

Analysing the Impact of Road Intersection's Safety Towards Road Accidents by Using Ordinary Least Square Regression

Adli, H. A. A. and Naharudin, N.*

Centre of Studies for Surveying Science and Geomatics, Faculty of Architecture, Planning and Surveying, Universiti Teknologi MARA, 40450 Shah Alam, Selangor, Malaysia

E-mail: nabilahnaharudin1290@uitm.edu.my

*Corresponding Author

DOI: <https://doi.org/10.52939/ijg.v18i4.2247>

Abstract

Road safety is a global issue that affects all countries because it is linked to the road management or operation as a whole. Roads are essential for transit and movement, whether driving, riding, walking, or using a bus. One of the most critical considerations for road users is road safety. This study aims to analyze the road intersection impact with roads accident by using GIS. Specifically, it investigated whether the perceived correlation between road intersection safety has an impact on the road accident. In this context, road intersection safety is translated based on parameters or factors that occur to be in an intersection area which are the road intersection geometry, traffic volume and waiting time at the traffic lights. The relationship was analysed by using correlation and ordinary least square regression. The results showed a positive correlation between two of the road intersection safety factors: higher traffic volume and longer traffic waiting time were associated with higher number of accident cases. These results show how the situations of Section 7, Shah Alam Road intersection in relation of road accidents. Traffic Light Waiting Time has the highest correlation with a value of 0.65131 while Road Geometry has the lowest with a value of 0.0031. On this basis, the concept of road intersection traffic safety and device timing can be considered in implementation of road intersection safety.

1. Introduction

Road safety is a global issue faced by all country since it is related to one of the important things in human's life, safety. This is because, human uses road for almost all their daily activities from going to school, work, stores, playground, and any others. In general, the mode of transportation on road uses by different age groups may differ as adults will use motor transportation such as cars and motorcycle while children will be most likely to use bicycle to commute. When travelling, there is always a possibility that each person may be exposed to various kind of risks related to their safety. The safety on road is always related to the users' behaviour, traffic conditions and road characteristics.

Road accidents in general have been reported to be the leading cause of death and the eighth leading cause of death among young people aged 5 to 29 years. The Ministry of Transportation (MOT) of Malaysia has released the 2020 Road Safety Programme, which was launched in 2014.

The proposal is a continuation of the 2006-2010 Road Safety Plan. In this plan, MOT stated that in identifying the cases that related to fatalities, a Death Calculation Index which has been accepted globally to measure the road safety will be used. As mentioned previously, road accidents always related to three things which are the behaviour of road user, traffic's conditions and the road's characteristics (WHO, 2016). Therefore, there may be a need to study the impact of these factors in road accidents.

Accidents often happens at road's intersections. This is due to the complex interaction of all types of road's users there. Although with the presence of traffic calming devices, accidents can still occur. The main problem might be the characteristics at the intersection itself, not only the presence of traffic lights for example. Other parameters and road intersection safety are also needed at the intersection to reduce the accidents risk. Accidents could occur for various reasons.

It could happen unconsciously, but sometimes even when the road's users are conscious and well-aware of their surroundings. There are many factors that can be the cause of accidents at the road intersections. It can be because of traffic lights, junction, crosswalk, careless behavior and more. All these could lead to any kind of accidents. With this study, it is hopeful the authority can do something to improve the traffic management.

2. Road Accidents in Malaysia

According to the statistics of health released on 2019 by Department of Statistics Malaysia (DOSM), road accidents was ranked as the 4th cause of death in Malaysia on 2017 and 2018 with 4499 and 4391 number of death recorded for both years respectively as shown in Figure 1 (DOSM, 2019). This proved that road accidents are becoming more worrying year by year. Therefore, something must be done to reduce these numbers. Figure 2 shows the top three areas with the most cases of road accidents. The city with the greatest number of cases is Petaling with a total of 82,058 cases followed by Johor Bahru with 45,970 and the last one is Ulu Langat with 26,784 cases.

The numbers might be related to the congestion of the area. Petaling of the busiest city that connects between Subang and Shah Alam, this might be the reason that the area has the highest number of road accidents. Increasing number of vehicles at an area exposed the users to more risk. Even though Petaling is a developed city, it still needs major improvement to lower the number of road accidents.

3. Road Intersection Characteristics

Three (3) road characteristics at intersection were taken into consideration in measuring the intersection safety which are road geometry, cycle time at traffic lights and traffic volume. The characteristics were adopted from previous literatures including a research conducted by Malaysia Institute of Road Safety Research (MIROS). This study is focusing on the cycle time of the traffic light. The geographic features of the road characteristics were collected first on site using mobile GIS data collection technique, then be used analysing the relationship between them and the traffic accidents by using Regression Analysis and Ordinary Least Square Analysis.

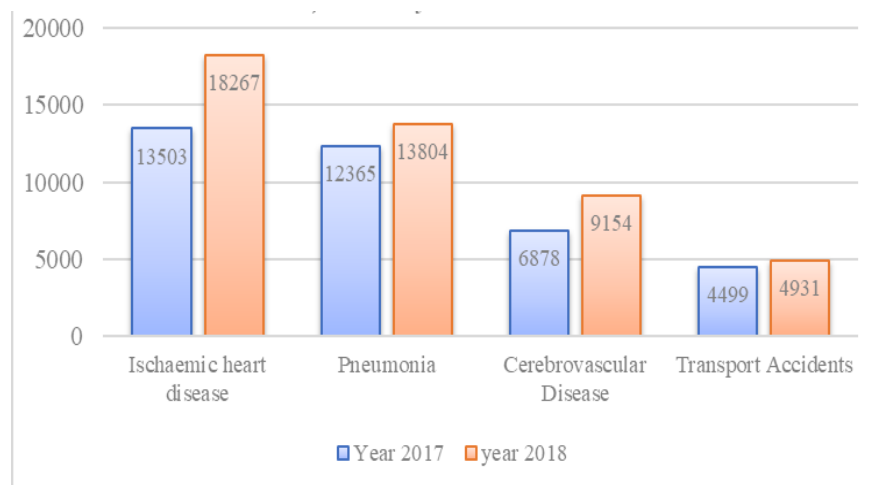


Figure 1: Top 4 principal causes of death in Malaysia for 2017 and 2018 (DOSM, 2019)

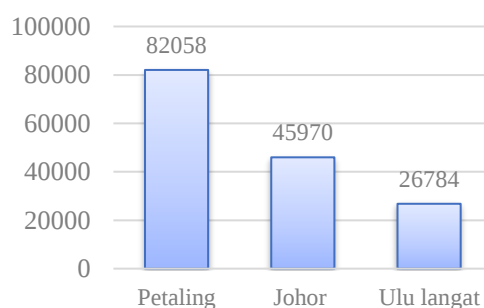


Figure 2: Number of Road Accident in Malaysia, 2019 (DOSM, 2019)

Traffic Calming devices is one of the measures that has been taken to slow down vehicles on the road as it moves around the neighbourhood, cities, commercial area. This measure helps drivers to slow down their vehicles and helps in decreasing number of accident occurring on the road (Wang et al., 2011 and Roslan et al., 2018). The calming devices can help the road user to feel safer while on the road. Different road requires a different types of road calming device. These devices depend on the road community of an area.

Traffic lights waiting time should not be too long and too short. A short cycle length (40-60 seconds) gave the best performance as it gave the lowest average delay. A longer cycle over 60 seconds will increase the volume the number of vehicles. The longer cycle length is at 120 second as the longer the timing leads to increasing capacity and rapid increase in the delay at the traffic light (Mohamed Eltayeb Zumrawi, 2016). The total length of a cycle is equal to the sum of all traffic phases. When comparing short cycle crossings to those with longer cycle lengths, there were 0.219 crashes per million cars (0.130) (Carmanah, 2018 and Kulanthayan et al., 2007). The length of traffic light cycle time has a positive influence for vehicle user to commit a traffic light violation There is significance relation of traffic light violation between long cycle time and short cycle time traffic lights. That study also discovered that 11.3% vehicles violated the short cycle time and 14.4% has violated the longer cycle time proving the length of the cycle time has influenced to cause a traffic light violation. The study also state that 8.1% vehicles violated a traffic light that is without countdown timer.

4. Previous Studies

GIS had been widely implemented previously in measuring road traffic safety. Spatial statistics is one of the most implemented GIS analysis techniques. Many studies had used GIS in an attempt to have a better road management system. Often the studies were conducted to measure the safety management by identifying the incident location and collect and visualize the road parameter. The GIS analysis helped in identifying the road safety which could later help in managing the consequence and eliminated the unwanted accidents by measuring the number of accidents, injuries, fatalities, and the accident cost the management.

5. Methodology

Figure 3 illustrates the flow of methodology conducted for this study. It comprises of four (4)

stages of background study, data acquisition, data processing and data analysis.

5.1 Data Acquisition

Data acquisition is the early stage in the methodology where the data need to be identified, types of data needed, and amount of enough data required for this research. There are two types of data needed for this study which is the primary and secondary data. The importance of collecting these two types of data is to make sure we have enough data to carried out the analysis. To obtain the data, both primary and secondary is used. The secondary data was obtained from the Royal Police Department in Shah Alam to get the updated statistics number of accidents occurred at the road intersections

5.1.1 Road accident cases in section 7 Shah Alam

The road accident location obtained from the local agencies is used to visualize the location of the accidents in a map. The names of location, the time of accidents, level of injury, causes of accidents all was stored as spatial database. The spatial location data was represented as using point features. The additional info was added in the attribute table of the point data.

5.1.2 Data collection for road intersection's characteristics

The traffic volume data for the Section 7 intersection is obtained from the TOMTOM Move website. The website offers various datasets of traffic stats at a different range of time. The data taken is a 24-hour traffic data of the intersection. The traffic volume is taken based on the probe count on the right-side of the image. The traffic data is then extracted and sorted in excel for other intersection volume data.

For primary data collection, the data was collected on-site by going to the road intersections in the study area to get the coordinate of the road intersections. Since a new intersection might have been constructed in the area. The geometry of the road also needs to be obtained to come up with the road geometry intersection value. In addition, the location of traffic lights, as well as its cycle time, were also acquired at the same time. The traffic light cycle time is taken manually by using a stopwatch/timer using a mobile phone. Primary data collection utilizes mobile GIS data collection techniques by using smart phones equipped with GPS sensors.

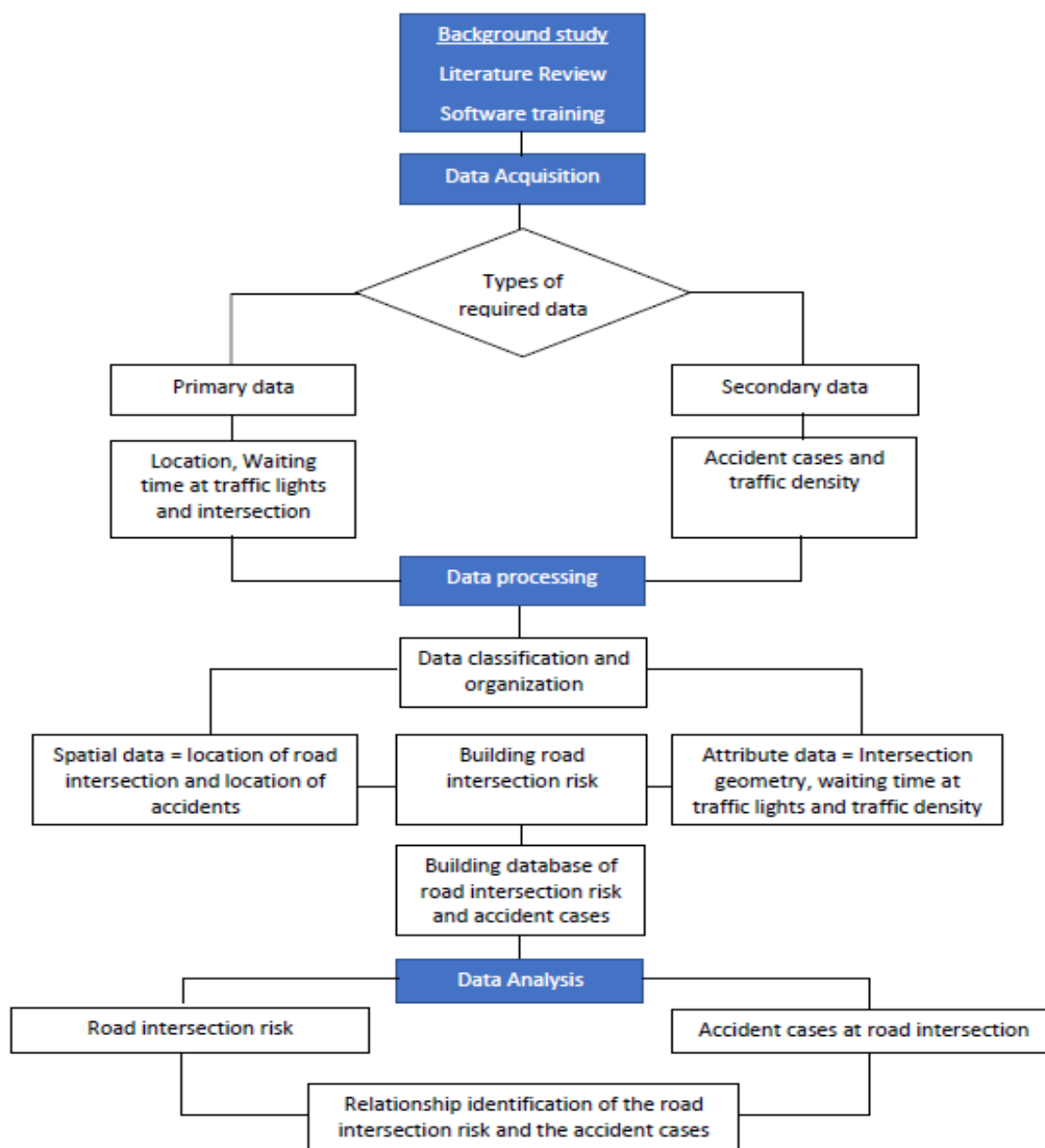


Figure 3: Methodology for indentification of road risks and accident cases

The location coordinate is obtained by using the GPS utilities in the mobile phone 5m tolerance will be obtained as the mobile GPS is being used 5m accuracy is enough to locate the intersection places due to the width of the intersection itself. The data processing will be done using ArcGIS software. The computer is being used to install and run the ArcGIS software for the data processing procedures.

The data was collected by using a Google Form that is integrated with a script to collect the location of the area. The Google Form was designed to allow the collection of the location of road intersections

and their characteristics (geometry and traffic lights' cycle time). By using the script (Figure 3), the spatial location was collected. The script was integrated by using a script editor in Google Form. The code and the index.html that was prepared for the location service. The location was collected once the submit button in the Google Form was used. It automatically obtains the location information and the intersection geometry in one spreadsheet. The Google Form that is integrated with the location service script is used to collect the road geometry and location of the intersection.

Number 1 indicates that the geometry exists and 0 indicates that the geometry does not exist. The geometry is collected in the form and the location is collected after the script link is clicked. The location will automatically be collected and stored in the google form spreadsheet data. The same method was also used in collecting the traffic light cycle time.

5.2 Data Processing

The data processing involves measuring the Road Intersection Risk which was measured based on the three (3) road characteristics which are road geometry, traffic lights' waiting time, and traffic volume. The first two is straight forward as it uses the value of the waiting time and volume to indicate the risk. But, for the first characteristic which is geometry, since there are nine (9) parameters included, a value needs to be calculated first to represent them in signifying the risk. The index taken represents the total number of the Road Intersection Geometry Index (RIGI).

The road intersection geometry was measured by using the data collected from previous steps. This index was measured by using compiling the existing road geometry of the intersections. Each intersection has a different type of road geometry. The geometry depends on the need of the road intersection. Since the road intersection has many types. This research focuses on the signalized and unsignalized road intersections. RIGI is a value that focuses on a road intersection area showing the number of total

geometries that occur. The value is calculated based on the number of Road Intersection parameters. Each of these parameters will be given its points to calculate the RIGI. These parameters also depend on the amount of traffic at the intersection and whether it is a major or minor intersection.

By using this formula of Road Intersection Geometry = Sum of the road intersection points. If all these parameters existed at the intersection location, the total for full points will be 9. So, each intersection might differ from the other. With the total of the RIGI combined with the Accident cases. Both data can be relating, and analysis can be carried out to see whether the RIGI has any impact on the road accident cases.

A heatmap was used to see which intersection has the most points which is the number of accidents. The point density is differentiated from each other by using a color variance. The darker the color shows more accidents at the intersection. Point density helps in focusing on the greatest number of accidents that happened. From here, the RIGI was used to analyze whether does the value contributes to the number of accident cases in the area. The credibility of the RIGI will be proven concerning the accident especially at the intersection with high-density cases. Based on the information of the RIGI and density of the accident at the road intersection, relationship and analysis can be done. The heatmap of road intersection index can be seen as Dashboard which had been developed as shown in Figure 4.

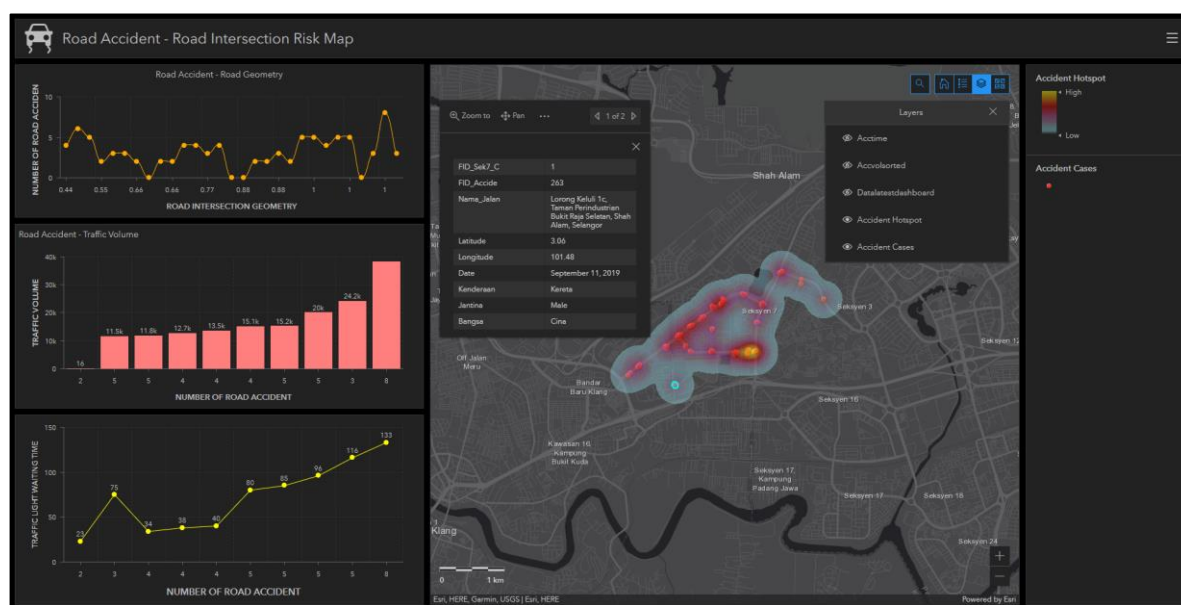


Figure 4: Heatmap of road intersection risks

5.3 Data Analysis

The regression analysis that is being done is to calculate the correlation between the factors of the data. In this case, the data that is used to see the correlation is between the number of accidents in the intersection and the RIGI. The correlation of the two data is explained with the value of the correlation is either close to 0 or close to 1. Strong correlation gave a value that is near to 1 and weak correlation shows the value close to zero. The dependant variables for the analysis if the number of accidents cases and the independent variable is the Road Intersection Risk. The Correlation analysis is conducted by using both Excel and ArcGIS data. The residual from both platforms can be compared to use as a double checker for the data such as the residual output, P-value Adjusted R, and the R-squared output.

Next, after all, three factors of the Road Intersection Risk have been analyzed the correlation

with the Road Accidents, a comparison will be done to identify which factor contributes the biggest impact toward road accidents. From the comparison, the level of risk also can be sorted to see the level of risk in those factors.

6. Results

Table 1 and Figure 5 shows the relationship between the all the Road Intersection Risk towards the Road Accidents in Section 7 Shah Alam. The results were taken from several output which is the Scatter plot pattern, R-Squared, Correlation R, P-value and Coefficient. The scatter plot shows the general relationship pattern of the results between the Road Accident and the Road Intersection Risk Factors. The results show that all Factors gave positive pattern. However, only two of them has strong positive correlation which is the Traffic Waiting Time and The Traffic Volume.

Table 1: Relationship between road intersection risk factors with road accident

Road Accident	Intersection geometry	Traffic Waiting Time	Traffic Volume
R-squared	0.003	0.651	0.562
Correlation R	0.055	0.807	0.749
P-value	0.775	0.005	0.013
Coefficient	0.563	0.034	0.0001
Scatter Plot Pattern	Slightly Positive	Positive	Positive

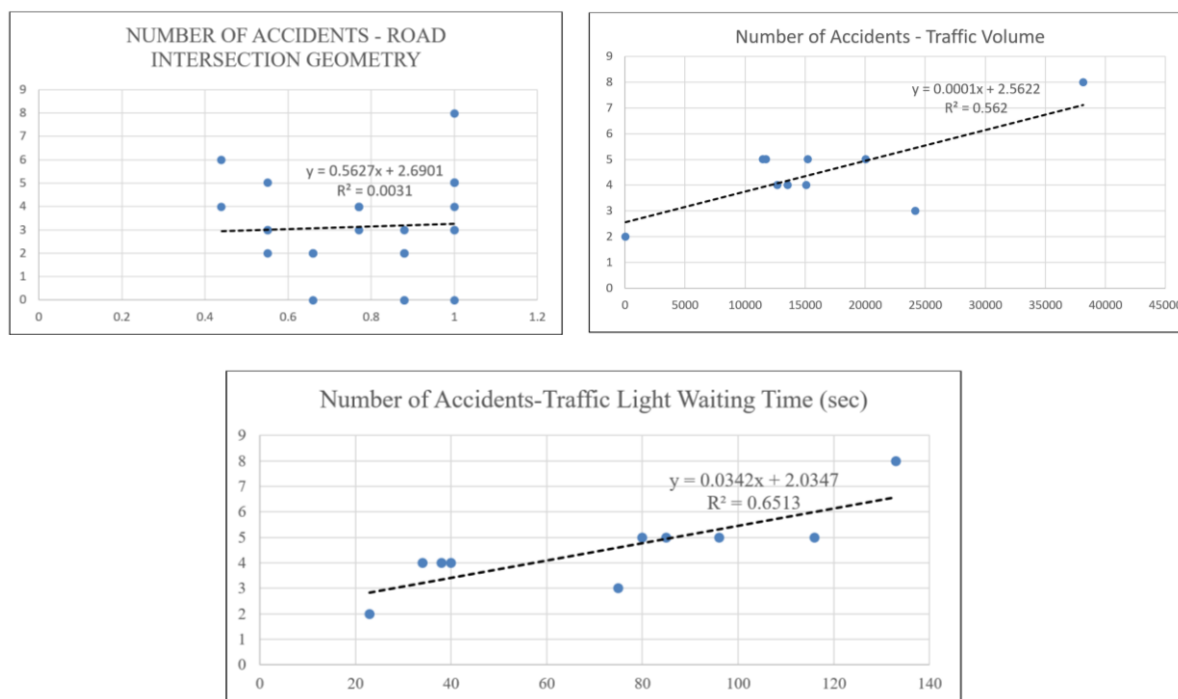


Figure 5: Regression analysis of road intersection's characteristics with road accidents

Because both factors have correlation in increasing Traffic Waiting Time/Traffic Volume results to increasing number of accidents. Both factors also occur at the same Intersection (Number 6) with total of 8 Accidents in the 40m buffer of the intersection. While on the other hand, the Road Intersection Geometry that results a positive linear is not valid as it is supposed to be contra wise. The positive linear is not valid because it indicates that as the Road Intersection Geometry increase, The number of Road Accidents also increases. It is supposed to be that the higher the number of Road Intersection Geometry, the lower the number of Road Accident which results to the negative pattern in the scatter plot.

Next is the R-squared and the Correlation R obtained from the Regression Analysis. The Road Intersection Risk Factors that gave the most positive correlation with the Road Accident is the Traffic Waiting Time factors with the value R-Squared = 0.65131 and Correlation $r = 0.807$ that is the closest to 1. The second factor is the Traffic Volume that show another positive correlation with R-squared = 0.562 Correlation $R = 0.749$. The Road Intersection Geometry gave the lowest correlation. The R-squared obtained is 0.0031 and Correlation $R = 0.055$ which is very low and making it has no correlation towards the road accident.

The p-value that is acceptable is also the Traffic Waiting Time and Traffic Volume which resulting < 0.05 . Traffic Waiting Time shows the lowest p-value = 0.004839. It is followed by Traffic Volume's p-value is 0.012. Showing that both analysis shows that it rejects the null hypothesis saying that there no relation between these both factors with the Road Accident. The Road Geometry Intersection is the only factors that obtains high p-value = 0.774679. The p-value exceed 0.05 making it the factors that shows no relation with the Road Accidents.

Therefore, based on the correlation analysis and OLS conducted between the three factors and the road accidents, it can be concluded that the factor which may contribute the most towards Road Intersection Risks is the Traffic Light's Waiting Time followed by Traffic Volume. There is very small probability to indicate that Road Intersection Geometry influence the Road Intersection Risks as the correlation between the two is too weak, almost no relation.

7. Conclusion

This section discusses a conclusion from the study between the road accidents and the factors influencing the road accident. The factors are road intersection geometry, traffic light waiting time and traffic volume at an intersection. An analysis is conducted to see the relationship between these factors and the road accident cases that is in Section 7, Shah Alam. From the result and analysis, it is found that two out of three of the factors has significant relation with road accident which are the traffic light waiting time and the traffic volume at intersection. It is proved by calculating the correlation and using Ordinary Least Square analysis. The road intersection geometry shows very small relationship with the road accident with the inconsistent data pattern. In a nutshell, the objectives of this study had been achieved. The road intersection risks were possible to be measured by using GIS. The road characteristics that influence road accidents the most is the traffic light waiting time followed by the traffic volume while road geometry shows minimal risk in influencing road accidents.

References

- Carmanah, 2018, *What Is Traffic Calming?*. Traffic Calming Devices and Signs. Carmanah.
- DOSM, 2019, *Statistics on Causes of Death, Malaysia, 2019*, Department of Statistics Malaysia Official Portal.
- Kulanthayan, S., Phang, W. K. and Hayati, K. S., 2007, Traffic Light Violation among Motorists in Malaysia. *IATSS Res.*, Vol. 31(2), 67–73, doi: 10.1016/s0386-1112(14)60224-7.
- Mohamed Eltayeb Zumrawi, M., 2016, Investigating Risk Factors Influencing Safety in National Highways in Sudan. *Am. J. Civ. Eng.*, Vol. 4(6), 276-281, doi: 10.11648/j.ajce.20-160406.12.
- Roslan, A., Jamil, H. M., Harun, N. Z. and Rohayu Sarani, S. A. S. M. R., 2018, *Risk of Road Crashes at 3-legged Intersection*. Malaysian Institute of Road Safety Research (MIROS).
- Wang, C., Quddus, M. A. and Ison, S. G., 2012, Factors Affecting Road Safety: A Review and Future Research Direction. *TRB 91st Annu. Meet.*, 1-17.
- World Health Organization (WHO), 2006, Unit 2. Risk Factors for Road Traffic Injuries in Road Traffic Injury Prevention Manual. WHO Press: Geneva.