Developing crash modification factors for roundabouts using a cross-sectional method

Mohammad Nour Al-Marafi*, Kathirgamalingam Somasundaraswaran, Ron Ayers

School of Civil Engineering and Surveying, Faculty of Health, Engineering and Sciences, University of Southern Queensland, Toowoomba, QLD 4350, Australia

Highlights

- Estimation of crash frequency at roundabouts using negative binomial (NB) error distribution.
- Application of the cross-sectional method to develop crash modification functions.
- Significant risk factors that raise safety issues at roundabouts were investigated.

Abstract

The objective of the current study was to evaluate traffic and geometric features and their influences on the safety performance of roundabouts by developing suitable crash modification factors (CMFs). The cross-sectional method can be applied as an alternative method to estimate the CMFs when before-and-after studies are impractical to apply, e.g., lack of data from the period after implementing treatments. To accomplish the study objective, CMFs were derived from generalised linear models (GLMs), i.e., negative binomial (NB) regression, using data collected on regional roundabouts in Toowoomba City, Australia. Six years of crash data from 49 roundabouts included all recorded crashes as well as traffic and geometric features for the entire roundabouts. Several candidate models were developed using the GLMs. Five models were selected based on statistical significance, goodness-of-fit (GOF) measures, and cumulative residual (CURE) analysis. The results show that increasing the number of entry lanes, entry width, entry radius, traffic volume, circulatory roadway width, weaving width, and speed limit have positive effects on roundabout safety. On the other hand, increasing the number of legs, number of exit lanes, exit width, exit radius, weaving length, central island diameters, and presence of fixed object on a central island have negative effects on roundabout safety. The study shows that quantifying the risk factors can support road safety stakeholders to identify safety improvements at roundabouts more effectively and efficiently.

© 2019 Periodical Offices of Chang'an University. Publishing services by Elsevier B.V. on behalf of KeAi Communications Co. Ltd. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).
1. Introduction

Roundabouts are usually associated with a positive impact on traffic safety compared to other types of at-grade intersections. Thus, the road authorities are considering roundabouts as the preferred choice over the other types of traffic control such as stop signs and traffic signals (Polders et al., 2015). In particular, roundabouts have a low number of potential conflict points and their geometry motivates to reduce the vehicle speeds to where it helps to reduce the delay and reduces the number of decision points for road users (Daniels et al., 2011). In regional areas where the traffic volume through an at-grade intersection is moderate the use of roundabouts has increased as an effective way of controlling traffic.

In Australia, roundabouts have been used widely in both urban and rural areas. However, with the number of roundabouts increasing in regional areas, it is important to make sure that both existing and new roundabouts are safer for the road users. In particular, there is a need to consider the traffic and geometric characteristics of roundabouts that significantly affect both crash frequency and severity.

The main objective of the current study is to estimate crash modification factors (CMFs) to identify the safety performance for various traffic and geometric characteristics at roundabouts in Toowoomba City. To accomplish this objective, initially, crash prediction models (CPMs) were developed using a negative binomial (NB) distribution with a log-linear function. In addition, several goodness-of-fit (GOF) statistics were employed to evaluate the suitability of the models. The study results apply to those regional roundabouts with similar geometric and traffic conditions.

The remainder of the study is organized as follows. The second section presents the previous studies related to the development CPMs and CMFs. The third section describes the data used in the analysis. The fourth section presents model development. The fifth section describes the CMFs estimation. The last section draws conclusions from the analysis performed in this study.

2. Literature review

A number of studies have been conducted to investigate the effects of the geometric elements and traffic conditions on safety at roundabouts (Anjana and Anjaneyulu, 2014; Daniels et al., 2011; De Brabander and Vereeck, 2007; Farag and Hashim, 2017; Kamla et al., 2016). In order to better understand crash causes and contributing factors, the researchers have paid considerable attention to developing different analytical approaches. The generalized linear model (GLM) approach (i.e., Poisson and negative binomial (NB) models) have proven to be a reliable method to reveal the relationship between the road crashes and explanatory variables. This is due to the fact that Poisson and NB distributions are able to describe adequately the random, non-negative, discrete, and typically sporadic events which are characteristics of crash frequency (Abdel-Aty and Radwan, 2000; Ackaaah and Salifu, 2011; Hadi et al., 1995). Abdel-Aty and Radwan (2000) stated that the Poisson distribution has some limitations, such that it is not able to handle the over-dispersion. The phenomenon of “over-dispersion” occurs when the observed variance is greater than the mean. In contrast, NB distribution does not require the assumption of observed variance being equal to the mean (Anjana and Anjaneyulu, 2014; Chin and Quddus, 2003; Kamla et al., 2016). In such a case, Kamla et al. (2015) developed a crash prediction model (CPM) using NB distribution to investigate the impacts of roundabout geometric and traffic characteristics on safety. Similar to this study, Daniels et al. (2011) employed Poisson distribution to identify the safety performance at roundabouts. In these studies, the NB or Poisson distributions were selected for safety analysis based on the dataset type, i.e., NB when the dataset was over-dispersed and Poisson when it was not.

Kim and Choi (2013) identified the major factors associated with road crashes at roundabouts in South Korea. In this study, NB distribution models were applied to analyse the impact of contributory factors on road safety using data from 14 roundabouts, where a total of eleven explanatory variables were examined. The results showed that six explanatory variables have significant impacts on roundabout safety including: number of approaches, circulating lane width, entry width, flare length, flare width, and circulating lane radius. Likewise, five explanatory variables have no significant impacts on roundabout safety including: inscribed circle diameter, central island diameter, number of entering lanes, entry lane radius, and number of circulating lanes. It is worth mentioning that this study has some limitations such as the use of a small sample size. Kamla et al. (2016) investigated the traffic and geometric characteristics and their impacts on the frequency of crashes, where crash records from a total of 70 roundabouts were used. The results indicated that the crash frequency tended to increase as the traffic volume and inscribed circle diameter increased.

Crash modification factors (CMFs) identify the change in road safety (crash frequency) resulting from implementing a particular treatment. This treatment may be in the form of design modification, change in traffic operations, or any countermeasures. The recognition of any change in geometric design features or traffic operation will increase or decrease crash frequency. There are several methods to estimate CMF values and these methods vary from a before-and-after study with a comparison group to relatively more sophisticated methods such empirical Bayes (EB) and full Bayes (FB) methods. These methods include estimating safety performance based on safety data before-and-after a specific treatment is implemented on either one or several sites (Shahdah et al., 2014). The EB and FB methods can be used to control for regression-to-mean (RTM) bias associated with observational studies (Gross et al., 2010; Hanley et al., 2000; Persaud and Lyon, 2007; Wood et al., 2015). Although EB and FB methods are considered as the more preferred methods for estimating CMFs, there are some practical limitations associated with these methods such as countermeasures or treatment implementation dates should be known to determine the before-and-after evaluation periods, sufficient years have to pass after treatments are implemented, and it
is difficult to distinguish safety effects when more than one treatment has been implemented at a specific site (Hauer, 1997; Persaud et al., 2010; Wood et al., 2015). Alternative safety evaluation methods are required to overcome these limitations, but they should be able to address RTM bias that is common to observational studies.

The cross-sectional method has been widely used in the recent years to overcome these issues (Anjana and Anjanyulu, 2014; Gross et al., 2010; Li et al., 2010; Park et al., 2015; Wu and Lord, 2016). In this method, the value of a CMF can be estimated directly from the coefficient of the variable associated with the proposed treatment. Thus, it is not necessary to have data on a specific treatment before-and-after implementation compared to other methods. It is worth mentioning that the cross-sectional method does not take into account the effects of factors that are not included in the analysis, i.e., external causal factors (Gross et al., 2010; Hauer, 2013). Another criticism is that a sufficient sample size is especially required when large explanatory variables are included in the developed model. Park et al. (2015) and Wood et al. (2015) evaluated the treatment effectiveness using both an EB observational before-and-after method and a cross-sectional method. The studies concluded that the results from the cross-sectional method seem to be consistent with the EB method results. However, AASHTO (2010) indicated that the cross-sectional method might be appropriate when observational before-and-after studies are not practical due to data restrictions (e.g., crash data in the before period are not available).

Ideally, it is not logical to assume a systematic safety effect for all treated sites with different characteristics. For instance, greater benefits of safety improvements may be obtained at the sites with high traffic volume. Thus, as a part of the cross-sectional method, a crash modification function (CMFunction) formula can be developed to estimate the variation in the values of CMF with different sites characteristics. This method has already investigated by researchers to estimate the effects of safety improvements (Elvik, 2011; Gross et al., 2010; Gross and Donnell, 2011; Lee et al., 2015; Park et al., 2015; Sacchi et al., 2014). However, it should be pointed out that there are few studies that have investigated the effects of safety improvements at roundabouts through using CMFunctions.

3. Data preparation

The current study is conducted using the crash data from 49 roundabouts in Toowoomba City, Australia. The selected roundabouts consist of 47 single-lane roundabouts and two multi-lane roundabouts. For each roundabout, the observed crashes, traffic volume, and geometric features were collected for the years 2010–2015. Crash data was obtained from the Department of Transport and Main Roads (DTMR), Queensland. Road geometric features were collected from historical design records, site visits, and Google Earth.

The datasets were divided into two groups. The first group of data was used to develop the crash prediction models based on three years (2010–2012). The second group was used for validation of the models against additional years (2013–2015) of crash data for the same roundabouts used in the development of the models. This validation was used to evaluate the models’ capability to predict crashes across time. Twenty-one explanatory variables describing traffic and road geometry were used in modelling as the most common factors which have been associated with road crashes at the roundabouts. A statistical summary of the explanatory variables considered in the development of safety models and the manner in which they are defined in the datasets is shown in Table 1.

Likewise, the roundabout geometric features include number of lanes entering and exiting, width of entry and exit lane, average radius of entry and exit path, width of circulating roadway, length and width of weaving section, and central island diameter, and other associated elements are identified in Figs. 1 and 2. The examples of roundabout layouts that were used in the study are presented in Fig. 3.

4. Model development

4.1. Model selection and estimation

The CPMs were developed using a generalized linear modelling (GLM) approach. Two types of GLM were identified for use in this study: negative binomial (NB) and Poisson distributions. As mentioned previously, these two types are more appropriate to analysing crash data (Abdul Manan et al., 2013;
In order to find which of these two models was suitable for estimating safety outcomes, the study adopted the over-dispersion assumption. The phenomenon of "over-dispersion" occur when the observed variance is greater than the mean of the datasets. Initially, the distributions of crash counts were assumed to follow a negative binomial distribution that deals with over-dispersion within the datasets. This assumption has been tested based on the value of the deviance divided by the degree of freedom (df) as well as the value of the Pearson chi-square ($\chi^2$) divided by the degree of freedom (df). If the result of these tests lies between 0.8 and 1.2, the NB model assumption will be accepted. Also, if it is out of this range the Poisson model will be used instead of NB model (Abdul Manan et al., 2013).

The general form of the predicted model by using Poisson or NB model assumption for the $i$th roundabout can be written in the form of Eq. (1).

\[ N_{\text{pre},i} = Q_{\text{major},i}^{a_1}Q_{\text{minor},i}^{a_2}e^{-\beta_0 - \sum_j \beta_j X_{ij}} \] (1)

where $N_{\text{pre},i}$ is predicted crash frequency at $i$th roundabout, $Q_{\text{major},i}$ and $Q_{\text{minor},i}$ are annual average daily traffic (AADT) on major and minor approach at $i$th roundabout, respectively, $X_{ij}$ is explanatory variable $j$ at $i$th roundabout, and $a_1$, $a_2$, $\beta_0$, and $\beta_j$ are model parameters.

Initially, the correlation among the explanatory variables were tested as they were useful to prevent the use of strongly correlated variables together within a model, i.e., strong correlation variables would strongly affect the other parameters in the same model. In particular, any two explanatory variables whose correlation test had between -0.49 and 0.49 (moderate correlation) was proposed in modelling. In addition, the variable parameters were considered to be statistically significant at 0.1 significance level (using 90% confidence).

The data analysis and model development was undertaken using SPSS software version 23 (IBM Corp., 2015). Different models were developed and fitness of results was assessed
based on the confidence level and the correlation values between the variables. Furthermore, a comparison of the developed models was performed using goodness-of-fit (GOF) measures including Akaike information criterion (AIC) and Bayesian information criterion (BIC). The smaller of the AIC and BIC values was considered better than the other models with higher values (Abdul Manan et al., 2013; Cafiso et al., 2010). After several trials of a different combination of variables, five models were identified and estimated using negative binomial (NB) error structure with log link function. The estimated regression parameters for the selected road safety models for the roundabouts are presented in Table 2. It is worth mentioning that some main explanatory variables (e.g., traffic volume on major approach) showed slightly stronger correlation with the other variables. Due to this correlation these variables have a p-value higher than 0.1.

The deviance and Pearson chi-square ($\chi^2$) statistics divided by its degrees of freedom (df) were estimated to be 0.916 and 0.860 for model I, 0.984 and 0.907 for model II, 0.856 and 0.871 for model III, 1.177 and 1.076 for model IV, and 1.086 and 1.081 for model V respectively as shown in Table 3. Specifically, the values of these two tests are within the allowable range (i.e., 0.8 and 1.2) implying that the NB distribution assumption is acceptable.

The GOF for the selected models was also investigated using the cumulative residuals (CUREs) plot. This method needed to achieve two conditions to indicate that the model fitted the data well: 1) the curve lies within two standard deviations ($\pm 2\sigma$ and $-2\sigma$ boundaries) of the mean and 2) the curve oscillate around zero. Fig. 4 shows the CURE plot, as a function of AADT, for all selected models. As noted in this figure, the CURE curve for all selected models are within the standard deviation boundaries which mean that all models are fitting the data well.

4.2. Model validation

The validation measures were used in this study to assess the models’ ability to predict road crashes over additional years. Four performance measures were used to validate the models including the mean squared prediction error (MSPE), mean absolute deviation (MAD), mean squared error (MSE), and Freeman–Tukey R-squared coefficient ($R^2_{FT}$). These performance measures can be calculated using the following equations (Washington et al., 2005).

\[
\text{MSPE} = \frac{1}{n} \sum_{i=1}^{n} (\hat{y}_i - y_i)^2 \tag{2}
\]

\[
\text{MSE} = \frac{1}{n-p} \sum_{i=1}^{n} (y_i^\prime - y_i)^2 \tag{3}
\]

\[
\text{MAD} = \frac{1}{n} \sum_{i=1}^{n} |y_i^\prime - y_i| \tag{4}
\]

\[
R^2_{FT} = \frac{\sum_{i=1}^{n} (\hat{f}_i - \hat{f})^2 - \sum_{i=1}^{n} e_i^2}{\sum_{i=1}^{n} (\hat{f}_i - \hat{f})^2} \tag{5}
\]
Table 2 – Negative binomial parameter estimates for selected models.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Model I</th>
<th></th>
<th>Model II</th>
<th></th>
<th>Model III</th>
<th></th>
<th>Model IV</th>
<th></th>
<th>Model V</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( \hat{\beta} )</td>
<td>p-value</td>
<td></td>
<td></td>
<td>( \hat{\beta} )</td>
<td>p-value</td>
<td>( \hat{\beta} )</td>
<td>p-value</td>
<td>( \hat{\beta} )</td>
<td>p-value</td>
</tr>
<tr>
<td>Intercept</td>
<td>-15.930</td>
<td>0.000</td>
<td>-15.471</td>
<td>0.000</td>
<td>-10.618</td>
<td>0.000</td>
<td>-10.616</td>
<td>0.003</td>
<td>-12.606</td>
<td>0.000</td>
</tr>
<tr>
<td>Ln1</td>
<td>-</td>
<td></td>
<td>0.467</td>
<td>0.121</td>
<td></td>
<td></td>
<td>-</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ln2</td>
<td>-</td>
<td></td>
<td>-</td>
<td></td>
<td>-0.222</td>
<td>0.233</td>
<td>-</td>
<td></td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>Le1</td>
<td>0.338</td>
<td>0.008</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-0.222</td>
<td>0.233</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>Le2</td>
<td>-</td>
<td></td>
<td>-</td>
<td></td>
<td>0.022</td>
<td>0.233</td>
<td>-</td>
<td></td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>En1</td>
<td>-</td>
<td></td>
<td>-</td>
<td></td>
<td>-</td>
<td></td>
<td>-</td>
<td></td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>En2</td>
<td>-</td>
<td></td>
<td>0.367</td>
<td>0.004</td>
<td>-</td>
<td></td>
<td>-</td>
<td></td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>Ex1</td>
<td>-0.068</td>
<td>0.000</td>
<td>-</td>
<td></td>
<td>-</td>
<td></td>
<td>-0.005</td>
<td>0.108</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>Ex2</td>
<td>-</td>
<td></td>
<td>-</td>
<td></td>
<td>-</td>
<td></td>
<td>-</td>
<td></td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>Rn1</td>
<td>-</td>
<td></td>
<td>-</td>
<td></td>
<td>-</td>
<td></td>
<td>-0.222</td>
<td>0.000</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>Rn2</td>
<td>-</td>
<td>0.035</td>
<td>0.000</td>
<td></td>
<td>-</td>
<td></td>
<td>-</td>
<td></td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>Rx1</td>
<td>-</td>
<td></td>
<td>-</td>
<td></td>
<td>-</td>
<td></td>
<td>-</td>
<td></td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>Rx2</td>
<td>-</td>
<td></td>
<td>-</td>
<td></td>
<td>-0.052</td>
<td>0.103</td>
<td>-</td>
<td></td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>Q_major</td>
<td>0.241</td>
<td>0.117</td>
<td>1.163</td>
<td>0.000</td>
<td>0.403</td>
<td>0.063</td>
<td>0.954</td>
<td>0.000</td>
<td>0.438</td>
<td>0.004</td>
</tr>
<tr>
<td>Q_minor</td>
<td>1.121</td>
<td>0.000</td>
<td>-</td>
<td></td>
<td>0.915</td>
<td>0.000</td>
<td>-</td>
<td></td>
<td>0.923</td>
<td>0.000</td>
</tr>
<tr>
<td>F</td>
<td>0.305</td>
<td>0.033</td>
<td>-</td>
<td></td>
<td>0.063</td>
<td>0.208</td>
<td>-</td>
<td></td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>CW</td>
<td>-</td>
<td></td>
<td>-</td>
<td></td>
<td>-</td>
<td></td>
<td>0.063</td>
<td>0.208</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>WL</td>
<td>-</td>
<td></td>
<td>-</td>
<td></td>
<td>-</td>
<td></td>
<td>-</td>
<td></td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>WW</td>
<td>-</td>
<td></td>
<td>0.035</td>
<td>0.005</td>
<td>0.023</td>
<td>0.023</td>
<td>0.138</td>
<td>0.000</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>CD</td>
<td>-0.005</td>
<td>0.001</td>
<td>0.012</td>
<td>0.037</td>
<td>-0.202</td>
<td>0.000</td>
<td>-</td>
<td></td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>V</td>
<td>0.038</td>
<td>0.057</td>
<td>0.079</td>
<td>0.068</td>
<td>0.005</td>
<td>0.103</td>
<td>0.005</td>
<td>0.108</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>Dispersion (k)</td>
<td>0.208</td>
<td></td>
<td>0.110</td>
<td></td>
<td>0.200</td>
<td></td>
<td>0.220</td>
<td></td>
<td>0.203</td>
<td></td>
</tr>
</tbody>
</table>

Note:

a Computed based on the Pearson chi-square.

b Significance at 0.1 level.

c Fixed object is 1 if present and is 0 if not present.

Table 3 – Goodness-of-fit tests for predicted models.

<table>
<thead>
<tr>
<th>Model</th>
<th>Parameter</th>
<th>Value</th>
<th>df</th>
<th>Value/df</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>Deviance</td>
<td>37.557</td>
<td>41</td>
<td>0.916</td>
</tr>
<tr>
<td></td>
<td>Pearson chi-square ( (x^2) )</td>
<td>35.266</td>
<td>860</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Akaike’s Info. criterion (AIC)</td>
<td>156.265</td>
<td>0.000</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Bayesian Info. criterion (BIC)</td>
<td>171.400</td>
<td>0.000</td>
<td></td>
</tr>
<tr>
<td>II</td>
<td>Deviance</td>
<td>40.348</td>
<td>41</td>
<td>0.984</td>
</tr>
<tr>
<td></td>
<td>Pearson chi-square ( (x^2) )</td>
<td>37.179</td>
<td>0.000</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Akaike's info. criterion (AIC)</td>
<td>153.512</td>
<td>0.000</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Bayesian info. criterion (BIC)</td>
<td>168.647</td>
<td>0.000</td>
<td></td>
</tr>
<tr>
<td>III</td>
<td>Deviance</td>
<td>35.937</td>
<td>42</td>
<td>0.856</td>
</tr>
<tr>
<td></td>
<td>Pearson chi-square ( (x^2) )</td>
<td>36.584</td>
<td>0.871</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Akaike’s info. criterion (AIC)</td>
<td>152.227</td>
<td>0.000</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Bayesian info. criterion (BIC)</td>
<td>165.470</td>
<td>0.000</td>
<td></td>
</tr>
<tr>
<td>IV</td>
<td>Deviance</td>
<td>48.262</td>
<td>42</td>
<td>1.177</td>
</tr>
<tr>
<td></td>
<td>Pearson chi-square ( (x^2) )</td>
<td>44.118</td>
<td>1.076</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Akaike’s info. criterion (AIC)</td>
<td>154.373</td>
<td>0.000</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Bayesian info. criterion (BIC)</td>
<td>169.508</td>
<td>0.000</td>
<td></td>
</tr>
<tr>
<td>V</td>
<td>Deviance</td>
<td>46.719</td>
<td>43</td>
<td>1.086</td>
</tr>
<tr>
<td></td>
<td>Pearson chi-square ( (x^2) )</td>
<td>46.490</td>
<td>1.081</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Akaike’s info. criterion (AIC)</td>
<td>147.967</td>
<td>0.000</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Bayesian info. criterion (BIC)</td>
<td>159.318</td>
<td>0.000</td>
<td></td>
</tr>
</tbody>
</table>

where \( y_i \) is predicted crashes number at \( i \)th roundabout, \( y_i \) is observed crashes number at \( i \)th roundabout, \( n \) is sample size of database, \( p \) is number of model parameters, \( f_i \) is Freeman–Tukey transform of \( y_i \), \( f_i = \sqrt{y_i + \frac{1}{4}} \), \( f' \) is sample mean of \( f_i \), \( \bar{f} \) is Freeman–Tukey deviate at \( i \)th roundabout \( (\bar{f}_i = f_i - \sqrt{4f_i^{2}} + 1) \).

MSPE is used to determine the variance of the difference between observed crashes and predicted crashes results. In addition, it is typically employed to evaluate error associated with a validation dataset. MSE is typically employed to evaluate error associated with an estimation dataset. Ideally, MSPE and MSE results can be used to reveal whether the models are over-fitted (MSPE > MSE) or under-fitted (MSPE < MSE) (Bissonette and Cramer, 2008; Washington et al., 2005; Young and Park, 2013). When the values of MSPE and MSE are similar, this indicates that the validation dataset fit the developed model similar to the estimation dataset. The MAD value provides a measure of the average magnitude of the prediction variability. In general, a smaller value (closer to zero) of MSPE, MAD, or MSE refers to lower prediction error. Likewise, the higher values of \( R^2_F \) indicate a better prediction performance and vice-versa (Washington et al., 2005).

Table 4 shows the results of the validation tests for the estimation dataset (2010–2012) and the validation dataset (2013–2015). The models were developed using the estimation dataset. The values of MSPE using validation dataset and MSE using estimation dataset are similar for all developed models, which represents a high level of transferability of the models. The same result was obtained for MAD where the estimation data and the validation data are similar for all developed models, whereas, the \( R^2_F \) test results were slightly lower for the validation data than that for the estimation data. This could be due to the difference of the datasets used to estimate and validate the models.
5. Estimating crash modification factors

5.1. Crash modification function

CMFunction method was employed in this study to estimate the road safety effect for each explanatory variable that was used in developing the CPMs at roundabouts. More specifically, this method was applied based on the parameter of the explanatory variable associated with the proposed treatment type. In this method, the value of CMF was estimated for a particular treatment type (i.e., variable) using Eq. (6) as follows (Lord and Bonneson, 2007).

\[
CMF_i = e^{\beta_i (X_i - X_0)}
\]  

(6)

where \(X_i\) is observed value for the variable \(i\), \(X_0\) is base condition for the variable \(i\), \(\beta_i\) is model parameter for the variable \(i\).

A CMF value of 1.0 represents no effect on safety, while CMF above 1.0 indicates a treatment resulting in a higher number of crashes. In contrast, a CMF below 1.0 indicates a treatment resulting in lower crash numbers.

The standard error (SE) of the CMF for each treatment type was also calculated using Eq. (7) as follows (Bahar, 2010; Harkey et al., 2008; Park et al., 2015).

\[
SE_i = \frac{e^{\beta_i (X_i - X_0) - SE_{\beta_i}} - e^{\beta_i (X_i - X_0) + SE_{\beta_i}}}{2}
\]

(7)

where \(SE_i\) is standard error of the CMF, \(SE_{\beta_i}\) is standard error of the model parameter \(\beta_i\).

It should be noted that when the value of standard error equals 0.1 or less, this indicates that a CMFunctio result is more reliable. The base condition values in this study were adopted from previous studies as well as the mean values of an individual explanatory variable. By definition, the base condition can be defined as the condition associated with CMF value 1.0.

Table 5 shows the CMFunctions used to estimate the values of CMF for safety effects of the traffic and geometric elements of a roundabout. CMFunctions were derived from the developed models (i.e., Models I–V) based on the presence of the explanatory variable and the goodness-of-fit for the model. It can be noted that the models were developed

Fig. 4 – Cumulative residual (CURE) plots for roundabout models. (a) Model I. (b) Model II. (c) Model III. (d) Model IV. (e) Model V.
based on the total entry lanes on major and minor approaches. Consequently the associated regression parameters (i.e., 0.564 for major and 0.022 for minor) have been doubled for both major and minor approaches. Therefore, the regression parameters were divided by two to estimate the CMFs for the number of entry lanes based on each entry approach (Li et al., 2010; Lord and Bonneson, 2007). The same method was used for exit lanes on major and minor approaches, where the associated regression parameters have been doubled.

5.2. Discussion of CMF results

The following sections discuss the safety effects of different traffic and geometric elements based on the values of CMF.

5.2.1. Number of legs

The 4-legged roundabout was adopted as a base condition to estimate CMFs. The results revealed that the 5-legged roundabout was associated with more crashes than 3-legged and 4-legged roundabouts. When the roundabout changed from 4-legged to 3-legged the number of crashes reduced by 37% and in the same way when the number of legs increased from 4-legged to 5-legged the number of crashes increased by 60%. This result was expected because the traffic volume and vehicle interactions at roundabouts will increase after adding more legs. A similar result has also been concluded in previous studies (Kim and Choi, 2013; Shadpour, 2012). It should be pointed out that the number of roundabout legs should preferably be limited to 4, as increased conflicts occur at multi-lane roundabout exits.

5.2.2. Number of entry lanes

The results indicate that the number of entry lanes was associated with more crashes for both major and minor approaches. For instance, after adding one entry lane on a major approach or a minor approach, probability of crashes increases by 25% and 1%, respectively. It can be noticed that the effect of the number of entry lanes at a major approach is found to be more significant than at a minor approach and this is probably due to the difference in traffic volume. Turner et al. (2009) also concluded that the multiple entry lanes are associated with greater crash frequency. In general, the number of entry roundabout lanes provided on major or minor approaches should be limited to the minimum number that meets the required capacity and operating requirements for the traffic volumes.

5.2.3. Number of exit lanes

The results indicate that road crashes increased by 18% and 4% after adding one exit lane on a major approach and a minor approach, respectively. This result was expected because the number of conflict points increases at the multi-lane entrances and exits when compared to the single-lane conditions. Therefore, the number of exit lanes should be limited by the number of circulating lanes to prevent the conflict between the merging and diverging vehicles.

5.2.4. Entry width

The results show that wider entry width at major and minor approaches was associated with higher road crash numbers.
compared with narrow width. This result is possible because the wider entry width is associated with higher vehicles speed at the entry of the roundabout. Designers should therefore aim to make entry lane widths no wider than necessary. Furthermore, the entry width must be able to accommodate the path of entering design vehicles. Fig. 5 represents the effect of entry width on road safety for both minor and major approaches.

5.2.5. Exit width

The study was also examined the effect of exit width in major and minor approaches at the roundabouts. The results revealed that a wider exit width for both major and minor approaches increased road safety. This result is possibly because the wider exit width increases comfort for drivers to exit the roundabout safely and to ensure that the exit width accommodates the swept path of the design vehicle (AUSTROADS, 2015). In roundabout design it is usually desirable to reduce entry width and entry path radius to slow vehicles, but to allow for vehicles to accelerate on the exit. Thus, the width of the exit must usually be wider than the entering width. Fig. 6 shows the relationship between exit width and road safety, where the exit width on minor approaches appears to have less impact on road safety compared to exit width on major approach.

5.2.6. Entry radius

The entry radius or entry path radius is one of the most important factors among geometric parameters at a roundabout, since it affects both safety and capacity (Montella et al., 2012). A large entry path radius usually results in faster entry speeds and results in additional road crashes. The larger entry path radius for both minor and major approach is associated with more road crashes at roundabout as shown in Fig. 7. Also, it can be noticed from the figure that the effect on CMF values of entry path radius for both minor and major approach is roughly the same.

![Fig. 5 – CMF for entry width.](image)

![Fig. 6 – CMF for exit width.](image)

### Table 5 – Estimated CMFs using a cross-sectional method.

<table>
<thead>
<tr>
<th>Roundabout feature</th>
<th>Base value a</th>
<th>CMFunction</th>
<th>SE of model parameter b</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lg</td>
<td>4 legs</td>
<td>e^(0.467(Lg-4))</td>
<td>0.050</td>
</tr>
<tr>
<td>LN₁</td>
<td>2 lanes per approach</td>
<td>e^(0.282(LN₁-2))</td>
<td>0.021</td>
</tr>
<tr>
<td>LN₂</td>
<td>2 lanes per approach</td>
<td>e^(0.011(LN₂-2))</td>
<td>0.014</td>
</tr>
<tr>
<td>LE₁</td>
<td>2 lanes per approach</td>
<td>e^(0.169(LE₁-2))</td>
<td>0.028</td>
</tr>
<tr>
<td>LE₂</td>
<td>2 lanes per approach</td>
<td>e^(0.040(LE₂-2))</td>
<td>0.236</td>
</tr>
<tr>
<td>En₁</td>
<td>4.2 m</td>
<td>e^(0.307(En₁-4.2))</td>
<td>0.106</td>
</tr>
<tr>
<td>En₂</td>
<td>4.2 m</td>
<td>e^(0.676(En₂-4.2))</td>
<td>0.030</td>
</tr>
<tr>
<td>Ex₁</td>
<td>4.2 m</td>
<td>e^(-0.068(Ex₁-4.2))</td>
<td>0.005</td>
</tr>
<tr>
<td>Ex₂</td>
<td>4.2 m</td>
<td>e^(-0.005(Ex₂-4.2))</td>
<td>0.065</td>
</tr>
<tr>
<td>Rn₁</td>
<td>60 m</td>
<td>e^(0.012(Rn₁-60))</td>
<td>0.010</td>
</tr>
<tr>
<td>Rn₂</td>
<td>60 m</td>
<td>e^(0.015(Rn₂-60))</td>
<td>0.008</td>
</tr>
<tr>
<td>Rx₁</td>
<td>60 m</td>
<td>e^(0.020(Rx₁-60))</td>
<td>0.010</td>
</tr>
<tr>
<td>Rx₂</td>
<td>60 m</td>
<td>e^(0.024(Rx₂-60))</td>
<td>0.014</td>
</tr>
<tr>
<td>Q_major</td>
<td>7000 veh/d</td>
<td>(Q_major/7000)⁻⁰.⁴³₈</td>
<td>0.034</td>
</tr>
<tr>
<td>Q_minor</td>
<td>4000 veh/d</td>
<td>(Q_minor/4000)^⁻⁰.⁴²₃</td>
<td>0.033</td>
</tr>
<tr>
<td>F</td>
<td>0 (no object)</td>
<td>e⁻⁰.⁵²(F-0)</td>
<td>0.272</td>
</tr>
<tr>
<td>CW</td>
<td>7 m</td>
<td>e⁻⁰.⁶³(CW-.⁷)</td>
<td>0.197</td>
</tr>
<tr>
<td>WL</td>
<td>15 m</td>
<td>e⁻⁰.⁰¹(WL-.⁵)</td>
<td>0.069</td>
</tr>
<tr>
<td>WW</td>
<td>7 m</td>
<td>e⁻⁰.³⁵(WW-.⁷)</td>
<td>0.143</td>
</tr>
<tr>
<td>CD</td>
<td>15 m</td>
<td>e⁻⁰.⁰²(CD-.⁵)</td>
<td>0.015</td>
</tr>
<tr>
<td>V</td>
<td>60 kph</td>
<td>e⁻⁰.⁰²(V-.⁶)</td>
<td>0.040</td>
</tr>
</tbody>
</table>

Note:

a Adopted from previous studies and from mean values of individual covariates.
b CMFunction result is more reliable when the SE equals or less than 0.10.
5.2.7. Exit radius

A smaller exit radius results in increased safety risk for both major and minor approaches at roundabouts. As mentioned early, the exit from the roundabout must be as comfortable and easy for a driver as possible. Entries of roundabouts are designed to decrease vehicle speeds, whilst exiting should be able to allow the vehicles to increase speed out of the circulating roadway. Thus, the exit radius should generally be greater than entry radius for safety and operational issues at roundabouts. The study found that a higher exit radius is associated with less crash risk as shown in Fig. 8. For instance, at the major approach, the percent of crash reduction after increasing exit radius by 10 m was 18%. This result agrees with the previous study done by Anjana and Anjaneyulu (2014).

5.2.8. Traffic volume (AADT)

Highway Safety Manual (AASHTO, 2010) uses traffic volume as a significant predictor in studying road safety. In this study, the base condition for a major approach was adopted at 7000 vehicles per day and for a minor approach at 4000 vehicles per day. These values were adopted based on the mean values of traffic volumes in the dataset. Fig. 9 shows that the crash risk increases with increasing traffic volumes. The results also show that the volume on the minor approach has a larger impact on safety than major approach at high traffic volumes. This may be due to the difference in geometric characteristics (i.e., lane width, number of lane, etc.) between minor and major approaches.

5.2.9. Fixed object

Fixed objects like trees may be placed within a central island area, provided it is large enough to ensure that clear zone requirements are met and the sightlines for drivers are not obstructed. In most cases, these fixed objects can be placed on the central island to reduce the entry speed of the vehicles and enhances the driver’s attention approaching the roundabout. The study found that roundabouts with fixed objects have about 5% less crashes than roundabouts without fixed objects.

5.2.10. Circulatory roadway width

The circulating roadway is the portion of roundabout between the inscribed circle and the central island used by vehicular traffic, as shown in Fig. 1. The circulating roadway width is recommended to be about 1.0–1.2 times the entry width to a roundabout (Montella et al., 2012). However, a wider circulatory roadway width should be avoided, especially at a single-lane roundabout, where drivers may think that two vehicles are allowed to drive side by side within the roundabout. Fig. 10 shows that the wider circulatory roadway width is associated with greater crash risk at roundabouts.

5.2.11. Weaving length

A weaving section is a dynamic portion in the roundabout, where vehicles carry out one or more lane changes to complete merging and diverging operations. The two significant parameters in the analysis of weaving sections, based on road safety and capacity, are weaving length and weaving width (Golob et al., 2004). This study has also investigated the impact of weaving length on road safety. The results revealed that an increase in weaving length results in a decrease in crash risk. This result was reasonable because a long distance of weaving length decreases the probability of crashes as a result of sufficient space and time to complete merging or diverging operations. Fig. 11 illustrates the relationship between weaving length and road safety.
5.2.12. Weaving width
As mentioned previously, one source of vehicles conflicts at the roundabout is the weaving section, where the merge and diverge occur between vehicles. The impact of weaving width on road safety was investigated in this study. The results showed that a wider weaving width results in an increase in crash risk, as shown in Fig. 12. The wider weaving width, as in the circulatory roadway width, can lead to attempts by vehicles to pass each other, resulting in high speed driving and therefore increased risk.

5.2.13. Central island diameter
The geometry of a central island should be designed to reduce high entry speeds to the roundabout. Likewise, the shape of central islands should preferably be circular because changes in curvature of the circulating carriageway lead to a variance in speeds and increases the complexity for drivers. Wider central island diameters are preferable, as it reduces of entry vehicle speeds. This is due to a reduction of the angle formed between the circulating and entering vehicle paths (AUSTROADS, 2015). The base condition in this study was an island diameter of 15 m and this value was adopted based on the mean values of the central island diameters in the dataset. Fig. 13 shows that the wider central island diameter roundabout was associated with lower crash risk. A similar result has been concluded by Kim and Choi (2013).

5.2.14. Speed limit
Speed limit is one of the most important parameters that significantly affect road safety at roundabouts (AUSTROADS, 2015). Ideally, lower operating speeds at roundabouts are associated with increased driver reaction time and thus reduce the number and severity of road crashes that do occur. In this study, the speed limits on major approaches were analysed and estimated the CMF values as shown in Fig. 14. The results indicated that the crash risk increases as posted speed limit increases. For instance, a 10 km/h increase in speed limit leads to a 26% increase in the expected number of crashes.

6. Summary and conclusions
The main objective in the current study is to evaluate the safety performance of different roundabout elements using a cross-sectional method. In this study, safety performance models are developed to predict the total number of crashes (i.e., fatal and injury crashes) at roundabouts in regional areas based on measurable explanatory variables. The negative binomial (NB) distribution with a log-function has been used to estimate the model parameters. The crash data used in this study observed over a six-year period from 49 roundabouts in the Toowoomba City, Australia. Five models were identified as recommended models based on statistical significance, GOF measures and CURE analysis. It is worth mentioning that the cross-sectional method used in this study does not consider crash risks that would be attributed to external causal factors. However, this method is a viable alternative method that can be adopted in cases where observational before-and-after studies are not practical due to data restrictions, e.g., dates of treatment installation are unknown or installation of more than one treatment at the same time to an entity. The results indicated that several roadway traffic and geometric elements affect the safety at roundabouts. It was found that increasing...
the number of legs, number of exit lanes, exit width, exit radius, weaving length, central island diameters, and the presence of a fixed object on a central island are associated with increased total crash frequency. On the other hand, increasing other variables such as number of entry lanes, entry width, entry radius, traffic volume, circulatory roadway width, weaving width, and speed limit are associated with reduced total crash frequency.

Some limitations to the current study must be taken into consideration. It is clear that the study models were estimated based on a sample of roundabouts in one particular city and can therefore not claim to be adequate for all roundabout in other situations. Hence, the values of CMF in this study are only applicable to those roundabouts with similar geometric and traffic conditions, i.e., within the range of the datasets used.

Further work may be needed to extend the present study. It is important to estimate the safety effects (i.e., CMFs) based on various severity levels and crash types. From this it may be possible to identify the impact of various treatment types on crash type and severity. In addition, studying additional roundabout geometric and operational features would extend the scope of the future studies to improve the overall safety at roundabouts.

Conflict of interest

The authors do not have any conflict of interest with other entities or researchers.

Acknowledgments

The first author gratefully acknowledges the financial support from a Tafila Technical University scholarship for carrying out his PhD study at the University of Southern Queensland. The authors acknowledge the data support by the Department of Transport and Main Roads and Toowoomba Regional Council.

REFERENCES


Mohammad Nour Al-Marafi obtained his master's degree in transportation engineering from the University of Jordan, Amman, Jordan in 2014 and his bachelor's degree in civil engineering from Tafila Technical University, Jordan in 2011. Mohammad is currently studying a doctorate degree in transport engineering at the Faculty of Health, Engineering and Sciences, University of Southern Queensland, Australia. Prior to migrating to Australia in 2016, he worked as a lecturer at Tafila Technical University, Jordan for two years.

Dr. Kathirgamalingam Somasundarawaran obtained his PhD degree from the Yokohama National University, Japan, master degree from the Asian Institute of Technology, Bangkok, Thailand, and bachelor degree from the University of Peradeniya, Sri Lanka. He is a chartered professional engineer in Australia. Soma is currently working as a senior lecturer in transport engineering at the School of Civil Engineering and Surveying, University of Southern Queensland, Australia, and teaching a range of road and transportation engineering courses. His research interests include road safety, traffic studies and pavement rehabilitation.

Prof. Ron Ayers obtained his master of engineering, bachelor of engineering and bachelor of science degrees from the University of New South Wales, Australia. Ron retired from full time involvement in the School of Civil Engineering and Surveying, University of Southern Queensland, Australia, in 2013, after 37 years of teaching and research in transportation, construction and management, and serving various periods as Head of Civil Engineering, and dean of the Faculty of Health, Engineering and Sciences. His current research interests include pavement management and materials, road safety and transport planning.