

**ORIGINAL ARTICLE**

## Analysis of Driving Parameters using Pearson Correlation and K-means Cluster: A case study of Sarawak, Malaysia

\*<sup>1</sup>Lim Kian Kiat, <sup>1,2</sup>Mastura Bujang and <sup>1,2</sup>Mohammad Shahril Osman

<sup>1</sup>School of Engineering and Technology, University of Technology Sarawak, 96000 Sibul, Sarawak, Malaysia

<sup>2</sup>Center of Research for Innovation and Sustainable Development (CRISD), University of Technology Sarawak, 96000 Sibul, Sarawak, Malaysia

**ABSTRACT** - Driving Cycle is widely adapted by automotive industry in evaluating vehicle fuel consumption and emission towards manufacturing of a sustainable and efficient vehicle. It has become a debatable issue in automotive industry due to tailorable features in line with driving and road condition of a specific place. This paper intends to clarify on the selection of representative driving parameters from the pool of microtrips based on the data gathered throughout Sarawak continent. Joint method of car chasing and circulation techniques were employed during data collection. Routes were selected by assessing the traffic volume extracted from Road Traffic Volume Malaysia (RTVM) and three traffic conditions were incorporated. Data collection had extended to collect three different periods of the week namely peak hour, off-peak hour and weekend to account potential traffic condition. Pearson product-moment correlation was applied to express the correlation between parameters and K-means cluster method in order to classify the microtrips. Through analysis, it was convinced that percentage of idle is most suited to be used to evaluate representativeness of microtrips. These findings could significantly contribute to developing a comprehensive and representative region-specific driving cycle.

**ARTICLE HISTORY**

Received: 03 Sept 2021

Revised: 08 July 2022

Accepted: 10 Aug 2022

**KEYWORDS**

*Driving Cycle*

*Pearson Correlation*

*K-mean Clustering*

*Percentage of Idle*

*Driving Parameters*

**INTRODUCTION**

In this era where technological advancements becoming more prominent, vehicle has become necessary as a means of transportation for individual to travel over a distance. The demand over this type of transportation shows only sign of inclination over the statistics. Sustainable and affordable energy has become a variable to be considered by vehicle industry as a way to improve their product. Countless effort has been invested in the trials to explore new and ground-breaking ways to enhance vehicle consumption and reduce emission. Driving cycle had been established to tackle such problem served as a standard for vehicle efficiency evaluation. Malaysia is one of the few countries that are currently implementing New European Driving Cycle (NEDC) for vehicle fuel consumption and carbon emission evaluation. Huge disparity appears on vehicle efficiency evaluation by using NEDC which was developed based on Europe conditions due to driving conditions heavily depends road condition and differs widely from one region to another. Hence, comprehensive studies on driving parameters extracted from kinematic segment are deemed necessary also served as a fundamental procedure for developing driving cycle.

In recent years, researcher had assessed through various vehicular parameters to construct a representative driving cycle that could explain realistic driving situation tailored to that region. Shi Qin et al. [1] used joint clustering method, Fuzzy C-means clustering method and principal component analysis to scrutinize the actual road conditions based on 10 parameters in Hefei City and developed typical driving conditions in Hefei urban area. Zhou et.al. [2] implement similar approach developed Shenyang

Driving Cycle with 8 driving parameters categorize into 4 driving mode namely acceleration, deceleration, idle and cruise. Lee and Filipi [3] had successfully captured 27 explanatory variables from raw cycle with transition probability matrices. The variables had categorized into velocity related, acceleration related, driving time and distance related, and driving characteristics related in which reduced using regression analysis into eight significant explanatory variables. There are also researchers who added new variable such as Vehicle Specific Power (VSP) which reflect vehicle emission better by incorporating engine power demand during driving but require extraordinary amount of data collection on dynamic variable of vehicle [4]. Ho et al. [5] developed Singapore Driving Cycle (SDC) by evaluating typical trip distance, proportion of expressway driving, proportion of arterial roads driving, peak and lull proportions, and trip duration and throughout the process, 20 driving parameters were shortlisted.

This paper aims to unearth the correlation between driving parameters extracted from raw kinematic segment taken from the actual road driving data throughout Sarawak region. Sarawak rank as the largest state and had cover a total road length of 35516.5 km in the year of 2016 in Malaysia [6]. The absence of expressway within the state had make it exceptional unique which contribute to perceptive driving parameters analysis. The result will be brought forward to K-mean clustering analysis for further insight.

## **DATA COLLECTION**

This research conducted with two unit of four stroke gasoline engine namely Perodua Bezza with engine capacity of 1.3L and Perodua Myvi with engine capacity of 1.5L as shown in Figure 1 and Figure 2. The test vehicles were provided by Malaysia Automotive Robotics and IOT Institute (MARI). Prior to the test, the vehicles were checked according to the manufacturer standards (tire pressure and wheel alignment) to ensure data consistency. The test parameters include time and velocity were captured in 1Hz frequency by control area network (CAN) bus data acquisition unit (NR-C512) with multiple modules include AC power supply (NR-U60), temperature measurement unit (NR-TH08) aid in measuring ambient and cabin temperature and data logger (NR-600). The connection ends with a laptop with proprietary software allow user to monitor acquisition process and to ensure it run smoothly. Raw data were saved in the format of Comma Separated Values (CSV) file. Data acquisition device configuration is as illustrated in Figure 3.



**Figure 1:** Perodua Bezza 1.3L



**Figure 2:** Perodua Myvi 1.5L

This research implemented joint method of chase car and circulation driving method to gather real world data. Chase car method is a method adapted from Johnson et al. [7] as a guideline for driving pace along the route. The driver is required to choose and follow the movement of the particular type of vehicle for as long as possible during testing to simulate the driving patterns of the targeted vehicle. Once target vehicle driven off the area of study, the driver has to choose a new target to follow. Circulation driving method adopted from Yu et.al. [8] to ensure that the data uniformity require driver drove on the same route more than once. It allowed data to be compared at the end of the process and possibly eliminate redundant data.



**Figure 3:** Data acquisition device

### Route Selection

In Sarawak, the huge number of possible routes are impractical to be conducted in term of real assessment of the vehicle speed characteristics on the proposed routes. A productive way to overcome this matter is to choose a number of routes that depict the dominant traffic situation throughout Sarawak. Traffic volume data from Road Traffic Volume Malaysia (RTVM) version 2016 and 2018, published by the Ministry of Works Malaysia were utilized to further enhance data accuracy superior as compared to relying Google Map derived traffic quality. In some of the cases, Google Map had been use to observe the traffic quality data. Nonetheless, Google Map data may suffer inaccurate information in regards with unusual traffic conditions due to lack of up-to-the-minute information. Hence, it was limited to identification of road length and direction.

The quality of traffic flow for each route is sorted based on the level of service (LOS). LOS is an approach presented by Highway Capacity Manual (HCM) to specify the quality level that can be obtained through various operation characteristics and traffic volume. There is a total of six LOS distinct by HCM which are A, B, C, D, E, and F. LOS F denotes the poorest quality meanwhile LOS A denotes the best quality of service. Should any of these selected routes is in the same network or chains, the road has had

to be discarded as described in [9]. In this study, seven routes in Sarawak had been chosen as illustrated in Table 1.

The samples were collected from three different periods of the day namely weekday morning and evening peak hour, off peak and weekends. Morning peak hour (0600 till 0800) and evening peak hour (1700 till 1900) as extracted from RTVM. The purpose of the inclusion of off-peak hour and weekends is to secure larger range of traffic condition.

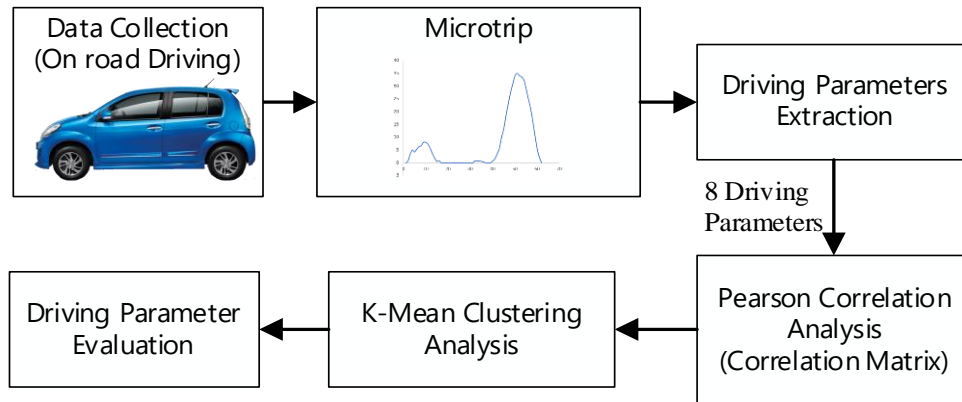
**Table 1.** Selected routes for Sarawak region

District	Census Station	Coordinate		Road Length (km)	LOS
		From	To		
Sibu	SR304	2.290048, 111.898950	2.322085, 111.837176	8.5	D
Miri	SR403	4.338328, 114.000513	4.399819, 113.995885	9.0	D
Miri	SR404	4.442542, 114.013836	4.532286, 113.986244	11.2	A
Kuching	SR102	1.418516, 110.159757	1.444044, 110.231882	9.8	F
Kuching	SR106	1.462454, 110.326631	1.567480, 110.321132	13.4	F
Kuching	SR108	1.447796, 110.330739	1.457254, 110.442484	13.8	F
Kuching	SR110	1.489663, 110.340352	1.545671, 110.379875	8.6	A

### Microtrip

The microtrip refers to segmented vehicle trip that begin and ends with two successive idle period. This segment of trip is be further analyze for vehicle parameters extraction. Filtration of microtrips performed to enhance statistical representativeness of microtrips [10]. The process flow of driving parameters evaluation is as illustrated in Figure 4. Criteria of the filtration are defined according to following rules:

- (1) Microtrip with duration less than 10s
- (2) Microtrip with maximum speed less than 3.33 m/s
- (3) Microtrip with acceleration higher than 4 m/s<sup>2</sup>
- (4) Microtrip with deceleration higher than 4.5 m/s<sup>2</sup>



**Figure 4:** The flowchart of parameters evaluation process

The selection of driving parameter is a vital step in measuring the representativeness of a microtrip. The number of parameters should not be too limited otherwise bias would likely to occur which could lead to improper representativeness evaluation. Nonetheless, utilizing numerous parameters would add unnecessary burden in analysis process. In this study, driving parameters are group into three categories namely velocity related, acceleration related and time related. The selection of parameters is shortlisted from [2, 5, 11, 12] illustrated as shown in Table 2. The average speed reported as parameter that have an important bearing on emission, hence primary target is to compute correlation between parameters and average speed [13]. Extensive study was carried out accordingly on the correlation between each parameter to express data consistency.

The equations of the selected parameters are defined:

$$a = \sum_{i=1}^n (v_i \cdot t_i) \div \sum_{i=1}^n (t_i) \tag{1}$$

$$v_{avg} = \left( \frac{d}{t} \right) 3600 \tag{2}$$

where  $a$  represents acceleration,  $v$  represents velocity,  $d$  represents distance,  $t$  represents time;  $a_i > 0$  m/s<sup>2</sup> and  $d_i < 0$  m/s<sup>2</sup> for all  $i = 1, 2, 3, \dots, n$  and  $v_i =$  instantaneous velocity at time  $i$

The time parameters defined as

$$T_{i,j} = T_{t,j} - T_{c,j} - T_{a,j} - T_{d,j} \tag{3}$$

$$T_{c,j} = T_{t,j} - T_{a,j} - T_{d,j} - T_{i,j} \tag{4}$$

$$T_{a,j} = T_{t,j} - T_{c,j} - T_{d,j} - T_{i,j} \tag{5}$$

$$T_{d,j} = T_{t,j} - T_{c,j} - T_{a,j} - T_{i,j} \tag{6}$$

where  $T_t$  represents total trip time,  $T_i$  represents total idle time,  $T_c$  represents total cruising time,  $T_a$  represents total acceleration time and  $T_d$  represents total deceleration time for individual microtrip  $j$  and  $j = 1, 2, 3, \dots, n$ .

The percentage of each time parameters are defined as

$$P_{i,j} = \frac{T_{i,j}}{T_j} \tag{7}$$

$$P_{c,j} = \frac{T_{c,j}}{T_j} \tag{8}$$

$$P_{a,j} = \frac{T_{a,j}}{T_j} \tag{9}$$

$$P_{d,j} = \frac{T_{d,j}}{T_j} \tag{10}$$

where  $P_i$  represents percentage of idle,  $P_c$  represents percentage of cruising,  $P_a$  represents percentage of acceleration and  $P_d$  represents percentage of deceleration time for individual microtrip  $j$  and  $j = 1, 2, 3, \dots, n$ .

**Table 2.** Driving Parameters

	Parameters	Connotation
Velocity related	$v_{avg}$	Average Velocity
	$v_{max}$	Maximum Velocity
Acceleration related	$a_{max}$	Maximum Acceleration
	$a_{min}$	Minimum Acceleration
Time related	$P_i$	Percentage of Idle
	$P_a$	Percentage of Acceleration
	$P_d$	Percentage of Deceleration
	$P_c$	Percentage of Cruising

## K-mean Clustering Analysis

K-mean clustering analysis allows data clustering to aggregate together due to certain resemblances. The Euclidean Distance between individual microtrip and microtrips cluster center is measured and taken as the main function of optimization. Thus, it is adopted to classify microtrips extracted. The calculation steps are as follows:

- (1) Decide number of K from the pool
- (2) Initialize the K cluster centers (randomly, if necessary)
- (3) Compute the Euclidean distance between each microtrip and cluster center and divide the corresponding microtrip based on the least Euclidean distance
- (4) Re-estimate cluster centers, by assuming the microtrip are in right position
- (5) Repeat step (3) and (4) until none of the microtrip change their position until last iteration

## RESULT AND DISCUSSION

Pearson product-moment correlation is a statistical evaluation tool to compute linear association between two parameters. The correlation expresses any value range from 1 to -1. Numbers greater than 0 indicate positive association and vice versa. Correlation is defined as

$$r = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum (x_i - \bar{x})^2 \sum (y_i - \bar{y})^2}} \quad (11)$$

where  $r$  represents correlation coefficient,  $x$  represents values of the x-variable of the sample,  $y$  represents values of the y-variable of the sample for all  $i = 1, 2, 3, \dots, n$ ,  $\bar{x}$  represents mean values of the x-variable and  $\bar{y}$  represents mean values of the y-variable.

### Average Speed Correlation

From Table 3, two parameter shows negative correlation and five positive correlations with  $v_{max}$ . According to the rules of thumb for correlation coefficient the correlation with  $v_{avg}$ ,  $v_{max}$  were categorize as strong correlation with value between 0.70-0.89;  $P_i$ ,  $P_a$  and  $P_d$  were classify as moderate correlation with correlation coefficient value of 0.40-0.69 and  $P_C$ ,  $a_{max}$  and  $a_{min}$  were at weak correlation with correlation coefficient of 0.10-0.39 [14]. Prior to correlation computation, significance test was performed with 95% confidence level. Significance test result shows  $p = <.05$  where significant correlation is concluded for all studied parameters. However,  $P_d$ ,  $P_C$ ,  $a_{max}$  and  $a_{min}$ , there were no evidence found regard to emission impact hence regression scattered plot were executed for visual inspection of relationship [15].

The correlation of  $P_i$  and  $P_a$  shows -0.664 and 0.642 respectively in which for  $P_i$  depicted declination correlation with  $v_{max}$  while  $P_a$  shows otherwise. Kent et al reported that idle time percentage have the greatest impact on emissions while Unal et.al had concluded that acceleration have direct impact to vehicle tailpipe emissions, thus both parameters will be scrutinized further in this study [16, 17].

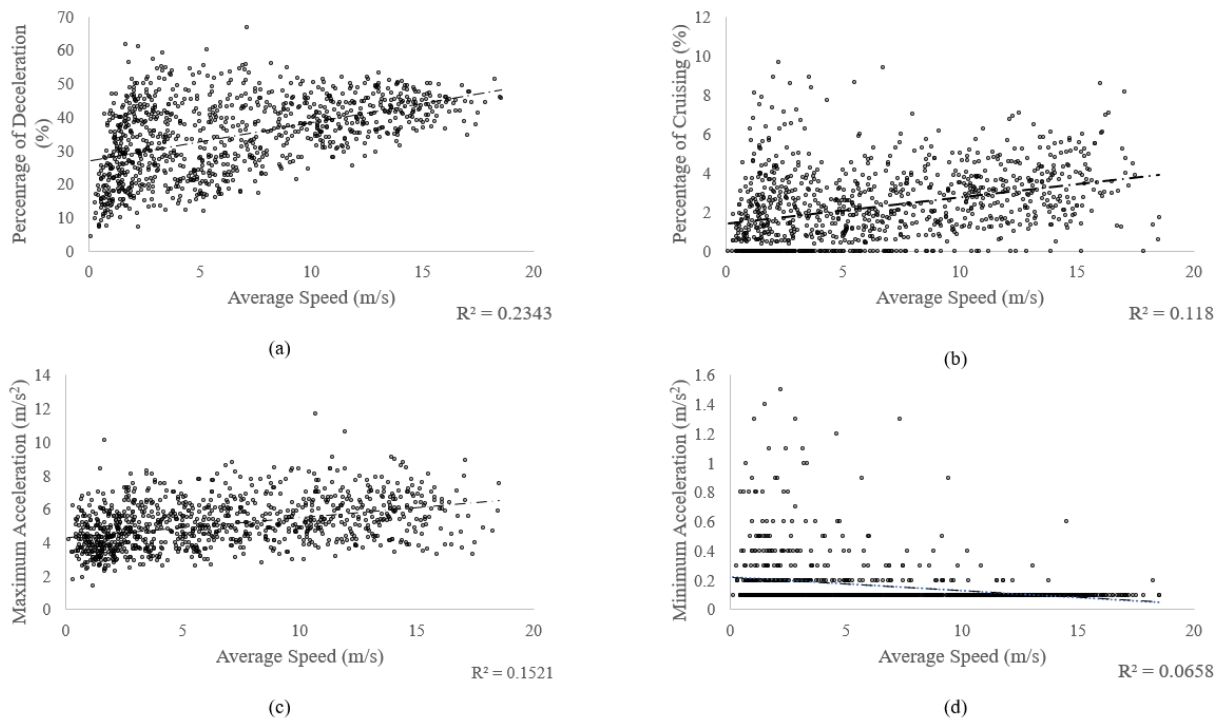
On the other hand, the correlation coefficient shows 0.867 between  $v_{max}$  and  $v_{avg}$  which is the highest correlation among parameters. Wang et al. studies the impact of  $v_{max}$  toward vehicle emission, however there are no evidence showing that  $v_{max}$  were utilized in developing driving cycle [18]. Hence, the parameter was only presence to evaluate data consistency in later stage.

**Table 3.** Correlation of average speed with studied parameters

Pearson product-moment correlation								
	$v_{avg}$	$P_i$	$P_a$	$P_d$	$P_C$	$a_{max}$	$a_{min}$	$v_{max}$
$v_{avg}$	1	-0.664	0.642	0.484	0.343	0.390	-0.257	0.867

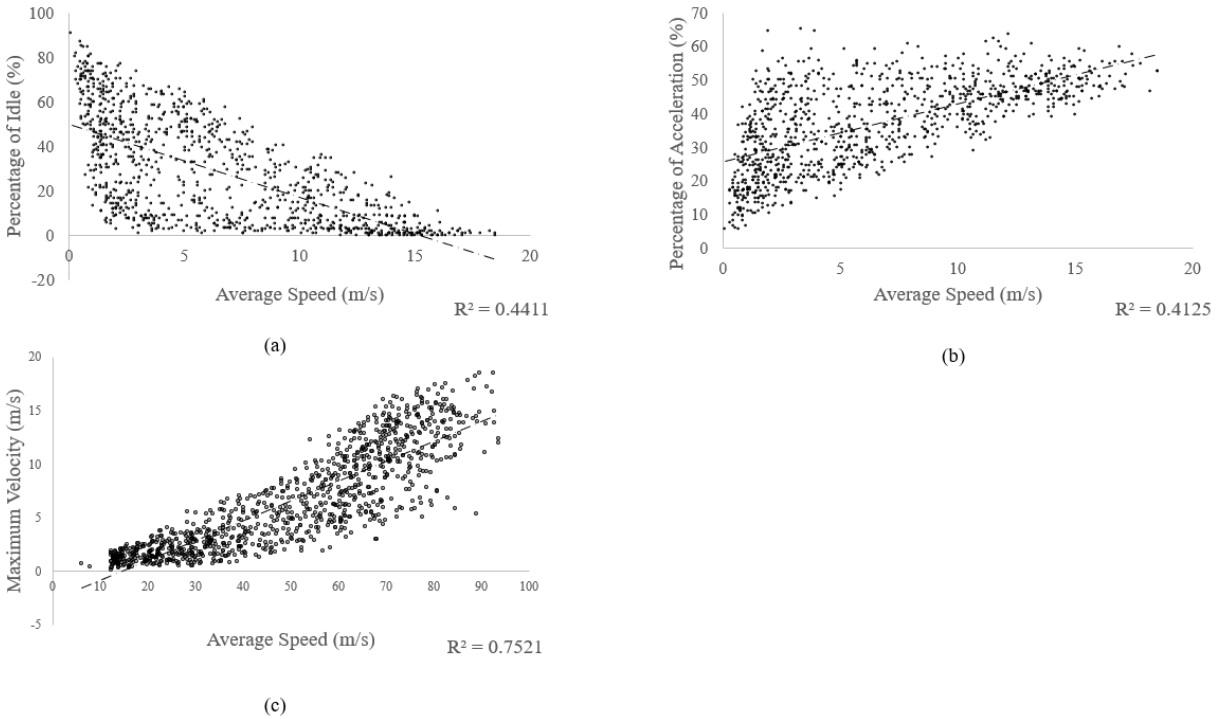
**Parametric Relationship Scatterplot**

The scatterplot from Figure 5 illustrated  $P_d$ ,  $P_C$ ,  $a_{max}$ , and  $a_{min}$  appears to possessed low  $R^2$  value. Furthermore, the scatterplot shows random patterns with scattered data which lead to undesirable for clustering purposes. Nonetheless, scatterplot for  $P_d$  was found to be resemble of scatterplot from  $P_a$  with higher  $R^2$ , thus operation was only focuses on acceleration rather than deceleration for greater segregation.



**Figure 5:** Scatterplot of (a)  $P_d$  versus  $v_{avg}$  (b)  $P_C$  versus  $v_{avg}$  (c)  $a_{max}$  versus  $v_{avg}$  (d)  $a_{min}$  versus  $v_{avg}$

Figure 6 illustrated the scatterplot of  $P_i$ ,  $P_a$  and  $v_{max}$  in which the show moderate relationship in term of  $P_i$  and  $P_a$  with  $R^2$  value of 0.4411 and 0.4125 respectively while strong relationship with  $v_{max}$  showing  $R^2=0.7521$ . Furthermore, both scatterplot of  $P_i$  and  $P_a$  depict greater segregation with “waterfall” pattern.



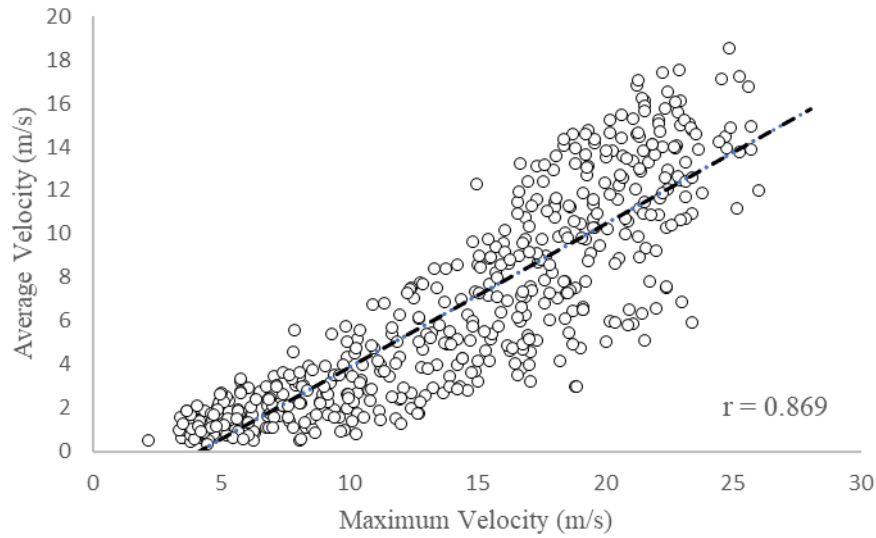
**Figure 6:** Scatterplot of (a)  $P_i$  versus  $v_{avg}$  (b)  $P_a$  versus  $v_{avg}$  (c)  $v_{max}$  versus  $v_{avg}$

### Maximum Velocity

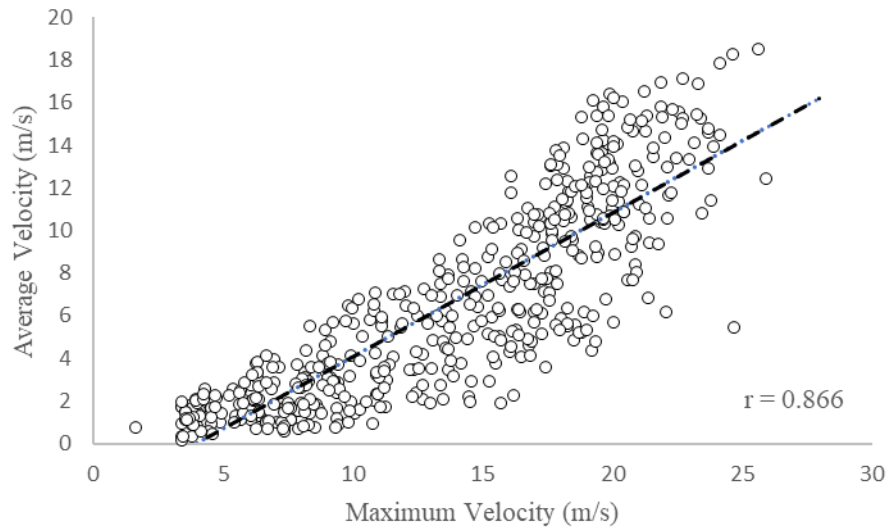
Since correlation between  $v_{max}$  and  $v_{avg}$  had been utilized for data consistency evaluation, the concluding theory shows that higher correlation indicates lesser acceleration spike microtrips. This is due to microtrips with abnormal acceleration gets filtered during microtrip selection to ensure data repeatability. The correlation value shows 0.869 for Perodua Bezza 1.3L and 0.866 for Perodua Myvi 1.5L as shown in Figure 7 and Figure 8. Hence, the data suggest consistent driving pattern during chase car operation and bias would unlikely to be occur throughout the study.

Despite from correlation for both percentage of idle and deceleration downscale to almost 17%, the value still depicts significant correlation as they fall within the limit.





**Figure 7:** Scatterplot of  $v_{avg}$  against  $v_{max}$  for Perodua Bezza 1.3L



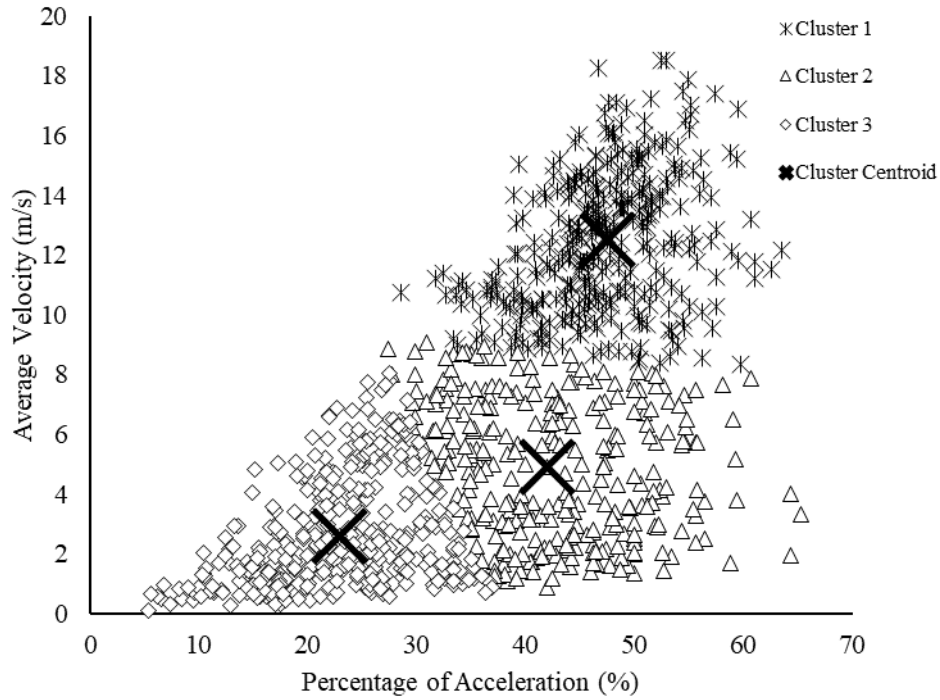
**Figure 8:** Scatterplot of  $v_{avg}$  against  $v_{max}$  for Perodua Myvi 1.5L

### Percentage of Idle and Acceleration

Both percentage of idle ( $P_i$ ) and percentage of acceleration ( $P_a$ ) sorted using K-mean clustering analysis and reported that three cluster is sufficient for distinction. For  $P_i$ , 443, 283 and 297 number of microtrips had been segregate into cluster 1, cluster 2 and cluster 3 respectively. On the other hand, 337, 304 and 382 number of microtrips had been segregated into cluster 1, cluster 2 and cluster 3 respectively for variable  $P_a$ . Table 4 shows microtrips sorted with  $P_i$  has better discreditable behavior as compare to microtrips sorted with  $P_a$  due to microtrips sorted with  $P_i$  variate 10.39% which is 12.83% lower than microtrips sorted with  $P_a$ . The lower the variance within cluster, the greater the variance between cluster. Despite from  $P_a$ . having larger within cluster variance, the segregation can be clearly visualized as shown in Figure 9.

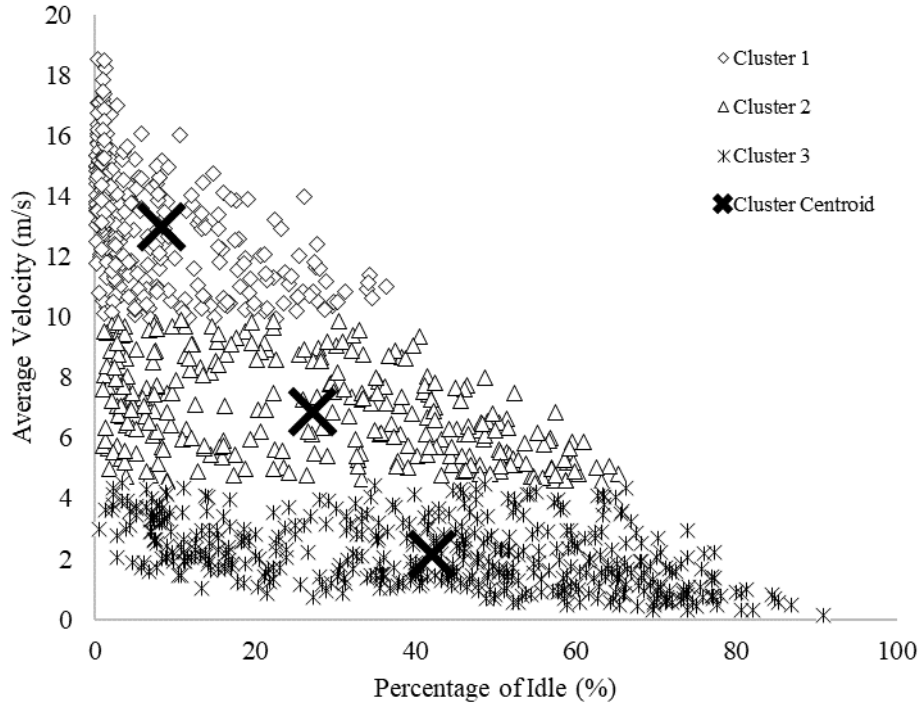
**Table 4.**  $P_i$  and  $P_a$  cluster centroid

Variance	$P_i$	$P_a$
Within-class	10.39%	23.22%
Between-class	89.61%	76.78%
Total	100	100



**Figure 9:** Data segregation scatterplot of  $v_{avg}$  against  $P_a$

Further segregation can be visualized from Figure 10, where the  $P_i$  for cluster 3 centroid are 69.12% and 80.08% lower as compare to cluster 2 and 1 respectively. Additionally, the  $v_{avg}$  discrepant at 68.72% for cluster 2 and 83.40% for cluster 1 when compare to cluster 3.



**Figure 10:** Data segregation scatterplot of  $v_{avg}$  against  $P_i$

For  $v_{avg}$  versus  $P_a$  cluster as shown in Figure8, the  $P_a$  of cluster 1 centroid are only at 11.72% difference as compare to cluster 2 despite from having 60.84% difference which is on par with  $v_{avg}$  versus  $P_i$  in term of  $v_{avg}$ . Thus, the differences in term of  $P_a$  between cluster 1 and 2 are insignificant.

### Average Speed and Percentage of Idle Correlation

The trendline of correlation between  $v_{avg}$  and  $P_i$  can be clearly visualized by clustering analysis in which for most of the microtrips with higher  $v_{avg}$  would result in lower  $P_i$  and vice versa. This clarification can be related with traffic congestion level of route. Congested traffic condition would result in lower  $v_{avg}$  and idle would likely to occur hence increase in  $P_i$ . Similarly, during free flow traffic condition, the  $v_{avg}$  of microtrips would be much higher and idle rarely occur which resulted in lower  $P_i$  as shown in Table 5. Moreover, vehicle emission had been reported to have increase with congestion level due to high idle frequency, higher congestion level would result in higher vehicle emission [19]. Thus,  $P_i$  is convinced to be more effective in term of vehicle emission evaluation.

**Table 5.**  $P_i$  and  $P_a$  cluster centroid

Cluster	$v_{avg}$ (m/s)	$P_i$ (%)
1	12.97	8.37
2	6.89	27.11
3	2.15	42.02

### CONCLUSION

This study had evaluated the representative driving parameters of a microtrip by the adaptation of Pearson product-moment correlation and K-mean clustering method. The data had been collected by using joint method of chase car and circulation method through three different period of the week in conjunction with three traffic condition throughout Sarawak, Malaysia region to reduce bias in the

evaluation of microtrips parameters. The trend observe from two method clearly indicates  $P_i$  among seven parameters is suitable to be utilized in construction of driving cycle for vehicle emission evaluation.

## ACKNOWLEDGEMENT

The authors gratefully acknowledge the support for this research from Malaysia Automotive Robotics and IOT Institute (MARii) and School of Engineering and Technology, University of Technology Sarawak (UTS).

## REFERENCES

- [1] Q. Shi, D. Qiu, and J. Wu, "Study on Driving Cycles based on Principal Component Analysis and Fuzzy C-M clustering," *Res. Environ. Sci.*, vol. 25, no. 1, pp. 70–76, 2012,
- [2] W. Zhou, K. Xu, Y. Yang, and J. Lu, "Driving Cycle Development for Electric Vehicle Application using Principal Component Analysis and K-means Cluster: With the Case of Shenyang, China," *Energy Procedia*, vol. 105, pp. 2831–2836, 2017, doi: 10.1016/j.egypro.2017.03.620.
- [3] T.-K. Lee and Z. Filipi, "Synthesis of real-world driving cycles using stochastic process and statistical methodology," *Int. J. Veh. Des.*, vol. 57, pp. 17–36, Nov. 2011, doi: 10.1504/IJVD.2011.043590.
- [4] Z. Dai, D. Niemeier, and D. Eisinger, "Driving cycles: a new cycle-building method that better represents real-world emissions," *U.C. Davis-Caltrans Air Qual. Proj.*, vol. 66, no. 66, p. 37, 2008.
- [5] S. H. Ho, Y. D. Wong, and V. W. C. Chang, "Developing Singapore Driving Cycle for passenger cars to estimate fuel consumption and vehicular emissions," *Atmos. Environ.*, vol. 97, pp. 353–362, 2014, doi: 10.1016/j.atmosenv.2014.08.042.
- [6] "Malaysia Road Length: Sarawak: Total," 2016. <https://www.ceicdata.com/en/malaysia/road-length-statistics/road-length-sarawak-total>.
- [7] T. M. Johnson, D. L. Formenti, R. F. Gray, and W. C. Peterson, "Measurement of Motor Vehicle Operation Pertinent to Fuel Economy," *SAE Trans.*, vol. 84, pp. 17–46, 1975, doi:10.4271/750003
- [8] L. Yu, Z. Wang, F. Qiao, and Y. Qi, "Approach to Development and Evaluation of Driving Cycles for Classified Roads Based on Vehicle Emission Characteristics," *J. Transp. Res. Board*, vol. 2058, pp. 58–67, 2008, doi: 10.3141/2058-08.
- [9] A. R. Mahayadin *et al.*, "Efficient methodology of route selection for driving cycle development," *J. Phys. Conf. Ser.*, vol. 908, no. 1, 2017, doi: 10.1088/1742-6596/908/1/012082.
- [10] M. Tutuianu *et al.*, "Development of a World-wide Worldwide harmonized Light duty driving Test Cycle (WLTC)," 2015.
- [11] B. Liu, Q. Shi, L. He, and D. Qiu, "A study on the construction of Hefei urban driving cycle for passenger vehicle," in *IFAC-PapersOnLine*, 2018, vol. 51, no. 31, pp. 854–858, doi: 10.1016/j.ifacol.2018.10.100.
- [12] Y. L. T. Nguyen, T. D. Nghiem, A. T. Le, and N. D. Bui, "Development of the typical driving cycle for buses in Hanoi, Vietnam," *J. Air Waste Manag. Assoc.*, vol. 69, no. 4, pp. 423–437, 2019, doi: 10.1080/10962247.2018.1543736.
- [13] A. Fotouhi and M. Montazeri-Gh, "Tehran driving cycle development using the k-means clustering method," *Sci. Iran.*, vol. 20, no. 2, pp. 286–293, 2013, doi: 10.1016/j.scient.2013.04.001.
- [14] P. Schober and L. A. Schwarte, "Correlation coefficients: Appropriate use and interpretation," *Anesth. Analg.*, vol. 126, no. 5, pp. 1763–1768, 2018, doi: 10.1213/ANE.0000000000002864.
- [15] F. Anscombe, "Graphs in statistical analysis," *Key Top. Surg. Res. Methodol.*, vol. 27, no. 1, pp. 441–475, 1973, doi: 10.1007/978-3-540-71915-1\_35.
- [16] J. H. Kent, G. H. Allen, and G. Rule, "A driving cycle for Sydney," *Transp. Res.*, vol. 12, no. 3, pp. 147–152, 1978, doi: 10.1016/0041-1647(78)90117-X.
- [17] A. Unal, H. C. Frey, and N. M. Roupail, "Quantification of Highway Vehicle Emissions Hot Spots Based upon On-Board Measurements," *J. Air Waste Manag. Assoc.*, vol. 54, no. 2, pp. 130–140, 2004, doi: 10.1080/10473289.2004.10470888.
- [18] X. Wang, Y. Xue, B. ling Cen, P. Zhang, and H. di He, "Study on pollutant emissions of mixed traffic

- flow in cellular automaton,” *Phys. A Stat. Mech. its Appl.*, vol. 537, no. X, p. 122686, 2020, doi: 10.1016/j.physa.2019.122686.
- [19] I. N. Anida, N. A. Abdul Latiff, and A. R. Salisa, “Driving cycle analysis for fuel rate and emissions in Kuala Terengganu city during go-to-work time,” *J. Eng. Sci. Technol.*, vol. 14, no. 6, pp. 3143–3157, 2019, doi: 10.11591/ijece.v9i3.pp1780-1787