-27, 1998.

[5] M. Abas, S. Rajoo and S. Zainal Abidin, "Development of Malaysian urban drive cycle using vehicle and engine parameters", Transportation Research Part D: Transport and Environment, vol. 63, pp. 388-403, 2018.

[6] G. Ganesh Sundarkumar, S. B. V., C. Munigety and A. Arora, "A time series clustering based approach for construction of real-world drive cycles", Transportation Research Part D: Transport and Environment, vol. 97, p. 102896, 2021.

[7] 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.

[8] S. Chugh, P. Kumar, M. Muralidharan, et al. "Development of Delhi driving cycle: a tool for realistic assessment of exhaust emissions from passenger cars in Delhi". SAE Technical Paper, 2012.

[9] J. Gallo, "Development of Representative Regional Delivery Drive Cycles for Heavy-Duty Truck Tractors", SAE International Journal of Commercial Vehicles, vol. 7, no. 1, pp. 337-347, 2014.

[10] 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. 115408, 2020.

[11] 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).