Pedestrians Under Influence: Findings from Correspondence Regression Analysis

Subasish Das, Ph.D.
(Corresponding Author)
Texas A&M Transportation Institute
1111 RELLIS Parkway, Room 4414, Bryan, TX 77807
Email: s-das@tti.tamu.edu
ORCID ID: 0000-0002-1671-2753

Sruthi Ashraf, Ph.D. Student
Texas A&M Transportation Institute
3135 TAMU, College Station, TX 77843
Email: s-ashraf@tti.tamu.edu

Anandi Dutta, Ph.D.
Computer Science and Engineering Dept., Ohio State University
2015 Neil Ave, Columbus, OH 43210
E-mail: dutta.34@osu.edu

Ly-Na Tran
Biomedical Sciences, Texas A&M University
660 Raymond Stotzer Pkwy, College Station, TX 77843
Email: lynattran22@tamu.edu

TOTAL WORDS: 7,142 words

5,642 words = text (including abstract and references)

1,500 = 6 tables

Submitted: August 01, 2019

ABSTRACT

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 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 factors that are associated with a high number of PUI crashes. In this study, the research team analyzed Louisiana traffic crash data from 2010 to 2016 to identify the key contributing factors and association patterns based on severity type for crashes involving PUI by applying correspondence regression analysis. The findings show that male and female pedestrians demonstrate different action traits that affect their crash involvement. Females are more likely to be involved in intersection crashes, while males are more likely to be involved in midblock crashes. In addition to these human factors, some roadway features like dark with no light, two-lane roadway with no physical separation are highly associated with PUI crashes.

Keywords: pedestrian crashes; pedestrian impairment; pattern recognition; correspondence regression; crash typing.

INTRODUCTION

Walking is a unique mode of transport as it provides several health benefits for individuals. In addition, it also imposes the lowest negative externalities to the transportation system. However, pedestrians in transportation networks are vulnerable to traffic crashes, often due to the poor design of the networks and a lack of protection against impact. Vehicle-pedestrian crashes are a major concern because these crashes have a unique inevitability and high severity level. Walking has become increasingly dangerous, and it is shown by the growing number of vehicle-pedestrian crashes that have caused pedestrian fatalities and serious injuries in recent years *(1)*. Historical data shows that 80,000 to 120,000 pedestrians are injured and about 4,600 to 4,900 pedestrians die in vehicle-pedestrian crashes each year in the United States *(2)*. In the U.S., 5,295 pedestrians died in traffic crashes in 2015. 2,022 (38%) of these 5,295 pedestrians had alcohol in their system at the time of the crash, and 1,800 (34%) had blood alcohol concentration (BAC) equals or greater than 0.08 g/dL in their body. From 2010 to 2016, out of the 564 pedestrian fatalities in Louisiana, 31% were pedestrians under influence (PUI) *(3)*.

Previous research concerning relations between the built environment and pedestrian crashes has primarily focused on the identification of key contributing factors. In regard to PUI involvement, little research has been conducted in crashes and how pedestrian actions interact with the built environment and other associated factors. The current study addresses this research gap by exploring why and how the aspects of the built environment and pedestrian actions affect PUI crashes. Further research needs to be conducted with additional resources and in newer directions. The current study shows that new and innovative analytical methods are necessary to comprehend the hidden patterns of association between crash attributes.

This paper is organized as follows. First, the existing literature relevant to this paper's topic 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 discussions and conclusions.

LITERATURE REVIEW

In 2016, a pedestrian was killed every 1.5 hours in traffic crashes. Majority of them occurred in urban settings (76%), at non-intersections (72%), and in the dark (75%) (*4*). Approximately 70 percent of the pedestrians killed in traffic crashes were male. Comparisons of the extent of alcohol involvement among various road user groups showed that alcohol involvement was most prevalent among pedestrians, and pedestrians were also more likely to have higher levels of BAC (*5*). A vast body of research has addressed alcohol impairment among drivers (*6*) and bicyclists (*7-10*). However, studies specifically focused on alcohol-impaired pedestrians are still in nuance stages (*11*). Alcohol impairment seems to be an extremely important factor when pedestrian (than any other user group) causalities are considered (*11*). 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 injury severity may reduce the observed effect by introducing alcohol-impaired pedestrians with lesser injuries (*12*).

In a New Zealand study conducted by Lindsay, alcohol impairment was found across all road user types, but it was particularly noted amongst pedestrians (55.8% of pedestrians that were tested) and drivers (24.3% of drivers that were tested) (*13*). 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, particularly in the case of pedestrians (50 percent of the total samples were eliminated).

An estimated 33 percent of fatal pedestrian crashes involved a pedestrian with a BAC of .08 g/dL or higher, and 5 percent involved a pedestrian with a BAC of .01 - .07 g/dL (*11*). The age groups with the largest number of pedestrian fatalities and BAC content above .01 g/dL were 45-to-54, 55-to-64, 25-to-34, and 35-to-44 (*1*, *9*). In Texas, pedestrians were found to be under the influence in 28 percent of crashes involving pedestrians on freeways (*11*). Furthermore, 86 percent of crashes involved at least one of the drivers under the influence of drugs or alcohol, and 89 percent of the crashes involved at least one of the pedestrians under the influence (*14*).

Using the fatality analysis reporting system (FARS), Ortiz and Ramnarayan evaluated 59,754 pedestrian fatalities from 2003 to 2015. Alcohol-involved crashes consist of any crash where a pedestrian's BAC was at least 0.01 g/dL (*15*). The most common age group to be involved in a drug or alcohol-related crash were young adults (21-34). The young age group also had the greatest odds of being in an alcohol-involved fatality, and alcohol use was not a common factor among the pedestrians of age less than 15 and greater than 75 (*9*). The crash location also had an effect as alcohol-involved fatalities were more likely to occur at metropolitan areas with most occurring on major arterial roads and at midblock locations (*13*, *15*) and on weekends (*12*).

The odds of an alcohol-impaired pedestrian being in a crash were more likely if the pedestrian was male, American Indian, or Hispanic. These results involving gender and ethnic minorities were supported by a study conducted by Demetriades et al. (*16*). 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, thus highlighting the relatively low involvement of young and middle-aged females in pedestrian crashes (*5*).

Gender role socialization and the association of masculinity with acceptance of risk, risk-taking behavior, and a disregard of pain and injury may be factors leading to more hazardous conduct from men (*17*). These actions include excessive consumption of alcohol, drug use, aggressive behavior to be in control of situations, and risky driving. Higher male pedestrian injury and fatality rates appear to hold regardless of time spent walking on the road and are attributed to alcohol use and risky behavior (*17*). Data consistently show that men are more likely than women to be driving or walking on the road under the influence of alcohol (*9*, *15*). Živkovic et al. (*18*) found evidence from previous literature that the risk of injury rises as a function of alcohol consumption. Moreover, alcohol use by pedestrians impairs judgment and cognitive function as they navigate their environment (*19*). Even though the intoxicated pedestrians frequently cannot fulfill the perceptual, cognitive, and physical skills required to cross safely in the complex traffic patterns seen in most urban areas (*20*), they are much more likely to engage in risky street-crossing behavior, including crossing against the signal or mid-block instead of at a designated crosswalk (*5*). In a simulator experiment, highly intoxicated participants with BAC levels of 0.07–0.10%, showed some lack of awareness of impairment, a tendency to engage in risky road-crossings, and difficulty integrating speed and distance information in a timely manner (*21*), necessary to select safe gaps in the traffic (*22*).

In areas with a greater population density and a greater density of bars, there is a greater likelihood of alcohol-impaired collisions (*20*). Posted speed limit, lighting condition, location (rural vs urban), and vehicle movement are significant factors when alcohol-impaired pedestrian fatalities are evaluated (*12*). 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) (*12*).

Hezaveh and Cherry used binary logistic regression modeling to analyze pedestrian crashes with respect to PUI. The independent variables considered in this study were pedestrian, road, and environmental characteristics. Males were more prevalent in the PUI cases (78% males compared to 22% females). Eighty percent of the PUI crashes occurred during the dark hours. The relative frequency of mid-block crashes (i.e., not at an intersection) for PUI crashes (69%) was higher than non- walking under influence (WUI) crashes (57%). Straight road alignments witnessed more PUI crashes than roads with curvature (95% vs. 92%). This may be because of the lack of driver expectancy of the pedestrians on the straight roads. The likelihood of being involved in WUI crashes were the highest for age groups between 40–54. Additionally, PUI crashes were 1.52 times likely to occur on weekends than weekdays (*14*). Fotios and Gibbons studied the importance of lighting among various road user groups and showed evidence from previous researches that ambient light offers significant benefit to a reduction in road collisions involving pedestrians (*24*). It refers that the lighting may help to increase the pedestrian visibility to drivers and hence reduce the frequency of road traffic collisions. Awareness campaigns are required to teach pedestrians about the risk of being alcohol-impaired because a new safety law in Serbia led to a significant decrease in the number of pedestrian fatalities, especially in the first year of its implementation. However, alcohol use in pedestrians remained a constant risk factor because the proportion of alcohol-intoxicated pedestrians showed no significant difference in the years preceding and following the new traffic safety law (*18*).

The literature review reveals that there is a need for conducting an in-depth study of PUI crashes. The current study demonstrates that cluster regression analysis can provide some significant insights on PUI crashes.

METHODOLOGY

Data Description

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

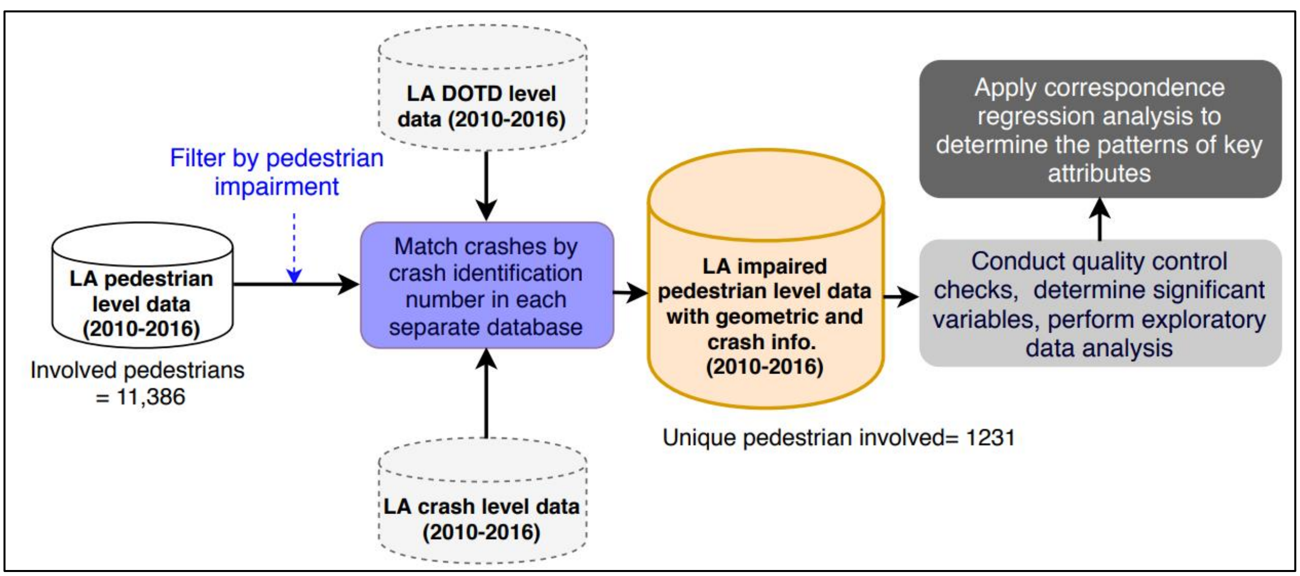

Figure 1 Data integration and analysis flowchart.

research team filtered the crashes to find those with pedestrians involving alcohol or drug impairments. Out of 11,386 pedestrian crashes, 1,231 crashes were identified as PUI crashes. Around 70 percent of these PUI crashes were fatal or severe injury crashes. The filtering and data analysis process is illustrated in Figure 1.

Exploratory Data Analysis

In Table 1, the frequency of PUI crashes in Louisiana for 2010-2016 are shown and sorted by severity level and in which month they occurred. Months that experienced a greater number of PUI crashes for a certain severity level are displayed in a darker green color. Severity level B shows the largest number of PUI crashes in every month except April and June. Additionally, severity level O shows the smallest number of PUI crashes in every month, excluding October and November. As shown in Table 1, there does not seem to be a consistent pattern in regard to the time of year in which a greater or lesser amount of crashes takes place; the number of crashes for each severity level does not vary greatly for the different months. Fatal crashes (K) are higher in number during the months of the Christmas and new year.

Table 1 PUI Severity by Months

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 contains data for parishes in Louisiana that have experienced a high number of PUI crashes from 2010-2016. The parishes are listed from greatest to least 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 was a significantly higher number than in the E Baton Rouge, which had the second greatest frequencies with 121 total crashes. There was also a significant difference between the totals of E Baton Rouge and the parish with the third highest total – Jefferson, which had a total of 76. Ascension had the lowest total for PUI crashes with 38 crashes. There were several other parishes that had totals similar to Ascension's; Terrebonne's total was 39, Lafayette's total was 40, and Rapides' total was 42. Although some parishes, such as Calcasieu and Rapides, showed an increase in the number of crashes per year, the data generally did not show a significant pattern from year to year.

Table 2 Parishes with High Number of PUI Crashes

Parish Name	2010	2011	2012	2013	2014	2015	2016	Grand Total
Orleans	29	25	28	31	37	39	32	221
E Baton Rouge	9	22	9	23	17	22	19	121
Jefferson	5	10	14	13	7	13	14	76
Calcasieu	6	6	13	7	10	14	13	69
Caddo	9	8	13	8	10	12	6	66
Ouachita	7	9	8	9	4	8	8	53
Rapides	1	2	5	6	9	8	11	42
Lafayette	5	9	4	7	3	8	4	40
Terrebonne	4	8	4	3	9	4	7	39
Ascension	6	6	10	3	5	2	6	38

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 sort of lighting on the roadway. Additionally, the percentage of crashes generally decreased during daylight and dawn or dusk conditions.

The greatest 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; one-way roads and roads with a physical separation or barrier experienced fewer crashes. A greater percentage of crashes occurred when there are males involved in comparison to female pedestrians. Additionally, the greatest percentage of crashes occurred when pedestrians did not cross at an intersection in comparison to the other pedestrian actions.

Table 3 Percentage of Variables by Severity

Variable	Attribute Code	Description	K	A	B	C	O
DOW	FSS	Friday, Saturday, Sunday	55.4%	60.0%	58.0%	53.8%	65.6%
DOW	MTWT	Monday, Tuesday, Wednesday, Thursday	44.6%	40.0%	42.0%	46.2%	34.4%
Light	Daylight	Daylight	10.0%	14.0%	16.9%	22.7%	22.2%
Light	Dark No Light	Dark with no Lighting	50.9%	29.3%	25.5%	24.9%	13.3%
Light	Dark w Light	Dark with Lighting	25.8%	43.3%	41.5%	40.1%	54.4%
Light	Dark Light at Int	Dark with Lighting at Intersection Only	12.2%	12.0%	13.8%	9.7%	7.8%
Light	Dawn/Dusk	Dawn/Dusk	1.1%	1.3%	2.3%	2.5%	2.2%
Weather	Clear	Clear	74.2%	80.7%	76.5%	78.0%	76.7%
Weather	Cloudy	Cloudy	14.4%	12.7%	16.3%	11.9%	18.9%
Weather	Rain/Fog/Snow	Rain/Fog/Snow	10.3%	6.0%	6.8%	9.4%	4.4%
Weather	Other Wea.	Other weather conditions	1.1%	0.7%	0.5%	0.7%	0.0%
Road	1-way	One-way roadways	3.0%	10.7%	10.6%	14.8%	18.9%
Road	2-way no Sepa	Two-way roadways with no Physical Separation	55.0%	63.3%	62.1%	61.4%	55.6%
Road	2-Way w Barrier	Two-way roadways with a Physical Barrier	4.4%	3.3%	1.8%	1.1%	1.1%
Road	2-way w Sepa	Two-way roadways with a Physical Separation	37.6%	22.0%	24.4%	20.9%	22.2%
Road	Other Roads	Other Roadway Types	0.0%	0.7%	1.1%	1.8%	2.2%
Int	At Int	Pedestrian is at Intersection	22.1%	28.0%	25.3%	30.0%	41.1%
Int	At Segm	Pedestrian is at Roadway Segment	77.9%	72.0%	74.7%	70.0%	58.9%
Loc	Business/Industrial	Business/Industrial	26.6%	29.3%	25.1%	31.8%	42.2%
Loc	Business/Residential	Business/Residential	29.5%	39.3%	40.2%	34.3%	30.0%
Loc	Residential	Residential	25.1%	22.7%	26.0%	26.4%	23.3%
Loc	Open Country	Open Country	17.7%	4.0%	6.5%	6.9%	3.3%
Loc	Other Locality	Other Locality	1.1%	4.7%	2.3%	0.7%	1.1%
Gen	Female	Female	26.6%	18.0%	23.0%	21.7%	17.8%
Gen	Male	Male	73.4%	82.0%	77.0%	78.3%	82.2%
Age	15-24	15-24	12.9%	18.0%	15.8%	8.7%	18.9%

Variable	**Attribute Code**	**Description**	**K**	**A**	**B**	**C**	**O**
Age	25-34	25-34	28.0%	22.7%	23.9%	24.5%	20.0%
Age	35-44	35-44	14.8%	19.3%	20.1%	15.5%	23.3%
Age	45-54	45-54	25.1%	22.7%	24.8%	34.7%	18.9%
Age	55-64	55-64	12.9%	12.7%	13.3%	13.4%	10.0%
Age	> 64	> 64	6.3%	4.7%	2.0%	3.2%	8.9%
Act	Cross at Int	Crossing at Intersection	10.7%	14.7%	15.3%	18.1%	12.2%
Act	Cross Not at Int	Crossing Not at Intersection	29.5%	34.7%	30.5%	30.3%	33.3%
Act	Stand in Road	Standing in Roadways	7.4%	8.7%	6.5%	5.1%	5.6%
Act	Walk Oppo Traff	Walking against the Traffic	5.5%	8.7%	7.9%	10.1%	7.8%
Act	Walk with Traff	Walking with the Traffic	15.1%	12.0%	19.6%	20.2%	13.3%
Act	Not in Road	Not in the Roadways	3.7%	0.7%	2.3%	2.9%	2.2%
Act	Others	Other Pedestrian Actions	28.0%	20.7%	17.8%	13.4%	25.6%
Contr	Move Prior Crash	Movement Prior to the Crash	4.4%	9.3%	9.9%	8.7%	12.2%
Contr	Ped Action	Pedestrian Actions	53.9%	41.3%	33.4%	31.4%	36.7%
Contr	Ped Cond	Pedestrian Conditions	9.2%	14.0%	16.5%	16.2%	18.9%
Contr	Violations	Pedestrian Violations	18.5%	26.7%	33.4%	39.0%	28.9%
Contr	Others	Other Violations	14.0%	8.7%	6.8%	4.7%	3.3%

Note: 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

Cluster Correspondence Analysis

Correspondence regression analysis is useful in analyzing the effects for a polytomous or multinomial outcome variable. This section presents a short overview of cluster correspondence analysis based on the work of Plevoets *(25).*

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:

$$\mathrm{Pr} = \frac{\mathrm{O} - \mathrm{E}}{\sqrt{\mathrm{O}}} \qquad (1)$$

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 et al. (*26*). 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.

$$\mathrm{D} = \mathrm{USV^T} \qquad (2)$$

Matrix U contains coordinates for explanatory variable X, and V contains response variable Y. The singular matrix 'S' is the diagonal matrix which, in turn, are the square roots of the eigen values. 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 as U+ and the corresponding sum of the total is referred to as r+, the standardized lower-order score /coordinate is indicated as $\mathrm{diag}\left(\frac{1}{\sqrt{r_+}}\right) \times U_+$. The contributions of the axes to the coefficients for the X and Y categories are in the rows of equation (3) and (4) respectively *(25).*

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

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

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).

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 equal 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 labels them 'principal axes.' The eigenvalues in the summary indicate the 'explanatory power' of each principal or latent axis (correspondence analysis also speaks of 'principal inertias,' which are the eigenvalues divided by N). 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.

Table 4 lists the actual eigenvalues, the percentages, and the cumulative percentages by first two axes with the observed measures together with lower confidence bound (in lower) and upper confidence bound (in upper). For example, the confidence interval for the first eigenvalue is [610.25; 814.87], the confidence interval for the second eigenvalue is [672.54, 751.23], and the confidence interval for the first relative eigenvalue is [0.223; 0.279].

Table 4 Values of Three Measures

Eigenvalues	**Axis 1**	**Axis 2**	**Total**
Value	1177.37	1172.21	2349.58
Lower	610.25	672.54	
Upper	814.87	751.23	
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	

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 representative uninformative ones. Because every data set is always a random sample, these uninformative axes can be considered as 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 5 shows that both K and O crashes have high total correlation values. A and C crashes show lower squared correlation value compared to other severity groups.

Table 5 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

Table 6 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.

Table 6 Contributions of the Points to Axes or Absolute Correlations

Attributes	**Contribution**		**Attributes**	**Contribution**	
DOW	**Axis 1**	**Axis 2**	**Axis 1**	**Axis 1**	**2**
FSS	0.00106	0.00002	15-24	0.00166	0.00023
MTWT	0.00142	0.00002	25-34	0.00098	0.00059
Light			35-44	0.00206	0.00128
Daylight	0.00276	0.00338	45-54	0.00162	0.00017
Dark Street light Int.	0.00076	0.00053	55-64	0.00045	0.00010
Dark w. Street Light	0.00783	0.00583	45-54	0.00162	0.00017
Dark No Light	0.01565	0.02284	55-64	0.00045	0.00010
Dawn/Dusk	0.00019	0.00091	>64	0.00211	0.00893
Weather			**Gender**		
Clear	0.00000	0.00011	Female	0.00094	0.00046
Cloudy	0.00142	0.00007	Male	0.00027	0.00013
Rain/Fog/Snow	0.00190	0.00135	**Act**		
Other Weather	0.00088	0.00058	Act= Cross at Int	0.00005	0.00267
Loc			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**		
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					
At Int	0.00604	0.00003			
At Segm	0.00225	0.00001			

Biplots are good visualization tools to illustrate the association between the key attributes. Figure 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 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 Figure 2.

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, and business/industrial locality. It is important to note that there is not much data for daylight crashes in the PUI crash database, as they are not very common. This cluster indicates the association between crashes involving these age groups and roadway crashes in daylight or in darkness, but with street lighting. One potential solution to crashes with these attributes is to increase the pedestrian crossing time phases at the traffic signals.

Cluster 2

This cluster contains several attributes, including 2-way roadways with no separation, residential and business/residential locality, male pedestrian, 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. Another potential solution is providing a curb extension or “bulbout” that extends the sidewalk or curb line into the street or parking lane, thus reducing the street width and improving sight distance between the driver and pedestrian. Encouraging ride-sharing during weekend nights would be useful to reduce these crashes.

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), female pedestrians, pedestrian’s location at the segment, weekdays (MTWT), dark with street light at an intersection only, and pedestrian standing on roadways. This cluster shows an association between several age groups of female 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. 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.

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 cluster 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.

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 contributing factor as others. 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 program and outreach activities on pedestrian safety can be helpful in reducing these occurrences.

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 is significantly deviated from 0 (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.

Figure 3 illustrates the association graph between pedestrian severity 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 eigen values with O having negative eigen values. 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. Only 'dark no lighting condition' has a negative eigenvalue. 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.

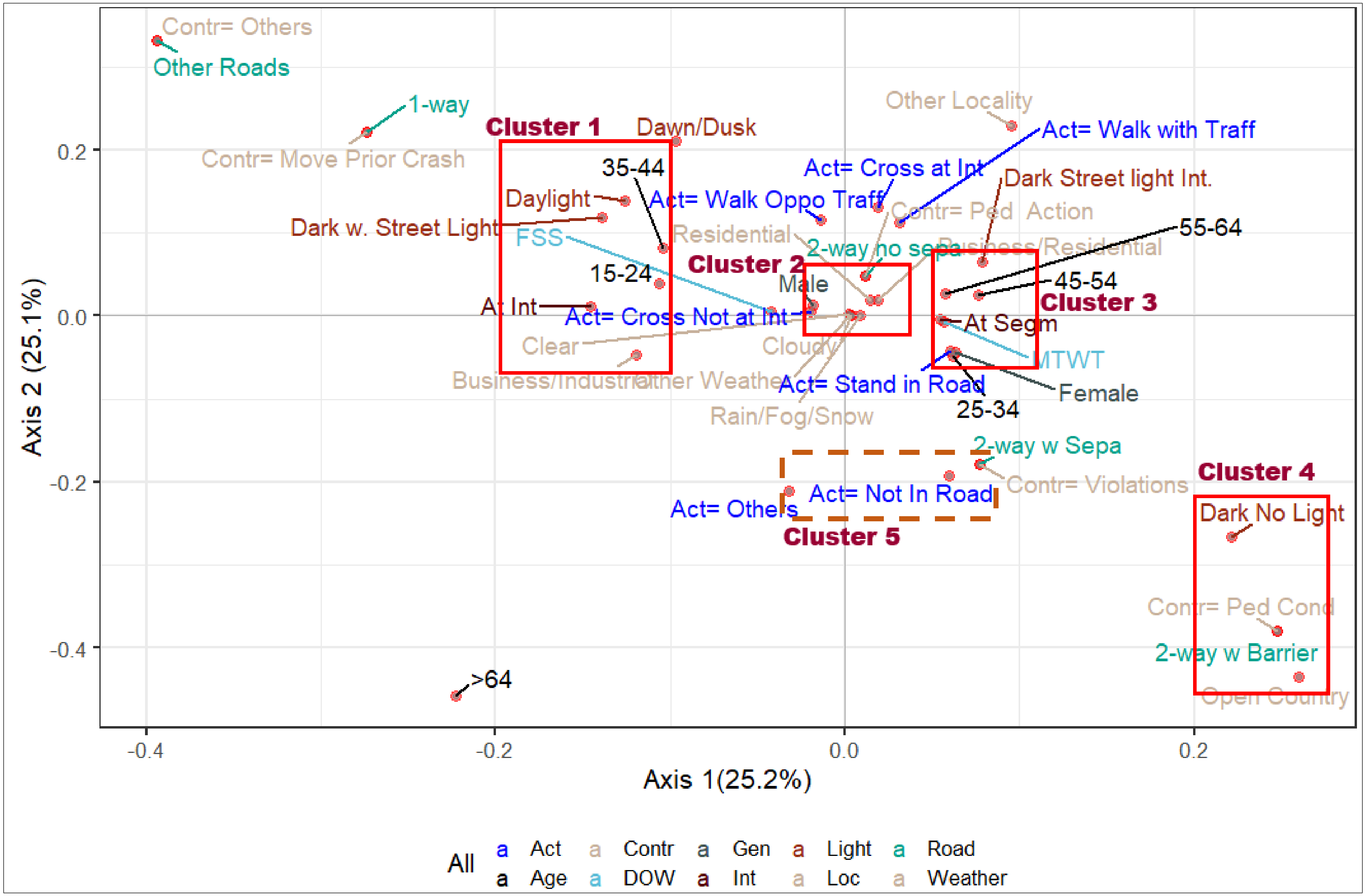

Figure 2 Biplot of exploratory variables.

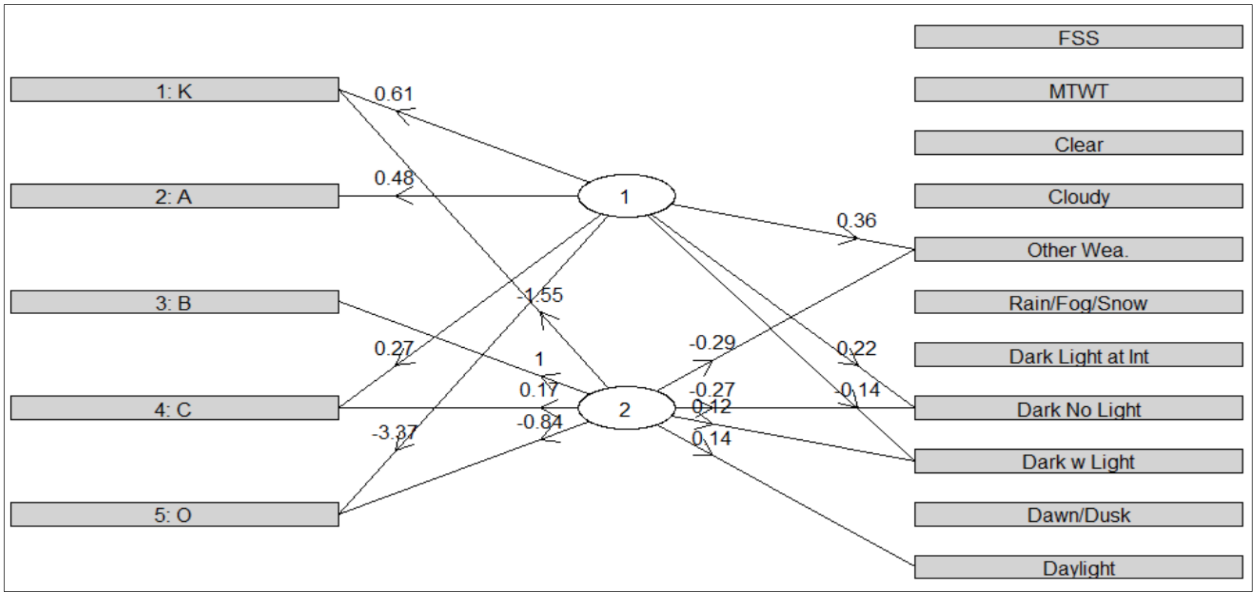

Figure 3 Association graph for temporal and environmental attributes.

Similar explanations can be developed for Figure 4 and Figure 5. Business/industrial and open country are associated with Axis 1 and business/residential and open country are associated with Axis 2. Axis 2 is more associated with roadway types than Axis 1, as Axis 1 is only associated with 1-way roads. The eigenvalue of 1-way 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’ is positive. As fatal and severe crashes have positive eigenvalues, 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.

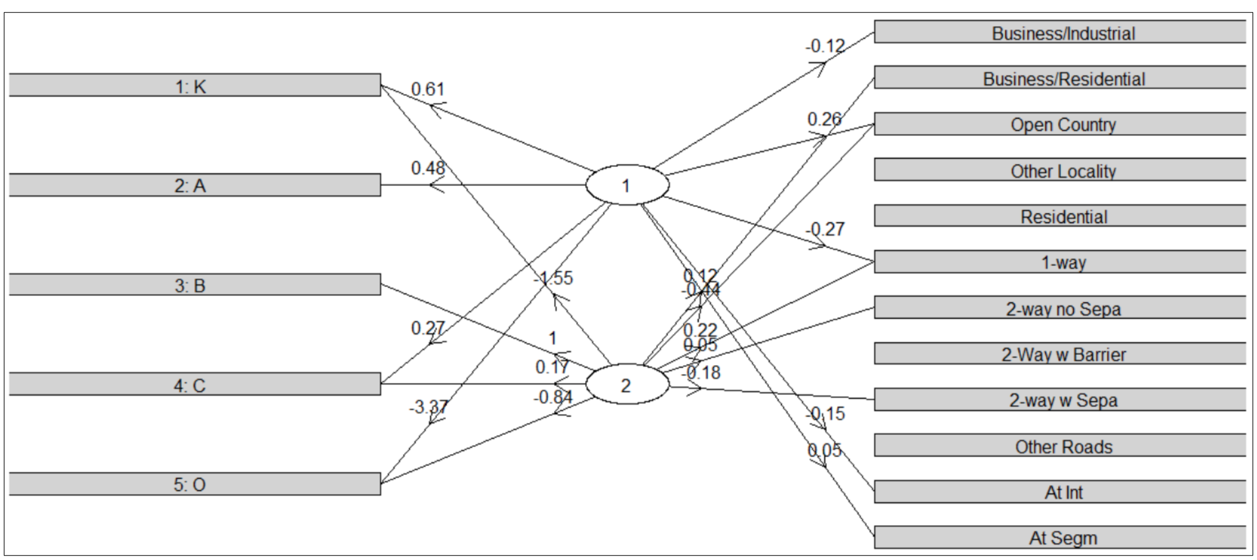

Figure 4 Association graph for site characteristics.

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. Axis 1 is associated with only one attribute (contributing factor as others), while Axis 2 is associated with all of the contributing factors. Two action types (crossing at intersection and others) are also associated with Axis 2.

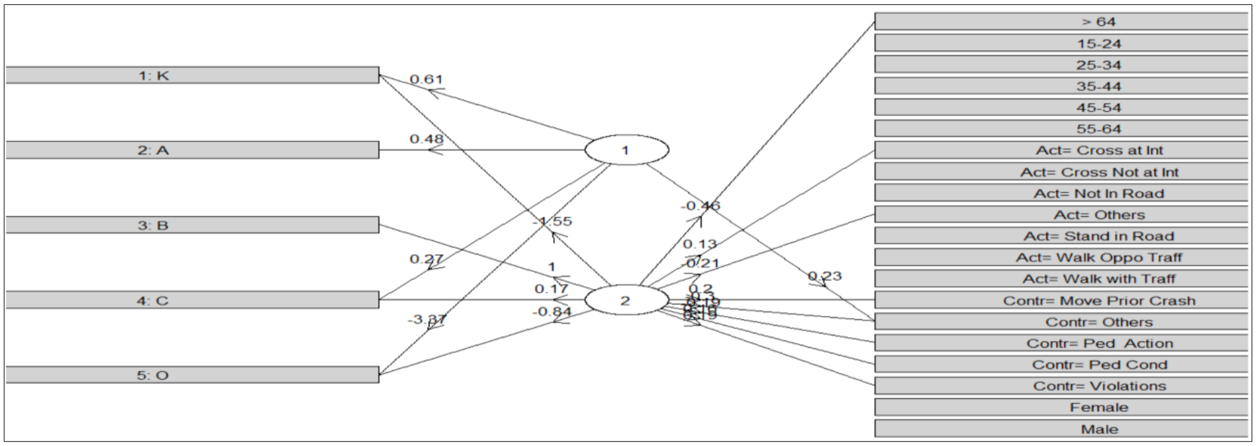

Figure 5 Association graph for pedestrian-related variables.

The coefficients of correspondence regression are essentially the coordinate scores of the categories on the latent axes. The fitted values use the coordinate scores in correspondence regression to compute predicted/expected frequencies for every cell in the cross table of the response and exploratory variables. Figure 6 illustrates the residual plots for different exploratory variables (by severity types) as indicated by number of the attributes (each number indicate an attribute; for example, age group 15-24 is an attribute) in the x-axis. The residual values indicate the difference between observed frequencies and predicted frequencies for every cell in the cross table of response and exploratory variables. The light grey shades indicate the 95% confidence boundaries developed for each crash severity group (K, A, B, C, and O). The residual plot indicates that there was no significant underestimation or overestimation relative to the original measures.

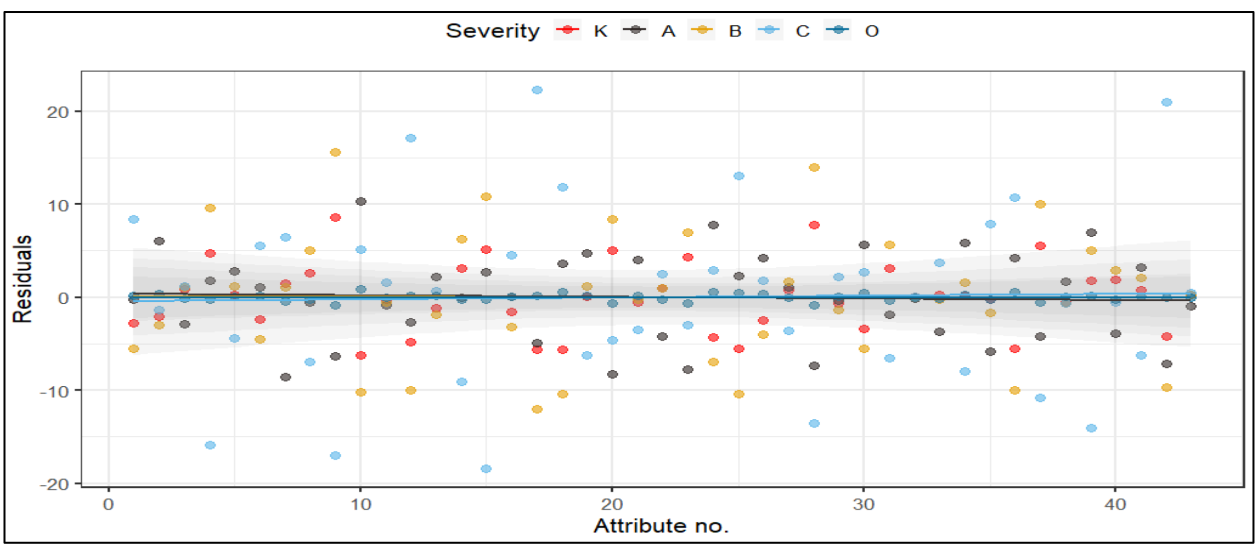

Figure 6 Residual plots by attributes.

CONCLUSIONS

Alcohol-induced impairment greatly affects human interaction and judgment, and it is one of the major contributing factors in roadway traffic crashes. Alcohol impairment can have dire consequences for both drivers and pedestrians on the roadway. Recent Louisiana crash data shows that PUI fatalities increased by approximately 120% from 2010 to 2016. The conventional approach to safety analysis has been to establish relationships between the geometric design of the roadways, traffic measures, and crash occurrence. However, these methods are limited due to the exclusion of human factors. Additionally, investigation of pedestrian crashes requires a deep understanding of the roadway features, pedestrian action, and the driver's reaction. This study applied correspondence regression to seven years of Louisiana PUI crash data to measure the effects of the key contributing factors. The findings show that male and female pedestrians demonstrate different action traits that affect their crash involvement. Females are more likely to be involved in intersection crashes, while males are more likely to be involved in midblock crashes. In addition to these human factors, some roadway features like dark with no light, two-lane roadway with no physical separation are highly associated with PUI crashes. The results of this study show that fatal PUI crashes are more associated with roadways with no lighting at night. The research team recommended several suitable countermeasures that can be used for reduction of PUI crashes. Examples of these countermeasures include pedestrian friendly intersection design, increasing pedestrian crossing phasing during the weekends at night, and encouraging ridesharing during weekends.

The current study is not without limitations. First, this study used a limited number of variables that are associated with PUI crashes; some variables, like roadway traffic condition, driver action, roadway visibility, were not considered in the analysis due to the high number of missing values in these variables. Second, the current study has not used the comprehensive crash typing data *(27)* 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. Limitations of the current study offer directions for future research in this domain.

DISCLAIMER

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

ACKNOWLEDGEMENT

The authors also appreciate the assistance provided by the students on this project: Magdalena Theel, and Bita Maraghehpour.

AUTHOR CONTRIBUTION STATEMENT

The authors confirm the contribution to the paper as follows: study conception and design: Subasish Das; data collection: Subasish Das; analysis and interpretation of results: Subasish Das; draft manuscript preparation: Subasish Das, Sruthi Ashraf, Anandi Dutta, and Ly-Na Tran; All authors reviewed the results and approved the final version of the manuscript.

REFERENCES

1. Zhang, G., K. K. Yau, X. Zhang. Analyzing Fault and Severity in Pedestrian–Motor Vehicle Accidents in China. Accident Analysis and Prevention, Vol. 73, 2014, pp. 141-150.

2. Retting, R. A., S. A. Ferguson, and A. T. McCartt. A Review of Evidence-Based Traffic Engineering Measures Designed to Reduce Pedestrian–Motor Vehicle Crashes. American Journal of Public Health, Vol. 93(9), 2003, pp. 1456-1463.
3. NHTSA. Traffic Safety Facts 2015 Data-pedestrians. Department of Transportation, National Highway Traffic Safety Administration. DOT-HS-812-375, Washington, DC, U. S. 2017. Available at https://crashstats.nhtsa. dot. gov/Api/Public/ViewPublication/812375. Accessed: August 2019.
4. 1NHTSA. Traffic Safety Facts 2016 Data - Pedestrians. https://crashstats.nhtsa.dot.gov/Api/Public/ViewPublication/812493.
5. Holubowycz, T. Age, Sex, and Blood Alcohol Concentration of Killed and Injured Pedestrians. p. 6.
6. Jones, R. K., and J. H. Lacey. State of Knowledge of Alcohol-Impaired Driving: Research on Repeat DWI Offenders. Publication DTNH22-98-C–05109. Mid-America Research Institute, 2000.
7. Crocker, P., O. Zad, T. Milling, and K. A. Lawson. Alcohol, Bicycling, and Head and Brain Injury: A Study of Impaired Cyclists' Riding Patterns R1. The American Journal of Emergency Medicine, Vol. 28, No. 1, 2010, pp. 68–72. https://doi.org/10.1016/j.ajem.2008. 09.011.
8. Hagemeister, C., and M. Kronmaier. Alcohol Consumption and Cycling in Contrast to Driving. Accident Analysis & Prevention, Vol. 105, 2017, pp. 102–108. https://doi.org/10.1016/j.aap.2017. 01.001.
9. Behnood, A., and F. Mannering. Determinants of Bicyclist Injury Severities in Bicycle-Vehicle Crashes: A Random Parameters Approach with Heterogeneity in Means and Variances. Analytic Methods in Accident Research, Vol. 16, 2017, pp. 35–47. https://doi.org/10.1016/j.amar. 2017. 08.001.
10. Helak, K., D. Jehle, D. McNabb, A. Battisti, S. Sanford, and M. C. Lark. Factors Influencing Injury Severity of Bicyclists Involved in Crashes with Motor Vehicles: Bike Lanes, Alcohol, Lighting, Speed, and Helmet Use. Southern Medical Journal, Vol. 110, No. 7, 2017, pp. 441–444. https://doi.org/10.14423/SMJ.0000000000000665.
11. Hezaveh, A. M., and C. R. Cherry. Walking under the Influence of the Alcohol: A Case Study of Pedestrian Crashes in Tennessee. Accident Analysis and Prevention, Vol. 121, 2018, pp. 64–70. https://doi.org/10.1016/j.aap.2018.09.002.
12. Miles-Doan, R. Alcohol Use among Pedestrians and the Odds of Surviving an Injury: Evidence from Florida Law Enforcement Data. Accident Analysis and Prevention, Vol. 28, No. 1, 1996, pp. 23–31. https://doi.org/10.1016/0001-4575(95)00030-5.
13. Lindsay, T. Characteristics of Alcohol Impaired Road Users Involved in Casualty Crashes, 2012.
14. Fitzpatrick, K., V. Iragavarapu, and M. A. Brewer. 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, p. 290.
15. Ortiz, N., and M. Ramnarayan. 1949 - Analysis of Alcohol-Involved Pedestrian Fatalities in the United States, 2003–2015. Journal of Transport and Health, Vol. 5, 2017, p. S93. https://doi.org/10.1016/j.jth.2017.05.250.
16. Demetriades, D., G. Gkiokas, G. C. Velmahos, C. Brown, J. Murray, and T. Noguchi. Alcohol and Illicit Drugs in Traumatic Deaths: Prevalence and Association with Type and

Severity of Injuries. Journal of the American College of Surgeons, Vol. 199, No. 5, 2004, pp. 687–692. https://doi.org/10.1016/j.jamcollsurg. 2004.07.017.

17. World Health Organization. WHO | Gender and Road Traffic Injuries. WHO. https://www.who.int/gender-equity-rights/knowledge/a85576/en/. Accessed: Jun. 14, 2019.
18. Živković, V., V. Lukić, and S. Nikolić. The Influence of Alcohol on Pedestrians: A Different Approach to the Effectiveness of the New Traffic Safety Law. Traffic Injury Prevention, Vol. 17, No. 3, 2016, pp. 233–237. https://doi.org/10.1080/15389588.2015.1054986.
19. Tagawa, M., M. Kano, N. Okamura, M. Itoh, E. Sakurai, T. Watanabe, and K. Yanai. Relationship between Effects of Alcohol on Psychomotor Performances and Blood Alcohol Concentrations. Japanese Journal of Pharmacology, Vol. 83, No. 3, 2000, pp. 253–260.
20. Oxley, J. A., E. Ihsen, B. N. Fildes, J. L. Charlton, and R. H. Day. Crossing Roads Safely: An Experimental Study of Age Differences in Gap Selection by Pedestrians. Accident Analysis & Prevention, Vol. 37, No. 5, 2005, pp. 962–971. https://doi.org/10.1016/j.aap.2005.04.017.
21. Clayton, A. B., M. A. Colgan, and R. J. Tunbridge. The Role of the Drinking Pedestrian in Traffic Accidents. p. 6.
22. Oxley, J., M. Lenné, and B. Corben. The Effect of Alcohol Impairment on Road-Crossing Behaviour. Transportation Research Part F: Traffic Psychology and Behaviour, Vol. 9, No. 4, 2006, pp. 258–268. https://doi.org/10.1016/j.trf.2006.01.004.
23. LaScala, E. A., F. W. Johnson, and P. J. Gruenewald. Neighborhood Characteristics of Alcohol-Related Pedestrian Injury Collisions: A Geostatistical Analysis. Prevention Science, 2001, p. 12.
24. Fotios, S., and R. Gibbons. Road Lighting Research for Drivers and Pedestrians: The Basis of Luminance and Illuminance Recommendations. Lighting Research & Technology, Vol. 50, No. 1, 2018, pp. 154–186. https://doi.org/10.1177/1477153517739055.
25. Plevoets, K. corregp: Open Source R Software Package. https://cran.r-project.org/web/packages/corregp/corregp.pdf . Accessed: July 2019.
26. Van der Heijden, P. G. M., A. De Falguerolles and J. De Leeuw. A Combined Approach to Contingency Table Analysis Using Correspondence Analysis and Loglinear Analysis. Applied Statistics, 38 (2), 1989, 249–292.
27. NHTSA. 2017 FARS/CRSS Pedestrian Bicyclist Crash Typing Manual A Guide for Coders Using the FARS/CRSS Ped/Bike Typing Tool, 2018.