International Journal of Road Safety

Journal homepage: www.miros.gov.my/journal

Red-Light Running Vehicles Behaviour Based on Linear Regression Approach at Traffic Lights along Bakau Condong Road, Batu Pahat, Johor

Joewono Prasetijo[1,*], Izuan Bin Musa[2], Nurfarhanna Ahmad Sulaiman[1], Maisara Ashran Mustafa[3], Leksmono Suryo Putranto[4], Nor Baizura Hamid[1], Alvin John Lim Meng Siang[5] & Muhd Faiz Abd Latif[1]
*Corresponding author: joewono@uthm.edu.my

[1] *Sustainable Transport and Safety Studies (STSS), Faculty of Engineering Technology, Universiti Tun Hussein Onn Malaysia, 84600 Panchor, Pagoh, Johor, Malaysia*
[2] *Zadra Consultant, Tabuan Stutong Commercial Centre, Jalan Setia Raja, 93350 Kuching, Sarawak, Malaysia*
[3] *Department of Road Safety Malaysia (JKJR) Johor, Suite 25.01, Johor Bahru City Square, 8000 Johor Bahru, Malaysia*
[4] *Faculty of Engineering, Universitas Tarumanagara, Jalan S. Parman No. 1, Grogol, Jakarta Barat 11440, Indonesia*
[5] *Faculty of Civil and Environmental Engineering, Universiti Tun Hussein Onn Malaysia, 86400 Parit Raja, Batu Pahat, Johor, Malaysia*

ABSTRACT

Red-light dilemma zone is widely known as an area on the high-speed intersection approach, where vehicles either safely stop before the stop line or proceed through the intersection during the red interval. Within such an area, a vehicle might be involved in a right angle crash or rear-end collision. The objective of this study is to develop a prediction model of red light running and determine red running distance and the dilemma zones. Data that have been collected by using a video camera and Traffic Data Collector (TDC) had been analyzed such as traffic flow, traffic speed that red-light running (RLR), the number of the vehicle that RLR and traffic light cycle time. The data was analyzed by using applications such as Microsoft Excel, Minitab and PETRAPro software. Based on the result of Multiple Linear Regression Analysis, it shows the parameters that involved in vehicles behaviours along the traffic lights has been proved by highly significant of the p-value that less or equal to 0.1 ($p \leq 0.1$) with 90% confidence level. The parameters such as the traffic flow and average speed of RLR vehicles with the p-value of 0.051 and 0.034 respectively proved that all these parameters cause the increasing rate of RLR toward traffic light intersections. The values of R^2 = 63.76 % and also show that the prediction model that was obtained is satisfied.

ARTICLE INFO

Article History:
Received 20 Oct 2019
Received in revised form
20 Feb 2020
Accepted
15 Mac 2020
Available online
01 May 2020

Keywords:
Red-light running
Running distance
Dilemma zones
Multiple regressions

1. Introduction

In Batu Pahat, Johor, the number of road accidents is increasing from 2014 to 2018 based on the data recorded. Statistic shows number of accidents 7,445; 7,695; 8,170; 8,445; 8,462 and resulted in road fatalities of 165, 197, 207, 165, 178, respectively. These numbers are considered high giving an index of 2.22, 2.56, 2.53, 1.95, 2.10 road fatalities per 100 accidents. Generally, Malaysia is taken place at seven (7) place out of nine (9) Southeast Asian countries with 24 accidents in 100,000 population. The number of registered vehicles in Malaysia is growing at a very high pace of average 6.6% per annum resulting in the total high number of registered vehicles. Unfortunately, this results in 2.9 road fatalities per 10,000 registered vehicles (RSD Johor, 2017).

Studies found several reasons cause crashes and injuries such as violating traffic rules; speeding and beating traffic lights. Report shows that 27% of crashes are due to risky driving and speeding is in 21% (RSD, 2014). Road crashes are reported by the Royal Malaysian Police (RMP) and as per their classifications, and traffic light violation is one of the major causes of crashes, deaths, and injuries at signalized intersections. Published crash statistics show that 136 Malaysians were killed and 155 were injured in 2002 due to traffic light violation related crashes. However, the report does not include unreported crashes, the possible misclassification of road crash caused by the police and the near misses from traffic light violation behaviour (RMP, 2006).

Intersection is known for high potential conflicts created by several traffic streams (signalized and unsignalized) which will produce high rate of accidents (Prasetijo et al., 2011, 2018; Prasetijo & Musa, 2016). Survey related to red-light running (RLR) signalized intersections in Metropolitan city shown that accidents and injuries were seen as unlikely as the result of running red-lights. Some studies have conducted analysis on the different type of vehicles speed approaching the traffic lights based on linear speed profiles. The study found that the speed changes are significantly close to the traffic lights (Prasetijo & Zainal, 2016). About 48% of the drivers who responded estimated that 10% or fewer cases of running red-light would result in a crash or other accident, with 56% of those who predicted accidents estimating a 10% or less chance of injury in such an accident (Retting et al., 1999).

Some behavioural research has clarified the role of two different factors on RLR in term of “dilemma zone” (Allos et al., 1992) and circumstances that make RLR possible (Bonneson et al., 2001; Fabio

et al., 2012). Interestingly, suggested that drivers' stop-go decision process should be considered the expectations due to previous knowledge of the intersection (especially waiting times at red-lights), the assessment of the consequences of a violation and the estimation of the consequences of stopping. The studies found common RLR behaviours along specific traffic light with homogeneous traffic that is typical in developed countries. However, studies may not be directly adopted for different traffic behaviours that consist of several types of vehicles in terms of the dimensions and engines (heterogeneous). Therefore, this study intention is to explore RLR vehicles performance with the typical heterogeneous traffic.

2. Method

2.1. Data Collection

Vehicles streams data were recorded at two four-legs signalized intersections along Jalan Bakau Condong at Batu Pahat District in Johor. The traffic lights were placed at the distances of 1.5 km in between. All streams volume were recorded with two cameras camcorder (at both intersections) and traffic lights phasing were counted. Road geometric design/dimensions data of all intersections were referred from the Manual (Arahan Teknik Jalan (13/87) by Public Works Department (1987)). Furthermore, the real dimensions were also measured from the site (primary data) with the dimensions of roads between 5.4 – 9.7 meters. The data were collected in two (2) working days and considered during the peak traffic hours (7.00 – 8.00 & 16.30 – 17.30).

2.1.1. Data Observation Method

This observation was carried out at the peak hour of the working days. Data were collected at morning peak (7.00 – 8.00 am) and evening peak (16.30 – 17.30 pm). The observations were made within 15 minutes of a time interval for one (1) hour for every intersection. Traffic flow data were collected by using a video camera at the intersections that are chosen. This was an effective method because the video will record all of the vehicles that enter the intersection. It also helps to prevent miscalculation of data. The videos were in perfect and clear for viewing. All of the videos were transfer from camera to computer for data analysis. The cameras were position at the corner of the intersection and with suitable distance to get a proper view of the intersection's video at approx.. of 1.00 – 1.50 meters from the edge of road (Figure 1 shows the camera's position). The road dimensions are shown in Table 1. Then, the data were recorded in a data observation form according to respective vehicle classifying according to Arahan Teknik Jalan (13/87).

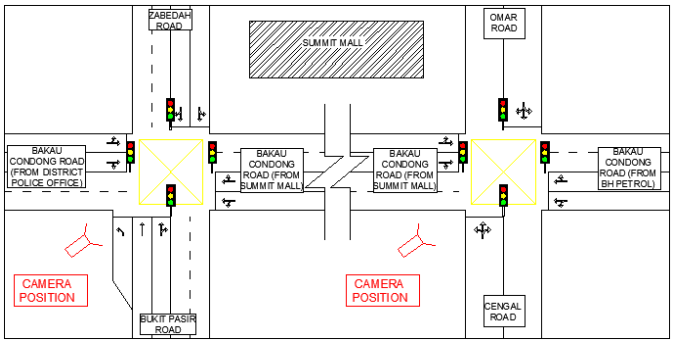

Figure 1: Layout plan of research and camera location.

The vehicle speed was collected by using Traffic Data Collector (TDC). By using this equipment, the time measurement was not required because it can automatically measure the speed based on the distance between stop line and the end line of the yellow box, (x) as in Figure 2. The observation was made for the vehicles RLR based on the vehicles passing through intersections during the yellow and red signs relating to criteria; type of vehicles, traffic volumes and speed. Furthermore, the number of vehicles violating the traffic lights could be identified.

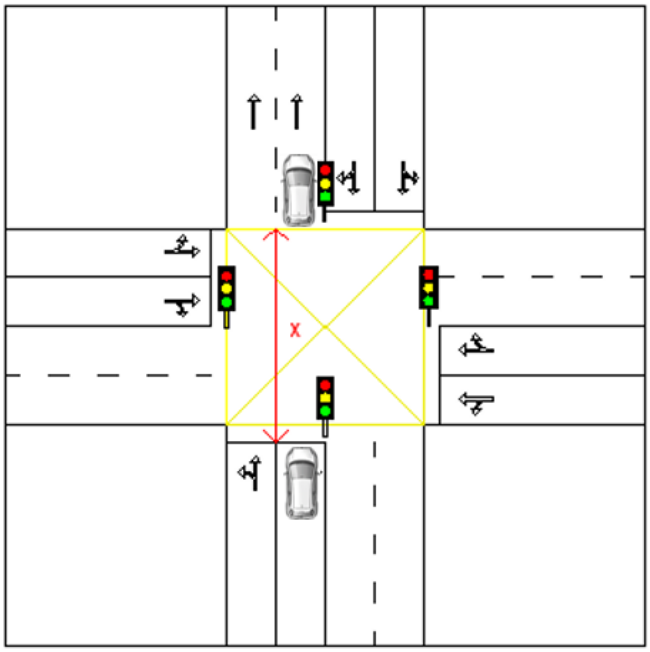

Figure 2: The sketch of on the distance between stop line and the end line of the yellow box, (x).

2.2. Data Analysis

Analysis of survey data includes data that has been observed in the study area such as the analysis of the flow of traffic, the speed and the total number of RLR vehicle analysis of traffic flow refers to the total number of vehicles passing through a signalized intersection in a study area. Data were analyzed with two methods that are Pearson Correlation and Multiple Linear Regression to produce Red-Light Running Model. Therefore, the relationship between dependent variable with one or more independent variable can be modelled. The obtained data from the analysis were revised in order to get R-square (R^2) to evaluate model fit. Data on the number of vehicles passing through this intersection is analyzed based on the intersection, vehicle class, the average speed of vehicles that violate the red-light and the number of lanes. The following analysis identifies the relationship between parameters that were observed during the data collection process.

3. Results

3.1. The Number of Red-Light Running Vehicles

Figure 3 shows the total traffic flow according to the segment lane at each intersection that is observed. From the chart, traffic flow at Zubaedah Road for day 2 is higher among the rest with 2,473 vehicles per hour while Cengal Road recorded the lowest traffic flow with 625 vehicles per hour.

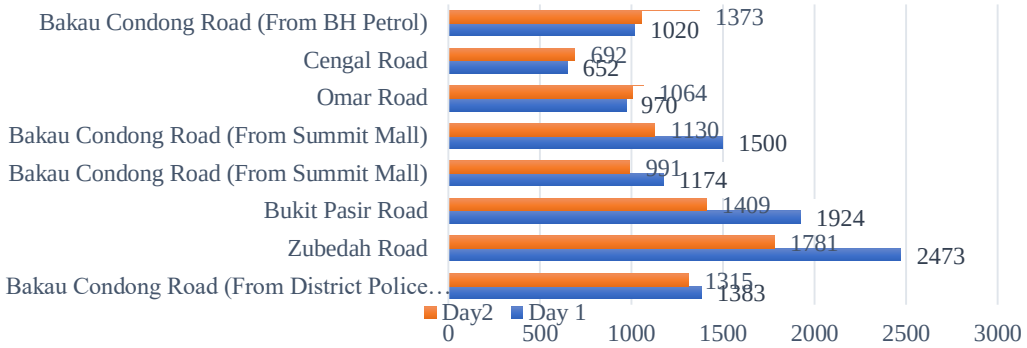

Figure 3: Total traffic flow according to segment lane at each intersection.

The following Figure 4 and Figure 5 show the traffic flows and RLR vehicles for each segment for two (2) days. From both charts, it shows that RLR increase when the traffic flows also increase.

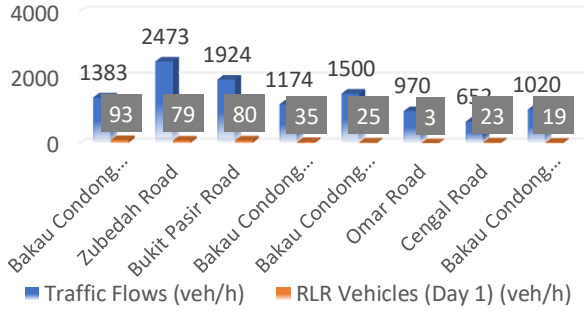

Figure 4: The traffic flow and RLR vehicles for each segment at day 1.

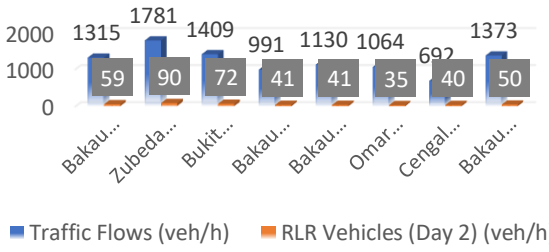

Figure 5: The traffic flow and RLR vehicles for each segment at day 2.

3.2. The Relationship between Number of Lane, Traffic Flow and Average Speed of RLR Vehicles with Number of RLR Vehicles

The study found significant parameters to be considered for good fit of number of RLR vehicles; number of lanes, traffic volumes and vehicles speed.

Table 1: Traffic and speed performances of two four-legs intersections.

Day	Intersection	Study Site	Number of Lane/ Width (m)	Traffic Flow (veh/hour)	Average Speed of RLR Vehicles (km/h)	Number of RLR Vehicles (veh/hour)
1	Bakau Condong / Bukit Pasir Road	Bakau Condong Road (from District Police Office)	2/7	1,383	49	93
		Zubedah Road	2/6.4	2,473	44	79
		Bukit Pasir Road	3/9.7	1,924	47	80
		Bakau Condong Road (from Summit Mall)	2/6.4	1,174	41	35
	Bakau Condong / Cengal Road	Bakau Condong Road (from Summit Mall)	2/6.7	1,500	40	25
		Omar Road	1/5.4	970	37	3
		Cengal Road	1/5.5	652	38	23
		Bakau Condong Road (from BH Petrol)	2/6.7	1,020	46	19
2	Bakau Condong / Bukit Pasir Road	Bakau Condong Road (from District Police Office)	2/7.0	1,315	43	59
		Zubedah Road	2/6.4	1,781	48	90
		Bukit Pasir Road	3/9.7	1,409	42	72
		Bakau Condong Road (from Summit Mall)	2/6.4	991	44	41
	Bakau Condong / Cengal Road	Bakau Condong Road (from Summit Mall)	2/6.7	1,130	44	41
		Omar Road	1/5.4	1,064	38	35
		Cengal Road	1/5.5	692	37	40
		Bakau Condong Road (from BH Petrol)	2/6.7	1,373	45	50

3.3. Pearson Correlation Level for RLR Vehicles

Correlation is a technique for investigating the relationship between two quantitative or continuous variables. Pearson's correlation coefficient (r) is a measure of the strength of the association between the two variables. Table 2 shows the result of Pearson Correlation Analysis based on the independent variable such as number of the lane, traffic flow, average speed of RLR vehicles and number of days. It shows that the strength of association between the variables is very strong for the average speed of RLR vehicles with the value 0.713 nearest to 1 and traffic flow is very high significant variable ($p = 0.03$). The p-value for the number of the lane, traffic flow and average speed of RLR vehicles with the value of 0.015, 0.003 and 0.017 respectively show very significant variable because the lower that 0.1 with confidence level 90%. Opposite with the number of days because the p-value is higher than 0.1 show that variable is not significant for RLR.

Table 2: Result of Pearson correlation analysis.

Variable	Pearson Correlation (r)	P-value
Number of lane	$0.594 \approx 1.f$	$0.015 \leq 0.1$
Traffic flow (veh/hour)	$0.693 \approx 1.0$	$0.003 \leq 0.1$
Average speed of RLR vehicles (km/h)	$0.713 \approx 1.0$	$0.017 \leq 0.1$
Number of days	$0.169 \approx 1.0$	$0.530 \geq 0.1$

4. Discussion

4.1. Multiple Linear Regression Level for RLR Vehicles

R-square (R^2) also known as the Coefficient of determination is a commonly used statistic to evaluate model fit. R-square is 1 minus the ratio of residual variability. Table 3 shows the result of Multiple Linear Regression analysis. The model must fulfill the requirement of p-value $\leq$ 0.1 (Confident level 90%) and the coefficient of determination is close to 1 ($R^2 \approx 1$). R-square (R^2) for this equation is 63.69% and adjusted R-square (R^2) is 58.11%. The value of R-square (R^2) showing that the model was obtained is satisfied with the value of 0.6369. It shows that all the variable has medium strength with each other. The following linear regression for RLR based on two (2) potential parameters as traffic flows and average speed can be shown in Table 3. The model found that number of lanes is not significant for RLR. Therefore, the model satisfies with traffic flow (TF) and average speed (AV).

The minimum yellow interval after considering the factor of dilemma zone is shown in Table 4. The increasing of the yellow interval is to reduce rate red light running so the driver can safely cross the intersection with their current speed. Based on Arahan Teknik Jalan (13/87) on a Guide to the Design of Traffic Signals, the maximum yellow interval is 5 seconds. The existing yellow interval is 5 seconds so that it was proposed to increase the all red interval by 2 seconds at Intersection Bakau Condong/Bukit Pasir Road. The all red interval will become 3 seconds for every segment of the intersections. Meanwhile, the existing yellow interval is 3 seconds for intersection Bakau Condong/Cengal Road is increasing by 2 seconds to become 5 seconds of yellow interval and also increasing the all red interval by 1 second to become 2 seconds.

Table 3: Multiple linear regression analysis for RLR (MINITAB).

Term	Coef	SE coef	T-value	P-value	VIF
Constant	-124.3	52.8	-2.35	0.035	
Traffic flow (TF) (veh/hour)	0.0251	0.0117	2.15	0.051	1.44
Average speed of RLR vehicles (AV)	3.30	1.39	2.36	0.034	1.44

The 2nd multiple linear regression found relationship of RLR, TF and AV as (Eq. 1).

$$RLR = -124.3 + 0.0251\,TF + 3.30AV \quad (1)$$

with model summary as S = 17.5090, R-sq = 63.69%, R-sq (adj) = 58.11%, R-sq (pred) = 42.62%, and analysis of variance is described as:

Table 4: Analysis of variance RLR.

Source	DF	Adj. SS	Adj. MS	F-value	P-value
Regression	2	6992	3495.8	11.40	0.001
Traffic flow (TF) (veh/hour)	1	1414	1413.9	4.61	0.051
Average speed of RLR Vehicles (AV)	1	1714	1714.2	5.59	0.034
Residual error	13	3985	306.6		
Total	15	10977			

4.2. Dilemma Zone Conflicts

A dilemma zone at an intersection is defined as an area approaching the stop line within which a driver finds him/herself is too close to stop safely and yet too far away to pass completely through the intersection at a legal speed before the red phase commences. The schematic of a dilemma zone can be seen in Figure 6.

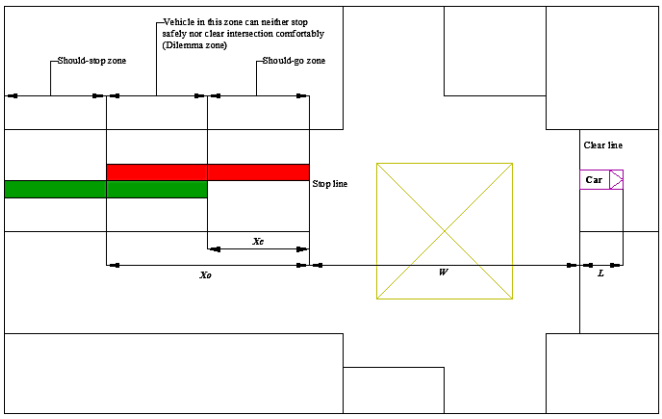

Figure 6: A schematic of a dilemma zone at an intersection (Traffic and Highway Engineering, 2014).

The minimum yellow interval after considering the factor of dilemma zone is considered as:

$$X_c = u_0\,(\tau_{min}) - (W + L) \quad (2)$$

$$X_0 = u_0\delta + \frac{u_0^2}{2a} \quad (3)$$

where X_c is the distance within which a vehicle travelling at the speed limit (u_0) during the yellow interval τ_{min} at the road dimensions of W width (m) and L length (m). X_0 is the minimum distance from the intersection for which a vehicle travelling at the speed limit u_0 with perception-reaction time δ (sec) and constant rate a (m/sec^2).

The dilemma zones can be eliminated to reduce the risk of safety reasons by having $X_0 = X_c$ therefore, the minimum yellow interval τ_{min} will be determined with:

$$\tau_{min} = \delta + \frac{W + L}{u_0} + \frac{u_0}{2a} \quad (4)$$

The minimum yellow interval and proposed all red interval of the intersections can be seen in Table 5. The maximum yellow interval is five seconds based on Arahan Teknik Jalan (13/87) on a Guide to the Design of Traffic Signals. Therefore, it is proposed to increase all red interval by 1 – 2 seconds at the intersections.

Table 5: Dilemma zone distance (m) and minimum yellow interval.

Intersection	Study Site	Dilemma Zone (m)	Minimum Yellow Interval (sec)	Proposed Increasing All Red Interval (sec)
Bakau Condong / Bukit Pasir Road	Bakau Condong Road (from District Police Office)	26.11	7.0	2
	Zubedah Road	19.71	6.5	2
	Bukit Pasir Road	19.71	6.5	2
	Bakau Condong Road (from Summit Mall)	26.11	7.0	2
Bakau Condong / Cengal Road	Bakau Condong Road (from Summit Mall)	41.85	5.5	1
	Omar Road	36.95	5.5	1
	Cengal Road	36.95	5.5	1
	Bakau Condong Road (from BH Petrol)	41.85	5.5	1

5. Conclusion and Recommendations

Based on the result of Multiple Linear Regression Analysis for RLR, it shows the parameters that involved in vehicles behaviours along the traffic lights has been proved by highly significant of the p-value that less or equal to 0.1 with confident level 90%. The parameter such as the traffic flow and average speed of RLR vehicles with the p-value of 0.051 and 0.034 respectively proved that all this parameter cause the increasing rate of RLR toward traffic light intersections. Multiple Linear Regression Analysis also shows the values of R^2 = 63.69%. It shows that the variables such as traffic flow and average speed of RLR for this prediction model of RLR that was obtained are satisfied. The dilemma zone distance determined after the value of X_0 – Xc. It proved that dilemma zones distance affects the rate of RLR by referring to Table 5 that rate of RLR vehicles. So, increasing yellow interval and all red interval can reduce the rate of RLR.

Acknowledgements

The authors would like to thank the Universiti Tun Hussein Onn Malaysia for supporting the study through UTHM Grant Vot. 1452 (Ref: FRGS/1/2014/TK07/UTHM/02/2), Malaysian Institute of Road Safety Research (MIROS) and all individuals/organization that have made this study possible. Thank you is also extended to the Sustainable Transport and Safety Studies (STSS) Focus Group, Faculty of Engineering Technology, Universiti Tun Hussein Onn Malaysia.

References

Allos, A. E., & Al-Hadithi, M. I. (1992). Driver behavior during onset of amber at signalized junctions. *Traffic Engineering and Control*, *33*(5), 312 – 317.

Bonneson, J., Brewer, M., & Zimmerman, K. (2001). *Review and evaluation of factors that affect the frequency of red-light-running* (No. FHWA/TX-02/4027-1). Austin, Texas: Texas Department of Transportation.

Fabio, G., Tullio, G., Margaret, B., Giovanni, T., & Tiziana, C. (2012). Traffic microsimulation model to predict variability of red-light running influenced by traffic light operations in urban area. *Procedia - Social and Behavioral Sciences*, *53*, 872 – 880.

Prasetijo, J., Pour, M. H., & Ghadiri, S. M. R. (2011). Capacity of unsignalized intersections under mixed traffic conditions. *Procedia-Social and Behavioral Sciences*, *16*, 676 – 685.

Prasetijo, J., & Musa, W. Z. (2016). Modeling zero-inflated regression of road accidents at Johor Federal Road F001. *MATEC Web of Conferences*, *47*, 03001.

Prasetijo, J., & Zainal, Z., F. (2016). Development of continuous speed profile using GPS at Johor Federal Roads F0050. *MATEC Web of Conferences*, *47*, 03024.

Prasetijo, J., G. Zhang, N. Guntor, A. Siang, B. Daniel, & M. Sanik. (2018). Change of road integrated design consistency due to antiskid transverse rumble strips on high-speed federal road FT050. *Advances in Civil Engineering Materials*, *7*(3), 460 – 472.

PWD (1987). *Arahan Teknik (Jalan) 13/87. A guide to the design of traffic signals*. Kuala Lumpur: Public Works Department (PWD).

Retting, R. A., Ulmer, R. G., & Williams, A. F. (1999). Prevalence and characteristics of red-light running crashes in the United States. *Accident Analysis & Prevention*, *31*(6), 687 – 694.

RMP (2006). *Statistical report road accidents Malaysia – 2006*. Kuala Lumpur: Traffic Branch, Bukit Aman, Royal Malaysian Police (RMP).

RSD (2014). *Road safety plan of Malaysia 2014 – 2020*. Putrajaya: Road Safety Department (RSD) Malaysia, Ministry of Transport.

RSD Johor (2017). *Road safety* statistics. Johor Bahru: Road Safety Department (RSD) Johor.