037 PREDICTION OF ACCIDENT TREND AT TWO-LANE FEDERAL HIGHWAYS USING STATISTICAL APPROACH

Cik Wan NORBALKISH Jusof¹, ISMAIL Yusof² and Mohd ERWAN Sanik³
¹, ², ³ Universiti Tun Hussein Onn Malaysia, Johor, Malaysia
balz0907@yahoo.com
ismaily@uthm.edu.my
erwans@uthm.edu.my

ABSTRACT: Road accident is an unfortunate event which is a matter of serious concern to the authority. An accident reduction measure taken in reducing the rate of accidents is by identifying hazardous locations for treatment. The purposes of this study are to determine the blackspot location along Federal Route 1, Federal Route 5 and Federal Route 24 in the Batu Pahat area and to develop an Accident Prediction Model by using multiple regression method as well as to carry out sensitivity analysis on the developed model. This study is limited to the two-lane federal road type and therefore multilane highway is not taken into account. Data for explanatory variables such as volume, speed, and road geometry are obtained using the Metrocount equipment and field measurements. Accident point weightage calculation is carried out in order to determine road accident trends and blackspot ranking using accident data provided by the Royal Malaysian Police. Using these data, an accident prediction model is successfully developed. The results from sensitivity analysis revealed that the increment of accident point weightage can be explained by either the rise in speed, number of major access point, and traffic volume. Meanwhile, an increment of lane width and shoulder width will reduce the weighting point rates.

Keywords: Accident Point Weightage, Blackspot ranking, Sensitivity analysis

1. INTRODUCTION

One of comprehensive definitions of road traffic accident is a rare, random, multi-factor event always preceded by a situation in which one or more road users have failed to cope with their environment, resulting in a collision on the public highway which should be recorded by the police (Mohamed Shafii et al., 1995). The first death involving a motor vehicle is said to have taken place in London in 1896 (PIARC, 2003). According to the World Health Organization, every year, nearly one million people killed, three million are severely disabled for life and thirty million are injured in road accidents. Road accidents lay in the ninth place out of a total of over a hundred separately identifiable causes of death or disability (Murray et al., 1996). And, by the year 2020, it has been predicted that road accidents will be the third leading cause of death worldwide (TRL, 1998).

It is possible to identify hazardous sections of the road network so that appropriate remedial measures can be undertaken to reduce the likelihood and severity of accidents at those locations. The term ‘blackspot’ is said to derive from the method that was originally used to identify hazardous sites. The terms
‘hazardous location’ and ‘high accident location’ often used as a synonym to blackspot. A blackspot was originally a road location of limited area with a high concentration of accidents. This definition has progressively evolved as several researches now recommend including the concept of potential for improvement (McGuigan, 1981; Elvik, 1988; Hauer, 1996).

An Accident Prediction Model (APM) is a mathematical equation that expresses the average accident frequency of an entity as a function of traffic flow and other road characteristics. It is important to examine the nature of relationships between roadway, environmental and operational factors and, accidents to understand the causal mechanisms involved in accidents on the one hand and to better predict their occurrence on the other hand. APMs are one path of inquiry often used to gain these insights (Reurings et al., 2005). The Multiple Linear Regression is a statistical methodology describing relationships between a continuous outcome and a set of explanatory variables (Kutner et al., 2005). This technique is appropriate to describe relationships between continuous accident outcomes and explanatory variables. A comprehensive study of road safety (Treat et al., 1977) found that human error was the sole cause in 57% of all accidents and was a contributing factor in over 90%. In contrast, only 2.4% were due solely to vehicle and only 4.7% were caused only by road and environment factor. Based on categorical representation of lane width, for two-lane rural, two-lane urban, four lane urban undivided, and urban freeways, widening lane width up to 4.0m, 3.7m, 4.0m and 4.0m respectively could be expected to decrease crash rates (Hadi et al., 1995). In Australia, excessive speed is an important factor in approximately 20 percent of fatal crashes (Haworth et al., 1993) and speed is a probable or possible cause in 25 percent of rural crashes (Armour et al., 1990). Meanwhile, roads with higher ADT and pedestrian traffic are associated with higher accident frequencies for all highway types (Berhanu, 2004).

2. METHODOLOGY

Four main stages in the methodology are data collection, analysis of accident data and identification of blackspot location, field work, and the development of an accident prediction model.
B.6 Data Collection
In this study, data collection involves two types of data namely primary data and secondary data. Secondary data are accident data obtained from the Royal Malaysian Police (RMP). Accident data was collected from the Batu Pahat Traffic Police Station, Batu Pahat Public Work Department, Bukit Aman RMP Headquater and Jabatan Keselamatan dan Jalan Raya (JKJR). The accident database is extracted from the POL27 and the accident record collected from the year 2007 until May 2010. Meanwhile, primary data includes geometrical data and field work data, specifically from field observation and Metrocount data.

B.7 Analysis of Accident Data
The purpose of preliminary analysis of the traffic accident data is to determine the broad nature of the accident problem. At this stage, the general patterns and trend of accident are established. Types of trend and statistics that take into account for the Federal Route 1 Johor Bahru – Labis (FT001), Federal Route 5 Johor Bahru – Melaka (FT005) and Federal Route 24 Yong Peng – Muar (FT024) are as follow:

j. Road accident statistics for year 2007 to May 2010
k. Accident and casualties for year 2007 to May 2010
l. Accident by hours of the day for year 2007 to May 2010 (every two hours)

B.8 Identification of Blackspot Location
Identification and prioritization of accident blackspot location is carried out through ranking by accident point weightage.

B.9 Field Work
The purpose of field work is to collect primary data at selected 8 locations from all route. Furthermore, the field works also important to document the site findings and photographs taking as evidence. The road sections selected for data collection are Km88, Km106 and Km112 at FT001, Km93 and Km122 at FT005 and Km5, Km11 and Km27 at FT024. These locations are selected based on the ranking of the Accident Point Weightage (APW). The Metrocount equipments are installed for 12 hours at each location in order to obtain data such as traffic volume and 85th percentile speed. Figure
Figure 1 depicts the Metrocount equipment used in the field work. The measurements of lane width, shoulder width and number of access point are carried out manually at the study locations. Figure 2 shows the odometer used to measure the road geometry.

Figure 1. Metrocount

Figure 2. Odometer
B.10 Accident Prediction Model and Sensitivity Analysis

In this study, Microsoft Excel and Minitab statistical software are used in analyzing all data and developing an APM, furthermore to validate the coefficients and the model. The dependent variable for this model is the APW and the explanatory variables are the 85th percentile speed, hourly traffic volume, lane width, shoulder width, total number of access point, and type of road geometric (straight, hilly, curvy). Later, the sensitivity analysis is carried out on the model to determine the effect of each independent variable on the dependent variable; for instance, the effect that changes in speed will have on the APW.

3. RESULTS AND DATA ANALYSIS

The blackspot locations are determined with the ranking of the APW. The data obtained from 9 selected locations which have been installed with the Metrocount equipment are analyzed using the Microsoft Excel before preceded with the development of the APM by using multiple regression method. Furthermore, sensitivity analyses based on the developed model is carried out to check the differences if changes in value made on the independent variables.

3.1 Road Accident Statistics at FT001, FT005 and FT024

Figure 3 shows the accident statistics at the FT001 (Km85-130), FT005 (Km88-151) and FT024 (Km1-36) from year 2007 until May 2010. Referring to the figure, the highest number of total accidents is recorded at the FT005 with 2774 cases followed by FT001 with 1033 cases and FT024 with 925 cases. Of these, 179 fatalities are recorded at all 3 routes in this period of time. The most frequent accident is the damage only cases.
3.2 Accident and Casualty

a) Accident and Casualty at FT001

In this study, records of FT001 involve Km85 to Km130 which is about 45 kilometers stretches. Figure 4 shows that in 2007 and 2008, the total number of accidents recorded along FT001 is the highest with 317 cases in each year. However, the total number of cases reduces in 2009 until May 2010. The total number of accidents recorded from 2007 until May 2010 at the route is 1033 cases. Of these, 45 cases are fatal, 3 cases are serious injuries, 63 cases are slight injuries and 922 are damage only cases.
b) Accident and Casualty at FT005

Records of FT005 involve Km88 to Km151 which is about 63 kilometers stretches. Figure 5 shows that in 2009 the total number of accidents recorded along FT005 are the highest with 844 cases. The total number of cases increased every year. The total number of accidents recorded from 2007 until May 2010 at this route is 2774 cases. Of these, 91 cases are fatal, 31 cases are serious injuries, 357 cases are slight injuries and 2295 cases are damage only.

![Figure 5. Accident and Casualty at FT005 (2007-May 2010)](image)

C) Accident and Casualty at FT024

In this study, records of FT024 involve Km1 to Km36 which is about 36 kilometers stretches. Figure 6 shows that in 2009 the total number of accidents recorded along FT024 are the highest with 295 cases. The total number of accidents recorded from 2007 until May 2010 at this route is 925 cases. Of these, 43 cases are fatal, 8 cases are serious injuries, 77 cases are slight injuries and 797 cases are damage only.

![Figure 6. Accident and Casualty at FT024 (2007-May 2010)](image)
3.3 Accident Data by Hours of the Day

Figure 7 through 9 show the total number of accidents for every 2 hours beginning midday to midnight throughout the year from 2007 to May 2010. The analyses are carried out for every route and the largest number of accidents recorded at 14:01 to 18:00 or 2:01 to 6:00 p.m. This pattern might have happened because the period especially 4:01 to 6:00 p.m. is the peak hours where the density of traffic is high on roads and since the federal roads are at the rural area, speeding may also be a factor contributing to accident occurrence.

a) Accidents by Hours of the Day at FT001

Figure 7 shows that the highest number of accidents occurred at FT001 recorded from 16:01 to 18:00 with 150 cases followed by 139 cases at 14:01-16:00. The least number of accidents happened is at 04:01 to 06:00 with 29 cases are recorded.

![Figure 7. Accident by Hours of the day at FT001](image)

b) Accidents by Hours of the Day at FT005

Figure 8 shows that the highest number of accidents occurred at FT005 recorded from 14:01 to 16:00 with 339 cases followed by 328 cases at 12:01-14:00. The least number of accidents recorded is at 04:01 to 06:00 with 54 cases are recorded.
c) Accidents by Hours of the Day at FT024

Figure 9 shows the highest number of accidents occurred at FT024 recorded from 16:01 to 18:00 with 130 cases followed by 123 cases at 14:01-16:00. The least number of accidents recorded is at 04:01 to 06:00 with 26 cases are recorded.

3.4 Accident Point Weightage Ranking

The calculation for APW is carried out using Equation (1) and ranked using the Microsoft Excel.

\[ APW = X_1(6.0) + X_2(3.0) + X_3(0.8) + X_4(0.2) \]

(1)

Where, \( X_1, X_2, X_3 \) and \( X_4 \) are the number of fatal, serious Injury, slight injury and damage only, respectively
3.5 Selected Study Location

Using Equation (1), APW for every kilometer in a route is calculated and the sections are accordingly ranked. From rank 1 to 5 of every route, 8 sections have been selected for further investigation. Table 1 shows the list of selected sections in every route with regards to type of accident and APW values.

Table 1. 8 selected sections for FT001, FT005 and FT024

<table>
<thead>
<tr>
<th>Route</th>
<th>Rank</th>
<th>Section No.</th>
<th>Fatal</th>
<th>Serious Injury</th>
<th>Slight Injury</th>
<th>Damage only</th>
<th>Total number of accidents</th>
<th>APW</th>
</tr>
</thead>
<tbody>
<tr>
<td>FT001</td>
<td>1</td>
<td>112</td>
<td>3</td>
<td>0</td>
<td>8</td>
<td>145</td>
<td>156</td>
<td>53.4</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>106</td>
<td>3</td>
<td>0</td>
<td>1</td>
<td>14</td>
<td>18</td>
<td>21.6</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>88</td>
<td>3</td>
<td>0</td>
<td>1</td>
<td>7</td>
<td>11</td>
<td>20.2</td>
</tr>
<tr>
<td>FT005</td>
<td>2</td>
<td>93</td>
<td>2</td>
<td>2</td>
<td>20</td>
<td>87</td>
<td>111</td>
<td>51.4</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>122</td>
<td>3</td>
<td>2</td>
<td>13</td>
<td>56</td>
<td>74</td>
<td>45.6</td>
</tr>
<tr>
<td>FT024</td>
<td>1</td>
<td>27</td>
<td>4</td>
<td>0</td>
<td>13</td>
<td>81</td>
<td>98</td>
<td>50.6</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>5</td>
<td>4</td>
<td>0</td>
<td>1</td>
<td>37</td>
<td>42</td>
<td>32.2</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>11</td>
<td>4</td>
<td>0</td>
<td>2</td>
<td>16</td>
<td>22</td>
<td>28.8</td>
</tr>
</tbody>
</table>

3.6 Number of Accidents at Study Locations

Figure 10 depicts the accident data at every selected section. Referring to the figure, for the fatal accident, Km 5, Km 11 and Km27 from FT024 recorded the highest number with 4 cases at each location while the lowest is recorded with 2 cases at Km93 (FT005). Furthermore, for the serious injury accident, Km93 (FT005) and Km122 (FT005) record 2 cases at each location while the other locations record none. Km93 (FT005) records the highest number of slight injury accident with 20 cases while Km106 (FT001), Km88 (FT001) and Km5 (FT024) record the lowest number with 1 case at each location. The highest number of damage only accident is at Km112 (FT001) with 145 cases and the lowest is at Km88 (FT001) with 7 cases. For overall, the location with highest total number of accidents is Km112 (FT001) with 156 cases while the lowest is Km88 (FT001) with 11 cases.
3.7 Model Development

In this study, multiple regression analysis is conducted using the Minitab. After considering the best fitted model and with transformation of level-level linear regression to log-level linear regression, the APM that has been developed is as follows:

\[
\ln (\text{APW}) = 0.000098(\text{HTV}) + 0.00469(85^{th} \text{ PS}) + 0.0476(\text{AP}) + 0.137(\text{ST}) - 0.4386(\text{HI}) - 0.170(\text{LW}) - 0.0602(\text{SW}) + 3.49
\]

(2)

Where,
APW = accident point weightage
HTV = hourly traffic volume
85\text{th} PS = 85\text{th} percentile speed
AP = number of access points per kilometer
ST = straight road geometry
HI = hilly road geometry
LW = lane width
SW = shoulder width

a) From the analysis, the developed model has an R-squared of 0.994 and adjusted R-squared of 0.993, which means 99.4% and 99.3% of the variation in the number of accidents has been explained the regression line. Table 2 shows the range of explanatory variables values can be used for this model.
Table 2. Range of values for independent variables in Accident Prediction Model developed

<table>
<thead>
<tr>
<th>Factor</th>
<th>Range of Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hourly Traffic Volume (HTV)</td>
<td>494 - 2415 veh/hr</td>
</tr>
<tr>
<td>85th Percentile Speed (85th PS)</td>
<td>60.8 – 98.3 km/hr</td>
</tr>
<tr>
<td>Number of access points (AP)</td>
<td>0 – 14</td>
</tr>
<tr>
<td>Lane width (LW)</td>
<td>3.4 – 4.2 m</td>
</tr>
<tr>
<td>Shoulder width (SW)</td>
<td>0.35 – 2.8 m</td>
</tr>
</tbody>
</table>

4. CONCLUSION

This study used the APW from the HPU to determine the black spot locations before ranking them in one kilometer at each location. Further analysis of this study is to develop predictive model relating accident weighting point with the selected independent variables has been achieved. Multiple regression techniques are used to build the model. The model is as shown in Section 3.8. The result of the sensitivity analysis provide sufficient evidence to support the hypothesis that the independent variables include Hourly Traffic Volume, 85th Percentile Speed, total number of access points, lane width and shoulder width are related to the accident weighting point. It can be conclude that for instance:

- Doubling the hourly traffic volume is expected to cause a 26.7% increase in percentage of the predicted APW.
- A 10% increment of 85th percentile speed is expected to cause a 4.7% increase in percentage of the predicted APW.
- Widening to lane by 10% increment from the existing width is expected to reduce the percentage of the predicted APW for about 6.9%.
- Widening to shoulder by 10% increment from the existing width is expected to reduce the percentage of the predicted APW for about 1.7%.
- An increase of 5 point in approach number of access point is expected to cause a 26.9% increase in the percentage of the predicted APW.

Finally, the developed Accident Prediction Model in this study is expected to be applicable in road safety improvement especially to reduce the number of accidents at two-lane federal roads by emphasizing into consideration the factors that predict the model at the design and planning stages.
5. ACKNOWLEDGEMENT

We would like to thanks the Batu Pahat Police Station and the Jabatan Kerja Raya, Batu Pahat Branch for their cooperation in getting the data.

6. REFERENCES


