ARTICLE IN PRESS

Journal of Safety Research xxx (xxxx) xxx

Contents lists available at ScienceDirect

Journal of Safety Research

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

Pedestrians under influence (PUI) crashes: Patterns from correspondence regression analysis

Subasish Das [a,*], Sruthi Ashraf [b], Anandi Dutta [c], Ly-Na Tran [d]

[a] Texas A&M Transportation Institute, 1111 RELLIS Parkway, Bryan, TX 77807, United States
[b] Texas A&M Transportation Institute, 3135 TAMU, College Station, TX 77843, United States
[c] Department of Computer Science, University of Texas at San Antonio, One UTSA Circle, San Antonio, TX 78249-0667, United States
[d] Biomedical Sciences, Texas A&M University, 660 Raymond Stotzer Pkwy, College Station, TX 77843, United States

ARTICLE INFO

Article history:
Received 14 October 2019
Received in revised form 14 May 2020
Accepted 16 July 2020
Available online xxxx

Keywords:
Pedestrian crashes
Pedestrian impairment
Pattern recognition
Correspondence regression
Crash typing

ABSTRACT

Introduction: Alcohol-related impairment is a key contributing factor in traffic crashes. However, only a few studies have focused on pedestrian impairment as a crash characteristic. In Louisiana, pedestrian fatalities have been increasing. From 2010 to 2016, the number of pedestrian fatalities increased by 62%. A total of 128 pedestrians were killed in traffic crashes in 2016, and 34.4% of those fatalities involved pedestrians under the influence (PUI) of drugs or alcohol. Furthermore, alcohol-PUI fatalities have increased by 120% from 2010 to 2016. There is a vital need to examine the key contributing attributes that are associated with a high number of PUI crashes. *Method:* In this study, the research team analyzed Louisiana's traffic crash data from 2010 to 2016 by applying correspondence regression analysis to identify the key contributing attributes and association patterns based on PUI involved injury levels. *Results:* The findings identified five risk clusters: intersection crashes at business/industrial locations, mid-block crashes on undivided roadways at residential and business/residential locations, segment related crashes associated with a pedestrian standing in the road, open country crashes with no lighting at night, and pedestrian violation related crashes on divided roadways. The association maps identified several critical attributes that are more associated with fatal and severe PUI crashes. These attributes are dark to no lighting, open country roadways, and non-intersection locations. *Practical Applications:* The findings of this study may be used to help design effective mitigation strategies to reduce PUI crashes.

1. Introduction

Walking is a unique travel mode as it provides many health benefits for individuals. In addition, it also enforces the lowest adverse externalities to the transportation network. However, pedestrians are the most vulnerable roadway users, often due to the lack of enough protection and poor facility design, leading to higher pedestrian exposure. Vehicle-pedestrian crashes are an important traffic safety concern because it has a pattern of a unique certainty and high severity level. The 2017 National Household Travel Survey (NHTS) estimated that about 16% of the U.S. population (around 5% increase from 2009 NHTS survey estimates) walked on their travel day for different reasons (USDOT, 2019). In the United States, 5,977 pedestrians died in traffic crashes in 2017; 2,141 (36%) of these 5,977 pedestrians had alcohol in their system at the time of the crash, and 1,903 (32%) had a blood alcohol concentration (BAC) equal to or greater than 0.08 g/dL in their body (NHTSA, 2019). An older study showed that 80,000–120,000 pedestrians are injured in vehicle–pedestrian crashes each year in the United States (Retting, Ferguson, & McCartt, 2003). According to the CDC's WISQARS™ (Web-based Injury Statistics Query and Reporting System), 952,122 nonfatal pedestrian injuries were reported (yearly average 190,000 injuries) during 2014–2018. In 2017, a pedestrian was killed every 1.5 hours in traffic crashes. Most of these crashes occurred in urban settings (80%), at non-intersections (73%), and in the dark (75%) (NHTSA, 2019).

Previous research concerning relations between the built environment and pedestrian crashes has primarily focused on the identification of key contributing factors. Little research has been conducted focusing completely on the 'pedestrians under the influence' (PUI) involvement in crashes and how pedestrian actions interact with the built environment and other associated factors. Further research must be conducted with new resources and in newer directions. From 2010 to 2016, out of the 564 pedestrian

* Corresponding author.
E-mail addresses: s-das@tti.tamu.edu (S. Das), s-ashraf@tti.tamu.edu (S. Ashraf), anandi.dutta@utsa.edu (A. Dutta), lynattran22@tamu.edu (L.-N. Tran).

https://doi.org/10.1016/j.jsr.2020.07.001
0022-4375/

Please cite this article as: S. Das, S. Ashraf, A. Dutta et al., Pedestrians under influence (PUI) crashes: Patterns from correspondence regression analysis, Journal of Safety Research, https://doi.org/10.1016/j.jsr.2020.07.001

fatalities in Louisiana, 31% were PUI (NHTSA, 2019). This study addresses the current research gap by exploring why and how the aspects of the built environment and pedestrian actions affect PUI crashes with the use of police-reported crash data from 2010 through 2016 in Louisiana, United States. This study used an innovative method (correspondence regression analysis) to determine the key risk patterns in PUI crashes. The current study shows that new and innovative analytical methods are necessary to comprehend the hidden patterns of association between crash attributes.

The rest of this paper is organized as follows. First, the relevant literature is summarized by pointing out representative examples of published work on PUI crashes. Then, a short review of correspondence regression is provided. In the methodology section, the data collection and descriptive statistics are provided. Finally, the results of this evaluation are shown, followed by the discussion and conclusions.

2. Literature review

Understanding the key attributes that influence PUI crashes is an important step toward overall safety enhancement because a thorough understanding of the patterns of the key attributes can significantly help to develop more efficient strategies and countermeasures. A number of studies have investigated PUI crashes. Comparisons of the extent of alcohol involvement among various road user groups showed that alcohol involvement was most prevalent among pedestrians and that pedestrians were also more likely to have higher levels of BAC (Holubowycz, Kloeden, & McLean, 1994). A vast body of research has addressed alcohol impairment among drivers (Jones & Lacey, 2000) and bicyclists (Crocker, Zad, Milling, & Lawson, 2010; Hagemeister & Kronmaier, 2017; Behnood & Mannering, 2017; Helak et al., 2017). However, studies specifically emphasized alcohol-impaired pedestrians are still in nuance stages. Alcohol impairment seems to be an extremely important factor when considering pedestrian causalities (Hezaveh & Cherry, 2018). However, the technique adopted to measure impairment plays an important role in the analysis and results obtained. Although measures of alcohol use are widely accepted as reliable, selection bias due to police judgments about crash severity may lessen the observed impact by introducing impaired pedestrians with lesser injuries (Miles-Doan, 1996).

In a New Zealand study conducted by Lindsay (2012), alcohol involvement was found for both drivers and pedestrians (55.8% of pedestrians that were tested and 24.3% of drivers that were tested). One major factor pointed out in this research was the requirement of data elimination because of the inability to confidently match the BAC test sample from the source. Hezaveh and Cherry (2018) found that around 33% of fatal pedestrian crashes involved a pedestrian with a high BAC (0.08 g/dL or higher), and 5% involved a pedestrian with a low BAC (0.01–0.07 g/dL). Pedestrians were found to be under the influence in 28% of crashes involving pedestrians on interstate roadways (Hezaveh & Cherry, 2018).

The age groups with the largest number of pedestrian fatalities and BAC content above 0.01 g/dL were 45-to-54, 55-to-64, 25-to-34, and 35-to-44 (Behnood & Mannering, 2017). Using the 2003–2015 fatality analysis reporting system (FARS), Ortiz and Ramnarayan (2017) showed that alcohol-impaired crashes consist of any crash where a pedestrian's BAC was at least 0.01 g/dL. Young adults (ages 21–34) were the most likely age group to be involved in a drug or alcohol-related crash. This age group also had the greatest chances of being in an alcohol-involved crash resulting in a fatality. Alcohol use was not a common factor in pedestrian crashes involving pedestrians younger than 15 or older than 75 (Behnood & Mannering, 2017). Another factor related to the likelihood of alcohol-impaired fatalities is crash location. Metropolitan areas were more likely to have alcohol-impaired crash fatalities, and the majority of crashes occurred on major arterial roads and at midblock locations (Lindsay, 2012; Ortiz & Ramnarayan, 2017). LaScala, Johnson, and Gruenewald (2001) showed that higher pedestrian injuries were observed in areas with dense and less solvent areas with a prevalence of younger or older age populations.

Using 931 trauma death results, Demetriades et al. (2004) found that Hispanic and African-American victims were more likely to have a positive screen for substance usage than Caucasians or Asians. Their study found that the fatality rates of males were at least three times those of females from the mid-teens to the mid-sixties. One of the earlier studies also showed relatively low involvement of young and middle-aged females in pedestrian crashes (Holubowycz et al., 1994).

Dangerous behavior may include the excessive consumption of alcohol, drug use, aggressive actions, and risky driving. Furthermore, the higher pedestrian injury and fatality rates among males stay consistent regardless of time exposure, which leads to the conclusion that these rates are attributed to alcohol impairment and risky behavior (WHO, 2019). However, this study did not account for travel related exposures. Other studies showed that mean are also more likely to drive or walk on the road under the influence of alcohol (Behnood & Mannering, 2017; Ortiz & Ramnarayan, 2017).

Tagawa et al. (2000) showed that impaired pedestrians have poor cognitive function resulting in them making poor judgments as they are walking. Impaired pedestrians are much more likely to take risky actions, including crossing against the signal or midblock instead of at a specified crosswalk (Holubowycz et al., 1994). This may be because impaired pedestrians are unable to demonstrate the perceptual, cognitive, and physical skills that are needed to cross streets safely in a complex urban environment (Oxley, Ihsen, Fildes, Charlton, & Day, 2005). In an experimental simulation, highly intoxicated participants with BAC levels of 0.07–0.10 g/dL, showed some lack of awareness of being alcohol-impaired, a tendency to take risky actions in crossing the road, and difficulty integrating speed and distance information in a timely manner (Clayton, Colgan, & Tunbridge, 2000).

There is a greater likelihood of an alcohol-impaired crash occurrence in areas with a greater population density and a greater density of bars (Oxley, Lenné, & Corben, 2006). Posted speed limit, lighting condition, location (rural vs. urban), and vehicle movement are significant factors when alcohol-impaired pedestrian fatalities are evaluated (Miles-Doan, 1996). The odds of dying in a crash where there is a higher posted speed limit is five to nine times more likely than if a lower speed limit is posted. The odds are three to almost five times more likely when the area was dark, and nearly two times more likely in rural areas. Alcohol use increased the odds of being killed or seriously injured relative to receiving only minor or no injuries from two to five times (depending on the model) (Miles-Doan, 1996).

Hezaveh and Cherry (2018) used binary logistic regression modeling to analyze pedestrian crashes involving PUI. The independent variables considered in this study were pedestrian, road, and environmental characteristics. Eighty percent of the PUI crashes occurred during the dark hours. The relative count of mid-block crashes for PUI crashes (69%) was higher than non-PUI crashes (57%). Roadways with straight alignments showed more PUI crashes than roads with curvature (95% vs. 92%). The likelihood of being involved in PUI crashes was the highest for age groups between 40 and 54 years. Fitzpatrick found that PUI crashes were 1.52 times more likely to occur on weekends than on weekdays (Fitzpatrick, Iragavarapu, & Brewer, 2014). Miles-Doan also found evidence to support that PUI crash fatalities were more likely to occur on weekends (Miles-Doan, 1996). Fotios and Gibbons

(2018) studied the importance of lighting among various road user groups and showed evidence from previous studies that ambient light is related to a decrease in traffic crashes involving pedestrians. This indicates that the lighting may help to increase pedestrian visibility to drivers and hence reduce the frequency of road traffic collisions.

The literature review reveals that there is a necessity for conducting a comprehensive study of PUI crashes. The current study demonstrates that innovative dimension reduction methods can provide some significant insights into PUI crashes.

3. Data description

3.1. Data integration

The research team used seven years (2010–2016) of crash data collected from the Louisiana Department of Transportation and Development (LADOTD). To focus on PUI crashes, the research team filtered the crashes to find those with pedestrians involving alcohol or drug impairments. Out of the 11,386 pedestrian crashes, 1,231 crashes were identified as PUI crashes. Around 70% of these PUI crashes were fatal or severe injury crashes. Fig. 1 illustrates the flowchart of the study design.

3.2. Exploratory data analysis

In Table 1, the frequencies of police-reported PUI crashes in Louisiana for 2010–2016 are shown and sorted by severity level as well as by which month they occurred. Months that experienced a greater number of PUI crashes for a certain severity level are displayed in a darker grey color. Severity level B (non-incapacitating injury) shows the largest number of PUI crashes in every month except April and June. Additionally, severity level O (property damage only or PDO) shows the smallest number of PUI crashes for every month. As shown in Table 1, there are slight fluctuations in PUI crash occurrences by month. It's important to note that Louisiana celebrates many festivals. Madri Gras and Jazz festivals take place from February to April. Fatal crashes (K) are higher in number during the months of the Christmas and New Year's Day.

Table 2 contains data for the top three parishes in Louisiana that have experienced a high number of PUI crashes from 2010 to 2016. The parishes are listed in order from the parish, with the highest number of total PUI crashes to the parish with the lowest number of total PUI crashes. The PUI crash totals ranged from 38 to 221. Orleans had the greatest total with 221 PUI crashes in the given time period. This total was significantly higher than the total in E Baton Rouge, which had the second greatest frequencies with 121 total crashes. There was also a significant difference between E Baton Rouge and Jefferson, the parish with the third-highest total, which had a total of 76 crashes. An interactive tool developed by a team member provides walk trip estimates from the five-year American Community Survey (ACS) data (https://subasish.shinyapps.io/ScRAM/). The exposures (in terms of population and walking trips) of these three Parishes are the highest in Louisiana.

3.3. Descriptive by injury levels of PUI crashes

Table 3 displays the percentage of crashes for each variable based on the severity level. For every severity level, a greater percentage of PUI crashes took place on the weekend (i.e., Friday, Saturday, and Sunday) than during weekdays. This could be due to a higher frequency of drug and alcohol use on weekends. For lighting conditions at K level severity, the greatest percentage of crashes took place when the road was dark with no lighting. However, for all other severity levels, the greatest percentage of crashes occurred when it was dark, but with some lighting on the roadway. Additionally, the percentage of crashes generally decreased during daylight and dawn or dusk conditions.

The highest percentage of crashes for every severity level occurred during clear weather conditions. This may be because there is likely to be a greater number of pedestrians when the weather is clear rather than when there is rain, snow, or other unfavorable weather conditions. Furthermore, the highest number of crashes for all severity levels occurred on two-way roadways with no physical separation in comparison to one-way roads and roads with a physical separation or barrier. It is mostly due to higher percentage (around 70% of the state roadway systems of Louisiana) of two-way undivided roadways. A greater percentage of crashes occurred when there were males involved in comparison to female pedestrians. It might be due to the difference between the trip exposures by gender. However, gender-based walking estimates are not readily available. Additionally, the greatest percentage of crashes occurred when pedestrians did not cross at an intersection in comparison to the other pedestrian actions. The sta-

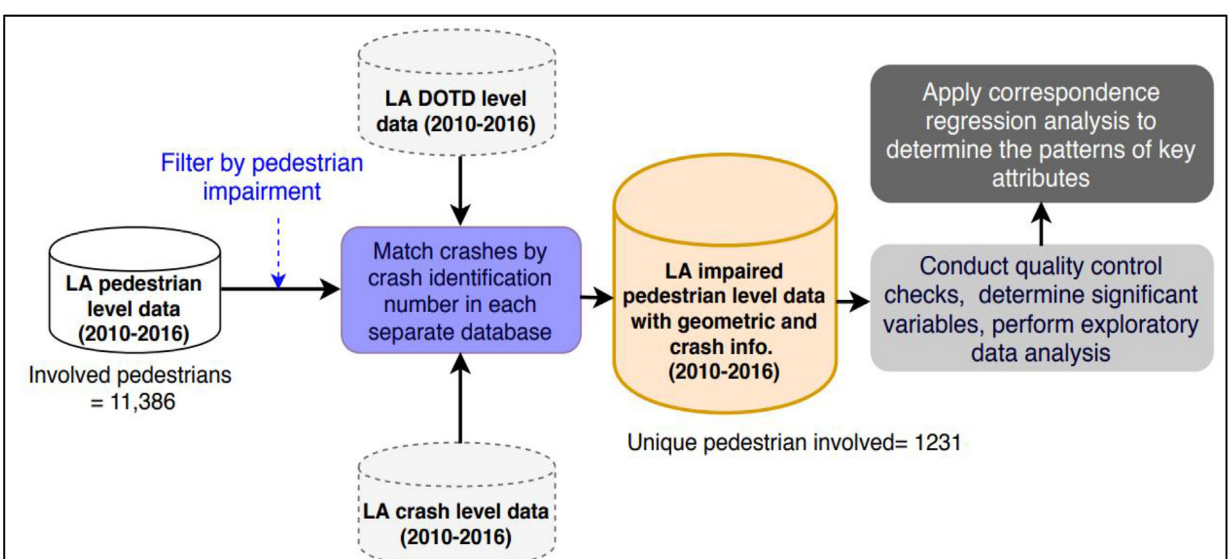


Fig. 1. Data integration and analysis flowchart.

Table 1
PUI severity by months and injury levels.

Severity	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
K[1]	31	29	19	29	15	23	23	25	21	17	15	24
A	13	10	9	16	17	8	7	13	10	16	16	15
B	35	45	37	24	37	22	38	42	51	39	36	37
C	21	20	27	17	13	25	25	18	26	36	29	20
O	5	5	11	9	5	7	8	4	6	12	10	8

Note: [1]K = Fatal, A = Incapacitating Injury, B = Non-Incapacitating Injury, C = Minor Injury O = Property Damage Only (PDO) or No injury.

Table 2
Top 3 Parishes with a high number of PUI crashes.

Parish Name	2010	2011	2012	2013	2014	2015	2016	Total Crashes	Population Estimates[1]	Walk Trip Estimates[1]
Orleans	29	25	28	31	37	39	32	221	376,738	8,078
E Baton Rouge	9	22	9	23	17	22	19	121	444,690	4,564
Jefferson	5	10	14	13	7	13	14	76	435,092	3,498

Note: [1]Source: ACS 2011–2015.

Table 3
Comparison results by injury levels.

Variable	Attribute Description (Code)	K	A	B	C	O	*p*-value[1]
DOW[2]	Friday, Saturday, Sunday (FSS)	55.4%	60.0%	58.0%	53.8%	65.6%	0.3025
	Monday, Tuesday, Wednesday, Thursday (MTWT)	44.6%	40.0%	42.0%	46.2%	34.4%	
Light	Daylight	10.0%	14.0%	16.9%	22.7%	34.4%	0.0094
	Dark with no Lighting	50.9%	29.4%	25.5%	24.9%	35.8%	
	Dark with Lighting	25.8%	43.3%	41.5%	40.2%	26.3%	
	Dark with Lighting at Intersection Only	12.2%	12.0%	13.8%	9.7%	3.4%	
	Dawn/Dusk	1.1%	1.3%	2.3%	2.5%	0.1%	
Weather	Clear	74.2%	80.7%	76.4%	78.0%	89.4%	0.0333
	Cloudy	14.4%	12.6%	16.3%	11.9%	8.2%	
	Rain/Fog/Snow	10.3%	6.0%	6.8%	9.4%	2.3%	
	Other weather conditions (Other Wea.)	1.1%	0.7%	0.5%	0.7%	0.1%	
Road	One-way roadways (1-way)	3.0%	10.7%	10.6%	14.8%	9.3%	0.0195
	Two-way roadways with no Physical Separation (2-way no Sepa)	55.0%	63.3%	62.1%	61.4%	29.1%	
	Two-way roadways with a Physical Barrier (2-way w Barrier)	4.4%	3.3%	1.8%	1.1%	1.1%	
	Two-way roadways with a Physical Separation (2-way w Sepa)	37.6%	22.0%	24.4%	20.9%	12.8%	
	Other Roadway Types	0.0%	0.7%	1.1%	1.8%	47.7%	
Int	Pedestrian is at Intersection (At Int)	22.1%	28.0%	25.3%	30.0%	41.1%	0.0024
	Pedestrian is at Roadway Segment (At Segm)	77.9%	72.0%	74.7%	70.0%	58.9%	
Loc	Business/Industrial	26.6%	29.2%	25.0%	31.8%	67.4%	0.0442
	Business/Residential	29.5%	39.3%	40.2%	34.3%	20.9%	
	Residential	25.1%	22.7%	26.0%	26.4%	9.5%	
	Open Country	17.7%	4.0%	6.5%	6.9%	1.3%	
	Other Locality	1.1%	4.7%	2.3%	0.7%	0.9%	
Gen	Female	26.6%	18.0%	23.0%	21.7%	17.8%	0.2434
	Male	73.4%	82.0%	77.0%	78.3%	82.2%	
Age	15–24	12.9%	18.0%	15.8%	8.7%	18.9%	0.0130
	25–34	28.0%	22.7%	23.9%	24.5%	20.0%	
	35–44	14.8%	19.3%	20.1%	15.5%	23.3%	
	45–54	25.1%	22.7%	24.8%	34.7%	18.9%	
	55–64	12.9%	12.7%	13.3%	13.4%	10.0%	
	>64	6.3%	4.7%	2.0%	3.2%	8.9%	
Act	Crossing at Intersection	10.8%	14.6%	15.4%	18.1%	4.6%	0.0001
	Crossing Not at Intersection	29.5%	34.6%	30.5%	30.3%	14.0%	
	Standing in Roadways	7.4%	8.7%	6.5%	5.0%	3.5%	
	Walking against the Traffic	5.5%	8.7%	7.9%	10.1%	3.5%	
	Walking with the Traffic	15.1%	12.0%	19.6%	20.2%	4.6%	
	Not in the Roadways	3.7%	0.7%	2.3%	2.9%	60.5%	
	Other Pedestrian Actions	28.0%	20.7%	17.8%	13.4%	9.3%	
Contr	Movement Prior to the Crash	4.4%	9.3%	9.9%	8.7%	12.2%	0.0001
	Pedestrian Actions	53.9%	41.3%	33.4%	31.4%	36.7%	
	Pedestrian Conditions	9.2%	14.0%	16.5%	16.2%	18.9%	
	Pedestrian Violations	18.5%	26.7%	33.4%	39.0%	28.9%	
	Other Violations	14.0%	8.7%	6.8%	4.7%	3.3%	

Note: [1]: p or p-trend value, [2]DOW = Day of the Week, Light = Lighting Condition, Weather = Weather Condition, Road = Roadway Types, Loc = Locality, Int = Location of Pedestrian, Gen = Gender, Act = Pedestrian Actions, Contr = Primary Contributing Factor.

Please cite this article as: S. Das, S. Ashraf, A. Dutta et al., Pedestrians under influence (PUI) crashes: Patterns from correspondence regression analysis, Journal of Safety Research, https://doi.org/10.1016/j.jsr.2020.07.001

tistical significance tests (chi-squared test) show that the majority of variable attributes have significant differences in crash injury levels (p-value <0.05). Day of the week and gender have some differences in crash injury types, although those do not reach statistical significance.

4. Methodology

4.1. Correspondence regression analysis

Correspondence regression analysis is useful in analyzing the effects of a polytomous or multinomial outcome variable. This section presents a short overview of correspondence regression analysis based on the work of Plevoets (2019).

The first step of correspondence regression is to create a frequency table by passing the answer vector (Y) with all feasible explanatory variables (X). Correspondence regression produces a triple way table for every viable combination of the conditional variables by crossing Y and X when the values of conditional variables Z. After this, the Pearson residuals (Pr) of the table is computed by the correspondence regression. By using the mutual independence of variables Y and X, the variable E is determined, given that no conditional variables are specified. Elsewhere, the variable E is computed using the standard formula of conditional independence of variables Y and X. Pr can be expressed as:

$$\mathrm{Pr} = \frac{\mathrm{O} - \mathrm{E}}{\sqrt{\mathrm{O}}} \tag{1}$$

For every variable X level 'i,' Y level 'j,' and Z level 'k,' the conditional variables with weights $\frac{n_{i+k}}{n_{i++}}$ and $\frac{n_{+jk}}{n_{+j+}}$ for X and Z, and Y and Z, respectively, is amassed on the triple way table. The resulting matrix and the three-way correspondence can be found in the study conducted by Van der Heijden, De Falguerolles, and De Leeuw (1989). The equation (2) below provides the singular value decomposition (SVD) of D, where variable D is the matrix of Pr. Matrix D is computed by the result of matrix multiplication of matrix U and S with the transpose of matrix V.

$$\mathrm{D} = \mathrm{USV}^{\mathrm{T}} \tag{2}$$

Matrix U contains coordinates for explanatory variable X, and V contains the response variable Y. The singular matrix 'S' is the diagonal matrix, which, in turn, are the square roots of the eigenvalues. The matrices U and V contain the standardized scores of X and Y groups, respectively, with conditional argument. For a certain lower-order category, the aggregated rows in U can be referred to as U+, and the corresponding sum of the total is referred to as r+, the standardized lower-order score /coordinate is indicated as $\mathrm{diag}\left(\frac{1}{\sqrt{r_+}}\right) \times \mathrm{U}_+$. The contributions of the axes to the coefficients for the X and Y categories are in the rows of equation (3) and (4), respectively (Plevoets, 2019).

$$\left(\mathrm{US}^2\right) \times \mathrm{diag}\left(\frac{1}{\sum_k \left(\mathrm{US}^2\right)_{ik}}\right) \tag{3}$$

$$\left(\mathrm{VS}^2\right) \times \mathrm{diag}\left(\frac{1}{\sum_k \left(\mathrm{VS}^2\right)_{jk}}\right) \tag{4}$$

The analysis in this paper is undertaken at the level of pedestrians involved in a crash. To perform the analysis, the research team used open source R software package 'corregp' (25).

5. Results and discussions

The synopsis of the correspondence regression provides some overall statistics and lists the distribution of the 'eigenvalues.' The Phi-squared value (3.79) is equivalent to the Chi-squared value (4666.19) divided by the total number of observations (N = 1231). Both the Phi-squared value and Chi-squared express the dependence between the response variable (severity of the pedestrian injuries) and the explanatory variables (e.g., pedestrian action, lighting cons). The underlying axes can also be thought of as latent variables, and correspondence analysis (CA) labels them 'principal axes.' The eigenvalues in the summary indicate the 'explanatory power' of each principal or latent axis. These measures can be exhibited in three ways: the first row (value) presents the actual eigenvalues, the second row (%) presents the relative values, and the third row (cum %) presents the cumulative relative values. The comprehension of this analysis indicates that the latent axes decompose the observed association of the response variable and explanatory variables into different sets.

On the basis of the eigenvalues, it can be established which latent axes are 'important.' As previously mentioned, the eigenvalues measure the observed association between response and explanatory variables, and the results always sort the eigenvalues from the smallest value to the largest one. Hence, the first eigenvalues represent important or informative axes, whereas the last eigenvalues represent uninformative ones. Because every data set is always a random sample, these uninformative axes can be considered to be reflecting the 'sampling noise' in the data. In practice, one inspects the eigenvalues for a certain cutoff point between the informative and the noisy axes.

There are two significant measures that can show the importance of the attributes and variables. The first measure, known as the contributions of the points to the axes or absolute contributions, illustrates how well each attribute represents (the corresponding inertia) a certain latent axis. The second measure is the contributions of the axes to the points or squared correlations, which represents how well each latent axis reflects a certain attribute. Table 4 shows that both fatal (K) and PDO crashes have high total correlation values. A and C crashes show lower squared correlation value compared to other severity groups.

Table 5 shows the contributions of the points to axes or absolute correlations. For example, 'dark condition with no streetlights' show the highest contributions in Axis 1 and Axis 2 when compared with other attributes in the lighting condition.

The values in Table 6 indicate the explanatory power by each axis. Biplots are good visualization tools to illustrate the association between the key attributes. Biplots are generated based on the concepts of CA. In recent years, many studies explored different variants of CA in solving transportation safety problems (Das & Sun, 2015, 2016; Jalayer, Pour-Rouholamin, & Zhou, 2018; Das, Minjares-Kyle, Wu & Henk, 2019; Das, Jha, Fitzpatrick, Brewer, & Shimu, 2019, Das & Dutta, 2020). Fig. 2 displays a biplot of exploratory variables in which the distances in the plot reflect the association between the attributes. Traits or attributes that are closer together are more

Table 4
Contributions of the axes to the points or squared correlations.

Severity	Axis 1	Axis 2	Total
K	0.11162	0.71768	0.82930
A	0.03426	0.00021	0.03447
B	0.00039	0.59828	0.59866
C	0.02306	0.00862	0.03168
O	0.93475	0.05803	0.99278

Note: The values indicate that Axis 1 explains O injuries significantly, and Axis 2 explains bot K and B crashes significantly.

Table 5
Contributions of the points to axes or absolute correlations.

Attributes	Contribution		Attributes	Contribution	
Light	Axis 1	Axis 2	Pedestrian Age	Axis 1	Axis 2
Daylight	0.00276	0.00338	15–24	0.00166	0.00023
Dark Street light Int.	0.00076	0.00053	25–34	0.00098	0.00059
Dark w. Street Light	0.00783	0.00583	35–44	0.00206	0.00128
Dark No Light	0.01565	0.02284	45–54	0.00162	0.00017
Dawn/Dusk	0.00019	0.00091	55–64	0.00045	0.00010
Weather			45–54	0.00162	0.00017
Clear	0.00000	0.00011	55–64	0.00045	0.00010
Cloudy	0.00142	0.00007	>64	0.00211	0.00893
Rain/Fog/Snow	0.00190	0.00135	Act (Pedestrian Action)		
Other Weather	0.00088	0.00058	Act = Cross at Int	0.00005	0.00267
Loc (Locality)			Act = Cross Not at Int	0.00013	0.00001
Business/Industrial	0.00429	0.00065	Act = Not in Road	0.00009	0.00097
Business/Residential	0.00013	0.00506	Act = Stand in Road	0.00025	0.00012
Residential	0.00006	0.00009	Act = Walk Oppo Traff	0.00002	0.00113
Open Country	0.00602	0.01698	Act = Walk with Traff	0.00018	0.00233
Other Locality	0.00018	0.00103	Act = Others	0.00022	0.00938
Road			Contr (Major Contributing Factor)		
1-way	0.00823	0.00543	Contr = Move Prior Crash	0.00258	0.00370
2-way no sepa	0.00008	0.00145	Contr = Ped Action	0.00150	0.01496
2-way w Barrier	0.00151	0.00358	Contr = Ped Cond	0.00223	0.00402
2-way w Sepa	0.00162	0.00874	Contr = Violations	0.00033	0.01134
Other Roads	0.00171	0.00122	Contr = Others	0.00453	0.00766
Int (Intersection)					
At Int	0.00604	0.00003			
At Segm	0.00225	0.00001			

Note: The values in the table indicate the explanatory power by each axis.

Table 6
Values of three measures.

Measures	Axis 1	Axis 2	Total
Percentages			
Value	0.252	0.251	0.503
Lower	0.223	0.236	
Upper	0.279	0.266	
Cumulative Percentages			
Value	0.252	0.505	0.7657
Lower	0.077	0.079	
Upper	1.177	1.176	

Note: This table lists, the percentages, and the cumulative percentages by the first two axes with the observed measures together with lower confidence bound (in lower) and upper confidence bound (in upper).

highly associated with each other. A cluster group with various traits or attributes indicate that these traits usually occur in groups. The following key clusters are displayed in Fig. 2.

5.1. Cluster 1

This cluster contains attributes such as daylight, dark with street lighting, two age groups (15–24, and 35–44), pedestrian location at the intersection (crosswalk), and business/industrial locality. It is important to note that there is not much data for daylight crashes in the PUI crash database (20% daytime PUI crashes vs. 78% daytime all crashes in Louisiana), as they are not very common. This cluster indicates the association between crashes involving these age groups and roadway crashes in daylight or darkness, but with street lighting. One potential solution to crashes with these attributes is to increase the rapid flash beacons (to alert the drivers) at the traffic signals (MnDOT, 2013).

5.2. Cluster 2

This cluster contains several attributes, including 2-way roadways with no separation, residential and business/residential locality, all weather conditions, pedestrian crossing at non-intersections, and pedestrian action as contributing factors. This cluster indicates crashes with male impaired pedestrians who are involved in midblock crashes on roadways with no physical separation. One potential solution to crashes with these traits is to provide in-street pedestrian crossing signs and high visibility crosswalks (Lin, Kourtellis, Zhang, Guo, & Bialkowska-Jelinska, 2017). The Pedestrian Hybrid Beacon (PHB) includes a sign instructing motorists to "Stop on Red" and a "Pedestrian Crossing" overhead sign. Fitzpatrick and Park (2010) showed that vehicle–pedestrian crashes were reduced by 69% after a PHB was installed. Encouraging ride-sharing during weekend nights would be useful to reduce the PUI crashes occurring on business locality areas. However, the association between ride-sharing and pedestrian impairment-related crash reduction has not been investigated yet.

5.3. Cluster 3

This cluster contains several attributes, including 2-way roadways with no separation, three age groups (25–34, 45–54, and 55–64 years), pedestrian's location at the segment, dark with street light at an intersection only, and pedestrian standing on roadways. This cluster shows an association between several age groups of pedestrians involved in midblock crashes. Safety improvement in regard to crashes with these traits could include in-street pedestrian crossing signs or a "bulbout" in the roadway design (Lin et al., 2017; MnDOT, 2013). Another option is to provide a pedestrian refuge island, which is typically constructed in the middle of a 2-way street and provides a place for pedestrians to stand and wait for motorists to stop or yield (Lin et al., 2017; MnDOT, 2013).

5.4. Cluster 4

This cluster contains several attributes, including 2-way roadways with a barrier, dark conditions with no lighting, pedestrian action as a contributing factor, and open country locality. This clus-

Please cite this article as: S. Das, S. Ashraf, A. Dutta et al., Pedestrians under influence (PUI) crashes: Patterns from correspondence regression analysis, Journal of Safety Research, https://doi.org/10.1016/j.jsr.2020.07.001

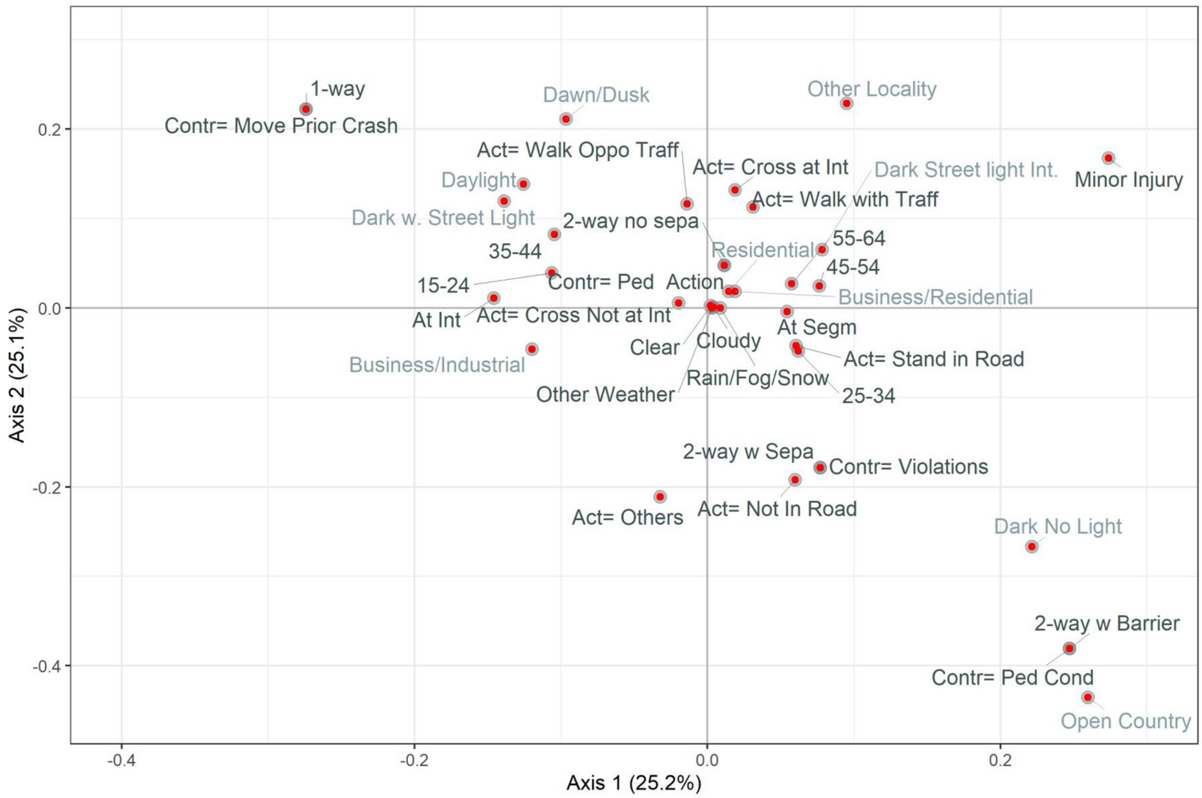


Fig. 2. Biplot of contributing attributes.

ter indicates that crashes taking place in dark lighting conditions are associated with PUI crashes in open country areas. The safety of these locations could be effectively improved by the installation of street lighting. Improvement of intersection lighting helps in improving pedestrian safety (Wang et al., 2016). High visibility crosswalks are effective in reducing pedestrian crashes on open country locations with high pedestrian activity (Lin et al., 2017).

5.5. Cluster 5

This cluster contains several attributes, including 2-way roadways with physical separation, actions as pedestrian not in road and others, pedestrian violations, and other contributing factors. This cluster illustrates PUI crashes that occur outside the roadway lanes when physical separation is present. The cluster trait violation indicates that the pedestrian violates the traffic rule at the specified location with other traits. Educational programs and outreach activities on pedestrian safety can help reduce these occurrences. Lin et al. (2017) explored several education and outreach programs such as WalkWise safety education, distribution of education tip cars, social media outreach, and enforcement roll call training.

The biplot clusters provide a sufficient understanding of the association patterns for the exploratory variables regardless of the nature of the response variables. However, it is essential to compute the relative importance of the explanatory variables to explain the variation in the response variable. Of the correspondence analysis, the regression aspect is the analysis of how strongly each explanatory variable is related to the response variable. As shown in the directed acyclic graph format, association graphs illustrate different latent axes, which are depicted as circles. Moreover, the individual characteristics of the response variable and explanatory variables are represented as boxes. An association graph draws an arrow from a specific latent axis to a specific category if, and only if, the score of that category on that latent axis deviates from 0 significantly (i.e., 0 does not lie within the confidence interval of that category on that latent axis). In other words, the lack of an arrow drawn between a certain category and a certain latent axis indicates that the score of that category on that latent axis is not significantly different from 0.

Fig. 3 illustrates the association graph between pedestrian injury levels and temporal and environmental attributes. Axis 1 is associated with four severity types (K, A, C, and O), and Axis 2 is associated with K, B, C, and O severity types. For Axis 1, K, A, and C have positive eigenvalues. For Axis 2, B and C have positive eigenvalues with and A and O having negative eigenvalues. The signs indicate that the latent Axis 1 is positively associated with fatal and severe injuries, and latent Axis 2 is positively associated with minor or no injuries. Axis 2 is associated with both daylight and dark lighting conditions. Axis 1 and Axis 2 are both associated with two lighting conditions– dark with no lighting and dark with lighting. For Axis 1, another associated temporal and environmental attribute is 'other weather.' For Axis 2, two other associated attributes are 'other weather' and 'daylight.' The same signs of K crash and 'dark no lighting' indicate that fatal crashes are highly associated with the 'dark no lighting' condition. The association graph shows that day of the week, dawn/dusk, 'dark with lighting at the intersection,' and weather conditions such as clear, cloudy, and rain/fog/snow, are not associated with any of the latent axes. The finding of this association graph (no arrows from weekend and weekday) complies with the statistical significance tests (see Table 3) that show that weekday and weekends do not show significant differences in crash injury levels (p-value >0.05).

Similar explanations can be given for Figs. 4 and 5. It is important to note that the left side of the association graphs (association of latent axes and injury levels) in Figs. 4 and 5 are the same. Business/industrial and open country are associated with Axis 1, and business/residential and open country are associated with Axis 2.

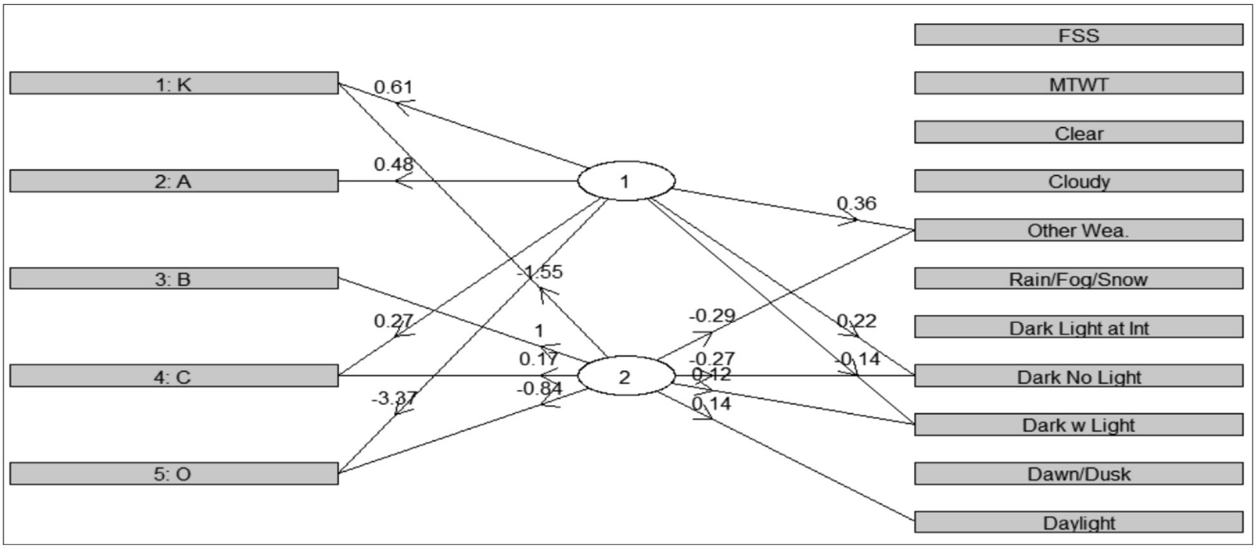


Fig. 3. Association graph for temporal and environmental attributes.

Axis 2 is associated more with roadway types than Axis 1, as Axis 1 is only associated with 1-way roads. The eigenvalue of 1-way (for Axis 1) is negative that indicates an association with O crashes. Axis 1 represents the signification of the location of the pedestrian with respect to the intersection. While the eigenvalue of 'at intersection' is negative, the eigenvalue of 'at segment' (not at intersection) is positive. As fatal and severe crashes have positive eigenvalues (for Axis 1), this association indicates that pedestrian crashes 'at segment' are more severe than 'at intersection' crashes. Other localities, residential localities, other roads, and 2-way with barrier are not significant for any of the axes.

Gender and age of the pedestrians are not associated with Axis 1. Only one age-group (pedestrians older than 64 years) is associated with Axis 2. Based on the eigen value measures, older pedestrians are associated with either fatal or PDO crashes. Axis 1 is associated with only one attribute (contributing factor as others), while Axis 2 is associated with several attributes. Two action types (crossing at intersection and others) are also associated with Axis 2. Several contributing attributes (movement prior to crash, pedestrian action, pedestrian condition, and pedestrian violation) are also associated with Axis 2. The signs of the eigen values indicate that these attributes are associated with moderate and complaint injuries. The finding of this association graph (no arrows from gender groups) complies with the statistical significance tests (see Table 3) that show that males and females have a difference in percentages. However, gender difference does not show significant differences in crash injury levels (p-value >0.05).

6. Conclusions

Alcohol-induced impairment greatly affects human interaction and judgment. It is one of the major contributing factors in roadway traffic crashes. Alcohol impairment can have serious consequences for both drivers and pedestrians on the roadway (Dultz & Frangos, 2013). Recent Louisiana crash data shows that PUI fatalities increased by approximately 120% from 2010 to 2016 (NHTSA, 2019). The conventional approach to pedestrian-related safety analysis establishes relationships between the geometric design of the roadways, traffic measures, and crash occurrences. However,

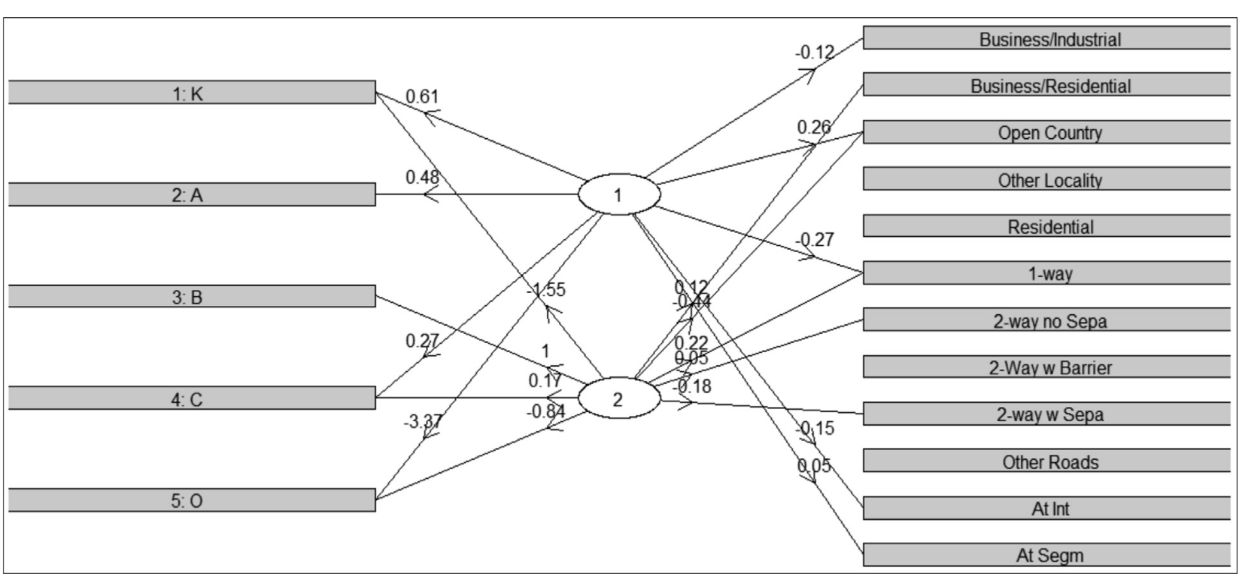


Fig. 4. Association graph for site characteristics.

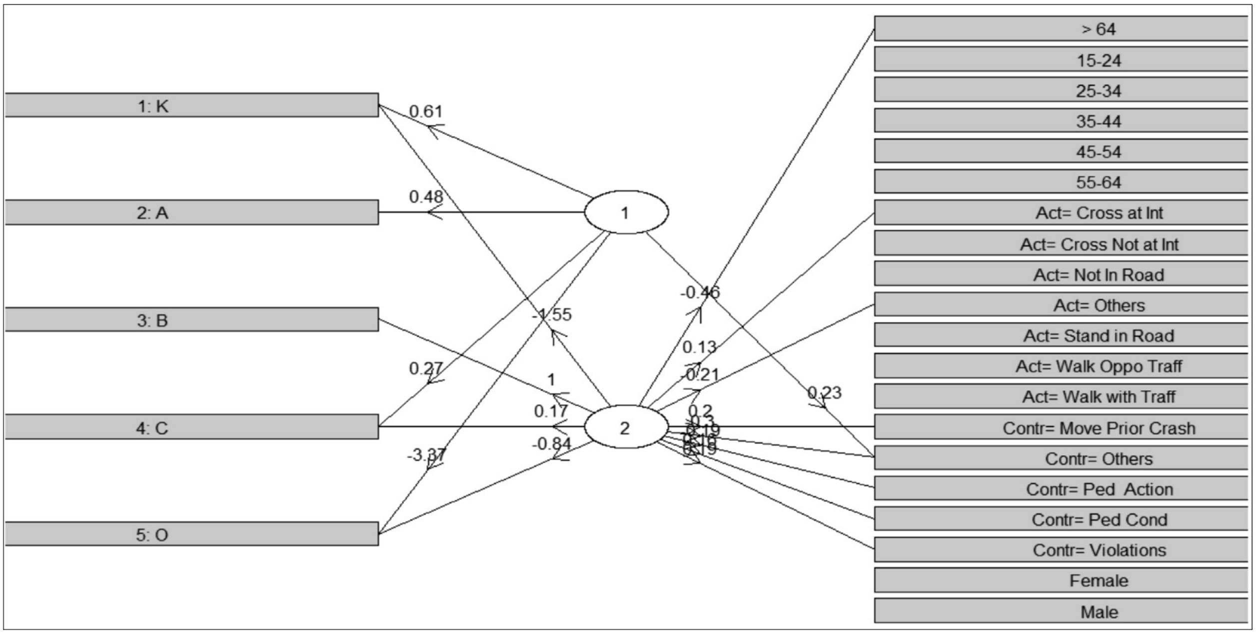


Fig. 5. Association graph for pedestrian-related variables.

these methods are limited due to the exclusion of pedestrian actions and behaviors (Das, Minjares-Kyle, et al., 2019; Das, Jha et al., 2019). Thus, the investigation of pedestrian crashes requires a deep understanding of roadway design, pedestrian action, and the driver's reaction.

This study applied correspondence regression to seven years of Louisiana PUI crash data. The findings identified five risk clusters: intersection crashes at business/industrial locations, mid-block crashes on undivided roadways at residential and business/residential locations, segment related crashes associated with a pedestrian standing in the road, open country crashes with no lighting at night, and pedestrian violation related crashes on divided roadways. The association maps identified several critical attributes that are more associated with fatal and severe PUI crashes. These attributes are dark at no lighting, open country roadways, and non-intersection locations. The research team recommended several suitable countermeasures that can be used to reduce the number of PUI crashes. Examples of these countermeasures include lighting at dark, pedestrian-friendly intersection design, high visibility crosswalks, adjustment of pedestrian crossing phasing during night time, and outreach and awareness activities such as the distribution of education tip cars, social media outreach, and enforcement roll call training.

The current study is not without limitations. First, this study used a limited number of variables that are associated with PUI crashes. Some variables such as roadway traffic condition, driver action, roadway visibility, were not considered due to the high number of missing values. Second, the current study has not used the comprehensive crash typing data (NHTSA, 2018) for fatal crashes that are available for all states as the current study considered all PUI crashes in the analysis instead of only fatal crashes. Third, the analysis is based on police reported crashes. There is a need to extend the study by using other demographic, economic, and land use data to mitigate unobserved heterogeneity. Fourth, Louisiana driving while intoxicated (DWI) penalties and fines have increased in recent years. However, BAC level ranges have not been changed over the years. The effect of strict laws on PUI crashes has not been examined in the current study. Limitations of the current study offer directions for future research in this domain.

7. Disclaimer

The contents of this paper reflect the views of the authors and not the official views or policies of the LADOTD.

References

Behnood, A., & Mannering, F. (2017). Determinants of bicyclist injury severities in bicycle-vehicle crashes: A random parameters approach with heterogeneity in means and variances. *Analytic Methods in Accident Research, 16*, 35–47.

Clayton, A., Colgan, M., & Tunbridge, R. The role of the drinking pedestrian in traffic accidents. Proceedings of the 15th international conference on alcohol, drugs and traffic safety in Stockholm, Sweden, May 2000.

Crocker, P., Zad, O., Milling, T., & Lawson, K. A. (2010). Alcohol, bicycling, and head and brain injury: A study of impaired cyclists' riding patterns. *The American Journal of Emergency Medicine, 28*(1), 68–72.

Das, S., & Sun, X. (2015). Factor association with multiple correspondence analysis in vehicle-pedestrian crashes. *Transportation Research Record: Journal of the Transportation Research Board, 2519*(1), 95–103.

Das, S., & Sun, X. (2016). Association knowledge for fatal run-off-road crashes by multiple correspondence analysis. *IATSS Research, 39*(2), 146–155.

Das, S., Minjares-Kyle, L., Wu, L., & Henk, R. (2019). Understanding crash potential associated with teen driving: Survey analysis using multivariate graphical method. *Journal of Safety Research*.

Das, S., Jha, K., Fitzpatrick, K., Brewer, M., & Shimu, T. (2019). Pattern identification from older bicyclist fatal crashes. *Transportation Research Record: A Journal of Transportation Research Board, 2673*(6), 638–649.

Das, S., & Dutta, A. (2020). Extremely serious crashes on urban roadway networks: Patterns and trends. *IATSS Research*.

Dultz, L., & Frangos, S. (2013). The impact of alcohol in pedestrian trauma. *Trauma, 15*(1), 64–75.

Demetriades, D., Gkiokas, G., Velmahos, G. C., Brown, C., Murray, J., & Noguchi, T. (2004). Alcohol and illicit drugs in traumatic deaths: Prevalence and association with type and severity of injuries. *Journal of the American College of Surgeons, 199* (5), 687–692.

Fitzpatrick, K., Iragavarapu, V., & Brewer, M. A. (2014). *Characteristics of Texas pedestrian crashes and evaluation of driver yielding at pedestrian treatments. Publication FHWA/TX-13/0-6702-1*. Texas A&M Transportation Institute.

Fitzpatrick, K., & Park, E. S. (2010). *Safety Effectiveness of the HAWK Pedestrian Crossing Treatment, FHWA-HRT-10-0421*. McLean, VA: FHWA.

Fotios, S., & Gibbons, R. (2018). Road lighting research for drivers and pedestrians: The basis of luminance and illuminance recommendations. *Lighting Research & Technology, 50*(1), 154–186.

Hagemeister, C., & Kronmaier, M. (2017). Alcohol consumption and cycling in contrast to driving. *Accident Analysis & Prevention, 105*, 102–108.

Helak, K., Jehle, D., McNabb, D., Battisti, A., Sanford, S., & Lark, M. C. (2017). Factors influencing injury severity of bicyclists involved in crashes with motor vehicles: Bike lanes, alcohol, lighting, speed, and helmet use. *Southern Medical Journal, 110*(7), 441–444.

Hezaveh, A. M., & Cherry, C. R. (2018). Walking under the influence of the alcohol: A case study of pedestrian crashes in Tennessee. *Accident Analysis and Prevention, 121*, 64–70.

Holubowycz, O., Kloeden, C., & McLean, A. (1994). Age, sex, and blood alcohol concentration of killed and injured drivers, riders, and passengers. *Accident Analysis & Prevention, 26*(4), 483–492.

Jalayer, M., Pour-Rouholamin, M., & Zhou, H. (2018). Wrong-way driving crashes: A multiple correspondence approach to identify contributing factors. *Traffic Injury Prevention, 19*(1), 35–41.

Jones, R. K., & Lacey, J. H. (2000). *State of knowledge of alcohol-impaired driving: Research on repeat DWI offenders. Publication DTNH22-98-C–05109*. Mid-America Research Institute.

LaScala, E. A., Johnson, F. W., & Gruenewald, P. J. (2001). Neighborhood characteristics of alcohol-related pedestrian injury collisions: A geostatistical analysis. *Prevention Science*, 12.

Lin, P., Kourtellis, A., Zhang, Y., Guo, R., & Bialkowska-Jelinska, E. (2017). *Application of demographic analysis to pedestrian safety. FDOT-BDV25-977-30-Rpt*. Tallahassee: Florida Department of Transportation.

Lindsay, T. Characteristics of alcohol impaired road users involved in casualty crashes, 2012.

Miles-Doan, R. (1996). Alcohol use among pedestrians and the odds of surviving an injury: Evidence from Florida law enforcement data. *Accident Analysis and Prevention, 28*(1), 23–31.

Minnesota Department of Transportation (MnDOT). Minnesota's best practices for pedestrian/bicycle safety. 2013. http://www.dot.state.mn.us/research/TS/2013/201322.pdf Accessed: March 2020.

National Highway Traffic Safety Administration (NHTSA). 2017 FARS/CRSS pedestrian bicyclist crash typing manual a guide for coders using the FARS/CRSS Ped/Bike Typing Tool, 2018.

National Highway Traffic Safety Administration (NHTSA). Traffic safety facts: Pedestrian (2017 Data), Washington DC, 2019.

Plevoets, K. corregp: Open source R software package. https://cran.r-project.org/web/packages/corregp/corregp.pdf Accessed: March 2020.

Ortiz, N., & Ramnarayan, M. (2017). 1949 - Analysis of alcohol-involved pedestrian fatalities in the United States, 2003–2015. *Journal of Transport and Health, 5*, S93.

Oxley, J. A., Ihsen, E., Fildes, B. N., Charlton, J. L., & Day, R. H. (2005). Crossing roads safely: An experimental study of age differences in gap selection by pedestrians. *Accident Analysis & Prevention, 37*(5), 962–971.

Oxley, J., Lenné, M., & Corben, B. (2006). The effect of alcohol impairment on road-crossing behaviour. *Transportation Research Part F: Traffic Psychology and Behaviour, 9*(4), 258–268.

Retting, R. A., Ferguson, S. A., & McCartt, A. T. (2003). A review of evidence-based traffic engineering measures designed to reduce pedestrian-motor vehicle crashes. *American Journal of Public Health, 93*(9), 1456–1463.

Tagawa, M., Kano, M., Okamura, N., Itoh, M., Sakurai, E., Watanabe, T., & Yanai, K. (2000). Relationship between effects of alcohol on psychomotor performances and blood alcohol concentrations. *Japanese Journal of Pharmacology, 83*(3), 253–260.

U.S. Department of Transportation (USDOT). Changing attitudes and transportation choices: 2017 National Household Travel Survey, 2019. https://nhts.ornl.gov/assets/FHWA_NHTS_Report_3E_Final_021119.pdf Accessed: March 2020.

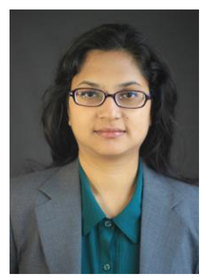
Van der Heijden, P. G. M., De Falguerolles, A., & De Leeuw, J. (1989). A combined approach to contingency table analysis using correspondence analysis and Loglinear analysis. *Applied Statistics, 38*(2), 249–292.

Wang, Z., Lin, P. S., Chen, Y., Hsu, P. P., Ozkul, S., & Bato, M. Safety effects of street illuminance on roadway segments in Florida. Transportation Research Board 96th annual meeting, Washington DC, 2016.

World Health Organization (WHO). Gender and road traffic injuries. https://www.who.int/gender-equity-rights/knowledge/a85576/en/. Accessed: March 2020.

Dr. Subasish Das is an Assistant Research Scientist at Texas A&M Transportation Institute (TTI). He received his Ph.D. in Civil Engineering from University of Louisiana at Lafayette in 2015 and holds a Master of Science in Civil Engineering from the same university in 2012. He has more than 10 years of national and international experience associated with transportation safety engineering research projects. His primary fields of research interest are roadway safety, roadway design, and associated operational issues. Dr. Das is the author or co-author of over 80 technical papers or research reports. He is the author of CRC Press book 'Artificial Intelligence in Highway Safety.'

Sruthi Ashraf pursues her Ph.D. in Transportation Engineering at the Texas A&M University, College Station, Texas. She completed her bachelor's in Civil Engineering and master's in Traffic and Transportation Planning from India. She has been working as a Graduate Research Assistant with different research groups of Texas A&M Transportation Institute (TTI) since 2016. Currently, she works on projects that are related to roadway safety, transportation economics and managed lanes and have published technical papers and research reports on the same.



Dr. Anandi Dutta has been working as an Assistant Professor of Practice at Department of Computer Science at University of Texas at San Antonio. She completed her Master of Science in Electrical Engineering from Louisiana State University in 2010. In 2016, she earned her Ph.D. in Computer Science and Engineering from the University of Louisiana at Lafayette. She has published many research papers in natural language processing and machine learning in transportation science, Systems-on-Chip, and FPGA.

Ly-Na Tran pursues her Bachelor of Science Degree in Biomedical Science at the Texas A&M University, College Station, Texas. She has been working as a Student Assistant at Texas A&M Transportation Institute (TTI) since 2018.