Accident Analysis and Prevention 125 (2019) 188–197

Contents lists available at ScienceDirect

Accident Analysis and Prevention

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

Multivariate copula temporal modeling of intersection crash consequence metrics: A joint estimation of injury severity, crash type, vehicle damage and driver error

Kai Wang[a,*], Tanmoy Bhowmik[b], Shamsunnahar Yasmin[b], Shanshan Zhao[a], Naveen Eluru[b], Eric Jackson[a]

[a] Connecticut Transportation Safety Research Center, Connecticut Transportation Institute, University of Connecticut, United States
[b] Department of Civil, Environmental & Construction Engineering, University of Central Florida, United States

ARTICLE INFO

Keywords:
Copula
Intersections
Driver error
Vehicle damage
Crash type
Crash severity
Temporal effects

ABSTRACT

This study employs a copula-based multivariate temporal ordered probit model to simultaneously estimate the four common intersection crash consequence metrics – driver error, crash type, vehicle damage and injury severity – by accounting for potential correlations due to common observed and unobserved factors, while also accommodating the temporal instability of model estimates over time. To this end, a comprehensive literature review of relevant studies was conducted; four different copula model specifications including Frank, Clayton, Joe and Gumbel were estimated to identify the dominant factors contributing to each crash consequence indicator; the temporal effects on model estimates were investigated; the elasticity effects of the independent variables with regard to all four crash consequence indicators were measured to express the magnitude of the effects of an independent variable on the probability change for each level of four indicators; and specific countermeasures were recommended for each of the contributing factors to improve the intersection safety.

The model goodness-of-fit illustrates that the Joe copula model with the parameterized copula parameters outperforms the other models, which verifies that the injury severity, crash type, vehicle damage and driver error are significantly correlated due to common observed and unobserved factors and, accounting for their correlations, can lead to more accurate model estimation results. The parameterization of the copula function indicates that their correlation varies among different crashes, including crashes that occurred at stop-controlled intersections, four-leg intersections and crashes which involved drivers younger than 25. The model coefficient estimates indicate that the driver's age, driving under the influence of drugs and alcohol, intersection geometry and control types, and adverse weather and light conditions are the most critical factors contributing to severe crash consequences. The coefficient estimates of four-leg intersections, yield and stop-controlled intersections and adverse weather conditions varied over time, which indicates that the model estimation of crash data may not be stable over time and should be accommodated in crash prediction analysis. In the end, relevant countermeasures corresponding to law enforcement and intersection infrastructure design are recommended to all of the contributing factors identified by the model. It is anticipated that this study can shed light on selecting valid statistical models for crash data analysis, identifying intersection safety issues, and helping develop effective countermeasures to improve intersection safety.

1. Introduction and motivation

In the United States, improving traffic safety by reducing crashes has continuously been a prominent goal of transportation agencies due to the huge societal and economic losses caused by motor vehicle crashes. Statistics show that motor vehicle crashes lead to more than 90 deaths per day (National Highway Traffic Safety Administration (NHTSA, 2015) and $230.6 billion losses per year (Governors Highway Safety Association (GHSA, 2009). Therefore, there is an urgent need to identify the relevant contributing factors of varied crash outcomes and to implement effective safety strategies and countermeasures to reduce crash occurrence and minimize crash consequences.

* Corresponding author.
E-mail addresses: kai.wang@uconn.edu (K. Wang), tanmoy78@Knights.ucf.edu (T. Bhowmik), shamsunnahar.yasmin@ucf.edu (S. Yasmin), shanshan.h.zhao@uconn.edu (S. Zhao), naveen.eluru@ucf.edu (N. Eluru), eric.d.jackson@uconn.edu (E. Jackson).

https://doi.org/10.1016/j.aap.2019.01.036
Received 12 October 2018; Received in revised form 28 January 2019; Accepted 29 January 2019
Available online 13 February 2019

In the past few decades, injury severity - determined based on the victim's responses, descriptions, and complaints, and the police officer's judgment after the crash has been considered as one of the most critical crash outcome indicators/metrics for motor vehicle crashes and has been modeled to identify the effects of contributing factors on driver injuries (Wang et al., 2015). One objective indicator – the extent of vehicle damage based on the destruction/deformation of the vehicle involved in the crash – has also been used to represent the crash consequence as a supplement of injury severity (Wang et al., 2015; Wang and Qin, 2014; Qin et al., 2013a, b). Furthermore, due to the expectation that specific collision types, such as head-on and sideswipe, have significant impacts on the injury level suffered by drivers or passengers (Yasmin et al., 2014a, 2014b), researchers have explored the usage of crash type as the crash outcome. Despite the different crash indicators used, it is well acknowledged that driver error is a key contributor to crashes and fatalities especially for intersections, due to the fact that approximately 90% of motor vehicle crashes are caused by driver errors (Wang and Qin, 2015). In addition, driver error is highly correlated with other crash contributing factors such as human characteristics, roadway characteristics, environmental factors and vehicle factors *etc*., therefore reducing the driver error by implementing countermeasures related to those factors can help mitigate preventable crashes. Considering the significant associations between driver error and injury severity, driver error has been increasingly modelled to identify the crash contributing factors, and used along with other three indicators to investigate the crash consequences and factors leading to crash occurrence (Wang and Qin, 2015). Although previous studies have been implemented to estimate one or two of these four metrics, no research has focused on simultaneously modelling them and exploring their complicated interrelationships due to the common unobserved attributes in crash data. In this study, we propose a multivariate copula approach to simultaneously model the four crash indicators – injury severity, crash type, vehicle damage and driver error – to identify the factors contributing to intersection crash consequences and explore the potential correlations among them using different copula formulation and parameterization strategies.

2. Previous research and current study in context

2.1. Injury severity literature

Discrete response models have extensively been employed by researchers to explore the relationship between driver, environmental, roadway, traffic and vehicle attributes and injury severity suffered by drivers/passengers. The multinomial logit (MNL) and multinomial probit (MNP) models are the two most common approaches when injury severity is treated as a non-ordinal indicator (Qin et al., 2013a, 2013b, Dissanayake, 2004; Ye and Lord, 2011; Ghulam et al., 2012; Wu et al., 2013). The ordered logit (OL), ordered probit (OP) and generalized ordered models are widely used when injury severity is treated as an ordinal indicator (Qin et al., 2013a, b; Eluru et al., 2008; Yasmin et al., 2014a, b; Eluru, 2013; Yasmin and Eluru, 2013).

Traditional logit/probit models have assumed constant parameters across all crashes, which might not always be true due to the unobserved heterogeneity in the crash data (Mannering et al., 2016; Savolainen et al., 2011). In order to account for the data heterogeneity, sophisticated statistical models have been developed to estimate injury severity, including the markov switching multinomial logit (MSMNL) model (Malyshkina and Mannering, 2009; Ivan and Konduri, 2018), the random parameter model (Qin et al., 2013a, b;Ye and Lord, 2011; Yasmin and Eluru, 2013; Milton et al., 2008; Kim et al., 2010; Wu et al., 2014; Lee et al., 2018), the latent segmentation model (Yasmin et al., 2014a, b), the finite mixture model (FMM) or latent class clustering (LCC) model (Ma et al., 2016; Depaire et al., 2008; Zhao et al., 2018) and spatio-temporal models (Liu and Sharma, 2018). In order to accommodate the influence of latent/unobserved variables on injury severity and capture the interrelationships among variables when they interact in indirect and complicated ways in crash data, the structural equation modeling (SEM) has also extensively been investigated in injury severity analysis (Wang and Qin, 2014; Ye and Lord, 2011; Hassan and Abdel-Aty, 2012; Hamdar et al., 2008; Lee et al., 2008; Xie et al., 2018). A detailed discussion of numerous discrete response models that are currently being used in injury severity estimation can be referred to in Savolainen et al. (2011) and Washington et al. (2011).

2.2. Vehicle damage literature

As mentioned earlier, the extent of vehicle damage mainly based on the destruction and deformation of the vehicle involved in the crash, which can be easily seen and measured – has been used as a complement to injury severity to characterize the crash outcome. Of these studies, Wang et al. (2015) indicated that the vehicle damage level is significantly associated with the environmental, roadway, traffic and crash attributes. A study conducted by Wang and Qin (2014) verified that the vehicle damage is highly correlated with the vehicle speed when the crash occurred in single-vehicle crashes. Quddus et al. (2002) found the factors that are most likely to increase the vehicle damage level in motorcycle crashes include increased engine capacity, headlight not turned on during daytime, collisions with stationary objects, driving during early morning hours, having a pillion passenger and when the motorcyclist is determined to be at fault for the accident. Huang et al. (2008) created a binary response variable by combining both injury severity and vehicle damage indicators to identify the significant factors affecting crash severity. The study found that crashes that occur in the night time, at T/Y type intersections, and in the right-most lane as well as those that occur at intersections where red light cameras are installed lead to more severe consequences. Although numerous studies succeeded in identifying the primary factors contributing to vehicle damage levels, a limited number of research studies have been conducted to explore the potential protection to vehicle occupants offered by the vehicle design. This may be due to the fact that many modern vehicles are designed to sustain a large percentage of damage to absorb energy to protect vehicle occupants. Qin et al. (2013a,b) found that the vehicle damage is an unbiased indicator of kinetic energy in crashes, and verified that the vehicle deformation in some crashes can serve to absorb or deflect the impact energy caused by collisions and minimize the forces transferred to occupants.

2.3. Crash type literature

Due to the expectation that specific collision types, such as head-on and sideswipe, have significant impacts on the injury level suffered by drivers or passengers (Yasmin and Eluru, 2014), researchers have explored the usage of crash type as the crash outcome and identified the contributing factors on crash type. The research conducted by Yasmin et al. (2014a,b) indicated that the crash type can be associated with driver, vehicle, roadway design and environmental characteristics. Ye et al. (2008) found that the presence of curves, traffic controls, time of day, weather conditions and driving under the influence of alcohol (DUI) are the main contributors to different crash types for two-vehicle crashes. Rana et al. (2010) modeled the crash type using the driver, roadway, environmental, crash and vehicle factors, and found the crash type propensity differs among different crash conditions. Bham et al. (2012) examined the contributing factors to angular, head-on, rear-end and sideswipe crashes, and illustrated that the risks associated with different crash types are significantly varied by different vehicle actions when the crash occurred, including whether the vehicle was changing lanes and merging on undivided and divided highways, travelling on a curve, driving on a wet road surface or avoiding objects.

2.4. Driver error literature

Despite the different crash indicators used, it is well acknowledged that driver error is a key contributor to crashes and fatalities, due to the fact that approximately 90% of motor vehicle crashes are caused by driver errors (Wang and Qin, 2015). In order to effectively reduce the possibility of crash occurrence, research has focused on studying driver errors to help identify countermeasures to mitigate crash consequence. Studies concluded that drivers over 60 years of age are prone to failing to stop or failing to yield the right-of-way at intersections (Keay et al., 2009; Braitman et al., 2007). Female, young, and high-income drivers were found to be more likely to be involved in distracted driving (Zhao and Khattak, 2017). Lord et al. (2007) indicated that older drivers are most likely to drive under the influence due to the medication use and disability. Yang and Najm (2007) found that younger drivers tend to violate red signals more often than others due to their more aggressive and risky driving behavior. A study conducted by Papaioannou (2007) confirmed that female drivers are less likely to violate the yellow signals and are more likely to driver slower than male drivers. Devlin et al. (2011) concluded that the probability of failing to notice traffic signals increases at both vertical and horizontal curves. Bonneson et al. (2003) suggested that heavy vehicles are prone to running the red light at intersections. Wang and Qin (2015) examined the key contributors to driver errors at uncontrolled, sign-controlled and signalized intersections, and provided countermeasures involving engineering, enforcement and education for each type of driver error. Shaon et al. (2018) explored the typical contributors to driver errors on highway segment, which can help researchers and safety professionals to develop cost-effective preventive countermeasures.

2.5. Copula model literature

Recently, extensive attention has been paid to the potential correlations between indicators due to both common observed and unobserved factors, when the indicators are simultaneously modelled. Ignoring the correlations may result in incorrect and biased coefficient estimates (Washington et al., 2011). The copula based model has been increasingly employed in transportation research to address this issue because the copula model can jointly model the indicators while also accounting for their interrelationships through a copula structure. Of these studies applying the copula based model, Pourabdollahi et al. (2013) jointly estimated the choice of freight mode and shipment size. Sener et al. (2010) examined the physical activity participation for all individuals within the same family. Eluru et al. (2010) modelled the injury severities suffered by all occupants involved in a crash. Rana et al. (2010) and Yasmin et al. (2014a, b) simultaneously investigated the crash type and injury severity. Wang et al. (2015) jointly modelled the injury severity and vehicle damage for two-vehicle crashes. Wali et al. (2018) simultaneously estimated the injury severity of at-fault and not-at-fault drivers in head-on collisions. Laman et al. (2018) extended the traditional two-dimensional copula framework to a three-dimensional model, and simultaneously modelled the reporting, response and clearance time after a crash occurred.

2.6. Current study in context

This study attempts to jointly model the four crash consequence metrics – injury severity, crash type, vehicle damage and driver error – and is mainly built upon several previous research studies conducted by the authors. These research studies examined the interrelationships between two of the four indicators. A previous study authored by Yasmin et al. (2014a, b) simultaneously modelled the injury severity and crash type by accounting for their potential correlations. The study verified that the injury severity and crash type are significantly correlated, and the correlation varies among different collisions types. The study conducted by Wang et al. (2015) jointly modelled the injury severity and vehicle damage for two-vehicle crashes. The study illustrated that the injury severity and vehicle damage are highly correlated, and their correlation varies among different crash characteristics, including manners of collision and collision types. Wang and Qin (2015) explored the interrelationships between the injury severity and driver error using both the statistical model and data mining technology, and found severe driver errors, such as reckless driving and disregard of traffic signals, are prone to lead to more severe injury severity.

Considering the varied and complicated interrelationships among the injury severity, crash type, vehicle damage and driver error, it is desirable to propose a joint approach to simultaneously model the four metrics to identify the factors contributing to intersection crash consequences by accounting for their potential correlations due to common observed and unobserved factors. Furthermore, research has extensively drawn attention to investigate the temporal instability of statistical models for crash data analysis over time. A recent study conducted by Mannering (2018) indicates that the temporal instability is likely to exist in crash prediction models using multi-year crash data, and should be accommodated in crash data analysis. To this end, this study presents a multivariate copula modeling with a four-dimensional dependent variable. To accommodate the correlation variations, the error correlation part of the copula model was examined using different copula formulations and parameterization strategies (*i.e.* the dependency parameter is allowed to vary across observations). To account for the temporal instability of model estimation, the model was estimated using not only the independent variables individually, but also the interaction variables between the year and each of the other variables.

The remainder of this paper is organized as follows. The next section presents the multivariate copula model and the estimation approach. The fourth section describes the data collection and preparation. The fifth section compares the model goodness-of-fit and prediction performance. The sixth section analyzes the model estimation results. The safety countermeasures are recommended in the seventh section. The concluding remarks are provided in the final section.

3. Methodology

3.1. Ordered probit model framework

Let i ($i = 1, 2, \ldots, I$) be the index for two-vehicle crashes at intersections, j ($j = 1, 2, \ldots, J$) be the index for the driver error, k ($k = 1, 2, \ldots, K$) be the index for the crash type, p ($p = 1, 2, \ldots, P$) be the index for the vehicle damage, and q ($q = 1, 2, \ldots, Q$) be the index for the injury severity. Using the ordered probit formulation, the driver error level (y_i) at intersection i can be specified with a propensity latent variable (y_i^*) as:

$$y_i^* = \boldsymbol{\alpha}'\boldsymbol{x_i} + \varepsilon_i\,,\;\; y_i = j,\; if\; \tau_{j-1} < y_{ij}^* < \tau_j \tag{1}$$

where $\boldsymbol{x_i}$ is a column vector of independent variables, $\boldsymbol{\alpha}$ is a column vector of parameters to be estimated, and ε_i is a random error term which follows a standard normal distribution. τ_j ($\tau_0 = -\infty\,,\; \tau_J = +\infty$) is the threshold associated with driver error level j, with the ordinal nature as: ($-\infty < \tau_1 < \tau_2 < \ldots < \tau_{J-1} < +\infty$). The resulting probability of the j^{th} driver error level in the i^{th} crash can be written as:

$$\Pr(y_i = j) = \phi(\tau_j - \boldsymbol{\alpha}'\boldsymbol{x_i}) - \phi(\tau_{j-1} - \boldsymbol{\alpha}'\boldsymbol{x_i}) \tag{2}$$

where $\phi(\bullet)$ is the cumulative standard normal distribution.

Similarly, the crash type level (u_i) for crash i can be specified with a propensity latent variable (u_i^*) as:

$$u_i^* = \boldsymbol{\beta}'\boldsymbol{x_i} + \xi_i\,,\;\; u_i = k,\; if\; \psi_{k-1} < u_{ik}^* < \psi_k \tag{3}$$

where $\boldsymbol{\beta}$ is the column vector of parameters to be estimated, ξ_i is also a random error term which follows a standard normal distribution, ψ_k is the threshold related to the k^{th} crash type level. The probability of the k^{th} crash type level can be written as:

$$\Pr(u_i = k) = \Lambda(\psi_k - \boldsymbol{\beta}'\boldsymbol{x_i}) - \Lambda(\psi_{k-1} - \boldsymbol{\beta}'\boldsymbol{x_i}) \tag{4}$$

where $\Lambda(\bullet)$ is the cumulative standard normal distribution.

The injury severity level (m_i) for crash i can be represented with a propensity latent variable (m_i^*) as:

$$m_i^* = \boldsymbol{\gamma}'\boldsymbol{x_i} + \sigma_i\ ,\ \ m_i = q,\ if\ \omega_{q-1} < m_{iq}^* < \omega_q \tag{5}$$

where γ is the column vector of parameters to be estimated, σ_i is a random error term which follows a standard normal distribution, ω_q is the threshold related to the q^{th} injury severity level. The probability of the q^{th} injury severity level can be written as:

$$\Pr(m_i = q) = \Sigma(\omega_q - \boldsymbol{\gamma}'\boldsymbol{x_i}) - \Sigma(\omega_{q-1} - \boldsymbol{\gamma}'\boldsymbol{x_i}) \tag{6}$$

where $\Sigma(\bullet)$ is the cumulative standard normal distribution.

The vehicle damage level (n_i) for crash i can be denoted with a propensity latent variable (n_i^*) as:

$$n_i^* = \boldsymbol{\rho}'\boldsymbol{x_i} + \vartheta_i\ ,\ \ n_i = p,\ if\ \pi_{p-1} < n_{ip}^* < \pi_p \tag{7}$$

where ρ is the row vector of parameters to be estimated, ϑ_i is a random error term which follows a standard normal distribution, π_p is the threshold related to the p^{th} vehicle damage level. The probability of the p^{th} vehicle damage level can be written as:

$$\Pr(n_i = p) = \Pi(\xi_p - \boldsymbol{\rho}'\boldsymbol{x_i}) - \Pi(\xi_{p-1} - \boldsymbol{\rho}'\boldsymbol{x_i}) \tag{8}$$

where $\Pi(\bullet)$ is the cumulative standard normal distribution.

3.2. Multivariate copula model

To simultaneously model the four indicators, the dependency among the four-dimensional dependent variable is accommodated by their error terms (*i.e.* ε_i, ξ_i, σ_i, and ϑ_i) in Eqs. (1,3,5, and 7). The joint probability function of involving the j^{th} driver error level, k^{th} crash type level, q^{th} injury severity level and p^{th} vehicle damage level for crash i can be expressed as (Laman et al., 2018):

$$\Pr(y_i = j,\ u_i = k,\ m_i = q,\ n_i = p\,) = Pr\begin{Bmatrix} [(\tau_{j-1} - \boldsymbol{\alpha}'\boldsymbol{x_i} < \varepsilon_i < \tau_j - \boldsymbol{\alpha}'\boldsymbol{x_i})], \\ [(\psi_{k-1} - \boldsymbol{\beta}'\boldsymbol{x_i} < \xi_i < \psi_k - \boldsymbol{\beta}'\boldsymbol{x_i})], \\ [(\omega_{q-1} - \boldsymbol{\gamma}'\boldsymbol{x_i} < \sigma_i < \omega_q - \boldsymbol{\gamma}'\boldsymbol{x_i})], \\ [(\pi_{p-1} - \boldsymbol{\rho}'\boldsymbol{x_i} < \vartheta_i < \pi_p - \boldsymbol{\rho}'\boldsymbol{x_i})] \end{Bmatrix} \tag{9}$$

The Eq. (9) can be written as (Laman et al., 2018):

$$\Pr(y_i = j,\ u_i = k,\ m_i = q,\ n_i = p\,) = \sum_{a=1}^{2}\sum_{b=1}^{2}\sum_{c=1}^{2}\sum_{d=1}^{2}(-1)^{a+b+c+d} \left[Pr\begin{pmatrix} \varepsilon_i < \tau_{j+a-1} - \boldsymbol{\alpha}'\boldsymbol{x_i},\ \xi_i < \psi_{k+b-1} - \boldsymbol{\beta}'\boldsymbol{x_i}, \\ \sigma_i < \omega_{q+c-1} - \boldsymbol{\gamma}'\boldsymbol{x_i},\ \vartheta_i < \pi_{p+d-1} - \boldsymbol{\rho}'\boldsymbol{x_i} \end{pmatrix}\right] \tag{10}$$

The copula is a function to link the dependency among multiple variables by generating a multivariate distribution, and can be defined as (Bhat and Eluru, 2009):

$$C_\theta(\eta_1, \eta_2, \eta_3, ..., \eta_t) = Pr(H_1 < \eta_1, H_2 < \eta_2, H_3 < \eta_3,\ ..., H_T < \eta_t\,) \tag{11}$$

where θ is a copula parameter representing the dependency among indicators.

The Eq. (10) can be written as (Laman et al., 2018):

$$\Pr(y_i = j,\ u_i = k,\ m_i = q,\ n_i = p\,) = \sum_{a=1}^{2}\sum_{b=1}^{2}\sum_{c=1}^{2}\sum_{d=1}^{2}(-1)^{a+b+c+d}[C_{\theta_i}(\eta_{j+a-1}, \eta_{k+b-1}, \eta_{q+c-1}, \eta_{p+d-1})] \tag{12}$$

In order to account for the correlation variations of the four indicators, in this study, the copula parameter θ_i is allowed to vary among different crashes and is estimated using the independent variables:

$$C_{\theta_i} = f_g(\boldsymbol{\zeta}'\boldsymbol{x_i}) \tag{13}$$

where ζ is the column vector of coefficients for the copula parameters to be estimated, and f_g is the functional form of copula structure. In this study, four commonly used Archimedean copula structures are tested. They are Frank, Clayton, Joe and Gumbel copulas (Bhat and Eluru, 2009; Yasmin et al. (2014a); (2014b)). Based on the permissible ranges of the dependency parameter in the copula model, different functional forms are assumed for the parameterization of the four copula structures in the analysis. The functional forms for them are assumed as $f_{Frank/Clayton} = exp(\boldsymbol{\zeta}'\boldsymbol{x_i})$ and $f_{Joe/Gumbel} = exp(\boldsymbol{\zeta}'\boldsymbol{x_i}) + 1$. Details about these copula structures can be found in Bhat and Eluru (2009); Sener et al. (2010); Eluru et al. (2010) and Yasmin et al. (2014a); (2014b).

Overall, the parameters to be estimated in the copula model are $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$, $\boldsymbol{\gamma}$, $\boldsymbol{\rho}$, τ_j, ψ_k, ω_q, π_p and $\boldsymbol{\zeta}$. The likelihood function can be expressed as:

$$L = \prod_{i=1}^{I}\left\{\prod_{j=1}^{J}\prod_{k=1}^{K}\prod_{q=1}^{Q}\prod_{p=1}^{P}[\Pr(y_i = j,\ u_i = k,\ m_i = q,\ n_i = p\,)]^{\varrho_{jkqp}}\right\} \tag{14}$$

where ϱ_{jkqp} is a dummy indicator variable with value 1 if the crash i has the j^{th} level of driver error, k^{th} level of crash type, q^{th} level of injury severity and p^{th} level of vehicle damage, and 0 otherwise. The coefficients are estimated by maximizing the likelihood function of Eq. (14) through the GAUSS programming package. A detailed discussion of the copula model estimation routines is available in Bhat and Eluru (2009).

4. Data preparation and analysis

In this study, the copula model was estimated using the two-year (2016–2017) intersection crash data collected from the Connecticut Crash Data Repository Connecticut Crash Data Repository (CTCDR), (2018) and only two-vehicle crashes were considered. After removing those with missing attributes, 20,917 intersection crashes were used in this study, of which 10,542 crashes occurred in 2016, and 10,375 occurred in 2017. With regard to the model estimation, eighty percent of crashes were randomly selected as the estimation datasets, which were used in estimating the model coefficients; the remaining twenty percent crashes were treated as the validation datasets, which were used to evaluate the model prediction performance. Table 1 presents the summary statistics of the variables selected in estimating the copula model. The last four rows of the table include the four crash consequence indicators which are treated as the four dependent variables in the copula model. To obtain sufficient observations in each level of different indicators, injury severities were categorized into three levels. They are **1)** PDO (property damage only) crashes; **2)** B (suspected minor injury) or C (possible injury) crashes; and **3)** A (suspected serious injury) or K (fatal) crashes. Crash types were also categorized into three ordinal levels based on the original travel direction of involved vehicles and the collision angle (Wang et al., 2017, 2018). They are **1)** same-direction crash (SDC) type, which includes turning-same direction crashes, sideswipe-same direction crashes and rear-end crashes; **2)** intersecting-direction crash (IDC) type, which includes turning-intersecting crashes and angle crashes; and **3)** opposite-direction crash (ODC) type, which includes turning-opposite direction crashes, sideswipe-opposite direction crashes and head-on crashes. According to the Model Minimum Uniform Crash Criteria (MMUCC) guideline (2017), vehicle damage levels were categorized into three levels, *i.e.* **1)** none (no damage) or minor (cosmetic damage) damage; **2)** moderate (broken or missing

Table 1
Description of Crash Data.

Variable Name	Variable Type	Description	Percentage
Driver Characteristics			
Age	Independent Variable	Younger Driver (Age < 25)	17.4%
		Middle-Aged Driver (Age 25-55)	57.1%
		Older Driver (Age > 55)	25.5%
Gender	Independent Variable	Female	46.6%
		Male	53.4%
DUI	Independent Variable	Drugs or Alcohol Involved	2.1%
Highway and Traffic Characteristics			
Crash Year	Independent Variable	2016	50.4%
		2017	49.6%
Time	Independent Variable	Afternoon Peak (4:00pm-6:59pm)	23.3%
		Night Time (7:00pm-6:59am)	15.7%
		Morning Peak (7:00am-9:59am)	35.7%
		Day Time (10:00am-3:59pm)	25.3%
Intersection Type	Independent Variable	Four-Leg Intersection	50.3%
		T Intersection	44.4%
		Y Intersection	3.8%
		Five- or More-Leg Intersection	1.4%
Traffic Control	Independent Variable	None/Yield	26.3%
		Stop Sign	7.1%
		Signal	66.6%
Speed Limit	Independent Variable	Low (< 35mph)	44.5%
		Middle (35mph - 45mph)	52.9%
		High (> 45mph)	2.6%
Horizontal	Independent Variable	Horizontal Curve Presence	4.9%
Vertical	Independent Variable	Vertical Curve Presence	17.2%
Environmental Characteristics			
Weather	Independent Variable	Clear	2.6%
		Cloudy/Windy	82.1%
		Rain	5.8%
		Snow	9.6%
Light	Independent Variable	Day	80.0%
		Night - With Lighting	17.4%
		Night - Without Lighting	2.6%
Road	Independent Variable	Dry	82.5%
		Wet	14.5%
		Snow/Slush	2.6%
		Ice	0.4%
Vehicle Characteristics			
Vehicle Type	Independent Variable	Passenger Car	89.7%
		Light Truck	9.5%
		Heavier Truck	0.9%
Crash Consequence Indicators			
Driver Error	Dependent Variable	No/Minor Error	40.2%
		Improper Driving Error	3.2%
		Careless Driving Error	15.7%
		Reckless Driving Error	40.9%
Crash Type	Dependent Variable	Same-Direction Crash (SDC)	61.5%
		Intersecting-Direction Crash (IDC)	35.4%
		Opposite-Direction Crash (ODC)	3.1%
Injury Severity	Dependent Variable	PDO Crash	74.0%
		B + C Crash	24.8%
		K + A Crash	1.2%
Vehicle Damage	Dependent Variable	None + Minor Damage	41.7%
		Moderate Damage	23.8%
		Disabling Damage	34.5%

parts) damage; and **3)** disabling (salvageable or total loss) damage. Referring to the study conducted by Wang and Qin (2015), driver errors were categorized into four ordinal levels based on the severity of violation. They are **1)** no or minor errors; **2)** improper driving errors which are traffic infractions that are punishable by a fine of no more than $500, including improper overtaking, improper turning, or driving too fast for the road conditions; **3)** careless driving errors which are normally defined as unintentionally operating a vehicle in an offensive and negligent manner, including following too close, failure to keep the vehicle under control, inattentive driving, left of center or unsafe backing; and **4)** reckless driving errors which are usually defined as intentionally break the traffic rules, including disregard of traffic controls, failure to yield and exceeding the speed limit.

The upper panel of Table 1 lists the contributing factors selected from the crash data. The factors were categorized into four groups. Driver characteristics include the age and gender of the driver, and whether the driver was driving under the influence of alcohol or drugs. Highway and traffic characteristics include the crash occurrence time and the intersection design related geometries. Environmental characteristics include the weather, light and road surface conditions when the crash occurred. Vehicle characteristics include the type of vehicle involved in the crash.

5. Model comparison and selection

The empirical analysis involves estimation of models in two different stages. In the first stage, we estimated independent models by considering the un-pooled sample and pooled sample. The un-pooled models were estimated by considering the data sample for different years separately, while the pooled models were estimated by considering data sample for both years together. Further, in the pooled independent models, interactions of exogenous variables with the year indicator dummy were utilized to control the effects of year varying variable. Then we compared the performance of these independent models based on the Bayesian Information Criterion (BIC). The BIC (log-likelihood, number of parameters) values at convergence for the un-pooled and pooled independent models are 113,449.56 (-56,163.58, 116) and 113,031.55 (-56,201.31, 65), respectively. The comparison exercise clearly highlights the superiority of the pooled independent models in the current study context. In the second stage of the empirical analysis, we estimated joint models building on the pooled independent models by employing the copula structure. For simplicity, in the following sections, we referred the pooled independent model as the independent model. As mentioned earlier, four types of copula models were estimated and compared with the independent model. In order to account for the variations of the correlations among the four crash consequence indicators, each copula model was estimated using both a constant and a parameterized copula structure. The upper panel of Table 2 presents the performance of both the independent model and four copula models using the estimation datasets. The Joe copula model with parameterized copula structure outperforms the other models based on the goodness-of-fit (with the highest Log-Likelihood value, and the lowest AIC and BIC values). The lower panel of Table 2 shows the comparison of model prediction performance between the independent model and the Joe copula model with parameterized copula structure, using the validation datasets. It verifies that the parameterized Joe copula model favors the independent model in terms of the prediction accuracy of the four indicators. These findings indicate that the injury severity, crash type, vehicle damage and driver error are significantly correlated due to the common observed and unobserved factors, and their correlation varies among different crashes. Accounting for their correlations can lead to more accurate model estimation results. The parameterized Joe copula model was selected to simultaneously estimate the four crash consequence indicators in this study.

Table 2
Model Performance Comparisons.

Model Estimation		Log-Likelihood	No. of Parameters	AIC	BIC
Independent Model		−56,201	65	112,533	113,032
Frank Copula	Without Parameterization	−56,175	66	112,482	112,988
	With Parameterization	−56,119	69	112,377	112,907
Gumbel Copula	Without Parameterization	−56,150	66	112,432	112,938
	With Parameterization	−56,093	69	112,323	112,853
Clayton Copula	Without Parameterization	−56,195	66	112,522	113,028
	With Parameterization	−56,167	68	112,469	112,991
Joe Copula	Without Parameterization	−56,145	66	112,422	112,929
	With Parameterization	**−56,084**	**69**	**112,306**	**112,836**
Model Validation		**Predicted Log-Likelihood**		**Predicted BIC**	
Independent Model		−17,644		35,842	
Joe Copula Model with Parameterization		**−17,616**		**35,820**	

Table 3
Joe Copula Model Estimation Results.

Variables		Driver Error		Crash Type		Vehicle Damage		Injury Severity	
		Coeff.	t-stat.	Coeff.	t-stat.	Coeff.	t-stat.	Coeff.	t-stat.
Driver Characteristics									
Age	Younger Driver (Age < 25)	Base Level		Base Level		NA		Base Level	
	Middle-Aged Driver (Age 25-55)	0.02	0.95	**0.13**	**4.57**	NA		**0.16**	**5.47**
	Older Driver (Age > 55)	**0.05**	**2.04**	**0.09**	**3.23**	NA		0.03	0.85
Gender	Female	Base Level		Base Level		NA		Base Level	
	Male	**−0.06**	**−2.66**	**0.05**	**2.19**	NA		−0.01	−0.47
DUI	Drugs or Alcohol Involved	**0.11**	**1.45**	**0.15**	**1.93**	NA		**0.56**	**7.98**
Highway and Traffic Characteristics									
Intersection Type	Y and Five- or More-Leg Intersection	Base Level		Base Level		Base Level		Base Level	
	Four-Leg Intersection	**−0.12**	**−5.02**	**0.32**	**6.05**	**0.24**	**5.01**	**0.20**	**3.78**
	Four-Leg Intersection * Year (2016)	0.00	−0.09	**0.12**	**3.24**	0.03	0.75	0.02	0.49
	T Intersection	−0.05	−1.08	**0.25**	**5.31**	**0.23**	**4.77**	**0.15**	**2.79**
Traffic Control	Signal	Base Level		Base Level		Base Level		Base Level	
	None/Yield	**−0.29**	**−8.20**	**0.34**	**9.12**	**0.44**	**17.26**	**0.20**	**7.52**
	None/Yield * Year (2016)	**−0.08**	**−1.77**	**0.10**	**1.97**	0.02	0.59	−0.01	−0.25
	Stop Sign	**0.25**	**4.73**	**0.58**	**12.67**	**0.22**	**5.64**	0.07	1.16
	Stop Sign * Year (2016)	**−0.11**	**−1.72**	0.02	0.23	0.08	1.16	−0.02	−0.28
Speed Limit	Low (< 35mph)	Base Level		Base Level		Base Level		Base Level	
	Middle (35mph-45mph) and High (> 45mph)	**0.18**	**8.40**	**−0.13**	**−5.41**	**0.19**	**9.16**	**0.08**	**3.61**
Vertical	Vertical Curve Presence	**0.11**	**3.93**	**−0.07**	**−2.30**	0.01	0.37	0.02	0.56
Environmental Characteristics									
Weather	Clear	Base Level		Base Level		Base Level		Base Level	
	Cloudy/Windy	**−0.37**	**−5.65**	**−0.19**	**−2.95**	**0.10**	**1.68**	**0.27**	**3.58**
	Cloudy/Windy * Year (2016)	**0.09**	**3.18**	**−0.12**	**−3.40**	−0.03	−0.96	−0.01	−0.19
	Rain	**−0.29**	**−3.79**	**−0.31**	**−3.87**	**0.14**	**1.80**	**0.32**	**3.67**
	Snow	**−0.25**	**−3.53**	**−0.21**	**−2.87**	**0.18**	**2.50**	**0.22**	**2.67**
Light	Day	Base Level		Base Level		Base Level		Base Level	
	Night - With/Without Lighting	**−0.22**	**−8.23**	**0.42**	**11.53**	**0.27**	**10.54**	**0.17**	**6.13**
	Night - With/Without Lighting * Year (2016)	0.06	1.62	**−0.14**	**−2.96**	−0.08	−1.38	0.08	1.49
Vehicle Characteristics									
Vehicle Type	Light Truck and Heavier Truck	Base Level		Base Level		Base Level		Base Level	
	Passenger Car	**0.12**	**3.71**	0.01	0.44	**−0.11**	**−3.43**	**0.13**	**3.43**
Threshold	*u1*	**−0.53**	**−7.01**	**0.57**	**7.10**	**0.33**	**3.93**	**1.34**	**13.65**
	u2	**−0.44**	**−5.89**	**2.19**	**27.00**	**0.96**	**11.40**	**2.95**	**28.92**
	u3	**−0.04**	**−0.51**	NA		NA		NA	
Copula Parameters	**Coeff. Estimates**	**S.E.**	**t-Stat.**	**P > |t|**					
Constant	**−4.20**	**0.20**	**−21.51**	**< 0.01**					
Stop Sign	**2.15**	**0.17**	**12.59**	**< 0.01**					
Four-Leg Intersection	**0.81**	**0.18**	**4.57**	**< 0.01**					
Younger Driver (Age < 25)	**1.01**	**0.16**	**6.14**	**< 0.01**					

Notes: LL=–56,084; No. of significant parameters = 69; AIC = 112,306; BIC = 112,836; "NA" represents "not applicable"; Bold coefficients are statistically significant at the 5% level of significance.

6. Model estimation results

Table 3 shows the coefficient estimation results of the Joe Copula model. The upper panel of the table presents the estimated coefficients for the four indicators. The middle panel of the table presents the estimated parameters for the copula structure. Bold coefficients are statistically significant at the 5% level of significance. A positive coefficient represents a propensity to increase the severity level of each indicator and vice-versa for a negative coefficient.

6.1. Driver error

Driver characteristics and behaviors are important factors contributing to the severity of driver errors, and driver error is highly associated with driver, highway and traffic, environmental and vehicle characteristics. Reducing the driver error by implementing countermeasures related to those factors may help traffic engineers and agencies mitigate preventable crashes. Older drivers are more prone to make severe driver errors than the other age groups. This indicates that older drivers usually have challenges to make correct decisions at intersections, due to their deteriorating vision, slower recovery from glare and misjudging of gap and speed of other vehicles (Wang and Qin, 2015). Males are less likely to be reckless drivers than females at intersections. This may be because females are more prone to making performance errors during the driving, such as failure to yield the right of way and failure to keep vehicle under control (Shaon et al., 2018). As expected, the severity of driver errors dramatically increases if a driver is driving under the influence of drugs or alcohol. Four-leg intersections are associated with less severe driver errors, and stop-controlled intersections are associated with more sever errors compared with the signalized intersections, which indicates that drivers are more likely to disregard the stop sign compared with the traffic signal. Higher speed limits and vertical curve tend to increase the driver error severity, due to the longer stop distance and obstructed driver's vision. All adverse environmental conditions are associated with decreased driver error severity, which indicates that drivers tend to drive more carefully under these conditions. Passenger car drivers are more likely to make more severe mistakes than truck drivers, which is possibly due to the higher travelling speed of passenger cars. Most of these findings are consistent with those in the study conducted by Wang and Qin (2015).

6.2. Crash type

Older drivers and male drivers are more likely to be involved in the opposite-direction and intersecting-direction crashes, which is consistent with the findings in the study conducted by Yasmin et al. (2014a,b). The possibility of opposite-direction and intersecting-direction crashes is higher at the four-leg intersections and T intersections, as well as yield- and stop-controlled intersections compared with the others. This verifies the conclusion made in Yasmin et al. (2014a,b) that the likelihood of head-on and sideswipe-opposite direction crashes is lower at signalized intersections and higher at T-type intersections. Higher speed limits and vertical curve are associated with decreased likelihood of opposite-direction crashes, and the possibility of opposite-direction crashes decreases with adverse weather conditions. These findings indicate that drivers might be more careful about the conflict vehicles on the road where the curve exists, with higher speed limit and under adverse weather conditions, and the possibility of opposite-direction crashes is expected to be lower. The possibility of opposite-direction and intersecting-direction crashes increases during the night times. This may due to the driver's poor vision at night.

6.3. Vehicle damage

The extent of vehicle damage based on the destruction/deformation of the vehicle involved in the crash can be used as an objective indicator to characterize the severity of a crash. The study conducted by Qin et al. (2013a,b) found that the effectively structural design of the vehicle can absorb or deflect the impact energy caused by collisions, and minimize the forces transferred to occupants and protect occupants from sustaining injuries. Therefore, investigating the vehicle damage may offer useful insight about occupant protection during the crash.

Driver characteristics were excluded in the vehicle damage model, based on the verification made by previous studies (Wang et al., 2015; Qin et al., 2013a, b) that vehicle damage is mainly affected by roadway, weather and vehicle factors other than driver factors. Four-leg and T intersections and sign-controlled intersections are associated with more severe vehicle damage, and the same conclusion is made for the intersections with higher speed limits. These findings are intuitive which is possibly due to the reason that the higher speed of these facilities leads to the larger physical collision force (Qin et al., 2013a, b). The vehicle damage severity increases under adverse weather conditions and during night times. This is likely due to the difficulty of properly controlling vehicles under adverse weather conditions, and the poor driver's vision at night times. Trucks suffer more severe damage than passenger cars, which might be due to their larger body size.

6.4. Injury severity

With regard to the injury severity component, middle-aged drivers and drivers under the influence of drugs and alcohol are more likely to suffer severe injuries. This is supported by the study conducted by Yasmin et al. (2014a,b) that the physiological strength of young drivers is higher than the other age groups. The injury severity level increases at four-leg and T intersections, yield-controlled intersections as well as intersections with higher speed limits. These findings are consistent with the vehicle damage component, which indicates that the drivers are more prone to driving faster and failing to yield the right-of-way at these types of intersections than the others. Severe crashes are more likely to occur under adverse weather and light conditions than the normal conditions. This might be counterintuitive to previous studies, as the drivers are expected to drive slower under adverse weather conditions. However, similar findings in the study conducted by Yasmin et al. (2014a,b) indicated that the injury risk propensities of some crash types including sideswipe and angle crashes are higher during adverse weather conditions, which may be due to unfavorable driving conditions resulting from the reduced visibility. In addition, the counterintuitive coefficient estimates may also be caused by the unobserved factors in crash data, such as intersection geometric features, vehicle condition and driver's maneuver against crashes. Passenger are associated with more severe injury severity than trucks, which might be due to their higher travelling speed.

6.5. Temporal instability of model estimation

The estimation of interaction variables yields the investigation of temporal instability of model estimation results over time. The results show that the interaction variables are only statistically significant in the driver error and crash type model components. Specifically, opposite-direction crashes at four-leg intersections are more likely to occur in 2016 than 2017. Crashes occurred at none- and yield-controlled intersections in 2016 are less likely to be associated with higher levels of driver errors, while more likely to be associated with opposite-direction crash type compared to the crashes occurred in 2017. Stop-sign controlled intersections in 2016 are associated with less severe driver errors than those in 2017. When the weather is cloudy or windy, the driver error severity reduces from 2016 to 2017 crash data, while the tendency of opposite-direction crashes increases from 2016 to 2017. The propensity of involving opposite-direction crashes increases at night times from year 2016 to 2017.

In general, the temporal instability is mainly associated with highway and traffic and environmental characteristics. In terms of the highway and traffic characteristics, the temporal instability may be due to the changes of driver's driving attitudes, which leads to the changes of driver error type and crash type at different time periods at these types of intersections, by continuously gathering information from both their own and other driver's experience, social networks, as well as the safety improvement countermeasures that have been implemented by the safety agencies (Mannering, 2018). With regard to the environmental characteristics, the temporal instability may be related to the different extents of the same weather condition at different time periods, which can further affect driver's behaviors at the same intersection. Although testing the temporal instability of the model estimation with only two-year crash data might not be ideal and so the temporal effects identified in this study should be considered with caution, the estimation of interaction variables highlights that the model estimation results vary over time, and the temporal instability of statistically models should be appropriately accommodated in crash data analysis.

6.6. Copula parameters

The estimated copula parameters offer additional insight about the potential correlations among injury severity, crash type, vehicle damage and driver error, and how the correlation varies among different crashes. A positive value represents a positive correlation among the common unobserved factors affecting the four crash consequence indicators, and a negative value indicates a negative dependency among the common unobserved factors affecting the four crash consequence indicators (Wang et al., 2015).

The estimation of copula parameters highlights the existence of dependencies among injury severity, crash type, vehicle damage and driver error, and the dependency varies among different crashes. Specifically, the dependencies among the four indicators caused by the common unobserved factors are all positive for the crashes that occurred at stop-controlled intersections and four-leg intersections, and the crashes with involved drivers younger than 25. This warrants that considering the effects of unobserved factors in crash data such as the intersection geometric features, vehicle condition and driver behaviors on crash outcomes, the propensities of risky driver errors, opposite-direction collisions, severe vehicle deformation and occupant injuries simultaneously increase or decrease. The magnitude of copula parameters indicates that the highest level of correlations among the four indicators is for crashes at stop-controlled controlled intersections, followed by the crashes with involved drivers younger than 25, and the crashes at four-leg intersections.

From the safety improvement perspective, these findings suggest that in the process of selecting or comparing countermeasures for an intersection, the reduction of crash or crash severity is not the unique target to be considered. The agencies should also account for the benefits of implementing the countermeasure both on the reduction of economic losses resulting from the vehicle damage and the prevention of risky driver behaviors, especially under the three aforementioned situations. From the crash estimation methodology perspective, these findings highlight that the four crash outcome related indicators are highly correlated, and their correlations should be accommodated when the outcome indicators are analyzed simultaneously to achieve the unbiased analysis results.

6.7. Elasticity effects

The estimated parameters alone in the copula model might not be straightforward to express the magnitude of the effects of an independent variable on the probability change for each level of four indicators. Therefore, the elasticity effects of the independent variables with regard to all four crash consequence indicators were measured and presented in the lower panel of Table 3. Note here that only the variables that are statistically significant across all four indicators were considered, and the elasticity effects of them were calculated for the top two levels of each crash consequence indicator. A detailed discussion of calculating the elasticity effects in the copula model can be found in Eluru and Bhat (2007).

In general, the elasticity effects of independent variables on the four indicators shown in Table 4 are consistent with those described in Table 3. Moreover, driving under the influence of drugs or alcohol increases the possibility of reckless driving errors, and ultimately leads to more opposite-direction crashes and severe injuries. Four-leg intersections have the lowest probability of sever driver errors, but the highest probability of intersecting-direction and opposite-direction crashes, as well as severe vehicle damage and injury severity levels. Drivers seem to drive more carefully and properly at yield- and non-controlled intersections, but the crashes occurred at these intersections are significantly higher than the others. Higher speed limit increases reckless driving behaviors and consequently contributes to severe crash consequences. Adverse weather conditions reduce unlawful driver behaviors, but increase the crash severity at intersections. Drivers are more prone to driving carefully at night, but the crashes occurred at night are more severe than day time.

7. Countermeasure recommendations

Based on the results of the copula model, we identified and categorized twelve contributing factors that are associated with severe crash consequences, and recommended specific countermeasures with regard to each of the category to improve the intersection safety in Table 5. The countermeasures are summarized mainly based on the types of contributing factors and their associations with different crash consequence indicators, where the law enforcement strategies and driving related education and training programs are recommended for the driver and vehicle characteristics, and infrastructure and intersection design related recommendations are provided to the highway and environmental characteristics. These countermeasures target on mitigating driver errors, and reducing crash occurrence and minimizing crash consequences for intersections. All of the countermeasures summarized in Table 5 are mainly recommended by the Crash Modification Factors (CMF) Clearinghouse (FHWA, 2019), Connecticut Strategic Highway Safety Plan (CTSHSP) (2017) and multiple previous studies including Wang and Qin (2015) and Devlin et al. (2011).

Specifically, with regard to particular driver groups such as older drivers, the countermeasures are related to enforced education and testing for driver license application, which can help reduce the risky driver errors and improve safety at intersections. Law enforcement and prosecution and conviction of DUI offenders are recommended for drivers that drive under the influence of alcohol. With respect to the unsignalized intersections, appropriate warning signs such as the "intersection ahead" and changing traffic control devices are recommended as an effort to increase the visibility of the intersections and the compliance of traffic right-of-way. Speed limit reconfiguration is recommended at locations with a higher speed limit to provide drivers with a sufficient stopping sight distance. Road surface friction improvement and road De-Icing chemicals are recommended for locations with a higher percentage of adverse weather conditions such as rain and snow. Appropriate roadway assistant technologies such as reflective pavement markers and strips are recommended to help

Table 4
Elasticity Effects of the Copula Model.

Elasticity Effects	Driver Error		Crash Type		Vehicle Damage		Injury Severity	
	Careless	Reckless	IDC	ODC	Moderate	Disabling	C + B	A + K
Drugs or Alcohol Involved	−0.19%	10.78%	12.59%	36.34%	NA	NA	69.04%	252.69%
Four-Leg Intersection	−0.39%	−10.74%	31.36%	86.53%	2.09%	24.41%	23.13%	52.07%
None/Yield-Controlled Intersection	−3.69%	−30.70%	33.34%	100.64%	−0.32%	46.72%	24.11%	56.97%
Middle and High Speed Limit	0.79%	16.44%	−10.76%	−27.34%	1.82%	19.13%	9.31%	20.10%
Cloudy/Windy	2.00%	−30.65%	−21.42%	−64.34%	1.17%	10.18%	30.16%	58.75%
Rain	−4.42%	−25.93%	−26.34%	−52.98%	0.45%	14.46%	39.10%	112.79%
Snow	−3.23%	22.65%	−17.49%	−38.75%	0.42%	18.57%	26.73%	69.83%
Night Time	−2.08%	−19.82%	29.54%	90.59%	0.39%	28.37%	19.93%	47.07%

Table 5
Countermeasure Recommendations.

Contributing Factors	Category	Potential Countermeasures
1. Middle-Aged Driver	Driver Behavior	Enforced vehicle control education and testing to specific driver groups for license application related to proper vehicle control, traffic rule and right-of-way compliance. Periodically enforced older driver vision and medical review practices.
2. Older Driver		
3. Drugs or Alcohol Involved	Law Enforcement	Law enforcement improvement for DUI prevention, such as high- visibility regional DUI enforcement, certified standardized field sobriety test (SFST) and drug recognition expert (DRE) training, and prosecution and conviction of DUI offenders.
4. Four-Leg Intersection	Intersection Design	Intersection geometry reconfiguration and appropriate traffic control design such as changing the partial-way stop-control to all-way stop-control at T intersections.
5. T Intersection		
6. None-Yield-Controlled Intersection	Intersection Traffic Control	Install "intersection ahead" warning signs. Increase the visibility and conspicuity of stop signs, and install rumble strips before intersections.
7. Stop-Controlled Intersection		
8. Middle and High Speed Limit	Speed Management	Speed limit reconfiguration. Install "reduced speed ahead" warning signs.
9. Cloudy/Windy	Adverse Weather Condition	Improve pavement friction such as resurfacing pavement with grooved pavement at intersections with heavy precipitation. Apply road De-Icing chemicals at intersections in winter.
10. Rain		
11. Snow		
12. Night Time	Advanced Technology	Install intersection lighting. Raise pavement markers and install reflective strips at intersections.

increase the visibility of driving lanes at intersections at nighttime.

8. Discussion and conclusions

This study attempts to simultaneously model the four crash consequence indicators – injury severity, crash type, vehicle damage and driver error to identify the crash contributing factors by accounting for the potential correlations among the four indicators caused by the common observed and unobserved factors, and the temporal instability of model estimation over time. A multivariable copula approach built upon the ordered probit model is implemented using the intersection crash data collected from the State of Connecticut. To identify the best model, four copula methodologies including the Frank, Clayton, Joe and Gumble copula models were tested and compared in this study. To investigate the variations of correlations among crash indicators, each copula model was then estimated using a constant copula parameter and the parameterized copula parameters, respectively. To verify the model stability over time, interaction variables between year and each of the other independent variables were incorporated into the model coefficient estimation. The model goodness-of-fit shows that the Joe copula model with the parameterized copula parameters outperforms the other models, which verifies that the injury severity, crash type, vehicle damage and driver error are significantly correlated due to the common observed and unobserved factors, and their correlation varies among different crashes. Accounting for their correlations can lead to more accurate model estimation results. The coefficient estimates of four-leg intersections, yield and stop-controlled intersections and adverse weather conditions varied over time, which indicates that the model estimation of crash data may not be stable over time and should be accommodated in crash prediction analysis.

Driver characteristics, highway characteristics, environmental characteristics and vehicle characteristics were used as the independent variables in the copula model. The model estimation results identify twelve important factors contributing to the crash consequences among them, including the driver's age, driving under the influence of drugs and alcohol, intersection geometry and control types, and adverse weather and light conditions. Based on the findings, specific countermeasures related to law enforcement strategies, driver education and training programs, intersection designs and engineering solutions were recommended for each of these contributing factors to improve the intersection safety. The copula parameter estimation offers additional insight about the variations of the correlations among injury severity, crash type, vehicle damage and driver error. The results show that the dependencies among the four indicators caused by the common unobserved factors are all positive for the crashes that occurred at stop-controlled intersections and four-leg intersections and the crashes with involved drivers younger than 25. In summary, this study offers a more accurate model structure to estimate different crash consequence indicators and identify the comprehensive contributing factors of different crash outcomes. It is anticipated that this study can shed light on selecting valid statistical models for crash data analysis, identifying intersection design issues and helping to develop effective countermeasures to improve intersection safety.

This paper is not without limitations. The copula model estimation results contain some variables that are counterintuitive to the findings from previous studies, such as the effects of adverse weather conditions on injury severity. This may indicate the presence of unobserved or omitted information in crash data. Future research is recommended to collect extra factors such as intersection geometrics and vehicle conditions to explore their influences on crash consequences. Another limitation of this paper is that it only employs the two-vehicle crashes in the analysis. The driver behaviors and crash consequences of single- and multi-vehicle crashes may not be same as two-vehicle crashes. Future research can also focus on developing models for these different crashes to identify their crash characteristics. The investigation of temporal effects with two-year crash data might not be ideal, multi-year crash data can be collected to test the temporal instability of model estimation over time in the future.

Acknowledgments

The authors would like to thank the Connecticut Crash Data Repository (CTCDR) for providing the intersection crash data.

References

Bham, G., Javvadi, B., Manepalli, U., 2012. Multinomial logistic regression model for single-vehicle and multivehicle collisions on urban U.S. Highways in Arkansas. J. Transp. Eng. 138 (6), 786–797.
Bhat, C.R., Eluru, N., 2009. A copula-based approach to accommodate residential self-selection effects in travel behavior modeling. Transp. Res. Part B Methodol. 43 (7), 749–765.
Bonneson, J., Zimmerman, K., Brewer, M., 2003. Guidelines to Reduce Red Light Running. Texas Transportation Institute Project -44 4027.
Braitman, K., Kirley, B., Ferguson, S., Chaudray, N., 2007. Factors leading to older drivers' intersection crashes. Traffic Inj. Prev. 8, 65–72.
Connecticut Crash Data Repository (CTCDR). University of Connecticut, 2018. https://www.ctcrash.uconn.edu/.
Connecticut Department of Transportation (CTDOT), 2017. Strategic Highway Safety Plan. https://www.ct.gov/dot/lib/dot/documents/dsafety/shsp.pdf.
Depaire, B., Wets, G., Vanhoof, K., 2008. Traffic accident segmentation by means of latent class clustering. Accid. Anal. Prev. 40 (4), 1257–1266.

Devlin, A., Candappa, N., Corben, B., Logan, D., 2011. Designing Safer Roads to Accommodate Driver Error. Curtin-Monash Accident Research Center Project 09-006RSC.
Dissanayake, S., 2004. Comparison of severity affecting factors between young and older drivers involved in single vehicle crashes. Int. Assoc.Traffic Saf. Sci. 28, 48–54.
Eluru, N., 2013. Evaluating alternate discrete choice frameworks for modeling ordinal discrete variables. Accid. Anal. Prev. 55 (1), 1–11.
Eluru, N., Bhat, C.R., 2007. A joint econometric analysis of seat belt use and crash-related injury severity. Accid. Anal. Prev. 39 (5), 1037–1049.
Eluru, N., Bhat, C.R., Hensher, D.A., 2008. A mixed generalized ordered response model for examining pedestrian and bicyclist injury severity level in traffic crashes. Accid. Anal. Prev. 40 (3), 1033–1054.
Eluru, N., Paleti, R., Pendyala, R., Bhat, C., 2010. Modeling multiple vehicle occupant injury severity: a copula-based multivariate approach. Transp. Res. Record 2165: J. Transp. Res. Board 1–11.
Federal Highway Administration (FHWA) (2019). Crash Modification Factors Clearinghouse. http://www.cmfclearinghouse.org/.
Ghulam, H., Bhanu, J., Uday, M., 2012. Multinomial logistic regression model for single-vehicle and multivehicle collisions on urban U.S. Highways in Arkansas. J. Transp. Eng. 138 (6), 786–797.
Governors Highway Safety Association (GHSA), 2009. A 10-Point Plan for Reauthorizing the Federal Highway Safety Program. Report No. 202-789-0942. .
Hamdar, S., Maahmassani, H., Chen, R., 2008. Aggressiveness propensity index for driving behavior at signalized intersections. Accid. Anal. Prev. 40, 315–326.
Hassan, H., Abdel-Aty, M., 2012. Exploring the safety implications of young drivers' behavior, attitudes and perceptions. Accid. Anal. Prev. 50, 361–370.
Huang, H., Chin, H.C., Haque, M.M., 2008. Severity of driver injury and vehicle damage in traffic crashes at intersections: a bayesian hierarchical analysis. Accid. Anal. Prev. 40, 45–54.
Ivan, J., Konduri, K., 2018. Chapter 15: crash severity methods. safe mobility: challenges, methodology and solutions. In: In: Lord, Dominique, Washington, Simon (Eds.), Transport and Sustainability Volume 11. Emerald Group Publishing, Binley, West Yorkshire, UK, pp. 325–350.
Keay, L., Jasti, S., Munoz, B., Turano, K., Munro, C., Duncan, D., 2009. Urban and rural differences in older drivers' failure to stop at stop signs. Accid. Anal. Prev. 41 (5), 995–1000.
Kim, J.K., Ulfarsson, G., Shankar, V., Mannering, F., 2010. A note on modeling pedestrian injury severity in motor vehicle crashes with the mixed logit model. Accid. Anal. Prev. 40 (5), 1695–1702.
Laman, H., Yasmin, S., Eluru, N., 2018. Joint modeling of traffic incident duration components (reporting, response, and clearance time): a copula based approach. Proceedings of the 97th Annual Meeting of the Transportation Research Board, Washington, DC.
Lee, J., Chung, J., Son, B., 2008. Analysis of traffic accident size for korean highway using structural equation models. Accid. Anal. Prev. 40, 1955–1963.
Lee, J., Abdel-Aty, M., Cai, Q., Wang, L., 2018. Effects of emergency medical services times on traffic injury severity: a random-effects ordered probit approach. Traffic Inj. Prev. 1–19.
Liu, C., Sharma, A., 2018. Using the multivariate spatio-temporal bayesian model to analyze traffic crashes by severity. Anal. Methods Accid. Res. 17, 14–31.
Lord, D., Schalkwyk, I., Chrysler, S., Staplin, L., 2007. A strategy to reduce older driver injuries at intersections using more accommodating roundabout design practices. Accid. Anal. Prev. 39, 427–432.
Ma, L., Wang, G., Yan, X., Weng, J., 2016. A hybrid finite mixture model for exploring heterogeneous ordering patterns of driver injury severity. Accid. Anal. Prev. 89, 62–73.
Malyshkina, N., Mannering, F., 2009. Markov switching multinomial logit model: an application to accident-injury severities. Accid. Anal. Prev. 41 (4), 829–838.
Mannering, F., 2018. Temporal instability and the analysis of highway accident data. Anal. Methods Accid. Res. 17, 1–13.
Mannering, F., Shankar, V., Bhat, C., 2016. Unobserved heterogeneity and the statistical analysis of highway accident data. Anal. Methods Accid. Res. 11, 1–16.
Milton, J.C., Shankar, V.N., Mannering, F.L., 2008. Highway accident severities and the mixed logit model: an exploratory empirical analysis. Accid. Anal. Prev. 40 (1), 260–266.
Model Minimum Uniform Crash Criteria Fifth Edition, 2017. U.S. Department of Transportation, Washington, DC. https://www.nhtsa.gov/mmucc.
National Highway Traffic Safety Administration (NHTSA), 2015. Traffic Safety Facts 2013: a Compilation of Motor Vehicle Crash Data From the Fatality Analysis Reporting System and the General Estimates System. DOT-HS-812-139. U.S. Department of Transportation.
Papaioannou, P., 2007. Driver behavior, dilemma zone and safety effects at urban signalized intersections in Greece. Accid. Anal. Prev. 39 (1), 147–158.
Pourabdollahi, Z., Karimi, B., Mohammadian, A., 2013. Joint model of freight mode and shipment size choice. Transp. Res. Record 2378: J. Transp. Res. Board 84–91.
Qin, X., Wang, K., Cutler, C., 2013a. Modeling large truck safety using logistic regression models. Transp. Res. Record 2392: J. Transp. Res. Board 1–10.
Qin, X., Wang, K., Cutler, C., 2013b. Analysis of crash severity based on vehicle damage and occupant injuries. Transp. Res. Record 2386: J. Transp. Res. Board 95–102.
Quddus, M., Noland, R.B., Chin, H.C., 2002. An analysis of motorcycle injury and vehicle damage severity using ordered probit models. J. Safety Res. 33, 445–462.
Rana, T., Sikder, S., Pinjari, A., 2010. A copula-based method to address endogeneity in traffic crash injury severity models: application to two-vehicle crashes. Transp. Res. Record 2147: J. Transp. Res. Board 75–87.
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. Accid. Anal. Prev. 43 (5), 1666–1676.
Sener, I., Eluru, N., Bhat, C.R., 2010. On jointly analyzing the physical activity participation levels of individuals in a family unit using a multivariate copula framework. J. Choice Model. 3 (3), 1–38.
Shaon, M., Qin, X., Zhi, C., Zhang, J., 2018. Exploration of contributing factors related to driver errors on highway segments. Transp. Res. Record: J. Transp. Res. Board 1–13.
Wali, B., Khattak, A., Jing, X., 2018. Contributory fault and level of personal injury to drivers involved in head-on collisions: application of copula-based bivariate ordinal models. Accid. Anal. Prev. 110, 101–114.
Wang, K., Qin, X., 2014. Use of structural equation modeling to measure severity of single-vehicle crashes. Transp. Res. Record 2432: J. Transp. Res. Board 17–25.
Wang, K., Qin, X., 2015. Exploring driver error at intersections: key contributors and solutions. Transp. Res. Record 2514 (1): J. Transp. Res. Board 1–9.
Wang, K., Yasmin, S., Konduri, K., Eluru, N., Ivan, J., 2015. Copula based joint model of injury severity and vehicle damage in two-vehicle crashes. Transp. Res. Record 2514: J. Transp. Res. Board 158–166.
Wang, K., Ivan, J., Nalini, R., Jackson, E., 2017. Multivariate poisson lognormal modeling of crashes by type and severity on rural two lane highways. Accid. Anal. Prev. 99 (Part A), 6–19.
Wang, K., Zhao, S., Jackson, E., 2018. Multivariate poisson lognormal modeling of weather-related crashes on freeways. Transp. Res. Record: J. Transp. Res. Board 1–15.
Washington, S., Karlaftis, M., Mannering, F.L., 2011. Statistical and Econometric Methods for Transportation Data Analysis, 2nd ed. Chapman and Hall/CRC, Boca Raton, FL.
Wu, Z., Sharma, A., Mannering, F.L., Wang, S., 2013. Safety impacts of signal-warning flashers and speed control at high-speed signalized intersections. Accid. Anal. Prev. 54, 90–98.
Wu, Q., Chen, F., Zhang, G., Liu, X., Wang, H., Bogus, S., 2014. Mixed Logit Model-Based Driver Injury Severity Investigation in Single- and Multi-Vehicle Crashes on Rural Two-Lane Highways. Accident Analysis & Prevention 72. 105-115.Solutions. Transp. Res. Record 2514: J. Transp. Res. Board 1–9.
Xie, K., Ozbay, K., Yang, H., 2018. Secondary collisions and injury severity: a joint analysis using structural equation models. J. Traffic Injury Prev. 19 (2), 189–194.
Yang, C., Najm, W., 2007. Examining driver behavior using data gathered from red light photo enforcement cameras. J. Safety Res. 38 (3), 311–321.
Yasmin, S., Eluru, N., 2013. Evaluating alternate discrete outcome frameworks for modeling crash injury severity. Accid. Anal. Prev. 59 (1), 506–521.
Yasmin, S., Eluru, N., Pinjari, A., Tay, R., 2014a. Examining driver injury severity in two vehicle Crashes-A copula based approach. Accid. Anal. Prev. 66, 120–135.
Yasmin, S., Eluru, N., Bhat, C.R., Tay, R., 2014b. A latent segmentation generalized ordered logit model to examine factors influencing driver injury severity. Anal. Methods Accid. Res. 1, 23–38.
Ye, F., Lord, D., 2011. Investigation of effects of underreporting crash data on three commonly used traffic crash severity models multinomial logit, ordered probit, and mixed logit. Transp. Res. Record 2241: J. Transp. Res. Board 51–58.
Ye, X., Pendyala, R.M., Al-Rukaibi, F.S., Konduri, K., 2008. A joint model of accident type and severity for Two-vehicle crashes. Proceedings of the 87th Annual Meeting of the Transportation Research Board, Washington, DC.
Zhao, S., Khattak, A.J., 2017. Factors associated with self-reported inattentive driving at highway-rail grade crossings. Accid. Anal. Prev. 109, 113–122.
Zhao, S., Iranitalab, A., Khattak, A., 2018. A clustering approach to injury severity in Pedestrian-train crashes at highway-rail grade crossings. J. Transp. Saf. Secur.