Analytic Methods in Accident Research 32 (2021) 100180

Contents lists available at ScienceDirect

Analytic Methods in Accident Research

journal homepage: www.elsevier.com/locate/amar

A bivariate extreme value model for estimating crash frequency by severity using traffic conflicts

Ashutosh Arun [a], Md. Mazharul Haque [a,*], Ashish Bhaskar [a], Simon Washington [b], Tarek Sayed [c]

[a] School of Civil and Environmental Engineering, Queensland University of Technology, Brisbane, Australia
[b] School of Civil Engineering, Faculty of Engineering, University of Queensland, Brisbane, Australia
[c] Department of Civil Engineering, University of British Columbia, Vancouver, Canada

ARTICLE INFO

Article history:
Received 1 February 2021
Received in revised form 17 June 2021
Accepted 17 June 2021
Available online 23 June 2021

Keywords:
Traffic conflict techniques
Crash severity
Crash-conflict relationship
Peak over threshold
Automated conflict extraction

ABSTRACT

Estimating crash frequency by severity levels using traffic conflicts remains relatively unexplored in conflict-based traffic safety assessment, limiting its application scope and appeal compared to traditional methods. No studies to date have predicted the frequency of severe and non-severe crashes utilizing traffic conflicts. This study aims to address this critical gap and stimulate discussion and development in this critical area. The study estimates the frequency of severe crashes and non-severe crashes by jointly modeling the indicators of crash frequency, namely, Time to Collision (TTC) and Modified Time to Collision (MTTC), and crash severity, namely, predicted post-collision change in velocity (Delta-V or Δv), using bivariate Extreme Value. Severe crashes here are defined as crashes with a Maximum Abbreviated Injury Scale rating of greater than or equal to 3. Rear-end conflict data (TTC $\leq$ 3.0 s) were collected for two days (12 h each day) from two four-legged signalized intersections in Brisbane, Australia. Bivariate peak-over-threshold models for both TTC and MTTC indicators, combined with Delta-V, were estimated. Alternatively, another univariate approach was also attempted where the probability of crash occurrence was estimated using the univariate peak-over-threshold model with TTC (or MTTC) and then multiplied with the injury probability estimated from Delta-V to estimate the frequencies of severe and non-severe injury crashes. The study results demonstrate that the bivariate approach is more advantageous than the univariate approach due to a superior statistical fit to the data and more precise estimations of crash frequencies by severity levels. Both TTC and MTTC indicators, in combination with Delta-V, provide comparable results using the bivariate approach owing to the weak asymptotic dependence between the frequency and severity indicators. Comparing the combined dataset model of the two intersections with the intersection-based models shows that sharing information between similar traffic sites improves the accuracy and precision of prediction.

1. Introduction

Advanced statistical methods for predicting crash frequency using traffic conflict measures are up-and-coming. Songchitruksa and Tarko (2006) introduced the Extreme Value Theory in the domain of conflict-based safety analysis, where

* Corresponding author.
E-mail addresses: ashutosh.arun@hdr.qut.edu.au (A. Arun), m1.haque@qut.edu.au (M.M. Haque), ashish.bhaskar@qut.edu.au (A. Bhaskar), simon@a-magroup.io (S. Washington), tsayed@civil.ubc.ca (T. Sayed).

https://doi.org/10.1016/j.amar.2021.100180
2213-6657/

they predicted right-angle crashes from Post Encroachment Time (PET). Subsequently, many studies utilized the univariate extreme value theory framework to estimate different types of crashes from various conflict indicators utilizing the two extreme value theory approaches, namely Peak-Over-Threshold and Block-Maxima (Tarko, 2012, 2018; Zheng et al., 2014a). These studies exploit the property of conflict indicators like Time to Collision (TTC), whose extreme values are reliable predictors of the occurrence of crashes. Thus, as TTC approaches a value of 0, the crash occurrence probability approaches 1. However, the studies involving univariate analysis often suffered from the imprecision of results as they had wide crash prediction intervals (Zheng et al., 2014b). Therefore, recent advancements in this regard argue for multivariate analysis using extreme value theory (Wang et al., 2019; Zheng and Sayed, 2019d; Zheng et al., 2019b; Fu et al., 2020), wherein multiple conflict indicators are used to predict crash frequencies. Given that every conflict indicator cannot be computed in all the situations, including more than one indicator in analysis captures more information regarding the crash generating process and improves the estimates' accuracy and precision. Further, studies like Zheng et al. (2019a) argue for a hierarchical modeling structure that uses combined conflict data from multiple similar sites for improved crash estimation compared to site-specific models.

However, unlike the tremendous amount of research being done in the field of conflict-based crash frequency analysis, the estimation of crash frequency by severity using conflict measures has received little attention (Laureshyn et al., 2017). The lop-sided focus on crash frequency estimation has caused significant practical problems in applying conflict methods for safety assessment. It is well-known that lower severity crashes are more frequent on the roads than higher severity ones (Hydén, 1987). Thus, a reduction in higher frequency-low severity crashes may not result in the same safety benefits as reducing lower frequency-high severity crashes. Arun et al. (2021b) have further pointed out that implementing site improvements solely based on a conflict-based assessment that does not provide information on crash severity runs the risk of reducing the number of low severity crashes but, counterproductively, increasing the probability of higher severity crashes. Therefore, recent review studies (Arun et al., 2021b; Zheng et al., 2021) make a case for extending the safety pyramid of traffic events (Hydén, 1987) to include severity measurement within the scope of conflict-based safety assessment, as shown in Fig. 1.

A few studies have attempted to estimate the severity of a crash event based on severity-specific conflict measures (Sobhani et al., 2011, Bagdadi, 2013). Delta-V (Δv)or the change in vehicle velocity as a result of a collision (assumed to be perfectly inelastic) has been shown to be a useful measure for crash severity (Shelby, 2011). Further, Evans (1994) developed the direct relationship between Delta-V and the resultant crash severity, thereby validating its use in conflict studies. Consequently, a few studies have used Delta-V with crash occurrence indicators to jointly estimate the crash occurrence and severity probabilities (Alhajyaseen, 2014, Wang and Stamatiadis, 2014).

However, the current severity estimation studies suffer from two significant drawbacks (Arun et al., 2021b). Firstly, as seen in Wang and Stamatiadis (2014), these studies directly adopted the power-law relationship given by Evans (1994) for severity estimation using delta-v. However, Evans (1994) only considered two-car homogeneous crash cases while utilizing crash data from 1981 to 82, meaning that the inferences drawn from those studies may be outdated and not representative of the current driving population and vehicle technology. Further, unlike in traditional crash-based analysis, these studies do not produce severity estimates consistent with the commonly used severity scales like the Injury Severity Scale (ISS) and the Abbreviated Injury Scale (AIS). Typical crash-based severity estimation studies (Savolainen et al., 2011) utilize such severity scales to enable unambiguous comprehension of estimation results across geographical regions and across var-

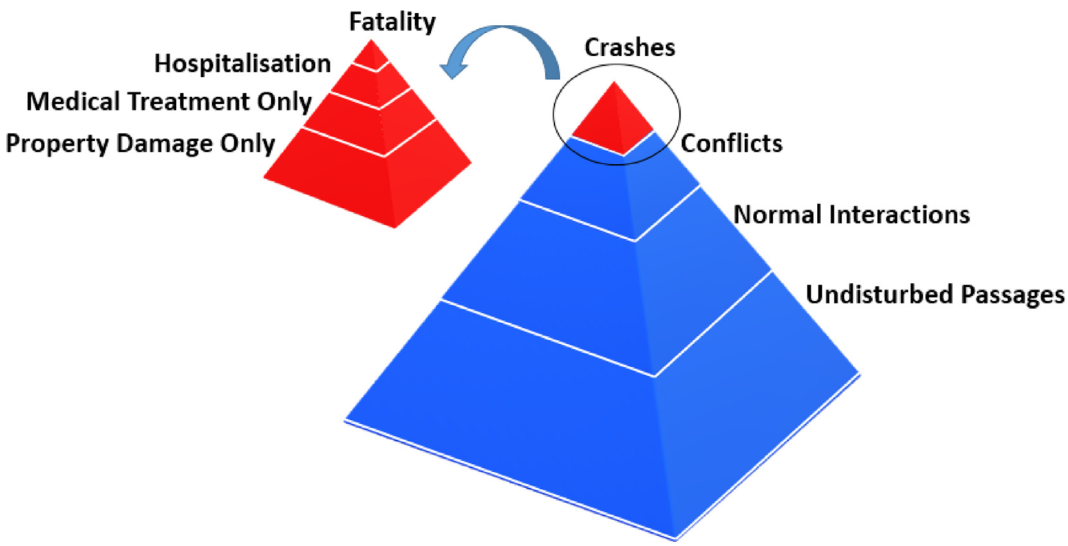

Fig. 1. Severity-integrated safety pyramid of traffic events [adapted from Hydén (1987)].

ious research disciplines. Moreover, since such scales are also used in other fields of research such as medical sciences and traffic psychology, their use can increase the utility of the study results in interdisciplinary research. Hence, the inability of the conflict-based safety assessment studies to estimate the severity risk per the above severity scales limits their applicability.

Thus, there is a significant need to incorporate the estimation of crash severity within conflict-based safety assessment in a way that is consistent with the current crash-based practice to allow for broader applicability. With this view, this study proposes to combine the conflict measures for crash frequency and severity for predicting crash frequency-by-severity, where the severity is defined according to the standard abbreviated injury scale. Given the demonstrated applicability of Extreme Value methods for crash prediction using conflicts, this study applies bivariate extreme value models for the joint modelling of crash frequency and severity for rear-end crashes at signalized intersections, benchmarking it against a univariate approach suggested by Tarko (2012). As per the authors' best knowledge, the present study is the first to attempt crash frequency-by-severity estimation using conflict measures.

The paper is organized in the following manner. Section 2 presents the modelling approaches for estimating crash frequency by severity using traffic conflicts, Section 3 presents the traffic conflict extraction method from video recordings and a summary of the extracted data, section 4 presents the estimation results, Section 5 provides model comparisons and interpretation of the results, and section 6 concludes with a summary of the significant results and contributions to the science.

2. Methodology

2.1. Modelling approaches

Recently, Zheng et al. (2021) discussed two methods of severity estimation using conflict measures. The first method was suggested by Tarko (2012), where the probability of crash occurrence is multiplied by the conditional probability of crash severity and then leveraging the Counterfactual Theory (Davis et al., 2011), the resultant product is multiplied by the number of conflicts to estimate the total number of crashes by severity-level. Therefore, the frequency of crashes for a given severity level can be computed as per Eq. (1):

$$N_k = Pr(k|crash).Pr(crash|conflict).n \tag{1}$$

where N_k is the frequency of crashes of the k^{th} severity level, n is the number of observed conflicts, $Pr(k|crash)$ is the probability of the severity outcome of the k^{th} severity level in the event of a crash, and $Pr(crash|conflict)$ is the probability of crash occurrence given the occurrence of a conflict. Tarko (2012) argued that while the crash occurrence probability ($Pr(crash|conflict)$) can be computed from common indicators of crash occurrence such as Time-to-Collision (TTC) using univariate extreme value theory models, the conditional probability of crash severity, on the other hand, can be estimated using the traditional discrete choice modeling methods like logistic regression. In the remaining paper, this approach is referred to as the univariate approach.

In the second method, the crash severity probability can be jointly estimated with the crash occurrence probability in a bivariate setting. Thus, the frequency of crashes for a given severity level in this approach can be calculated using Eq. (2):

$$N_k = Pr(X \geq X_{Occurrence} \wedge Y \geq Y_k).n \tag{2}$$

where X is an indicator of nearness-to-collision with $X_{Occurrence}$ being the limit representing crash occurrence. Thus, if time-to-collision (TTC) is used as the indicator of nearness-to-collision X, then $X_{Occurrence} = 0$, which is the TTC threshold corresponding to crash occurrence. For Deceleration Required to Avoid Collision (DRAC) indictor, a suitable value of $X_{Occurrence}$ is the mean of the truncated normal distribution N(8.45, 1.42)I(4.23, 12.68) of Maximum Available Deceleration Rate (MADR) of vehicles (Fu and Sayed, 2021). Y in Eq. (2) is the indicator of crash severity with Y_k being the limit of observing a crash of severity level k. Zheng et al. (2021) argued that this joint exceedance probability could be estimated using a bivariate extreme value distribution. In the remaining paper, this approach is referred to as the bivariate approach.

2.2. Crash frequency and severity indicators

Time-to-Collision (TTC) is one of the most prevalent conflict measures of crash frequency (Arun et al., 2021a) and has also been used successfully in the extreme value theory framework for crash prediction (Jonasson and Rootzen, 2014, Zheng et al., 2018b). Therefore, TTC was used as the indicator for crash frequency prediction in this study. TTC can be computed using the following equation:

$$TTC = \frac{x_l - x_f - D_l}{v_f - v_l}; \forall (v_f - v_l) > 0 \tag{3}$$

where x_l and x_f are the positions of the front bumpers of the leading and the following vehicles at the time of observation, respectively, D_l is the length of the leading vehicle, and v_l and v_f are the speeds of the leading and following vehicles at the time of observation, respectively. However, time-to-collision definition assumes that the conflicting vehicles continue moving on their path with constant velocity from the point of observation. This property is a problem because, in many cases, drivers take some evasive action like braking, swerving, acceleration, or a combination of those on detecting the occurrence of conflict. Therefore, an improvement to the time-to-collision indicator was introduced by Ozbay et al. (2008), where the vehicles are assumed to continue moving on their respective paths with constant acceleration from the point of observation. This indicator is called Modified Time-to-Collision (MTTC) and can be computed as per Eq. (4):

$$MTTC = \frac{\Delta s \pm \sqrt{\Delta s^2 + 2\Delta a(x_l - x_f - D_l)}}{\Delta a} \tag{4}$$

where $\Delta s = v_f - v_l$ is the relative speed of the conflicting vehicles and $\Delta a = a_f - a_l$ is their relative acceleration. Zheng and Sayed (2019a) compared several conflict indicators for rear-end conflicts and found that MTTC was the best conflict measure among MTTC, TTC, Post-Encroachment Time (PET), and Deceleration Rate to Avoid Collision (DRAC). Thus, MTTC was also investigated in this study.

As discussed before, Delta-V can be a suitable candidate for a severity indicator in a bivariate extreme value theory model along with frequency indicators like TTC to estimate crash frequency-by-severity levels. Delta-V can be estimated according to the procedure mentioned by Laureshyn et al. (2017). Specifically, the absolute value of Delta-V for the leading and following vehicles involved in a perfectly inelastic rear-end collision is given by Eq. (5):

$$\Delta v_f = \frac{m_l}{m_f + m_l}\sqrt{v_f^2 + v_l^2 - 2v_f v_l \cos\theta}$$

$$\Delta v_l = \frac{m_f}{m_f + m_l}\sqrt{v_f^2 + v_l^2 - 2v_f v_l \cos\theta} \tag{5}$$

where m_l *and* m_f are the masses of the leading and the following vehicles, respectively; v_l and v_f are their speeds; θ is their approach angle. The severity of the entire interaction can be taken as the larger of the two Delta-v values.

A significant challenge toward such a joint estimation is the absence of Delta-V thresholds for distinguishing among the various severity levels. Unlike crash occurrence, where a time-to-collision value of 0 represents crash occurrence, there are no natural upper thresholds on Delta-V that correspond to the various crash severity levels (Zheng et al., 2021). However, some crash reconstruction studies have estimated relationships between Delta-V and the probability of injury and fatality in the case of a crash (Bahouth et al., 2014; Evans, 1994). Estimating severity probabilities from Delta-V is helpful in our case because probability values have natural thresholds for whether a crash of a given severity level k is obtained (k^{th} severity probability = 1) or not (k^{th} severity probability = 0). Given that the relationships developed by Bahouth et al. (2014) are recent, they have been used in this study to arrive at suitable thresholds for the Delta-V indicator for use in the bivariate approach. Moreover, the estimated severity probabilities can be directly used in the univariate approach for multiplication with the expected number of crashes as per Eq. (1).

In particular, Bahouth et al. (2014) investigated the probability of crash severity per the standard injury-severity scales, namely Maximum Abbreviated Injury Scale (MAIS) and Injury Severity Scale (ISS). They developed logistic regression models for estimating the probability of severe injury crashes (MAIS3 + or ISS > 15) according to the point of impact, namely, frontal, nearside, far side, and rear, as a function of Delta-V alongside other predictors such as indicator variables for multiple effects, belt use and age of the drivers. Given that it was not possible to collect data regarding the drivers' age and the presence of multiple points of impact from video observations of conflicts, the values of corresponding indicator variables were kept as 0 in this study. Therefore, specifically, for rear-end crashes, the following simplified version of the equation from Bahouth et al. (2014) was used to estimate the logit component in logistic regression for MAIS3 + crash probability:

$$\eta = -5.7077 + 0.1492 \times \Delta v - 1.8892 \times Belt_{use} \tag{6}$$

where Δv is delta-v (in km/h) and $Belt_{use}$ is the dummy variable for whether the crash victims were using the seat belts. Subsequently, the probability of severe (MAIS3 +) and non-severe crashes can then be obtained using Eq. (7):

$$p_{severe} = \frac{\eta}{1+\eta}; p_{non-severe} = 1 - p_{severe} \tag{7}$$

Thus, Eqs. (6) and (7) have been utilized to estimate the MAIS3 + crash probability from Delta-V values in this study.

2.3. Extreme value analysis

Extreme value theory (EVT) is used to analyze the tail of distributions and extrapolate the information from the observed levels in the data to the unobserved (extreme) levels (Coles, 2001). In conflict-based safety assessments, EVT corresponds to analyzing the tail of the distribution of traffic events that contain the traffic conflicts (observed levels in the data) to predict the traffic crashes (unobserved levels). Given the rare and stochastic nature of crash occurrence, EVT, in conjunction with comparatively more frequent traffic conflicts, aids faster analysis and designing safety interventions (Zheng et al., 2014a; Zheng and Sayed, 2019b).

EVT analysis can be carried out following two popular approaches: The Block-Maxima method and the Peak-Over-Threshold method. Details regarding these two methods can be found elsewhere (Coles, 2001; Beirlant et al., 2004). A body of work has reported that the peak-over-threshold method performs better than the block-maxima method for conflicts-based crash estimation, primarily due to better utilization of limited data in the former approach (Zheng and Sayed, 2019d, 2019c). Therefore, the peak-over-threshold method was applied for both univariate and bivariate analysis in this study. The modeling approaches are briefly explained below.

3. Univariate Peak-over-threshold model

In a peak-over-threshold model, a generalized Pareto distribution is fitted to the data points that exceed a specified threshold. Such points are called exceedances of that threshold, and the threshold is chosen high enough so that the exceedances are in some sense extreme, justifying the application of extreme value statistics.

Let $\{X_1, X_2 \cdots X_n\}$ be a set of independent random observations with the unknown cumulative distribution function $F(x) = \Pr(X_i \leq x), i = 1 \cdots n$. Then the distribution function of exceedances $X = X_i : X_i > u$ over a "large" threshold u can be approximated by generalized Pareto distribution and is given by Eq. (8):

$$F_u(x) = \Pr(X - u \leq x|X) = 1 - \left(1 + \frac{\xi(X-u)}{\sigma}\right)^{-\frac{1}{\xi}} \tag{8}$$

where $\sigma \in (0, \infty)$ and $\xi \in (-\infty, \infty)$ are the scale and shape parameters, respectively, of the Generalized Pareto distribution.

4. Bivariate Peak-over-threshold model

Bivariate extreme value models are analogous to the univariate models in that the marginal distributions of the two random variables X and Y are modeled as independent generalized Pareto distributions, with the added complexity of defining the dependence structure for the two margins.

Let $\{(X_1, Y_1), (X_2, Y_2) \cdots (X_n, Y_n)\}$ be a set of independent random observations of a bivariate vector (X, Y) with the unknown cumulative distribution function $F(X, Y)$. For suitably large marginal thresholds u_x and u_y, each marginal distribution of exceedances $x > u_x$ and $y > u_y$ is approximated by Eq. (9) below, with respective parameter sets $(\zeta_x, \sigma_x, \xi_x)$ and $(\zeta_y, \sigma_y, \xi_y)$, where $\zeta_x = \Pr\{x > u_x\}$ and $\zeta_y = \Pr\{y > u_y\}$.

$$G(z) = 1 - \zeta\left(1 + \frac{\xi(z-u)}{\sigma}\right)^{-\frac{1}{\xi}}, z > u \tag{9}$$

Transforming the variable (X, Y) to $\left(\tilde{X}, \tilde{Y}\right)$ such that

$$\tilde{X} = -\left(\log\left\{1 - \zeta_x\left[1 + \frac{\xi_x(X-u_x)}{\sigma_x}\right]^{-\frac{1}{\xi_x}}\right\}\right)^{-1} \text{and } \tilde{Y} = -\left(\log\left\{1 - \zeta_y\left[1 + \frac{\xi_y\left(Y-u_y\right)}{\sigma_y}\right]^{-\frac{1}{\xi_y}}\right\}\right)^{-1}$$

Then the joint distribution function $\tilde{F}\left(\tilde{X}, \tilde{Y}\right)$ has margins that are approximately standard Fréchet distributed for $X > u_x$ and $Y > u_y$. Coles (2001) explained that in such a case, the joint distribution of threshold excesses $F(x, y)$ can be derived from Eq. (10):

$$F(x, y) = \tilde{F}\left(\tilde{x}, \tilde{y}\right) \approx G(x, y) = \exp\left\{-V\left(\tilde{x}, \tilde{y}\right)\right\}, x > u_x, y > u_y \tag{10}$$

where $\tilde{x}$ and $\tilde{y}$ are the transformations for $x > u_x$ and $y > u_y$, respectively, and $V\left(\tilde{x}, \tilde{y}\right)$ gives the dependence function of the Fréchet distributed margins and can be estimated as

$$V\left(\tilde{x}, \tilde{y}\right) = 2\int_0^1 \max\left(\frac{w}{\tilde{x}}, \frac{1-w}{\tilde{y}}\right) dH(w) \tag{11}$$

with $H(w)$, called spectral measure, being any distribution function on $w \in [0, 1]$ that satisfies the mean constraint.

$$\int_0^1 w dH(w) = 1/2 \tag{12}$$

An alternative representation of $V\left(\tilde{x}, \tilde{y}\right)$ is as below:

$$V\left(\tilde{x},\tilde{y}\right) = 2\left(\frac{1}{\tilde{x}}+\frac{1}{\tilde{y}}\right)\int_0^1 \max\left(\frac{w\tilde{y}}{\tilde{x}+\tilde{y}},\frac{(1-w)\tilde{x}}{\tilde{x}+\tilde{y}}\right)dH(w)$$

Then, let

$$A(t) = \int_0^1 \max\left[w(1-t),(1-w)t\right]dH(w)$$

and $F(x,y)$ can be written as:

$$F(x,y) \approx G(x,y) = \exp\left\{-\left(\frac{1}{x}+\frac{1}{y}\right)A\left(\frac{x}{x+y}\right)\right\} \quad (13)$$

where $A(t)$ is the Pickands dependence function that is convex and satisfies the relation $(1-t)\vee t \le A(t) \le 1, t \in [0,1]$.

In parametric estimation procedure, several distribution families can be used for the function $H(.)$. The most popular one that has been used in the previous studies is the logistic distribution function (Zheng et al., 2018a, 2019b; Zheng and Sayed, 2019d). The list of other popular parametric models is given elsewhere (Beirlant et al., 2004). The parameters of the dependence functions can measure the strength of dependence between the two coordinates. For example, in the logistic distribution function, the parameter $\varphi \in (0,1)$ determines the strength of dependence. Thus, $\varphi = 1$ and $\varphi \downarrow 0$ correspond to independence and perfect dependence, respectively, as given in Table 1 (Beirlant et al., 2004).

The logistic model is a symmetric distribution model as both the coordinates x and y are interchangeable. In this study, some other symmetric models, namely negative logistic and Husler-Reiss, have also been investigated for the dependence structure. The details of the specific models used in this study are given in Table 1.

5. Threshold selection

The discussion on peak-over-threshold models' application stresses that such models are valid only if the chosen thresholds are 'large' enough for modeling the extreme values located in the tail of a distribution. Thus, threshold selection is an essential component of peak-over threshold analysis and needs to be undertaken carefully. Choosing a very low threshold will result in even expected values to be characterized as extremes and, therefore, will likely violate the generalized Pareto model's asymptotic basis. On the other hand, a very high threshold will lead to an inadequate number of exceedances resulting in higher variances in the parameter estimates. Therefore, the standard practice is to choose a threshold as low as possible, subject to the condition that the limit model provides a reasonable approximation (Coles, 2001).

In the univariate approach, the validity of approximation to the threshold excess distribution can be checked based on two graphs, namely Mean Residual Life Plot and Threshold Stability Plot. The theory behind these two plots is that if the generalized Pareto distribution is valid for excesses over a threshold u_0, then it should be equally valid for excesses over all thresholds $u > u_0$ subject to the appropriate change of scale parameter σ_u (Coles, 2001). The shape parameter ξ is invariant to the threshold. Thus, for $u > u_0$, the mean of threshold excesses $\mathrm{E}(X-u|X>u)$ is a linear function of u. Thus, the mean residual life plot should be linear above the appropriate threshold. Moreover, the scale parameter σ_u for thresholds $u : u > u_0$ is a linear function of the threshold u as given below:

$$\sigma_u = \sigma_{u_0} + \xi(u-u_0) \quad (14)$$

Rearranging the terms in the above equation:

$$\sigma_u - \xi u = \sigma_{u_0} + \xi u_0 or, \sigma^* = \sigma_{u_0} + \xi u_0$$

where σ^* is the reparameterized scale parameter, which is constant with respect to the threshold u. Since the shape parameter remains unchanged, the estimates of both σ^* and ξ should be constant for any threshold $u > u_0$ after allowing for variability due to sampling errors.

The above two plots can also be utilized to obtain the suitable threshold intervals for the margins in the bivariate approach (Zheng et al., 2018a). Further, Beirlant et al. (2004) developed a Spectral Measure Plot that leverages the spectral

Table 1
Parametric distributions used for Bivariate peak-over-threshold models in this study.

Distribution	Functional form	Strength of dependence measure
Logistic	$\boldsymbol{G}(\boldsymbol{x},\boldsymbol{y}) = \exp\left\{-\left(\boldsymbol{x}^{1/\varphi}+\boldsymbol{y}^{1/\varphi}\right)^{\varphi}\right\}$ $0<\boldsymbol{\varphi}\le 1$	Independence: $\boldsymbol{\varphi}=1$ Perfect dependence: $\boldsymbol{\varphi}\to 0$
Negative Logistic	$\boldsymbol{G}(\boldsymbol{x},\boldsymbol{y}) = \exp\left\{-\boldsymbol{x}-\boldsymbol{y}+\left(\boldsymbol{x}^{-\varphi}+\boldsymbol{y}^{-\varphi}\right)^{-1/\varphi}\right\}$ $0<\boldsymbol{\varphi}<\infty$	Independence: $\boldsymbol{\varphi}\to 0$ Perfect dependence: $\boldsymbol{\varphi}\to\infty$
Husler-Reiss	$\boldsymbol{G}(\boldsymbol{x},\boldsymbol{y}) = \exp\left\{-\boldsymbol{x}\Phi\left[\boldsymbol{\varphi}^{-1}+\frac{1}{2}\boldsymbol{\varphi}\left(\log\left\{\frac{\boldsymbol{x}}{\boldsymbol{y}}\right\}\right)\right]-\boldsymbol{y}\Phi\left[\boldsymbol{\varphi}^{-1}+\frac{1}{2}\boldsymbol{\varphi}\left(\log\left\{\frac{\boldsymbol{y}}{\boldsymbol{x}}\right\}\right)\right]\right\}$ $0<\boldsymbol{\varphi}<\infty$ and $\Phi(.)$ is the standard normal distribution	Independence: $\boldsymbol{\varphi}\to 0$ Perfect dependence: $\boldsymbol{\varphi}\to\infty$

decomposition of a bivariate peak-over-threshold model's dependence structure for threshold selection. Specifically, the margins are transformed to standard Fréchet distributions using the following formula:

$$x_{*ij} = -\frac{1}{\log u_{ij}}, \quad i = 1, 2, ..., n; j = 1, 2 \tag{16}$$

Then the pseudo-polar coordinates of the transformed vector (x_{*i1}, x_{*i2}) with both norms equal to the sum-norm can be written as:

$$r_i = x_{*i1} + x_{*i2}, \quad \omega_{ij} = x_{*ij}/r_i \tag{17}$$

Arranging the radial coordinates r_i in ascending order like $r_{(1)} \leq \cdots \leq r_{(k)} \leq \cdots \leq r_{(n)}$, the value of k (denoted k_0) that provides an estimate of the spectral measure $\widehat{H}([0,1]) = (k_0/n)r_{(n-k_0)}$ close to two (2), which is the theoretical value of $H([0,1])$ for the bivariate case, can be used to select the threshold. Thus, a suitable threshold can be estimated using Eq. (18):

$$u_j = x_{(n-k),j}, \quad j = 1, 2; k \leq k_0 \tag{18}$$

Importantly, as Beirlant et al. (2004) noted, the above threshold selection methods are not absolute and may not always lead to an optimal choice according to some criterion. Indeed, in the bivariate case, the marginal and dependence structure considerations may result in different thresholds. Therefore, there are other methods of threshold selection in the broader literature. For instance, Zheng et al. (2019b) mentioned a "fixed quantile rule" method wherein a high upper quantile such as 5% is used as the threshold. Therefore, in this study, these threshold selection methods were used to select initial threshold values for the variables, and the optimal thresholds were iteratively obtained through model estimation based on improvements in Akaike Information Criterion (AIC) values as well as comparisons of the estimated severe crash frequency with the observed value as suggested by Cavadas et al. (2020).

5.1. Crash frequency-by-severity estimation and validation

In the univariate approach, the expected number of crashes was estimated using Eq. (1). There are two critical components of this estimation. The first is the estimation of $Pr(crash|conflict)$, which in this case was estimated using the univariate peak-over-threshold model given by Eq. (8). Specifically, given that $TTC(orMTTC) = 0$ indicates the occurrence of a crash, the probability of crash given a conflict can be obtained as the probability of the negated TTC (or MTTC) indicator exceeding the value of 0 (Zheng and Sayed, 2019d), that is:

$$\Pr(crash|conflict) = \Pr(Z \geq 0) = 1 - F_u(0) \tag{19}$$

where Z is either negated TTC or negated MTTC.

The second component required in Eq. (1) is the conditional estimate of the crash severity probability ($Pr(k|crash)$), which can be obtained as the conditional probability of a severe (maximum abbreviated injury scale value ≥ 3; MAIS3 +) or a non-severe crash estimated per Eqs. (6) and (7). Thus, the number of severe ($N_{T,severe}$) and non-severe crashes ($N_{T,non-severe}$) in the period of interest (T) can be estimated according to the following formulae:

$$\begin{aligned} N_T &= \tfrac{T}{\tau} \times n_\tau \\ n_\tau &= \textstyle\sum_{i=1}^{m} \Pr(Z_i \geq 0) \times p_i \forall m : Z \geq u_Z \end{aligned} \tag{20}$$

whereN_T is the number of crashes, τ is the conflict observation period, p is either p_{severe} or $p_{non-severe}$, the probabilities of severe and non-severe crashes (from Eq. (7)), respectively, u_Z is the selected threshold of the negated TTC (or MTTC) indicator, and m is the number of extreme values higher than the selected threshold. The number of severe (maximum abbreviated injury scale value $\geq$ 3; MAIS3 +) or non-severe crashes can be calculated by substituting the suitable probability value (p_{severe} or $p_{non-severe}$) in the above equation.

In the bivariate approach, an injury crash is obtained if both the crash frequency and severity indicators simultaneously exceed their respective thresholds; that is, the quantity of interest is the joint exceedance probability of the two variables, which can be modelled as a Bivariate generalized Pareto distribution $F(x,y)$ given by Eq. (10). Again, using negated TTC (or MTTC), the threshold of exceedance is 0. For severity indicator Delta-V, the exceedance threshold can be computed from the severity probability. In this study, if the severity probability is significant ($p_{severe} \geq 0.5$) then the resultant crash is a severe (maximum abbreviated injury scale value $\geq$ 3; MAIS3 +) crash; otherwise, it is labelled a non-severe crash. Accordingly, by substituting the value of logit $u = 1$ in Eq. (6), corresponding to $p_{severe} = 0.5$, and assuming all the occupants of the conflicting vehicles were belted at the time of the conflict, ($Belt_{use} = 1$), the delta-v is computed to be 16.006 m/s (rounded to three decimal places). The assumption regarding all the conflicting vehicle occupants being belted flowed from the fact that seatbelt non-compliance is very low (5%) in Queensland (Minister for Transport and Main Roads, 2018) and that all the crash victims in the crashes analyzed in this study were reported wearing seatbelts. Therefore, 16 m/s was used as the delta-v threshold in the bivariate approach. Accordingly, the crash estimates can be estimated using the formula adapted from Zheng et al. (2018a) that is given below:

$$N_{Severe,T} = \frac{T}{\tau} \times \Pr(Z \geq 0 \wedge Delta - V \geq 16) = 1 - F(0,16) \tag{21}$$

$$N_{Non-severe,T} = Totalcrashes - SevereCrashes = N_T - N_{Severe,T}$$

where N_T can be computed from the estimated marginal distribution of the frequency indicator (negated TTC or negated MTTC) per Eq. (20).

Validation of the above frequency-by-severity estimates is not simple since the observed crash frequency is a point measure with no variance. Songchitruksa and Tarko (2006) suggested a novel method of constructing Poisson Confidence Intervals over the observed crash frequency and investigating if the estimated crash frequency lies within that interval. This approach has subsequently been adopted in several other studies (Wang et al., 2019, Zheng and Sayed, 2019d, Zheng et al., 2019b). The Poisson confidence intervals can be estimated according to the following formula:

$$\left\{ \lambda : \frac{1}{2T}\chi^2_{2y_0,1-\alpha/2} \leq \lambda \leq \frac{1}{2T}\chi^2_{2(y_0+1),\alpha/2} \right\} \tag{23}$$

where y_0 is the observed crash frequency in time period T, λ is the estimated crash frequency for time period T, and α is the level of significance.

5.2. Comparison between site-specific and combined data models

Another methodological aspect investigated in this study was whether there was any advantage in combining the data of similar sites for extreme value analysis. Extreme value models are known to be data-hungry (Zheng et al., 2014a), and evidence from previous literature suggests some benefits of combining the data from similar sites in terms of better prediction accuracy and precision compared to site-specific analysis (Zheng et al., 2019a). Given that this aspect has remained under-investigated, this study proposes to compare the prediction results from both the modeling approaches for site-specific and combined dataset models.

6. Data

The data used in this study were collected in August 2019 from two four-legged signalized intersections located in South-East Queensland, Australia. The intersections were selected based on their crash history so that they had recorded a severe injury (maximum abbreviated injury scale ≥ 3; MAIS3 +) (labeled "Hospitalization") crashes in the past five years. Moreover, the two sites are similar in terms of geometric layout (four-legged intersections), operations (signal-controlled), and location as both are located within the same jurisdiction (Brisbane Metropolitan Region), situated only 8 km apart along the same road corridor.

This study's primary data consisted of traffic movement footage recorded using a dedicated overhead camera assembly mounted on a 6.5 m high mast installed at the intersections. Multiple camera positions were required to record the traffic movements at the intersections. Each intersection required four cameras to obtain unobstructed views of all the targeted road user movements. The cameras collected video data for four weekdays from Tuesday to Friday at each intersection. Given that traffic conditions on Friday may differ from those on other weekdays, the conflict data from the various weekdays were investigated to ensure the absence of any temporal trends before aggregating them together. Table 2 gives the details of the study intersections and the data collection schedule, while Fig. 2 shows a plan view of one of the intersections (Beaudesert

Table 2
Details of study intersections and the data collection schedule.

Intersection name	Camera Number	Camera position	Data collection	Total Duration
Beaudesert Rd – Granard Rd (BG) Intersection	1	View 1	20/08/2019 (Tuesday; 6 am to 6 pm) 21/08/2019 (Wednesday; 6 am to 6 pm)	24 h
		View 2	22/08/2019 (Thursday; 6 am to 6 pm) 23/08/2019 (Friday; 6 am to 6 pm)	24 h
	2	View 1	20/08/2019 (Tuesday; 6 am to 6 pm) 21/08/2019 (Wednesday; 6 am to 6 pm)	24 h
		View 2	22/08/2019 (Thursday; 6 am to 6 pm) 23/08/2019 (Friday; 6 am to 6 pm)	24 h
Kessels Rd – Logan Rd (KL) Intersection	1	View 1	20/08/2019 (Tuesday; 6 am to 6 pm) 21/08/2019 (Wednesday; 6 am to 6 pm)	24 h
		View 2	22/08/2019 (Thursday; 6 am to 6 pm) 23/08/2019 (Friday; 6 am to 6 pm)	24 h
	2	View 1	20/08/2019 (Tuesday; 6 am to 6 pm) 21/08/2019 (Wednesday; 6 am to 6 pm)	24 h
		View 2	22/08/2019 (Thursday; 6 am to 6 pm) 23/08/2019 (Friday; 6 am to 6 pm)	24 h

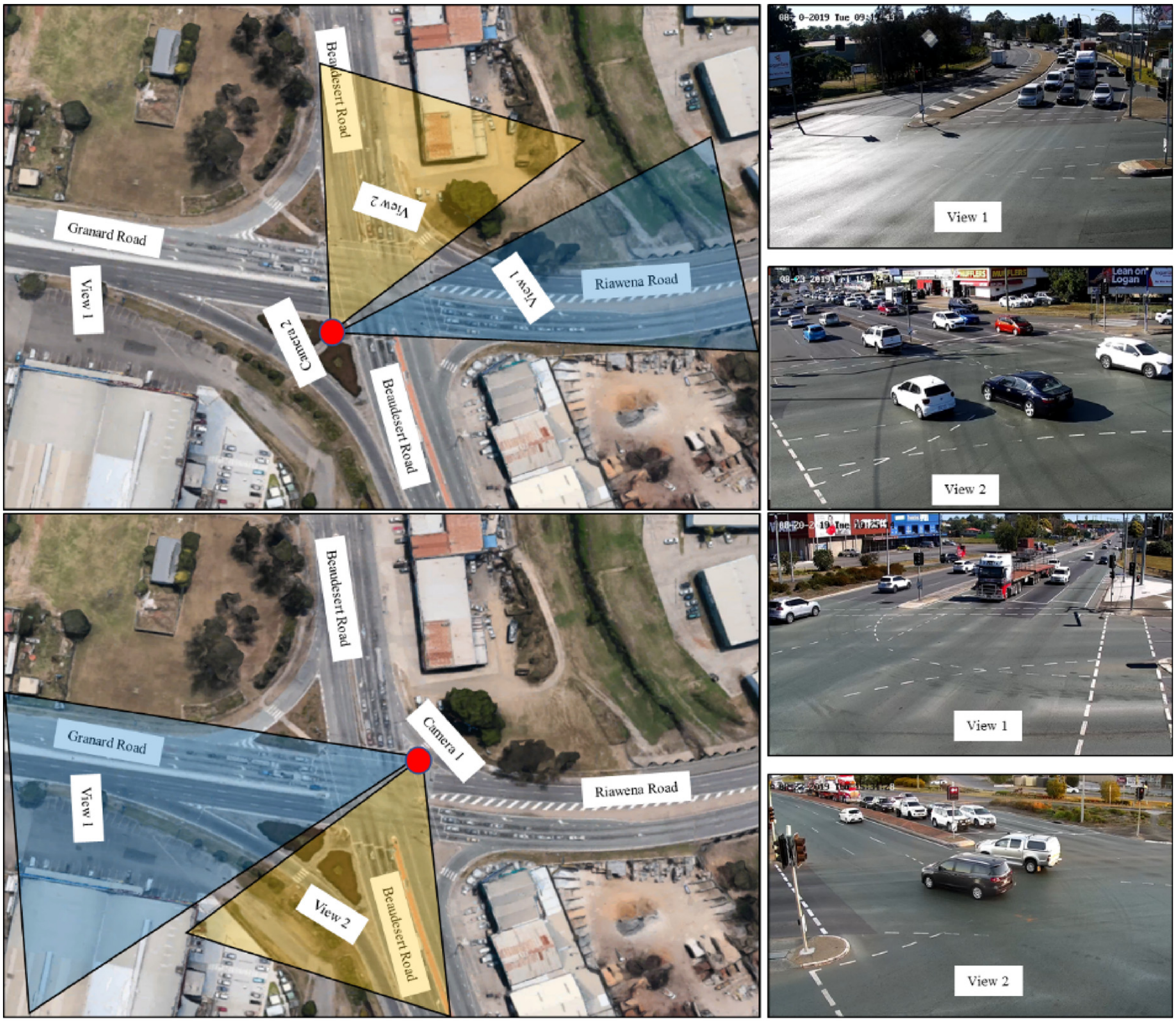

Fig. 2. Camera positions and views from the cameras at the Granard Rd – Beaudesert Rd Intersection.

Rd – Granard Rd intersection) and the camera placement and intersection areas covered in the camera views. As shown in Fig. 2, one camera-mast was placed on the pedestrian refuge island at the corner of Beaudesert Road (North) and Riawena road, and the other camera-mast was placed on the pedestrian refuge island at the corner of Granard Road and Beaudesert Road (South). Each mast was used to cover two different views, as reported in Table 2. The shaded areas show the approximate coverage of each camera used at this site. Similar practices were followed for the Kessels Rd – Logan Rd intersection to collect the traffic movements data.

The raw video data were then processed using a deep neural network-based automated conflict extraction method to extract the conflict measures of interest, namely, time-to-collision (TTC), modified time-to-collision (MTTC), and Delta-V. This method involves six main procedures: camera calibration, object detection and tracking, prototype generation, prototype matching, event generation, and conflict identification. Fig. 3 gives an overview of this system.

The system was developed at the Queensland University of Technology based on earlier research conducted at the University of British Columbia (Saunier and Sayed, 2008, St-Aubin et al., 2015). However, unlike the feature tracking and grouping algorithm followed in the previous studies, the system developed at the Queensland University of Technology employs the YOLOv3 algorithm for object detection in the traffic scene and then the DeepSORT algorithm to track all the road users (motorized as well as non-motorized) for greater efficiency in the object detection and tracking module. The overlapping of the predicted vehicle trajectories would generate a safety-relevant event, and all the events with time-to-collision $\leq$ 3.0 s were identified as traffic conflicts and subsequently used in this study for developing the crash-conflict relationships. Modified time-to-collision and Delta-V values were subsequently calculated for these identified conflicts. The severity probability (p_{severe}) was calculated according to Eq. (7). In this study, only the rear-end conflicts data were utilized for crash prediction modeling to limit the scope of the study.

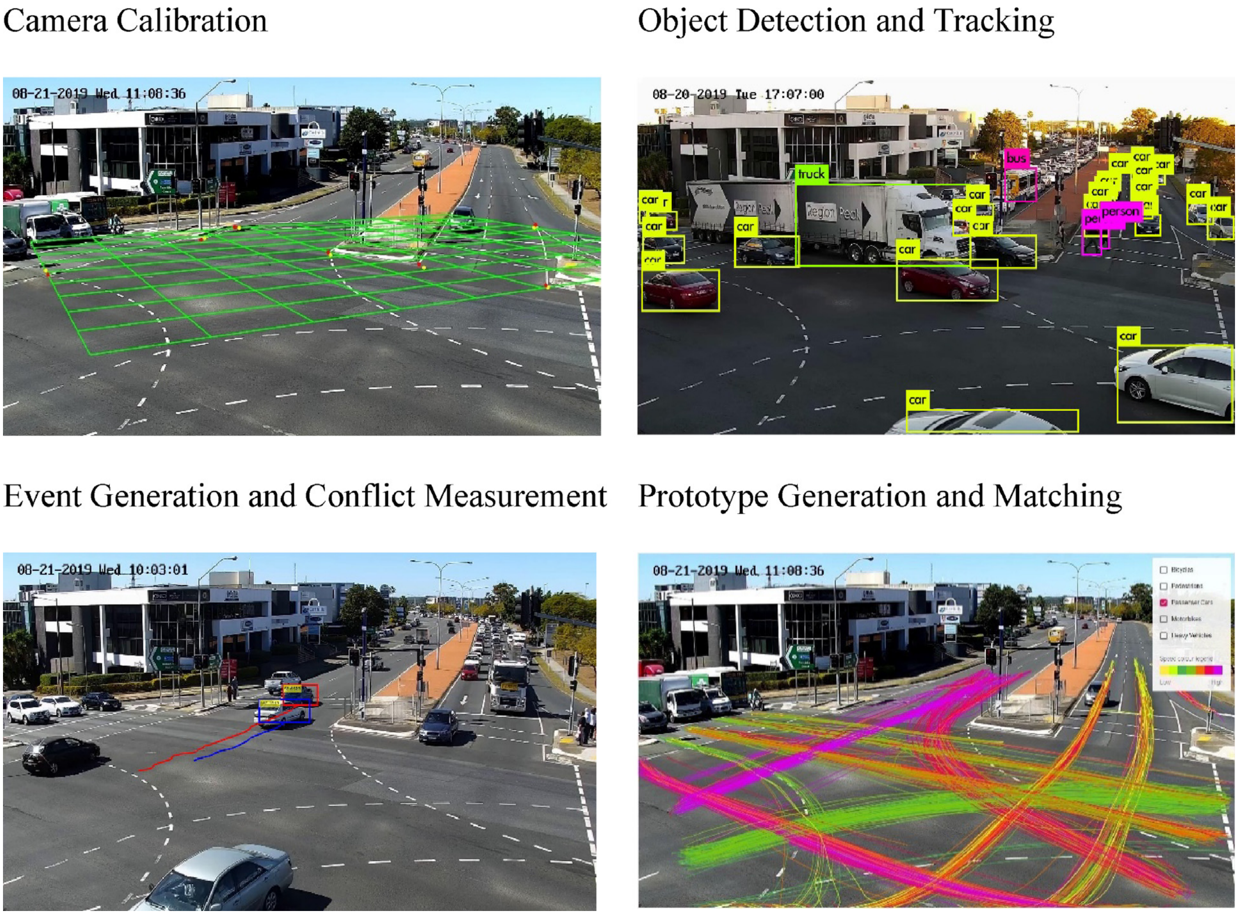

Fig. 3. Computer Vision-based automated conflict data extraction method.

Table 3 provides the descriptive statistics of the conflict measures used in this study for the study intersections and the combined dataset of the two intersections. An important thing to note is the correlations between the frequency indicators (TTC and MTTC) and the severity indicator Delta-V. The correlations are weak (absolute value of Kendall's tau < 0.5) for both combinations (TTC and Delta-V, and MTTC and Delta-V) for both intersections. However, the correlation coefficient between MTTC and Delta-V is significantly different from zero for both intersections considered individually (p-value < 0.05). For TTC and Delta-V combination, the correlations are significant in Beaudesert Rd - Granard Rd intersection dataset. Neither of the combinations is significantly correlated in the combined dataset. Such weak correlations between the two types of indicators are expected because they measure fundamentally different properties, namely the nearness to a collision and the severity consequence of that collision, respectively. Indeed, Delta-V estimation is independent of TTC (or MTTC) estimation and can be calculated even in the case of non-conflict traffic interactions.

The crash data given in Table 3 are required for validating the crash prediction models developed in this study. The five-year (2015–2019) crash data for the study intersections were provided by the Department of Transport and Main Roads, Queensland Government (DTMR). The crash data's key information parameters were time, location, severity, road user information, and collision type of each crash. It was necessary to filter the crash data so that it corresponded to the conditions under which the conflicts were observed for accurate comparisons with the crashes predicted by the peak-over-threshold models. Hence, the crash data were filtered to include only the cases corresponding to the following criteria: the crashes occurring during weekdays (Monday - Friday), during the daytime (6 am – 6 pm), and during fair weather conditions.

An issue with the crash data was that they did not report crash severity per the abbreviated injury scale (AIS). In the Queensland crash data, the crash severity classes, namely Minor Injury, Medical Treatment-Only, Hospitalization, and Fatality in the increasing order of severity, do not directly correspond to the AIS levels. This issue is critical because, as mentioned before, the Delta-V-to-severity relationships only exist for the standard AIS scale. However, Henley and Harrison (2016), while investigating the trend of road traffic crash severity in Australia, noted that the hospitalization and fatal crash severity levels in the Australia New Zealand Trauma Registry corresponded to the maximum abbreviated injury scale ≥ 3 (MAIS3 +) severity level. Thus, hospitalization severity crashes in Queensland crash data could be used for model validation. The DTMR-provided crash data does not report any fatal crashes at the study intersections in the last five years and contains only two

Table 3
Descriptive statistics of the conflict measures.

Intersection name	No. of rear-end conflicts	No. of severe rear-end crashes (2015–19)	No. of non-severe rear-end crashes (2015–19)	Conflict measures	Mean	Median	Std. Dev.	Min.	Max.	Kendall's Tau (p-value)
Beaudesert Rd – Granard Rd (BG) Intersection	3829	1	15	TTC	1.67	1.70	0.68	0.02	2.97	
				MTTC	1.13	1.01	0.55	0.02	2.87	
				Delta-V	6.89	6.71	3.41	0.23	23.5	
				p_{severe}	0.07	0.02	0.14	0.0	0.99	
				TTC, Delta-V						−0.04 (0.01)
				MTTC, Delta-V						−0.04 (0.01)
Kessels Rd – Logan Rd (KL) Intersection	2772	1	8	TTC	1.41	1.34	0.62	0.01	2.88	
				MTTC	0.91	0.82	0.43	0.01	2.69	
				Delta-V	6.52	6.41	3.2	0.0	21.8	
				p_{severe}	0.05	0.02	0.12	0.0	0.98	
				TTC, Delta-V						0.03 (0.15)
				MTTC, Delta-V						0.07 (0.00)
Combined Dataset	6601	2	23	TTC	1.56	1.52	0.67	0.01	2.97	
				MTTC	1.03	0.91	0.52	0.01	2.87	
				Delta-V	6.73	6.57	3.33	0.0	23.5	
				p_{severe}	0.06	0.02	0.13	0.0	0.99	
				TTC, Delta-V						0.002 (0.79)
				MTTC, Delta-V						0.013 (0.63)

Legend:
TTC: Time-to-Collision.
MTTC: Modified Time-to-Collision.
p_{severe}: Probability of obtaining a severe crash, see Eq. (7).

rear-end hospitalization crashes, one each reported at Beaudesert Rd – Granard Rd Intersection and Kessels Rd – Logan Rd Intersection. Hence, only the number of hospitalization crashes at these two intersections was utilized to validate the severe crash frequency model. The number of non-hospitalization crashes was used to validate the non-severe crash frequency model.

7. Results

The univariate extreme value models were estimated with the maximum likelihood estimation method. In contrast, the bivariate extreme value models were estimated using the censored likelihood estimation approach, where only the values above the respective thresholds were utilized in estimating the likelihood function (Beirlant et al., 2004). The goodness-of-fit of these models was examined by inspecting the reasonableness of quantile and density plots. The procedure for developing the specific models is explained in detail below.

7.1. Data declustering

The peak-over-threshold models assume that individual excesses are independent of each other and are at odds with serial dependence commonly observed in traffic events like conflicts (Zheng et al., 2014a). In the case of rear-end conflicts, for instance, a hard braking incident by a lead vehicle in a group of vehicles closely following each other may lead to successive hard braking by the following vehicles, and, as such, all these vehicles can come into conflict with each other. Thus, the independence assumption under the peak-over-threshold approach is likely to be violated. Therefore, in this study, the extremes in the data occurring in clusters were treated using a runs-based de-clustering method recommended by Coles (2001). Specifically, in chronologically arranged data, if more than three successive values within 5 s were found to be lying above a chosen threshold, they were identified as belonging to a cluster, and only the maximum of those clustered observations was selected for subsequent analysis. The optimum run-length of three was identified based on estimating the extremal index (Smith and Weissman, 1994), which gave the level of clustering in the data, and auto-tail dependence function (Coles, 2001). Given that 3 consecutive values could be in different time periods, the data were grouped in several time intervals ranging from 5-seconds to one hour, and the optimum interval length was obtained by comparing the extremal index and auto-tail dependence function values for the different time intervals. Consequently, 5-second was identified as the opti-

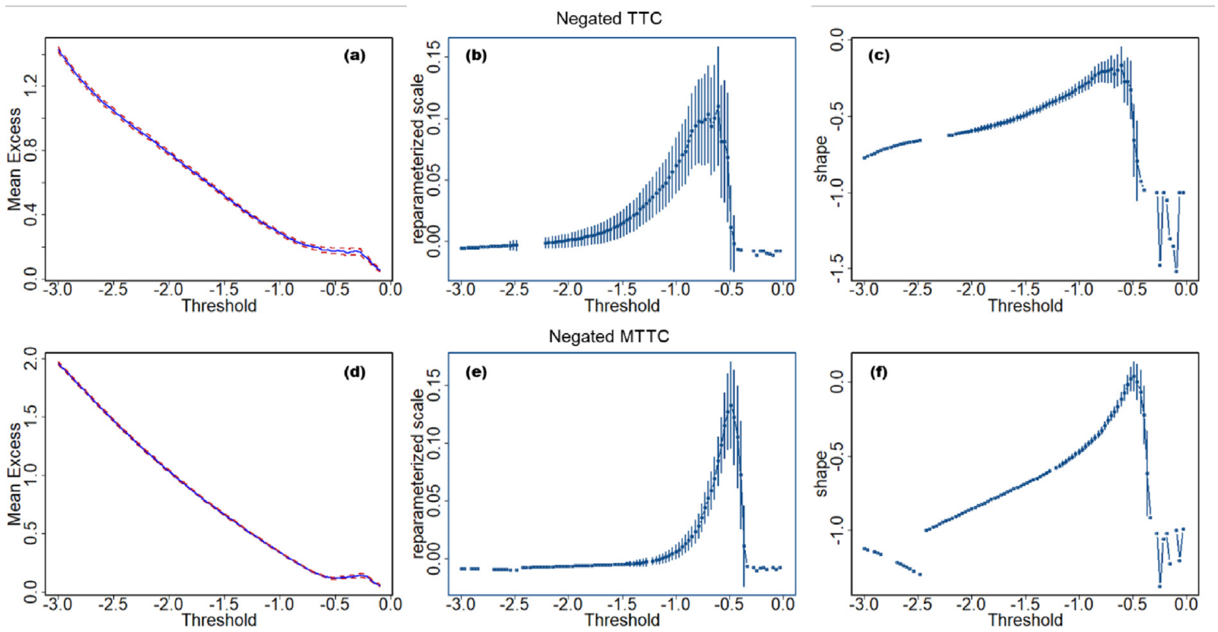

Fig. 4. Mean Residual Life (MRL; plots (a) and (d)) and Threshold Stability (TS; plots (b), (c), (e) and (f)) plots for Negated TTC (top) and Negated MTTC (bottom) indicators.

mum time interval. The threshold was kept as −3.0 s for both negated time-to-collision and negated modified time-to-collision indicators to make the most efficient use of data.

7.2. Univariate approach with combined dataset

The first step in this approach is threshold selection for the conflict measures. As mentioned previously, the analysis was performed for the negated values of time-to-collision (TTC) and modified time-to-collision (MTTC). The mean residual life and threshold stability plots for both the indicators are given in Fig. 4. For negated TTC, the mean residual life plot is almost linear until a threshold of −0.5 s, while the threshold stability plot indicates that the scale and the shape parameters of the generalized Pareto distribution are somewhat stable below −1.7 s. Thus, [-3.0 s, −1.7 s] was chosen as the threshold range of the negated TTC indicator. Similarly, for negated MTTC, the mean residual life plot is linear until a threshold of −0.5 s, while the scale parameter plot is almost stable below −1.2 s, rendering [-3.0 s, −1.2 s] as the threshold range for this indicator. These plots helped select the initial thresholds; the final thresholds for both the indicators were iteratively obtained based on lower Akaike information criterion (AIC) and higher prediction accuracy of their respective generalized Pareto distribution models. For instance, for the TTC indicator, the model failed to converge until a threshold of −2.05 s, and the AIC and prediction error values corresponding to that threshold were 6030.29 and −0.221, respectively. The threshold values were then successively increased from −2.05 s until −1.7 s at incremental steps of 0.1 s, and the AIC and prediction errors were computed at each step. The final threshold was selected as −2.02 s, which had the lowest AIC of 5752.89 and the lowest prediction error of −0.055. For comparing the prediction performance, the estimated annual crash frequency of severe crashes—which was estimated by substituting $T = 1$ year in Eq. (20) along with the severity probability values—was compared with the observed severe crash frequency. The non-severe crash frequency was not utilized for model checking because the non-severe crashes might be underreported since it is a typical characteristic of police-reported crash data (Savolainen et al., 2011).

The estimation results are given in Table 4 for both TTC and MTTC-based peak-over-threshold models, along with the log-likelihood (*LL*) and Akaike information criterion (AIC) values. The 95% Poisson confidence intervals of the observed crash frequencies were constructed using Eq. (23). From Table 4, in terms of model fit, the negated MTTC model has a better fit to the data than the negated TTC model owing to much lower AIC (higher log-likelihood value). However, in terms of prediction accuracy, indicated by whether the estimated severe and non-severe crashes lay within their respective observed confidence intervals, negated TTC model is better as both severe and non-severe crash estimates lie within the confidence bounds. Conversely, in the negated MTTC model, only the estimated severe crashes lie in the observed confidence intervals. The TTC model estimates are also more precise than those obtained from the MTTC model. Another critical observation is regarding the conflict threshold determined by the peak-over-threshold procedure. The conflict thresholds for the negated TTC model is −2.02 s, which is considerably higher than the average response time of 1.56 s of the Australian driving population in low-speed emergency situations (Sharma et al., 2019). Thus, it is questionable if the negated TTC model is effectively modeling

Table 4
Estimation results for the univariate GPD modeling with Negated TTC and Negated MTTC.

Parameter	Negated TTC	Negated MTTC
$\hat{\sigma}$(SE)	1.211 (0.012)	0.589 (0.009)
$\hat{\xi}$(SE)	−0. 599 (0.006)	−0.53 (0.009)
$\boldsymbol{u}$	−2.02	−1.11
Exceedances	4854	4288
LL	−2874.443	254.021
AIC	5752.886	−504.043
Estimated $\boldsymbol{N_{Severe}}$ (95% CI)	0.455 (0.0–472.478)	0.512 (0.0–480.415)
Observed $\boldsymbol{N_{Severe}}$ (95% CI)	0.4 (0.048–1.44)	
Estimated $\boldsymbol{N_{Non-severe}}$ (95% CI)	6.465 (0.0–6709.93)	7.298 (0.0–6849.613)
Observed $\boldsymbol{N_{Non-severe}}$ (95% CI)	4.6 (2.916–6.902)	

Legend:
TTC: Time-to-Collision.
MTTC: Modified Time-to-Collision.
LL: Log-likelihood.
AIC: Akaike Information Criterion.
$\boldsymbol{N_{Severe}}$: Number of severe crashes.
$\boldsymbol{N_{Non-severe}}$: Number of non-severe crashes.
SE: Standard Error.
95% CI: Confidence Interval at 95% level of significance.

truly extreme traffic conflicts. Nevertheless, both the univariate models were highly imprecise because the widths of the prediction intervals for both severe and non-severe crashes were substantially large compared to the observed confidence intervals.

7.3. Bivariate approach with combined dataset

The bivariate extreme value theory approach involves the estimation of the marginals as well as a dependence structure. For estimating the dependence structure, the relative dependence between the two variables needs to be investigated first. Table 3 indicates that crash frequency indicators (TTC or MTTC) and the severity indicator (Delta-V) are weakly correlated. A more formal test of dependence was done using the chi plot method suggested by Coles (2001). The resultant plots (Fig. 5) for both the bivariate vectors are interesting because they suggest that the two variables in both the combinations, (negated TTC, Delta-V) and (negated MTTC, Delta-V), are almost independent except for weak asymptotic dependence between them at very high quantiles from the chi-bar plots. The chi-bar value for (negated TTC, Delta-V) vector at the highest quantile ≈ 1 is

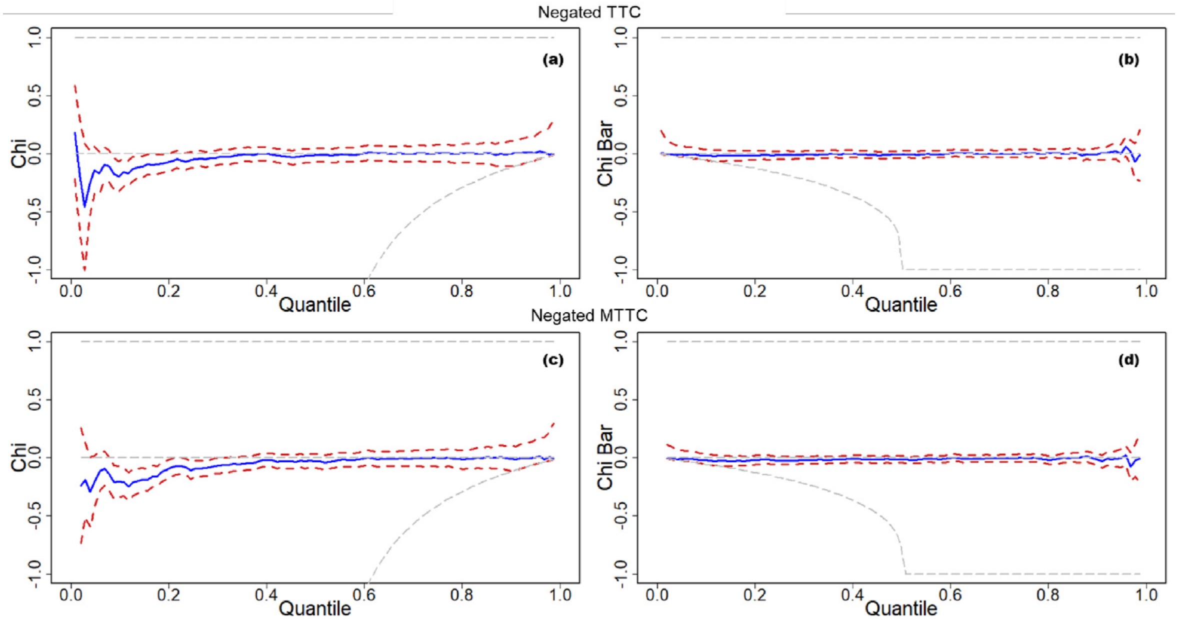

Fig. 5. Chi ((a) and (c)) and Chi Bar ((b) and (d)) Plots for Negated TTC and Delta-V (top) and Negated MTTC and Delta-V (bottom).

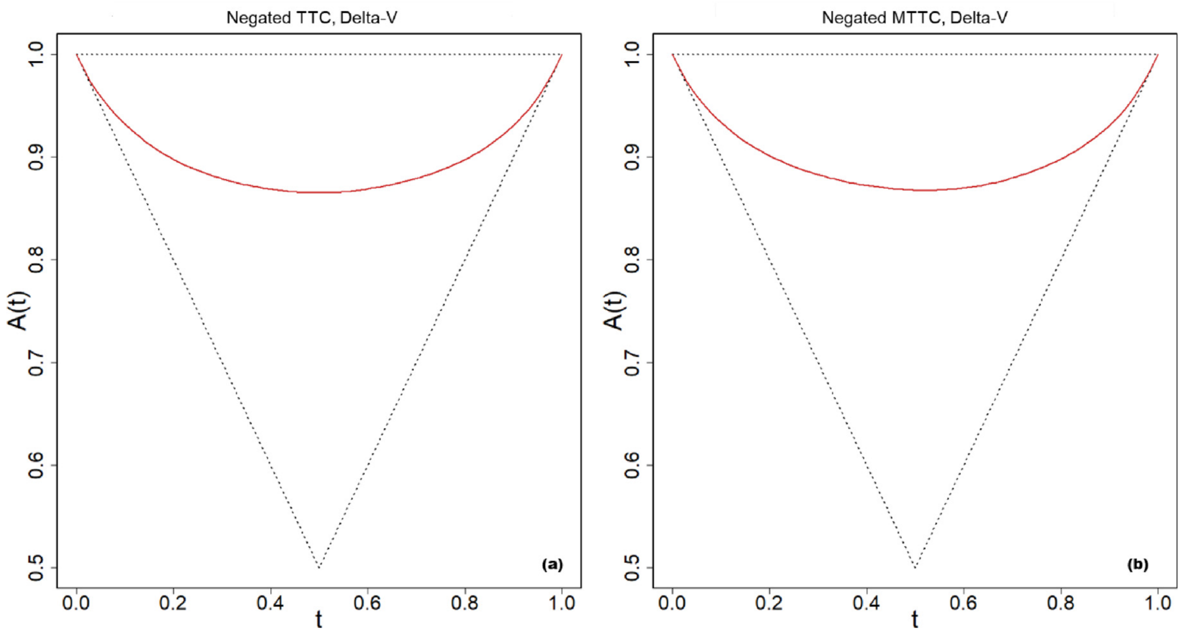

Fig. 6. Nonparametrically estimated Dependence Function Plot for (a) Negated TTC and Delta-V, and (b) Negated MTTC and Delta-V.

−0.013 (95% confidence interval: −0.232, 0.206), while that for the (negated MTTC, Delta-V) vector is −0.012 (95% confidence interval: −0.23, 0.209), indicating that very weak asymptotic dependence may be present among the two variables in both the bivariate vectors.

Consequently, the Pickands' dependence function was nonparametrically estimated for both the bivariate vectors through peak-over thresholds method that only uses the k-largest values for estimation (Stephenson, 2002), where k is determined following the spectral decomposition of the joint distribution function per Eq. (18) (Beirlant et al., 2004). The resultant plots are given in Fig. 6 where $t \in [0, 1]$ is the parameter of the Pickands' Dependence function $A(t)$ (see Eq. (13)). This process returned an estimate of 0.866 for (negated TTC, Delta-V) vector and a slightly higher 0.868 for (negated MTTC, Delta-V) vector, indicating that each bivariate vector's variables are weakly dependent and further justifying the use of the bivariate extreme value model at least for very high quantiles (Dutfoy et al., 2014). There is also no evidence of asymmetry in the resultant plots meaning that symmetric bivariate dependence structures like logistic models are suitable for both the bivariate models.

Threshold selection for the bivariate data was carried out with the help of the mean residual life and threshold stability plots for each margin and the spectral measure plot for the combined data. The mean residual life and threshold stability plots for TTC and MTTC were given earlier in Fig. 4, while those for Delta-V are given in Fig. 7. The spectral measure plots for both the bivariate vectors are given in Fig. 7, where k is the index of the ordered bivariate vector, $H([0, 1])$ is the spectral measure estimate, and k_0 is the optimum value of k at which $H([0, 1]) \approx 2$ (see discussion for Eqs. (17) and (18)). Based on the analysis of the marginal threshold choice plots of negated TTC and negated MTTC indicators, initial thresholds of −1.7 s and −1.2 s were chosen, as explained earlier. For Delta-V, the mean residual life plot is almost linear until 12 m/s, while the threshold stability plots are somewhat stable in the range [10 m/s, 12 m/s). The spectral measure plots (Fig. 8), on the other hand, indicated a $k_0$value of 3,253 for the (negated TTC, Delta-V) vector and 3,282 for the (negated MTTC, Delta-V) vector that corresponded to the thresholds (−1.47 s, 6.81 m/s) and (−0.89 s, 6.77 m/s), respectively. Thus, taking the intersections of thresholds suggested by the marginal and spectral measure plots, the initial thresholds were chosen as (−1.47 s, 10.0 m/s) and (−0.89 s, 10.0 m/s) for the (negated TTC, Delta-V) and the (negated MTTC, Delta-V) vectors, respectively.

Considering the absence of asymmetry in the nonparametric dependence plots (Fig. 5), only symmetric parametric distributions like Logistic, Negative Logistic and Husler-Reiss distributions were considered to model the dependence structure, while the marginals were modeled as generalized Pareto distributed. Several thresholds were iteratively tested for optimality based on improvements in Akaike information criterion (AIC) values and prediction accuracy. It was attempted to keep the number of joint exceedances the same in all three models for comparison. Fig. 9 provides the density and quantile plots of the final bivariate models for the negated MTTC (mttc.neg) and Delta-V (delta_v) vector. The quantile curves indicate that the best fitting bivariate GPD models are obtained only at very high quantiles that confirm the findings from chi-bar plots in Fig. 5. Moreover, the final distributions with all the bivariate dependence structures fit the data well as all the quantile curves are convex and have similar shapes (de Haan and de Ronde, 1998). Similar observations were made from the density and quantile plots of the (negated TTC, Delta-V) models, which are not presented here due to space constraints.

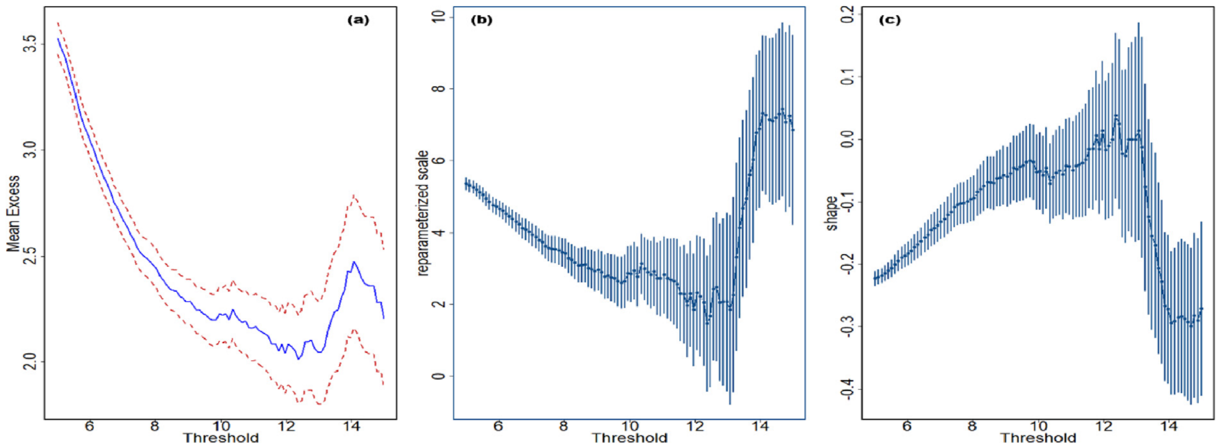

Fig. 7. Mean Residual Life (MRL) Plot (a) and Threshold Stability (TS) Plots ((b) and (c)) for Delta-V.

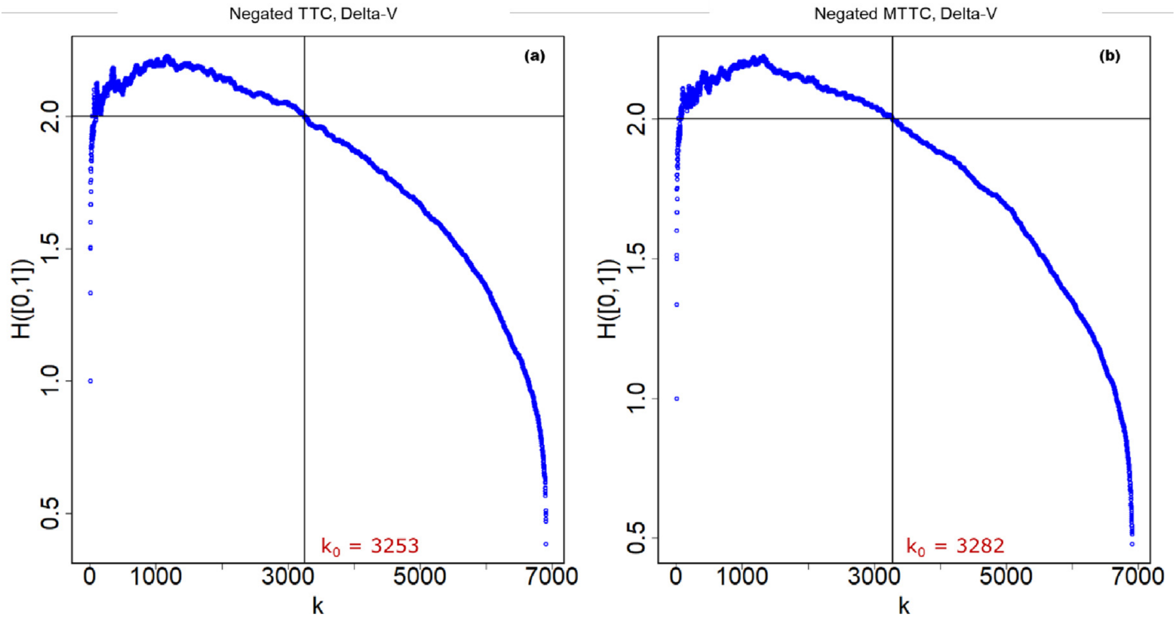

Fig. 8. Spectral Measure Plots for Negated TTC and Delta-V (left) and Negated MTTC and Delta-V (right).

Table 5 gives the estimation results of the best fitted Logistic (Log), Negative Logistic (NL) and Husler-Reiss (HR) models for the two bivariate vectors. Notably, unlike in the univariate models, the means of the estimated annual severe and non-severe crash frequencies are within the 95% confidence interval of their observed counterparts for all the bivariate models. More importantly, the estimated prediction intervals from the bivariate models are much narrower than the univariate models and closer to the observed confidence interval, indicating that considerably better precision is obtained using the bivariate approach.

Among the negated TTC models, the negative logistic model ($NL_{TTC,\ Delta\text{-}V}$) fits the data best, having the lowest AIC value, followed by the logistic ($Log_{TTC,\ Delta\text{-}V}$) and the Husler-Reiss ($HR_{TTC,\ Delta\text{-}V}$) models in that order. However, in terms of prediction accuracy, the $Log_{TTC,\ Delta\text{-}V}$ model was the most accurate with prediction errors of −0.151 for severe crash frequency and 0.049 for total crash frequency (sum of severe and non-severe crash frequencies). The $HR_{TTC,\ Delta\text{-}V}$ model was the next most accurate but more precise than the logistic model because the prediction intervals for both severe and non-severe crash frequencies were narrower. The $NL_{TTC,\ Delta\text{-}V}$ model was the least accurate and most imprecise of the three. However, overall, all the TTC models performed comparably in terms of fit to the data and the prediction accuracy and precision of crashes.

Among the negated MTTC models, the logistic model ($Log_{MTTC,\ Delta\text{-}V}$) had the best prediction performance in terms of accuracy and precision, although its goodness-of-fit was lower than the Husler-Reiss model ($HR_{MTTC,\ Delta\text{-}V}$). The $NL_{MTTC,}$

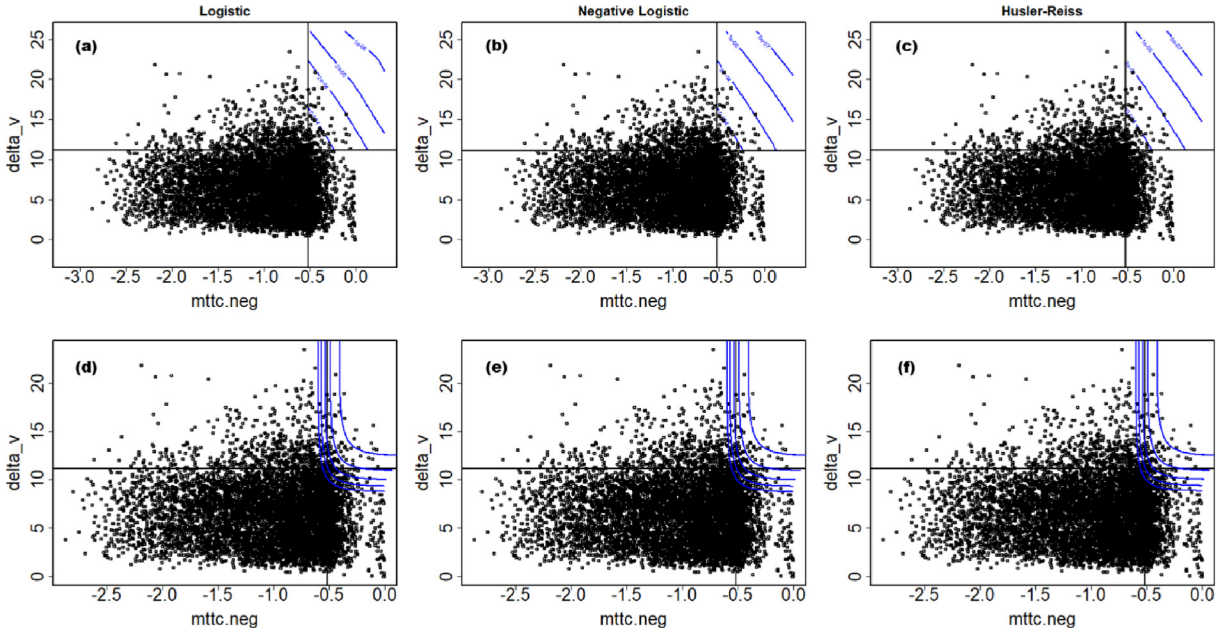

Fig. 9. Density ((a), (b) and (c)) and Quantile ((d), (e) and (f)) Plots for the best fitted Bivariate GP distribution for Negated MTTC and Delta-V with Logistic, Negative Logistic and Husler-Reiss dependence structures.

$_{\text{Delta-V}}$ model provided the worst fit to the data, and its prediction performance was comparable to the HR $_{\text{MTTC, Delta-V}}$. Based on the observations from both TTC and MTTC models, therefore, it can be concluded that the logistic dependence structure is the most suitable for modeling traffic conflicts for crash frequency-by-severity estimation.

Comparing among all the six models, the TTC models generally fit their data better with lower AIC values. However, they all had lower accuracy than their MTTC counterparts and were also more imprecise, as indicated by their wider prediction intervals for both severe and non-severe crashes. Moreover, the best-fit TTC models had fewer joint exceedances than the MTTC models, meaning that fewer traffic interactions can be identified as serious conflicts using the TTC and Delta-V combination. Thus, the higher precision of the MTTC models may be attributable to the additional information available from the higher number of serious conflicts identified using the negated MTTC and Delta-V combination. Overall, based on the higher accuracy and predictions of the crash frequency-by-severity estimates, the Log$_{\text{MTTC, Delta-V}}$ model outperforms the other models while also providing a good fit to the data. Nevertheless, all the models systematically overestimated the crash frequency estimates for severe crashes; the estimated means were higher than the respective observed means.

7.4. Intersection-wise crash Frequency-by-Severity estimation

As stated before, the univariate and bivariate models estimated using the combined dataset of the two intersections were compared with those estimated separately for each intersection. Only the best performing univariate and bivariate models were estimated for intersection-wise analysis based on their combined dataset performance. Thus, the negated MTTC model was estimated using the univariate approach, while the Log$_{\text{MTTC, Delta-V}}$ model was estimated using the bivariate approach. Table 6 provides the result of the estimation of intersection-wise models using both approaches.

For both intersections, again, both the univariate and bivariate approaches provided accurate frequency predictions for severe crashes as the estimated means lay within the observed 95% confidence interval. Moreover, except the univariate model for Kessels Rd-Logan Rd intersection, all the models accurately estimated the non-severe crash frequency. However, the bivariate approach's real advantage is apparent in the much-improved precision of the frequency estimates. The prediction intervals of the bivariate models for both severe and non-severe crashes and both the intersections are substantially smaller than those of both the corresponding univariate models. Comparing among the bivariate models for both intersections themselves, the Kessels Rd-Logan Rd intersection model was more accurate and precise than the Beaudesert Rd-Granard Rd intersection model in terms of prediction performance.

For comparing the combined dataset models and the intersection-wise models, the intersection-wise crash estimates from the two models can be utilized. From Table 5, the Log$_{\text{MTTC, Delta-V}}$ model predicts the two intersections to have observed 0.417 severe crashes and 4.656 non-severe crashes in the last five years. Apportioning the crashes to individual intersections based on the relative numbers of observed conflicts from Table 3, the combined dataset model thus predicted 0.242 severe and 2.701 non-severe crashes at BG intersection and 0.175 severe and 1.955 non-severe crashes at Kessels Rd-Logan Rd intersection. Thus, the combined dataset model was more accurate for Beaudesert Rd-Granard Rd intersection than the

Table 5
Estimation results of the bivariate peak-over-threshold models with Logistic, Negative Logistic and Husler-Reiss dependence structures.

	Parameter	Negated TTC and Delta-V			Negated MTTC and Delta-V		
		Logistic ($Log_{TTC,\ Delta\text{-}V}$)	Negative Logistic ($NL_{TTC,\ Delta\text{-}V}$)	Husler-Reiss ($HR_{TTC,\ Delta\text{-}V}$)	Logistic ($Log_{MTTC,\ Delta\text{-}V}$)	Negative Logistic ($NL_{MTTC,\ Delta\text{-}V}$)	Husler-Reiss ($HR_{MTTC,\ Delta\text{-}V}$)
Marginals	$\hat{\boldsymbol{\sigma}}_1$(SE)	0.245 (0.016)	0.243 (0.016)	0.24 (0.015)	0.141 (0.008)	0.115 (0.006)	0.115 (0.006)
	$\hat{\xi}_1$(SE)	−0.236 (0.048)	−0.229 (0.049)	−0.216 (0.047)	−0.078 (0.043)	0.028 (0.042)	0.028 (0.042)
	$\boldsymbol{u}_1$	−0.66	−0.66	−0.67	−0.51	−0.51	−0.51
	ME1	542	542	570	821	821	821
	$\hat{\boldsymbol{\sigma}}_2$(SE)	2.208 (0.135)	2.202 (0.136)	2.207 (0.136)	1.811 (0.098)	2.291 (0.129)	2.279 (0.129)
	$\hat{\xi}_2$(SE)	−0.042 (0.044)	−0.037 (0.045)	−0.038 (0.045)	0.061 (0.043)	−0.055 (0.04)	−0.052 (0.041)
	$\boldsymbol{u}_2$	11.45	11.46	11.45	11.16	11.12	11.14
	ME2	566	564	566	637	642	638
Dependence	$\boldsymbol{\varphi}$	0.999 (2×10^{-6})	0.051 (2×10^{-6})	0.2 (2×10^{-6})	0.999 (2×10^{-6})	0.051 (2×10^{-6})	0.202 (2×10^{-6})
	JE	46	46	49	72	72	72
LL		−4501.964	−4492.936	−4553.325	−4852.483	−4865.57	−4847.529
AIC		9013.928	8995.872	9116.65	9714.965	9741.141	9705.058
Estimated $\boldsymbol{N_{Severe}}$ (95% CI)		0.551 (0.0–2.998)	0.567 (0.0–3.039)	0.556 (0.0–2.882)	0.417 (0.019–1.721)	0.551 (0.061–1.857)	0.553 (0.062–1.865)
Observed $\boldsymbol{N_{Severe}}$ (95% CI)		0.4 (0.048–1.445)					
Estimated $\boldsymbol{N_{Non-severe}}$ (95% CI)		4.4 (0.0–15.965)	4.556 (0.0–16.115)	4.482 (0.0–15.324)	4.656 (0.386–12.66)	5.03 (1.071–11.196)	5.028 (1.071–11.19)
Observed $\boldsymbol{N_{Non-severe}}$ (95% CI)		4.6 (2.916–6.902)					

Legend:
TTC: Time-to-Collision.
MTTC: Modified Time-to-Collision.
LL: Log-likelihood.
AIC: Akaike Information Criterion.
$\boldsymbol{N_{Severe}}$: Number of severe crashes.
$\boldsymbol{N_{Non-severe}}$: Number of non-severe crashes.
$\left(\hat{\boldsymbol{\sigma}}_1, \hat{\xi}_1, \boldsymbol{u}_1\right)$ are parameters for negated TTC or negated MTTC margin; $\left(\hat{\boldsymbol{\sigma}}_2, \hat{\xi}_2, \boldsymbol{u}_2\right)$ are parameters for Delta-V margin.
ME1: Marginal Exceedances of negated TTC negated MTTC margin; ME2: Marginal Exceedances of Delta-V margin.
JE: Joint Exceedances.
SE: Standard Error.
95% CI: Confidence Interval at 95% level of significance.

Table 6
Estimation results for the intersection-wise bivariate and univariate GPD models.

Intersection		Parameter	Bivariate (Negated MTTC and Delta-V)	Univariate (Negated MTTC)
Beaudesert Rd-Granard Rd intersection	Marginals	$\hat{\sigma}_1$(SE)	0.111 (0.008)	0.404 (0.01)
		$\hat{\xi}_1$(SE)	0.021 (0.056)	−0.428 (0.013)
		u_1	−0.53	−0.94
		ME1	466	1820
		$\hat{\sigma}_2$(SE)	1.888 (0.125)	–
		$\hat{\xi}_2$(SE)	0.025 (0.049)	–
		u_2	11.18	–
		ME2	417	–
	Dependence	φ	0.999 (2×10^{-6})	–
		JE	48	–
	LL		−2967.505	607.937
	AIC		5945.011	−1211.873
	Estimated N_{Severe} (95% CI)		0.343 (0.007–1.634)	0.229 (0.0–203.622)
	Observed N_{Severe} (95% CI)		0.2 (0.005–1.114)	
	Estimated $N_{Non-severe}$ (95% CI)		3.515 (0.172–11.075)	2.689 (0.0–2389.45)
	Observed $N_{Non-severe}$ (95% CI)		3 (1.679–4.948)	
Kessels Rd–Logan Rd intersection	Marginals	$\hat{\sigma}_1$(SE)	0.177 (0.008)	0.718 (0.016)
		$\hat{\xi}_1$(SE)	−0.164 (0.03)	−0.598 (0.014)
		u_1	−0.62	−1.2
		ME1	803	2214
		$\hat{\sigma}_2$(SE)	2.155 (0.206)	–
		$\hat{\xi}_2$(SE)	−0.011 (0.075)	–
		u_2	11.1	–
		ME2	219	–
	Dependence	φ	0.999 (2×10^{-6})	–
		JE	53	–
	LL		−2138.906	−157.08
	AIC		4287.812	318.161
	Estimated N_{Severe} (95% CI)		0.194 (0.001–1.23)	0.194 (0.0–547.335)
	Observed N_{Severe} (95% CI)		0.2 (0.005–1.114)	
	Estimated $N_{Non-severe}$ (95% CI)		1.809 (0.05–6.088)	3.628 (0.0–10215.11)
	Observed $N_{Non-severe}$ (95% CI)		1.6 (0.691–3.153)	

Legend:
TTC: Time-to-Collision.
MTTC: Modified Time-to-Collision.
LL: Log-likelihood.
AIC: Akaike Information Criterion.
N_{Severe}: Number of severe crashes.
$N_{Non-severe}$: Number of non-severe crashes.
$(\hat{\sigma}_1, \hat{\xi}_1, u_1)$ are parameters for negated TTC margin; $(\hat{\sigma}_2, \hat{\xi}_2, u_2)$ are parameters for Delta-V margin.
ME1: Marginal Exceedances of negated TTC margin; ME2: Marginal Exceedances of Delta-V margin.
JE: Joint Exceedances.
SE: Standard Error.
95% CI: Confidence Interval at 95% level of significance.

intersection-wise model and less so for the Kessels Rd-Logan Rd intersection. However, given that the sum of severe crashes from the intersection-wise models is 0.537 and that of non-severe crashes is 5.324, the combined dataset models provided more accurate prediction results than the intersection-wise models when comparing the total severe and non-severe crashes at the two intersections.

8. Discussion

This study has demonstrated the estimation of crash frequency-by-severity from conflict measures—addressing a critical scientific gap and, thereby, paving the way for integrating crash severity estimation within conflict-based safety assessment frameworks. Importantly, this study has provided a novel procedure to utilize the Delta-V indicator in extreme value theory analysis for estimating crash frequencies by severity levels. It is recognized that the results are contingent upon the analysis of a small sample of locations. Thus, repeating these experiments is recommended to build confidence in the generalizability of the result discussed here.

Firstly, there are distinct advantages of using the bivariate generalized Pareto distribution for predicting crash frequency-by-severity over the univariate method suggested by Tarko (2012) and Zheng et al. (2021). The joint prediction of crash frequency-by-severity employing bivariate peak-over-threshold models yields more accurate (estimated mean lies within the Poisson confidence interval over observed crash frequency) and more precise (prediction interval is narrower) estimates than the univariate approach (comparing the results in Table 4 and Table 5). In the univariate models, the upper limits of the prediction intervals for both severe and non-severe crash frequencies and employing both TTC and MTTC were several times higher than the respective upper limits of the confidence intervals of observed crashes. The bivariate models' improved prediction performance is likely attributable to the additional information present in the dependence structure between crash frequency and severity indicators. For the bivariate models given in Table 5, the dependence measure estimates were small but statistically significant, indicating a dependence between crash frequency and severity risks at extreme quantiles. Physically, this result means that very serious conflicts could potentially carry a higher risk of serious injury to involved persons. This dependence information is not utilized in the univariate modeling approach. For the intersection-wise analysis, the bivariate models provided more precise estimates with narrower prediction intervals than the univariate models for both the study intersections, further indicating the effects of dependence between crash frequency and severity at extreme quantiles.

A significant observation from the study results is the systematic overestimation of the severe crash frequency estimates by all the models. However, this observation may not be a study limitation for the following reason. The number of crashes, specifically in Queensland crash data, is underreported, as the Queensland crash database does not contain property damage only crashes. Thus, the total number of non-severe crashes predicted by the models, which comprises all medical treatment-only, minor injury, and property damage only crashes, should realistically be higher than the observed crash frequencies. Indeed, Watson et al. (2015) found a 68.6% discordance between the injury crash data collected from hospital records and the Queensland crash database for the year 2009 across all levels of crash severity, meaning that there was significant evidence of underreporting of injury crashes in the Queensland road crash database. Assuming this discordance rate to be consistent even for our data and accordingly inflating the observed number of crashes, the new values for observed severe (95% confidence intervals) and observed non-severe crashes (95% confidence intervals) for the combined dataset are 0.674 (0.218, 1.75) and 7.76 (5.55, 10.4), respectively, which are much closer to the estimated numbers. Since the exact amount of underreporting is unknown for the crash dataset utilized in this study, model validation is challenging. The strength of the proposed bivariate extreme value models is that they do not rely on crash data for model estimation and instead model the crash mechanism distribution through conflict measures to predict crash frequencies by severity. Thus, this study's crash estimates could probably be better indicators of the true crash frequencies than the observed crashes in the Queensland crash database.

Compared to TTC, MTTC was found to be a superior conflict indicator for estimating crash frequency-by-severity. The MTTC models in both univariate and bivariate approaches were more precise and had narrower prediction intervals than the TTC models. This result is in line with prior studies on rear-end conflicts at signalized intersections that showed that MTTC is a better indicator of crash frequency than TTC (Zheng and Sayed, 2019a). The superiority of MTTC is due to its utilization of the conflicting vehicles' acceleration information.

The current Delta-V parameter estimation relies on the constant velocity assumption, like in TTC. Laureshyn et al. (2017) have introduced an improved version of Delta-V that utilizes the acceleration information of the conflicting vehicles in estimation; however, it could not be used in this study due to the unavailability of information regarding the relationship between the Extended Delta-V indicator and crash severity probability. Thus, the Extended Delta-V for joint frequency and severity estimation is an important future research direction that can improve the models' accuracy and precision developed in this study.

Another significant topic concerning the Delta-V indicator is the validity of the severe (maximum abbreviated injury scale ≥ 3 or MAIS3 +) crash threshold used in this study. This threshold of 16.0 m/s was derived from the relationship specified by Bahouth et al. (2014) (Eq. (6)) with an assumption that the values of the variables corresponding to multiple crash impact points and ages of the vehicle occupants were 0 since data for such variables cannot be obtained from video-based conflict observation process. Given that Bahouth et al. (2014) did not explicitly investigate the presence of multicollinearity in their models, it is possible that ignoring these variables might bias the estimation of the Delta-V severe crash threshold. To illustrate this point, the proportions of Queensland driving license holders in the age groups 55–74 (27.58%) and 75+ (6.27%) were calculated for the year 2019 from the official data (Queensland Government, 2020)) and substituted as surrogates for the vehicle occupants' age-related variables in the Bahouth et al. rear-end crash severity model to capture the possible average effect of these variables on the Delta-V threshold. Correspondingly, the Delta-V severe crash threshold was computed as 15.84 m/s, which is roughly equivalent to a + 10% error in the model coefficients in Eq. (6). Therefore, a sensitivity analysis test was conducted by introducing errors of ± 5% and ± 10% to the model coefficients in Eq. (6) and re-estimate the peak-over-threshold models. Only the best performing $\text{Log}_{\text{MTTC, Delta-V}}$ model was re-estimated for this analysis and the results of the analysis are shown in Fig. 10. A ± 10% error in the model coefficients results in approximately 10% variation in the predicted severe crashes in the opposite direction. Thus, an added error of 10% to the model coefficients resulted in a threshold of 15.8 m/s and predicted average severe crash frequency of 0.691 – approximately 10% lower than the predicted frequency 0.769 corresponding to 16 m/s threshold. The upper and lower bounds of the prediction interval of severe crashes also did not vary significantly in the analysis. Significantly, the predicted severe crash frequencies in the sensitivity analysis are still all higher than the observed severe crash frequency, highlighting the presence of bias and discordance in the Queensland Crash data discussed above. Thus, the study results do not seem to be affected significantly by variables omitted in Eq. (6).

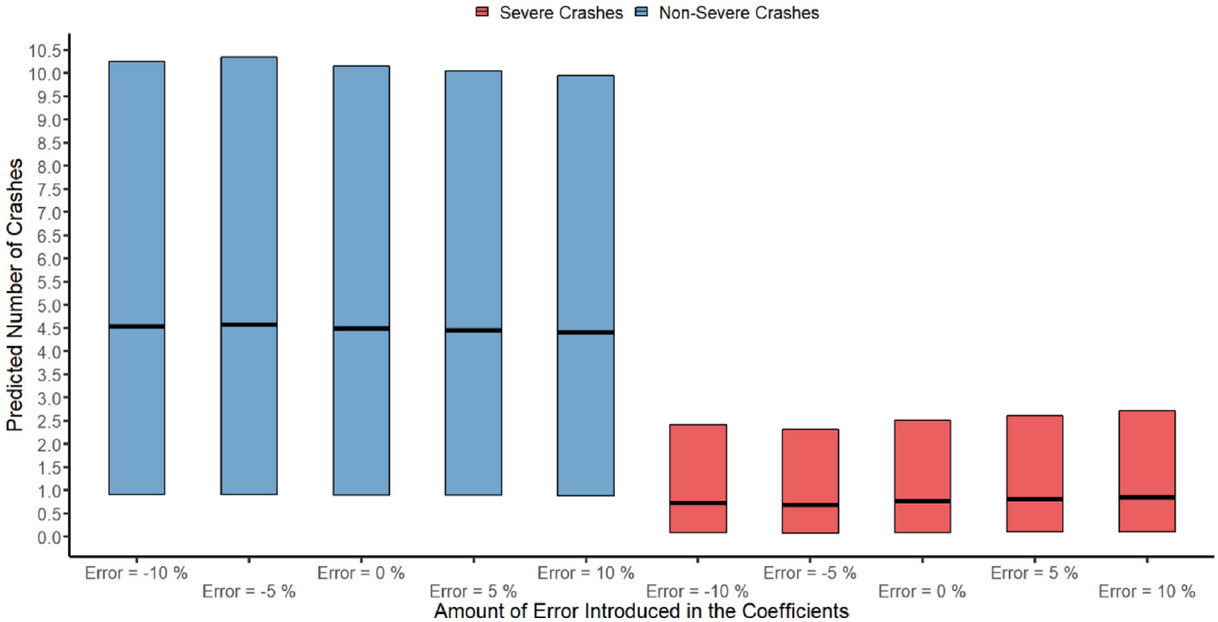

Fig. 10. Sensitivity Analysis of Delta-V crash threshold on predicted severe and non-severe crashes considering ±5% and ±10% errors. The black horizontal bar in the middle denotes the mean crashes; the lengths of the vertical bars denote the 95% confidence intervals.

From Table 6, the Kessels Rd-Logan Rd intersection model was more accurate and precise than the Beaudesert Rd-Granard Rd intersection model. However, when the datasets were combined, the sharing of information between the two similar sites substantially improved the accuracy of the combined model. Investigating the results further by obtaining the site-wise predictions from the combined model, the Beaudesert Rd-Granard Rd intersection predictions were much more accurate than those from the corresponding at-site model. Thus, combining the dataset resulted in the model borrowing the additional pertinent information available from the Kessels Rd-Logan Rd intersection data to improve the overall accuracy and precision of the estimates. This result is significant and in line with the observations from Zheng et al. (2019a), who, through a Bayesian hierarchical analysis of traffic conflicts, also concluded that combining the data from similar sites generally leads to much better crash predictions from the extreme value models.

This study underscores the importance of observing conflicts for longer durations. Compared to previous studies employing bivariate extreme value theory for crash prediction from conflict measures (Zheng et al., 2018a, Zheng and Sayed, 2019d, Zheng et al., 2019b), the crash estimates obtained in this study are generally more precise with narrower prediction intervals. The main reason is that 24 h of conflict observations were analyzed (12 h each for two days), which is significantly higher than previous studies where the observation periods were typically a few hours. Note that the prediction intervals of the crash estimates in the study by Zheng and Sayed (2019d), where they developed Bivariate peak-over-threshold models using Time-to-Collision and Post-Encroachment Time conflicts observed for 15–17 h, are roughly in the range of [0,11) comparable to the prediction intervals estimated in this study.

Although the bivariate models in this study have shown some promising results, the prediction intervals are still wider than the confidence intervals of the observed number of crashes. Larger sample sizes may improve this prediction performance. Conflict observations for more extended periods, such as 24 h for 7 days, may yield an adequate number of conflicts of various types over a typical week during day and night and improve model prediction. Besides, data from similar transport facilities may be combined to improve the models' estimation precision. Moreover, in this study, the conflict data were assumed to be time stationary, which is not ordinarily the case, probably causing the imprecision of the estimates. An issue with the crash data used for validation in this study is the low number of severe (maximum abbreviated injury scale $\geq$ 3 or MAIS3 +) crashes. Although it is not expected to affect the extreme value theory-based crash estimation approach given that crash data is not used for model development, the low sample means issue has not been explicitly investigated in this context and presents another future research opportunity.

9. Conclusions and recommendations

Many conclusions and recommendations arise from the findings of this study:

1. This research should be repeated at other locations to build momentum for this line of research inquiry and test the generalizability of these results.

2. Estimating crash frequency by severity levels will significantly improve the applications of conflict-based safety assessments, leading to a better understanding of critical conflicts that lead to serious injuries.
3. Based on this study's results, it was found that Modified Time to Collision (MTTC) is a better indicator of crash frequency than Time to Collision (TTC) and should be preferred over the latter for estimation of crash frequency-by-severity along with the severity indicator (Delta-V).
4. Future research should consider the non-stationarity that is ordinarily present in conflict data to further improve the joint frequency and severity crash-conflict models.
5. Multi-site models using combined conflict data are vital to advancing the practice of conflict-based crash estimation. Adopting a Bayesian hierarchical approach demonstrated by Zheng et al. (2019a) could be a way forward for developing such models. The accumulation of conflict data should serve to build and improve conflict-based models for the road safety community and profession
6. This modeling work contributes to an increasing body of evidence focused on and building the case for proactive safety management methods—which should replace reactive methods moving forward.
7. This study automated the entire process of estimating crash frequency-by-severity using Computer Vision methods. This automation capability will gain higher significance in the impending era of Connected and Automated Vehicles in the traffic streams. It can provide the road authorities with tools to manage the safety of their facilities in real-time by leveraging vehicle-to-everything (V2X) communication technologies.

References

Alhajyaseen, W.K.M., 2014. The integration of conflict probability and severity for the safety assessment of intersections. Arabian Journal for Science and Engineering 40 (2), 421–430.

Arun, A., Haque, M. M., Washington, S., Sayed, T. and Mannering, F. L. 2021a. A Systematic Review of Conflict-Based Safety Measures with a Focus on Application Context. Analytic Methods in Accident Research, under review.

Arun, A., Haque, M.M., Bhaskar, A., Washington, S., Sayed, T., 2021b. A systematic mapping review of surrogate safety assessment using traffic conflict techniques. Accident Analysis and Prevention 153, 106016.

Bagdadi, O., 2013. Estimation of the severity of safety critical events. Accident Analysis and Prevention 50, 167–174.

Bahouth, G., Graygo, J., Digges, K., Schulman, C., Baur, P., 2014. The benefits and tradeoffs for varied high-severity injury risk thresholds for advanced automatic crash notification systems. Traffic Injury and Prevention 15 (Suppl 1), S134–S140.

Beirlant, J., Goegebeur, Y., Segers, J., Teugels, J., 2004. Statistics of extremes: Theory and applications. Wiltshire, UK, John Wiley and Sons Ltd.

Cavadas, J., Azevedo, C.L., Farah, H., Ferreira, A., 2020. Road safety of passing maneuvers: A bivariate extreme value theory approach under non-stationary conditions. Accident Analysis and Prevention 134, 105315.

Coles, S., 2001. An introduction to Statistical Modeling of Extreme Values. UK, Springer-Verlag, London, London.

Davis, G.A., Hourdos, J., Xiong, H., Chatterjee, I., 2011. Outline for a causal model of traffic conflicts and crashes. Accident Analysis and Prevention 43 (6), 1907–1919.

de Haan, L., de Ronde, J., 1998. Sea and wind: Multivariate extremes at work. Extremes 1 (1), 7–45.

Dutfoy, A., Parey, S., Roche, N., 2014. Multivariate extreme value theory – A tutorial with applications to hydrology and meteorology. Dependence Modeling 2, 30–48.

Evans, L., 1994. Driver injury and fatality risk in two-car crashes versus mass ratio inferred using Newtonian mechanics. Accident Analysis and Prevention 26 (5), 609–616.

Fu, C., Sayed, T., Zheng, L., 2020. Multivariate Bayesian hierarchical modeling of the non-stationary traffic conflict extremes for crash estimation. Analytic Methods in Accident Research 28, 100135.

Fu, C., Sayed, T., 2021. Comparison of threshold determination methods for the deceleration rate to avoid a crash (DRAC)-based crash estimation. Accident Analysis and Prevention 153, 106051.

Henley, G., Harrison, J., 2016. Trends in Serious Injury Due to Road Vehicle Traffic Crashes, Australia: 2001 to 2010. Australian Institute of Health and Welfare, Canberra.

Hydén, C., 1987. The Development of a Method of Traffic Safety Evaluations: The Swedish Traffic Conflicts Technique. Lund University, PhD.

Jonasson, J.K., Rootzen, H., 2014. Internal validation of near-crashes in naturalistic driving studies: a continuous and multivariate approach. Accident Analysis and Prevention 62, 102–109.

Laureshyn, A., De Ceunynck, T., Karlsson, C., Svensson, A., Daniels, S., 2017. In search of the severity dimension of traffic events: Extended Delta-V as a traffic conflict indicator. Accident Analysis and Prevention 98, 46–56.

Minister for Transport and Main Roads. 30 August, 2018. Finding out what makes people click in regional Queensland [Online]. Available: https://statements.qld.gov.au/statements/85424 [Accessed 19/04/2021].

Ozbay, K., Yang, H., Bartin, B., Mudigonda, S., 2008. Derivation and validation of new simulation-based surrogate safety measure. Transportation Research Record 2083, 105–113.

Queensland Government. 2020. Licensing statistics [Online]. Brisbane, Queensland: Department of Transport and Main Roads. Available: https://www.tmr.qld.gov.au/safety/transport-and-road-statistics/licensing-statistics.aspx [Accessed 19/04/2021].

Saunier, N., Sayed, T., 2008. Probabilistic framework for automated analysis of exposure to road collisions. Transportation Research Record 2083, 96–104.

Savolainen, P.T., Mannering, F.L., Lord, D., Quddus, M.A., 2011. The statistical analysis of highway crash-injury severities: a review and assessment of methodological alternatives. Accident Analysis and Prevention 43, 1666–1676.

Sharma, A., Zheng, Z., Kim, J., Bhaskar, A., Haque, M.M., 2019. Estimating and comparing response times in traditional and connected environments. Transportation Research Record 2673, 674–684.

Shelby, S. G. Delta-V as a measure of traffic conflict severity. Transportation Research Board 90th Annual Meeting, 23–27 January 2011 2011 Washington DC, USA.

Smith, R.L., Weissman, I., 1994. Estimating the Extremal Index. Journal of the Royal Statistical Society. Series B (Methodological) 56 (3), 515–528.

Sobhani, A., Young, W., Logan, D., Bahrololoom, S., 2011. A kinetic energy model of two-vehicle crash injury severity. Accident Analysis and Prevention 43, 741–754.

Songchitruksa, P., Tarko, A.P., 2006. The extreme value theory approach to safety estimation. Accident Analysis and Prevention 38 (4), 811–822.

St-Aubin, P., Saunier, N., Miranda-Moreno, L., 2015. Large-scale automated proactive road safety analysis using video data. Transportation Research Part C 58, 363–379.

Stephenson, A.G., 2002. Evd: Extreme value distributions. R News 2 (2), 31–32.

Tarko, A.P., 2012. Use of crash surrogates and exceedance statistics to estimate road safety. Accident Analysis and Prevention 45, 230–240.

Tarko, A.P., 2018. Estimating the expected number of crashes with traffic conflicts and the Lomax Distribution - A theoretical and numerical exploration. Accident Analysis and Prevention 113, 63–73.
Wang, C., Stamatiadis, N., 2014. Derivation of a new surrogate measure of crash severity. Transportation Research Record 2432, 37–45.
Wang, C., Xu, C., Dai, Y., 2019. A crash prediction method based on bivariate extreme value theory and video-based vehicle trajectory data. Accident Analysis and Prevention 123, 365–373.
Watson, A., Watson, B., Vallmuur, K., 2015. Estimating under-reporting of road crash injuries to police using multiple linked data collections. Accident Analysis and Prevention 83, 18–25.
Zheng, L., Ismail, K., Meng, X., 2014a. Freeway safety estimation using extreme value theory approaches: a comparative study. Accident Analysis and Prevention 62, 32–41.
Zheng, L., Ismail, K., Meng, X., 2014b. Traffic conflict techniques for road safety analysis: open questions and some insights. Canadian Journal of Civil Engineering 41 (7), 633–641.
Zheng, L., Ismail, K., Sayed, T., Fatema, T., 2018a. Bivariate extreme value modeling for road safety estimation. Accident Analysis and Prevention 120, 83–91.
Zheng, L., Sayed, T., Tageldin, A., 2018b. Before-after safety analysis using extreme value theory: A case of left-turn bay extension. Accident Analysis and Prevention 121, 258–267.
Zheng, L., Sayed, T., 2019a. Comparison of traffic conflict indicators for crash estimation using peak over threshold approach. Transportation Research Record 2673, 493–502.
Zheng, L., Sayed, T., 2019b. Application of extreme value theory for before-after road safety analysis. Transportation Research Record 2673, 1001–1010.
Zheng, L., Sayed, T., 2019c. Bayesian hierarchical modeling of traffic conflict extremes for crash estimation: A non-stationary peak over threshold approach. Analytic Methods in Accident Research 24, 100106.
Zheng, L., Sayed, T., 2019d. From univariate to bivariate extreme value models: Approaches to integrate traffic conflict indicators for crash estimation. Transportation Research Part C 103, 211–225.
Zheng, L., Sayed, T., Essa, M., 2019a. Bayesian hierarchical modeling of the non-stationary traffic conflict extremes for crash estimation. Analytic Methods in Accident Research 23, 100100.
Zheng, L., Sayed, T., Essa, M., 2019b. Validating the bivariate extreme value modeling approach for road safety estimation with different traffic conflict indicators. Accident Analysis and Prevention 123, 314–323.
Zheng, L., Sayed, T., Mannering, F., 2021. Modeling traffic conflicts for use in road safety analysis: A review of analytic methods and future directions. Analytic Methods in Accident Research 29, 100142.