Estimation of Crash Risk for Vehicles Behind Buses in Mixed Traffic

PAPER NUMBER 15-0634

SECOND SUBMISSION

Kelvin Chun Keong Goh, PhD Student, Public Transport Research Group, Institute of Transport Studies, Department of Civil Engineering, Building 60, Monash University, Clayton, Victoria 3800, AUSTRALIA
Phone: +61 3 9905 1848, Fax: +61 3 9905 4944, Email: kelvin.goh@monash.edu

Graham Currie, Professor, Chair in Public Transport, Public Transport Research Group, Institute of Transport Studies, Department of Civil Engineering, Building 60, Monash University, Clayton, Victoria 3800, AUSTRALIA
Phone: +61 3 9905 5574, Fax: +61 3 9905 4944, Email: graham.currie@monash.edu

Majid Sarvi, Associate Professor, Institute of Transport Studies, Department of Civil Engineering, Building 60, Monash University, Clayton, Victoria 3800, AUSTRALIA
Phone: +61 3 9905 9696, Fax: +61 3 9905 4944, Email: majid.sarvi@monash.edu

David Logan, Dr, Senior Research Fellow, Monash University Accident Research Centre, Building 70, Monash University, Clayton, Victoria 3800, AUSTRALIA
Phone: +61 3 9905 4376, Email: david.logan@monash.edu

*Corresponding author

Submitted for Presentation Only

Committee number ANB70 Truck and Bus Safety

Words: 5,284 + ((6 Figures + 3 Tables)×250=2,250) = 7,534 (limit = 7,500)

Initial Submission Version: Sunday, July 13, 2014
Post Review Submission Version: Tuesday, October 21, 2014
ABSTRACT

This paper summarises findings from a three-stage modelling approach used to estimate the crash risk of a vehicle that is behind a slowing or stationary bus in a mixed traffic configuration.

This approach involves the development of regression and neural network models to represent drivers’ lane changing behaviour, followed by an establishment of crash risk probability and estimation of crash risk through a Monte Carlo simulation approach using time-to-collision and accident data. Through a case study of a road corridor, results showed that speed differences between the subject and lead vehicles in the current and adjacent lanes, distances between the subject and lead or lag vehicle in the adjacent lane as well as whether the bus is a lead vehicle were significant factors that influence lane change. The Monte Carlo simulation results revealed that average crash risk of vehicles that performed the lane change (LC) and those remained in the current lane (NLC) differ (0.0185% vs. 0.0062%). Overall crash risk was found to be 0.0154% (with a standard error of 0.0063%).

The risk estimates serve as important findings for bus safety and bus priority research as well as policy-makers in road and transit agencies, as they provide new knowledge of the quantum of risk involved in designing bus stops in mixed traffic as well as benefits delivered by bus priority schemes that segregate buses from mainstream traffic.

Keywords: Time-to-collision, Regression, Neural network, Monte Carlo, Bus priority

Abstract = 220 words (limit = 250 words)
INTRODUCTION

In a mixed traffic configuration, buses that decelerate or stop at bus stops for boarding and alighting passengers risk being hit or ‘rear-ended’ by vehicles approaching from the rear. The risk of a rear-end collision also exist for any of the vehicles approaching from behind, as drivers may be caught unaware of the slowing or stationary vehicle ahead. Evidence from the existing literature show that rear-end collision ranks as one of the highest risks for buses and that bus stops is a common location where collisions occur [1, 2].

Numerous studies have been conducted to investigate rear-end crash risks. They have however been mainly confined to applications in work zones, freeway (or highway) and intersection locations. As such, our understanding on rear-end crash risks involving buses at bus stop locations remains unclear.

This paper aims to estimate the crash risk of a vehicle that is behind a slowing or stationary bus at a bus stop in a mixed traffic configuration. A key motivation is to understand the safety benefits of having bus priority lanes that segregate buses from the mainstream traffic\(^a\). To understand these benefits we must better understand the risks of having buses operating in mixed traffic.

This paper is structured as follows; the next section reviews previous research on bus safety and rear-end accident risks. Details of the methodology are then provided including a description of the modelling approach to estimate crash risks for vehicles behind a bus. This is followed by an application of the methodology on a selected site in Melbourne, Australia. Model results and their implications are then presented before the paper concludes with a summary and recommendations on future research.

RESEARCH BACKGROUND

Previous studies on bus safety have showed that certain accident types are common for buses. One of the earliest studies found that the two most common collision types are side-swap and rear end [3]. Findings revealed that a high percentage of automobile occupant injuries occurred when automobiles ‘rear-ended’ the bus, which led the author to suggest that stationary buses (either stopped for a queue of vehicles or to process passengers) pose the greatest risk to automobile occupants. In examining bus and coach occupant injuries, Björnstig et al. [4] found that approximately half of the occupant injuries were due to buses or coaches being ‘rear-ended’ by other vehicles. Zegeer et al. [5] found likewise that rear-end accidents in which one vehicle stopped and sideswipe accidents to be the most common accident type in commercial bus crashes across five states in the U.S. A similar finding was obtained by Rey et al. [6] when they investigated transit bus crashes in Florida, U.S. Following an analysis of school bus crashes and injuries, Yang et al. [7] also found cases of vehicles ‘rear-ending’ buses as well as vehicles hitting buses when the latter were turning to be most common. With regard to accident location, published evidence suggests that bus stop locations and intersections are the most accident prone areas [1, 3].

There have been numerous studies that investigated rear-end crash risks, with a number focussing on work zone areas. Meng et al. [8] developed a probabilistic quantitative risk assessment model to estimate crash frequency based on the work zone characteristics. Similarly, Harb et al. [9] developed conditional logistic regression and multiple logistic regression models to identify key work zone freeway crash characteristics. An analysis of rear-end accidents in work zones was also done by Qi et al. [10], from which truncated count models based on historical crash data in New York were developed to identify work zone characteristics that are associated with crash frequency. Using crash data from work zones in California, U.S., Khattak et al. [11] developed negative binomial models, which revealed that crash frequencies increase with increasing work zone duration, length, and average daily traffic. Whilst the above studies relied on historical crash data, a recent study leveraged on video

\(^a\) This research is part of the wider Australian Research Council Industry Linkage Program project LP100100159, ‘Optimising the Design and Implementation of Public Transport Priority Initiatives’ done by Institute of Transport Studies, Monash University in association with the Transport Research Group, University of Southampton, UK. The Principal Chief Investigator is Professor Graham Currie the Chief Investigator is Associate Professor Major Sarvi and the Partner Investigator is Professor Nick Hounsell. Mr Goh is one of three APAI PhD students on the project. The Industry Sponsors include VicRoads and Public Transport Victoria.
technology to analyse and evaluate rear-end crash risk at a work zone area in Singapore [12]. Its approach is based on the traffic conflict technique, in which a surrogate safety measure (Deceleration Rate to Avoid a Crash) is used to measure rear-end crash risk, following which crash risk models were developed to examine the relationship between rear-end crash risk and its contributing factors.

A number of studies had also investigated rear-end crash risk at freeway (or highway) and intersection locations. Wang and Abdel-Aty [13] utilized generalized estimating equations with the negative binomial link function to model rear-end crash frequencies at signalized intersections. Results showed that heavier traffic, additional right and left-turn lanes and high speed limits on the major roadway, a large number of phases per cycle and high population areas are correlated with high rear-end crash frequencies. Hourdos et al. [14] focussed on high-crash locations at a freeway by analysing video data collected by detection and surveillance infrastructure. Along with visual observations, an identification of the most relevant real-time traffic was done and subsequently incorporated into a model to estimate crash likelihood. Pande and Abdel-Aty [15] developed probabilistic neural network models to identify traffic conditions that are associated with higher risks of rear-end crashes on a highway in the U.S. A similar approach of using data collected from inductive loop detectors was adopted by Oh et al. [16] to develop a methodology to identify rear-end collision potentials on freeways. Key in this methodology is the formulation of a rear-end collision risk index based on the safety distance in car-following situations to reflect freeway rear-end traffic collisions. Using sensor and communication technology, Oh et al. [17] followed up with a study that utilized time-to-collision (TTC) and stopping distance index to develop a methodology to detect hazardous traffic events and evaluate the real-time safety performance of a freeway. In a subsequent study, Oh and Kim [18] developed another methodology to estimate rear-end crash probabilities on freeways based on real-time vehicle trajectory data. Through the development of a binary logistic regression model on lane-changing and derivation of crash probability based on TTC values, a crash risk index was developed in the analysis to establish rear-end crash potential for each subject vehicle along the freeway.

In summary, there have been numerous studies conducted to investigate rear-end crash risks. However, their focus had been on work zones, freeway and intersection locations. At present there is only limited knowledge regarding rear-end crash risks involving buses at bus stop locations. This is surprising as the evidence from existing literature show that rear-end collision ranks as one of the highest risks for buses, and that collisions occurring at bus-stops are just as common.

RESEARCH AIM

This study aims to estimate the crash risk potential of a vehicle that is behind a slowing or stationary bus at a bus stop in a mixed traffic configuration. A key motivation in this study is to understand the quantum of such risk involved as it provides an appreciation of the safety benefit delivered by bus priority measures that segregate buses from mainstream traffic.

METHODOLOGY

Drivers of vehicles travelling along a road with a mixed traffic configuration are faced with two options when they find themselves behind a bus that is slowing down to call or stationary at a bus stop ahead. To avoid a collision, they can choose to either: (1) slow down and wait for the bus to move off after processing passengers or (2) switch lanes to overtake the bus. Figure 1 shows vehicle n in such a situation, in which collision risks exist if the driver of vehicle n fails to decelerate in time in the current lane. Should the driver decide to switch lane to overtake the bus, there exists the risk of collision with the lead or lag vehicle in the adjacent lane. In both situations, a key factor in whether a crash would occur is the time-to-collision (TTC) value between the subject and lead or lag vehicles, i.e. lower TTC values are associated with higher likelihood of collision.

Given the possibilities above, a three-stage approach is adopted in this study to estimate the vehicle’s crash risk potential:

1. Calculation of lane change probability;
2. Calculation of crash probability given a TTC value; and
3. Estimation of crash risk potential based on values obtained in (1) and (2)
Lane Change Probability

With lane changing essentially involving decision-making between two choices, discrete choice modelling approach can be adopted to establish lane change probability. In this field, two widely adopted analytical approaches are binary logit regression (BLR) and artificial neural network (ANN) modelling. For this study, both approaches as well as a third incorporating both BLR and ANN (hybrid approach) were used to model lane change probability. A key step in the methodology is the selection of the best performing lane-change model for the subsequent estimation of crash risk.

In the BLR approach, the lane change probability can be formulated as:

\[
p(\text{LC}_n | X_n) = \frac{\exp(X_n^\mathbf{\beta})}{1 + \exp(X_n^\mathbf{\beta})} \tag{1}
\]

\[
p(\text{NLC}_n | X_n) = 1 - p(\text{LC}_n | X_n) \tag{2}
\]

where \(p(\text{LC}_n | X_n)\) and \(p(\text{NLC}_n | X_n)\) are the probabilities that the subject vehicle \(n\) will and will not switch lane respectively under traffic conditions \(X_n\). In both equations, \(X_n\) represents a vector of explanatory variables affecting the decision of subject vehicle \(n\). As part of the BLR model development, the “linktest” function in STATA (2005) was employed and inspection of Variance Inflation Factors (VIF) values done to ensure the final model was free from specification errors and heteroscedasticity respectively.

The ANN approach leverages on the ability of artificial neural networks to model complex relationship and solve problems that require forecasting, pattern recognition and classification. Many types of artificial neural networks are available in the literature. In this study, a three-layer feed-forward neural network based on the back-propagation approach using the Lavenberg-Marquardt [19] algorithm is adopted. The ANN model structure is shown in Figure 2, where \(X_n\) are the input neurons that represent the traffic conditions, \(Z_k\) the hidden neurons and \(Y\), the output neuron in the model that represents the lane change probability. A key disadvantage of using a neural network approach is model over-fitting, which results when the network is strong in fitting the random error (noise) in the data but not the underlying relationship. To address this issue and still ensure good generalization of the model, the “early stopping” technique was applied when training the network. Likewise, the dataset was also randomly separated into two parts (in a 3:1 proportion) for the purpose of training and testing the model.
A typical approach in ANN modelling is the application of an algorithm for selection of the input variable(s). In the hybrid BLR-ANN approach, a similar principle was used in which variables found to be significant as well as the predicted probability from the BLR model will be used as inputs to the BLR-ANN model. Figure 3 presents the key steps involved in the BLR-ANN approach.

**FIGURE 3** Key steps in the hybrid BLR-ANN approach.
Probability of a Crash and Crash Risk Estimation

The crash risk modelling was done using TTC values that were extracted from the video data. According to Amundsen and Hyden [20], the TTC value is the time that remains from an instant \( t \) before a collision between two vehicles takes place (assuming both vehicles’ direction and speed remain unchanged). If the subject vehicle \( n \) decides to remain in the lane in which the bus ahead has slowed down, then the TTC value can be derived as follows:

\[
TTC_{n-\text{Lead}}(t) = \frac{D_{1-\text{Lead}}(t)}{V_n(t) - V_{1-\text{Lead}}(t)} \quad \forall V_n(t) > V_{1-\text{Lead}}(t) \tag{3}
\]

where \( D_{1-\text{Lead}} \) is the gap between the subject and lead vehicle, while \( V_n \) and \( V_{1-\text{Lead}} \) are the speeds of the subject and lead vehicle respectively at time \( t \). If the driver of the subject vehicle decides to switch lane, the corresponding TTC values between the subject and lead or lag vehicles can be determined as:

\[
TTC_{n-\text{Lead}}(t) = \frac{D_{2-\text{Lead}}(t)}{V_n(t) - V_{2-\text{Lead}}(t)} \quad \forall V_n(t) > V_{2-\text{Lead}}(t) \tag{4}
\]
\[
TTC_{n-\text{Lag}}(t) = \frac{D_{2-\text{Lag}}(t)}{V_{2-\text{Lag}}(t) - V_n(t)} \quad \forall V_{2-\text{Lag}}(t) > V_n(t) \tag{5}
\]

In the above equations, \( D_{2-\text{Lead}} \) and \( D_{2-\text{Lag}} \) are the gaps between the subject and lead, and subject and lag vehicles respectively, while \( V_{2-\text{Lead}} \) represents the speed of the lead vehicle and \( V_{2-\text{Lag}} \) speed of the lag vehicle in the adjacent lane.

It is generally accepted that TTC values are closely linked to crash potential [18, 21]. Following previous research [18, 21], the probability of a crash based on a TTC value \( p \) can be assumed to take on the following exponential decay relationship in which:

\[
p = e^{-\frac{\text{TTC}}{\lambda}} \tag{6}
\]

where \( \lambda \) is a parameter that reflects crash propensity of a given road segment, i.e. its value differs across roads with different characteristics. \( \lambda \) can be computed based on the study site’s historical crash records and TTC profile obtained from the video data. The probability of a crash occurring in an hour can thus be computed as follows:

\[
\text{Crash risk / hour} = \sum_{i=1}^{N} e^{-\frac{\text{TTC}_i}{\lambda}} \tag{7}
\]

where \( \text{TTC}_i \) is the time-to-collision value recorded for vehicle \( i \) and \( N \) is the number of vehicles with TTC values recorded in an hour on a selected day. Following this, the probability of a crash between subject vehicle \( n \) and lead vehicle based on a given TTC can be estimated by:

\[
p(Crash_{n-\text{Lead}})(t) = p(NLC_n | X_n)(t).p(C_{n-\text{Lead}} | \text{TTC})(t) \tag{8}
\]

Using the same approach, the probabilities of a crash between the subject vehicle \( n \) and the lead or lag vehicles in the adjacent lane can be computed. The establishment of the probabilities of lane changing and crash based on a given TTC sets the stage for the estimation of the rear-end crash risk. Figure 4 shows an overview of the variables involved in estimating the crash risk.
Given that the TTC information come from a sample of traffic data, it was important to account for uncertainty in the analysis. For this reason, the bootstrapping technique based a resampling size of 500 was employed to obtain the mean and variance of $\lambda$. The uncertainty in the final crash risk estimation was also accounted for through the use of a stochastic analysis tool available in the Crystal Ball software. This was used to analyse and generate the best-fit probability or frequency distribution for each variable with inherent uncertainty. The distribution information was then fed into a Monte Carlo simulation model to compute the mean and variance of the final crash risk estimate from 1,000 trials. Through this approach, a quantifiable degree of uncertainty is incorporated in the final crash risk estimate to reflect the likelihood that drivers at times in reality base their driving decisions on imprecise perceptions of the surrounding traffic.

APPLICATION AND DATA COLLECTION

For this study, a three-lane divided arterial road with a speed limit of 70kph and annual average daily traffic volume of 17,000 in Melbourne (Blackburn Road) was selected for the rear-end crash risk estimation. The focus is on the area upstream of a bus stop located on the northbound carriageway of Blackburn Road (Figure 5), where vehicles movement in relation to buses can be tracked. The bus stop of interest serves three different bus services, with each operating at a service frequency that ranges between ten to sixty minutes.

Using video recording equipment mounted on top of a thirteen-story building, 24-hour video recordings of the traffic in the vicinity of the bus stop were undertaken over 2 weeks in December 2012. From the video recordings, vehicle trajectory data on weekdays were extracted at intervals of 0.2s using the Tracker software. This software facilitated axis definition in terms of orientation and origin setting for measurement purposes, thus allowing calibration to be done prior to extraction of vehicle trajectory information. The 3-year accident record of the site was also extracted from CrashStats for the purpose of computing the $\lambda$ value and its statistical properties. The final dataset consisted of a total of 338 sets of trajectory information sets, with each set comprising individual position (in x-y coordinates) for the subject, lead and lag vehicles in the current and adjacent lanes. This was used to extract information for the following variables, which were chosen based on existing literature [26-28]:

1) $V_n$ - Speed of subject vehicle $n$;
2) $dV_{1,\text{Lead}}$ - Speed difference between subject and lead vehicle in the current lane;
3) $dV_{1,\text{Lag}}$ - Speed difference between subject and lag vehicle in the current lane;

FIGURE 4 A Monte Carlo simulation approach to estimate crash risk.
4) \( dV_{2,lead} \) - Speed difference between subject and lead vehicle in the adjacent lane;
5) \( dV_{2,lag} \) - Speed difference between subject and lag vehicle in the adjacent lane;
6) \( D_{1,lead} \) - Distance between subject and lead vehicle in the current lane;
7) \( D_{1,lag} \) - Distance between subject and lag vehicle in the current lane;
8) \( D_{2,lead} \) - Distance between subject and lead vehicle in the adjacent lane;
9) \( D_{2,lag} \) - Distance between subject and lag vehicle in the adjacent lane; and
10) \( BA \) - Dummy variable to indicate whether bus is directly ahead in the current lane.

FIGURE 5 Road segment with video recording equipment (inset) for data collection.

RESULTS AND DISCUSSION

Lane Change Probability
Table 1 presents results of the parameter estimates for the BLR model while Table 2 captures the performance of the BLR, ANN and BLR-ANN models based on sensitivity, specificity, correct classification rate (CCR) and area under ROC curve (AUC).

TABLE 1 Results of BLR model on lane change probability (based on training dataset)

<table>
<thead>
<tr>
<th>Variable</th>
<th>( \beta )</th>
<th>S.E.</th>
<th>Wald Statistic</th>
<th>Odds Ratio</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>( dV_{1,lead} )</td>
<td>0.074</td>
<td>0.016</td>
<td>20.119</td>
<td>1.077</td>
<td>0.000</td>
</tr>
<tr>
<td>( dV_{2,lead} )</td>
<td>-0.059</td>
<td>0.027</td>
<td>4.861</td>
<td>0.943</td>
<td>0.027</td>
</tr>
<tr>
<td>( D_{2,lead} )</td>
<td>0.042</td>
<td>0.019</td>
<td>4.861</td>
<td>1.043</td>
<td>0.027</td>
</tr>
<tr>
<td>( D_{2,lag} )</td>
<td>0.044</td>
<td>0.009</td>
<td>21.331</td>
<td>1.045</td>
<td>0.000</td>
</tr>
<tr>
<td>( BA )</td>
<td>1.146</td>
<td>0.383</td>
<td>8.970</td>
<td>3.147</td>
<td>0.003</td>
</tr>
<tr>
<td>Intercept</td>
<td>-2.511</td>
<td>0.536</td>
<td>21.916</td>
<td>0.081</td>
<td>0.000</td>
</tr>
</tbody>
</table>

\(-2LL\) | 187.87 | 387.78 | (intercept only)
\(AIC\) | 199.87 | 389.78 |
\(BIC\) | 221.89 | 393.45 |
\(LR\) chi-square | <0.001 |
\(Wald\) chi-square | <0.001 |
From Table 1, the BLR model indicates that speed differences between the subject and lead vehicles in the current \((dV_{1-lead})\) and adjacent lanes \((dV_{2-lead})\) are significant factors that influence lane change. Just as significant are the distances between the subject and lead vehicle \((D_{2-lead})\) as well as subject and lag vehicle \((D_{2-lag})\) in the adjacent lane. The coefficient signs for these variables were as expected and similar to previous findings [27], i.e. lane change was more likely when the speed of lead vehicle in the current lane was smaller or lead vehicle in adjacent lane was greater. Results from a previous study also found that a larger gap between the subject and lead or lag vehicles in the adjacent lanes were associated with lane changing [28]. An interesting result found in this research is that drivers were more likely to switch lanes if the bus was directly ahead \((BA)\) as compared to being a few vehicles ahead. While such a finding was as expected, it suggests that drivers in Melbourne have good lane discipline as they are unlikely to switch lanes until the ones ahead of them (and behind a slowing or stationary bus) had already done so. This mirrors what has been observed from the video recordings in that lane changing was done in an orderly and sequential manner on most occasions.

**TABLE 2 Performance of BLR, ANN and BLR-ANN models**

<table>
<thead>
<tr>
<th>Measure</th>
<th>BLR Training Dataset (290)</th>
<th>BLR Test Dataset (48)</th>
<th>ANN Training Dataset (290)</th>
<th>ANN Test Dataset (48)</th>
<th>BLR-ANN Training Dataset (290)</th>
<th>BLR-ANN Test Dataset (48)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sensitivity</td>
<td>0.8421</td>
<td>0.9204</td>
<td>0.9204</td>
<td>0.7143</td>
<td>0.8571</td>
<td>0.8571</td>
</tr>
<tr>
<td>Specificity</td>
<td>0.9034</td>
<td>0.8644</td>
<td>0.8588</td>
<td>0.8824</td>
<td>0.8529</td>
<td>0.8824</td>
</tr>
<tr>
<td>CCR (%)</td>
<td>0.8790</td>
<td>0.8860</td>
<td>0.8830</td>
<td>0.8330</td>
<td>0.8540</td>
<td>0.8750</td>
</tr>
<tr>
<td>AUC</td>
<td>0.9290</td>
<td>0.9442</td>
<td>0.9458</td>
<td>0.9430</td>
<td>0.9097</td>
<td>0.9945</td>
</tr>
<tr>
<td>MSE</td>
<td>0.0989</td>
<td>0.0928</td>
<td>0.0851</td>
<td>0.0973</td>
<td>0.1153</td>
<td>0.0843</td>
</tr>
</tbody>
</table>

Note: Shaded figures indicate the best performing model for each measure and dataset

Results from Table 2 show that the ANN and the BLR-ANN approach resulted in better performing models, as they were able to correctly classify an additional 0.7% to 4.2% of lane changing decisions as compared to the BLR model. Although this represents only a marginal improvement in model performance, the results point to the potential of adopting the neural network approach as an alternative in modelling binary outcomes and its usefulness when variables are nonlinear or have a non-specific functional form.

For this study, the BLR-ANN model was considered to have the best performance and its lane change probability prediction was thus selected as inputs for the subsequent computation of crash risk estimation.

**Probability of a Crash and Crash Risk Estimation**

Table 3 captures results of the parameter estimates that were obtained when the Crystal Ball software was used to find the best-fitted distributions of the variables involved in the crash risk estimation, while Figure 6 presents results of estimated crash risk (expected value and range based on one standard error) that emerge from the Monte Carlo simulation.

As reflected by statistics in Table 1, it was observed that the average TTC value for vehicles that do not change lanes \((TTC_{n}^{1-lead})\) is generally lower than those that do \((TTC_{n}^{2-lead} \text{ or } TTC_{n}^{2-lag})\). However, \(TTC_{n}^{2-lead}\) values in the lower range were found to be smaller in comparison to \(TTC_{n}^{1-lead}\), as reflected by the location (or shift) parameter values. This difference showed up in the final crash risk estimates, where the average crash risk for vehicles that changed lanes (LC) was found to be higher than those that do not (NLC).

Simulation results show the average crash risk of vehicles in the NLC group and LC groups are 0.0062% (with a standard error of 0.0008%) and 0.0185% (with a standard error of 0.0065%) respectively. When both groups of vehicle are considered collectively, the crash risk was found to be 0.0154% (with a standard error of 0.0063%). Based on the latter crash risk value and assumption that an average of thirty (30) TTC conflicts occur a day, it works out that there is an approximate 80% chance of one or more accidents taking place on an annual basis as a result of buses slowing down or being stationary at bus stops.
The risk estimates represent important findings for bus safety and in particular bus priority research, as the risk of rear-end crashes due to buses slowing down or being stationary at bus stops is totally eliminated when bus priority measures that segregate buses from mainstream traffic are implemented. This study represents the first attempt to quantify such risks and highlights the importance of considering safety implications in bus priority strategies. In the authors’ earlier studies, it was found that the implementation of bus priority in Metropolitan Melbourne had led to an approximate 14% and 53% reduction in reported injury accidents [29] and bus-involved accidents [30] respectively. Given that these reductions were recorded at the aggregate-level, it was not possible then to identify and quantify any specific safety effect at play. Findings from this research are therefore significant because a component of the safety benefits delivered by bus priority is now known. In this regard, the results present an opportunity for policy-makers to account for safety benefits as part of the overall cost-benefit analyses typically done prior to bus priority implementation.

### TABLE 3 Best-fit distributions for variables (used as inputs in Monte Carlo simulation)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Best-Fit Distribution</th>
<th>Parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td>λ</td>
<td>Beta</td>
<td>Min.= 0.10; Max.= 0.44; $\alpha = 41.48$; $\beta = 53.44$</td>
</tr>
<tr>
<td>NLC</td>
<td>Beta</td>
<td>Min.= 0.0; Max.= 1.0; $\alpha = 0.3$; $\beta = 0.42$</td>
</tr>
<tr>
<td>$TTC_{1-lead}^n$</td>
<td>Log-normal</td>
<td>Location = 1.52; Mean = 11.33; S.D. = 21.11</td>
</tr>
<tr>
<td>$TTC_{2-lead}^n$</td>
<td>Log-normal</td>
<td>Location = 0.51; Mean = 26.15; S.D. = 46.44</td>
</tr>
<tr>
<td>$TTC_{2-lag}^n$</td>
<td>Gamma</td>
<td>Location = 1.74; Scale = 30.31; Shape = 0.643</td>
</tr>
</tbody>
</table>

![Figure 6: Monte Carlo simulation results](image-url)

FIGURE 6 Monte Carlo simulation results (with dot and arrows representing the expected value and range based on 1 standard error) for crash risk.
CONCLUSIONS AND RECOMMENDATIONS

In this paper, a three-stage modelling approach was adopted to estimate the crash risk of a vehicle that is behind a slowing or stationary bus on a selected road with mixed traffic configuration in Metropolitan Melbourne. The key motivation in establishing the quantum of risk involved was to gain an appreciation of the safety benefit that is delivered by bus priority schemes that segregate buses from the mainstream traffic. The first stage involved the development of competing regression and neural network models to represent drivers’ lane changing behaviour behind buses, while the second and third stages entailed the establishment of crash risk probability and estimation or crash risk. For the latter, a Monte Carlo simulation approach was adopted using time-to-collision and accident data collected from a selected road corridor.

Results in the first stage revealed that speed differences between the subject and lead vehicles in the current ($dV_{1,lead}$) and adjacent lanes ($dV_{2,lead}$), distances between the subject and lead ($D_{2,lead}$) or lag vehicle ($D_{2,lag}$) in the adjacent lane as well as whether the bus is a lead vehicle ($BA$) are significant factors that influence lane changes. The latter finding was interesting and likely to be reflective of driving behaviour in Melbourne, as it indicated that drivers are unlikely to switch lanes until the ones ahead that are behind the bus had already done so. Results also showed that the hybrid regression-neural network approach yielded the best performing model. As such, predictions from this model were used as inputs in the second stage. Following a calculation of crash probability based on TTC values in stage 2, the Monte Carlo simulation results in stage 3 revealed that the average crash risk of vehicles that performed the lane change (LC) and those remaining in the current lane (NLC) are 0.0185% (with a standard error of 0.0065%) and 0.0062% (with a standard error of 0.0008%) respectively. The overall crash risk was found to be 0.0154% (with a standard error of 0.0063%).

The risk estimates serve as important findings for bus safety and bus priority research, as an estimate of the safety benefit delivered by bus priority that segregate buses from mainstream traffic is now available. In practice, this estimate could serve as an important consideration for policy-makers given this new knowledge of the quantum of risk involved in designing bus stops in a mixed traffic configuration as well as bus priority schemes where buses are segregated from mainstream traffic. An interesting observation from this study is that drivers behind a slowing bus are three times more likely to be involved in accidents when they change lane. Such information may be useful for the motoring public and could be used as part of an education program on safe driving for drivers.

Whilst findings from this study can act as a useful planning tool for road agencies, there remain limitations that policy makers should be aware of. Firstly, the lane change modelling and fitting of TTC distributions were done based on a sample of (2-week) data. As such, additional data can be collected to improve the model performance and reliability. Secondly, this study was based on a bus stop that is located along a three-lane divided road (with an average annual daily traffic volume of 17,000 and speed limit of 70kph) in Metropolitan Melbourne. Although such roads are considered typical in Melbourne, results could differ when roads with different characteristics are considered. In this regard, further research could be done to establish a more precise value for the $\lambda$ parameter and additional ones for different road types. Finally, a linear bus stop (mixed traffic configuration) was considered in this study. Hence, there exists much scope to investigate crash risks on roads with other bus stop configurations, e.g. indented bus bay. Further research could also examine the effects of improved visibility of stop locations or even new collision avoidance technology for buses braking.

ACKNOWLEDGEMENT

The authors would like to acknowledge the financial support provided by the Australian Research Council and Land Transport Authority of Singapore for this study. Appreciation also goes to our industry partners - Victorian Department of Transport and VicRoads. Any omissions or errors in the paper are the responsibilities of the authors.

LIST OF FIGURES

FIGURE 1: Vehicle n behind a slowing or stationary bus
FIGURE 2: Three-layer feed-forward artificial neural network with one output neuron
LIST OF TABLES

TABLE 1: Results of BLR model on lane change probability (based on training dataset)

TABLE 2: Performance of BLR, ANN and BLR-ANN models

TABLE 3: Best-fit distributions for variables (used as inputs in Monte Carlo simulation)

REFERENCES