Accident Analysis and Prevention 111 (2018) 43–55

Contents lists available at ScienceDirect

Accident Analysis and Prevention

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

Full length article

Investigation on the wrong way driving crash patterns using multiple correspondence analysis

Subasish Das[a,*], Raul Avelar[a], Karen Dixon[a], Xiaoduan Sun[b]

[a] Texas A&M Transportation Institute (TTI), 3135 TAMU, College Station, TX 77843-3135, United States
[b] Department of Civil Engineering, University of Louisiana at Lafayette, Lafayette, LA 70504, United States

A R T I C L E I N F O

Keywords:
Wrong way driving crashes
Multiple correspondence analysis
Dimensionality reduction
Contributing factors

A B S T R A C T

Wrong way driving (WWD) has been a constant traffic safety problem in certain types of roads. Although these crashes are not large in numbers, the outcomes are usually fatalities or severe injuries. Past studies on WWD crashes used either descriptive statistics or logistic regression to determine the impact of key contributing factors. In conventional statistics, failure to control the impact of all contributing variables on the probability of WWD crashes generates bias due to the rareness of these types of crashes. Distribution free methods, such as multiple correspondence analysis (MCA), overcome this issue, as there is no need of prior assumptions. This study used five years (2010–2014) of WWD crashes in Louisiana to determine the key associations between the contribution factors by using MCA. The findings showed that MCA helps in presenting a proximity map of the variable categories in a low dimensional plane. The outcomes of this study are sixteen significant clusters that include variable categories like determined several key factors like different locality types, roadways at dark with no lighting at night, roadways with no physical separations, roadways with higher posted speed, roadways with inadequate signage and markings, and older drivers. This study contains safety recommendations on targeted countermeasures to avoid different associated scenarios in WWD crashes. The findings will be helpful to the authorities to implement appropriate countermeasures.

1. Introduction

Wrong way driving (WWD) crashes on different roadways are considered as constant traffic safety problems. Although wrong way crashes are not large in numbers, the outcomes of these crashes tend to involve disproportionally higher number of fatalities or serious injuries. According to Pour-Rouholamin and Zhou (2016), "WWD crashes happen when a driver, inadvertently or deliberately, drives against the main direction of traffic flow on a controlled-access highway". A study conducted by Friebele et al. (1971) mentioned that "the wrong-way driver, travelling head-on into an unsuspecting traffic stream, is simply a time bomb ticking off the seconds toward a possible disaster". Pour-Rouholamin et al. (2014) found 1.34 fatalities per fatal WWD crashes in the U.S. from 2004 to 2013, while for other crashes the fatalities per fatal crash rate is 1.10 during the same time period. According to National Highway Traffic Safety Administration (NHTSA) statistics, around 350 people are killed each year nationwide due to WWD crashes (NHTSA, 2013). In Louisiana, around 300 WWD crashes (0.2% of total crashes) happened every year. Around 0.45% of total crashes in Louisiana are fatal crashes, but for wrong way crashes this percentage is higher (around 1.6% of the total WWD crashes). Thus, it is crucial to identify key risk factors associated with WWD crashes.

The Federal Highway Administration (FHWA) Highway Safety Improvement Program (HSIP) includes a project to monitor WWD crashes and identify hot spots of WWD crashes. It includes a wrong-way study warrant based on total crash and fatal crash rates. The National Transportation Safety Board (NTSB) recommends that the FHWA develop a HSIP policy memorandum for use by state department of transportation agencies to establish wrong-way monitoring programs (NTSB, 2012). The outcomes of the monitoring programs can help in developing improved signage and marking as well as technology like wrong way navigation alerts on vehicles. For an effective monitoring program, determining key association factors in WWD crashes would be particularly helpful.

One of the major tasks in highway safety analysis is the identification of the key contributing factors for different types of crashes. Multiple Correspondence Analysis (MCA) is a dimensionality reduction method useful to describing the significance of co-occurrence of groups of variables or variable categories from a high dimension dataset. This method is also referred to as the pattern recognition method that treats

* Corresponding author.
E-mail addresses: s-das@tti.tamu.edu (S. Das), r-avelar@tti.tamu.edu (R. Avelar), k-dixon@tti.tamu.edu (K. Dixon), xsun@louisiana.edu (X. Sun).

https://doi.org/10.1016/j.aap.2017.11.016
Received 21 November 2016; Received in revised form 3 November 2017; Accepted 13 November 2017
Available online 21 November 2017
0001-4575/

arbitrary data sets as combination of points in a larger dimensional space. It uniquely simplifies complex data into knowledge extraction in a completely different way than parametric estimation does. In MCA analysis, the objective is to investigate associations between multiple variables, as opposed to the more traditional characterization of associations between a set of predictor variables and a single response variable of interest (i.e., number of crashes).

The study team used five years (2010–2014) of Louisiana WWD (for the remainder of this paper Louisiana wrong way crashes, both driving and cycling, will be referred as WWD crashes for consistency) crashes to determine the relationship of the variables and their significance. The objectives of this study are: (1) to identify the relative closeness of the key association factors to determine meaningful co-occurrence, and (2) to recommend countermeasures when appropriate. The findings of this study could help authorities to determine effective and efficient crash countermeasures.

2. Literature review

Although traffic safety research includes an extensive array of research areas, the most prominent are- crash frequency analysis, and crash severity analysis. Lord and Mannering (2010) provided a detailed overview of the properties of crash-frequency data and associated methodological alternatives and limitations for examining such data. Savolainena et al. (2011) provided a similar assessment on crash-severity analysis. Recently, Mannering and Bhat (2014) bridged and extended the previous studies of Lord and Mannering (2010) and Savolainena et al. (2011) by overviewing both count data models and crash severity models. Interested readers can consult these studies as well as a hyperlinked webpage developed by Das (2016) for further information. That webpage lists 592 research papers on statistical and algorithmic methods as well as hyperlinks to all corresponding papers.

The literature review reveals a surge of research on WWD crashes since 2014. Table 1 described the research efforts conducted on WWD crashes starting from 1971. Most of the studies used freeway as the main interest group. Few studies focused on all roadways or divided roadways (Ponnaluri, 2016; Kemel, 2015). Many studies performed descriptive statistics to describe the nature of the factors in WWD crashes (Friebele et al., 1971; Copelan, 1989; Cooner et al., 2004; Braam, 2006; Scaramuzza and Cavegn, 2007; SWOV, 2009; Morena and Leix, 2012; Finley et al., 2014; Xing, 2014; Zhou et al., 2015; FDOT, 2015). In many cases, simple descriptive statistics should not suffice to explaining the impact of the contributing factors. It is also important to note that some of these studies focused more on operational considerations than safety (Friebele et al., 1971; Copelan, 1989; Cooner et al., 2004; Braam, 2006; Finley et al., 2014). Several authors explored the analysis of crash outcomes and crash types using a modeling approach. Some studies simply used logistic regression models to differentiate between WWD and non-WWD crashes (Kemel, 2015; Ponnaluri, 2016). As WWD crashes are very small in numbers compared to non-WWD crashes, this small sample size problem is likely to significantly influence the outcomes and statistical power of the models. Pour-Rouholamin et al. (2014) used Firth's penalized-likelihood logistic regression to control the influence of all confounding variables on the probability of WWD crashes while considering the rareness of the WWD event. Pour-Rouholam and Zhou (2016) used generalized ordered logistic regression to perform crash severity analysis using WWD crashes.

The idea of MCA begins in 1970 with French Statistician Jean-Paul Benźecri (Roux and Rouanet, 2010), though there are similarities with Principal Component Analysis (PCA) and Factor Anaysis (FA), two well documented multivariate statistical methods. PCA mainly deals with numerical data, and MCA is a well-accustomed tool for multi-dimensional categorical data.

MCA has been reinvented many times under different frameworks while keeping the goals similar (De Leeuw, 1973; Hoffman and De Leeuw, 1992). A limited number of studies has been conducted in applying MCA in the transportation safety research. Hoffman and De Leeuw (1992) interpreted MCA as multidimensional scaling method and associated different vehicle models with crash severities. Fontaine (1995) performed MCA on one year of pedestrian crash data to determine the statistical proximity of the significant factors. This study identified few distinctive groups as a basis for more in depth analysis. Factor et al. (2010) applied MCA in determining the association between driver's social characteristics and their involvements in crash severities. This study exposed new facets in the social organization of fatalities. Das and Sun (2015) used eight years (2004–2011) of pedestrian crash data in Louisiana to determine key associations between risk factors. This study determined several significant groups of factors that require deeper exploration in future. Xu et al. (2016) used quasi-indiced exposure method to identify the key factors contributed to pedestrian crashes in Las Vegas from 2004 to 2008. This study later used MCA to determine the interaction between different factors. Das and Sun (2016) applied MCA on eight years (2004–2011) of fatal run-of-road (ROR) crashes in Louisiana to examine the degree of association between risk factors. Das et al. (2017) recently applied MCA on the second Strategic Highway Research Program's (SHRP 2) Washington Roadway Inventory Database (RID) to identify the key association factors for inclement weather crashes. The finding revealed some specific factor groups that require careful attention from the safety professionals.

Table 2 shows variables used in previous studies addressing wrong-way driving crashes using different methods. These past studies will be used to inform the exploratory analysis presented in the following sections.

3. Theory of multiple correspondence analysis

MCA is an unsupervised learning algorithm. In MCA, one does not need to distinguish between explanatory variables and the response variable. It requires the construction of a matrix based on pairwise cross-tabulation of each variable. For example, the dimension of the final dataset of this study is: 1203 × 24. For a table of qualitative or categorical variables with dimension 1203 × 24, MCA can be explained by taking an individual record (in row), *i [i = 1 to 1203],* where 24 categorical variables (represented by 24 columns) have different sizes of categories. MCA can generate the spatial distribution of the points by different dimensions based on these 24 variables.

Let P be the number of variables (i.e., columns) and I is the number of transactions (i.e., rows). This will generate a matrix of *'I multiplied by P'*. If L_p is the number of categories for variable p, the total number of categories for all variables is, $L = \sum_{p=1}^{P} L_p$. It will generate another matrix *'I multiplied by L'*. In this matrix, each of the variables will contain several columns to show all of their possible categorical values.

The cloud of categories is considered as a weighted combination of J points. Category j is represented by a point denoted by C^j with weight of n_j. For each of the variables, the sum of the weights of category points is n. In this way, for the whole set J the sum is nP. The relative weight w_j for point C^j is $w_j = n_j/(nP) = f_j/P$. The sum of the relative weights of category points is 1/P, which makes the sum of the whole set as 1.

$$w_j = \frac{n_j}{nP} = \frac{f_j}{P} \quad \text{with} \ \sum_{j\in J_q} w_j = \frac{1}{P} \qquad \text{and} \ \sum_{j\in J} w_j = 1$$

Here, $n_{jj'}$ represents the number of individual records which have both categories k and k'. The squared distance between two categories C^j and $C^{j'}$ can be represented by

$$\left(C^j C^{j'}\right)^2 = \frac{n_j + n_{j'} - 2n_{jj'}}{n_j n_{j'}/n} \tag{4}$$

The numerator of Eq. (4) is the number of individual records associating with either j or j' but not both. For two different variables, p and p', the denominator is the familiar "theoretical frequency" for the cell (j, j') of the $J_p \times J_{p'}$ two-way table.

Table 1
Studies on WWD Crashes.

No.	State	Study Period	Roadway	Factors	Method	Ref
1	Texas	1967–1970	Freeway	Entrance by exit ramp; Diamond interchange; partial interchange; less than 1000 feet of sight distance; impaired driver; improper signing	Descriptive statistics	Friebele et al. (1971)
2	California	1983–1987	Freeway	Darkness; Intoxicated drivers; Time of the day; Urban areas; Interchanges with short sight distance; Partial cloverleaf interchanges; Half and full diamond interchanges; Trumpet interchanges; Slip ramps; Buttonhook ramps; Scissors exit ramp; Left-side exit ramp; Five-legged intersections near exit ramps	Descriptive statistics	Copelan (1989)
3	Texas	1997–2000	Freeway	Early morning hours; Male drivers; Drivers less than 34 years old; Intoxicated drivers; Left-side exit ramps; Urban areas; One-way street transitioned into freeway	Descriptive statistics	Cooner et al. (2004)
4	North Carolina	2000–2005	Freeway	Alcohol-related; Younger drivers; Older drivers; Interstate routes; Rural areas; Time of day (midnight to 5:59 a.m.); Months of February and June; Two-quadrant parclo interchanges; Full diamond interchanges	Descriptive statistics	Braam (2006)
5	Switzerland	2003–2005	Motorways	Younger drivers; Intoxicated drivers; Older drivers; Time of day; Female drivers; Lighting condition	Descriptive statistics	Scaramuzza and Cavegn (2007)
6	Netherlands	1996–1998	Motorways	Older drivers; Younger drivers; Inexperienced drivers; Intoxicated drivers	Descriptive statistics	SWOV (2009)
7	New Mexico	1990–2004	Freeway	Darkness; Intoxicated drivers; Older drivers; Male drivers; Passenger cars; Month of November; Non-Hispanic and native Americans	Comparison group	Lathrop et al. (2010)
8	Michigan	2005–2009	Freeway	Darkness; Intoxicated drivers; Time of the day (late night and early morning); Younger drivers; Older drivers; Male drivers; Parclo interchanges; Trumpet interchanges; Tight diamond interchanges	Descriptive statistics	Morena and Leix (2012)
9	Texas	2007–2011	Freeway	Time of day (7:00 p.m. to 12:00 p.m.); Younger drivers (16–24 years); Male drivers; Impaired drivers;	Descriptive statistics	Finley et al. (2014)
10	Japan	2005–2009	Motorways	Older drivers; Younger drivers; Darkness; Type of interchange; Making a U-turn on carriageway; Dementia; Time of day (4:00 p.m. to 10:00 p.m.);	Descriptive statistics	Xing (2014)
11	Illinois	2004–2009	Freeway	Darkness; Older Driver; Younger drivers; Male drivers; Local drivers; Intoxicated drivers; Time of day (midnight to 5:00 a.m.); Head-on crashes; Weekends; Urban areas; Type of interchange; Passenger cars; Single-occupant vehicles;	Descriptive statistics	Zhou et al. (2015)
12	Alabama	2009–2013	Freeway	Time of the day (evening and afternoon); Older drivers; Intoxicated drivers; Physically impaired drivers; Driver residency distance (local drivers); Vehicles older than 15 years; Roadway condition; Months of March, May, and November	Firth's Penalized-Likelihood Logistic Regression	Pour-Rouholamin et al. (2014)
13	France	2009–2012	Divided roadways	Darkness; Older drivers; Intoxicated drivers; Local drivers; Driving older vehicles; Passenger cars; Single-occupant vehicles; Unlicensed drivers;	Logistic regression	Kemel (2015)
14	Alabama	2009–2013	Freeway	Time of the day (evening and afternoon); Older drivers; Intoxicated drivers; Physically impaired drivers; Driver residency distance (local drivers); Vehicles older than 15 years; Roadway condition; Months of March, May, and November	Generalized ordered logit	Pour-Rouholamand Zhou (2016)
15	Florida	2009–2013	Freeway	Head-on crashes; Months of January through April, June, and July; Weekend; Impaired drivers; Darkness; Younger drivers;	Descriptive statistics	FDOT (2015)
16	Florida	2003–2010	All roadways	Driver age; Driver gender; Driver condition (eyesight, fatigue, illness, seizure, epilepsy); Intoxicated drivers; Time of the day; Urban areas; Darkness; Rainy and foggy weather; Vehicle use; Day of week; AADT	Logistic regression	Ponnaluri (2016)

Table 2
Significant variables used in past studies.

Variable	Variable used in Studies
Driver	
Age	Cooner et al. (2004), Braam (2006), Scaramuzza and Cavegn (2007), Lathrop et al. (2010), Scaramuzza and Cavegn (2007), SWOV (2009), Morena and Leix (2012), Finley et al. (2014), Xing (2014), Zhou et al. (2015), FDOT (2015), Ponnaluri (2016), Pour-Rouholamin et al. (2014), Pour-Rouholamin and Zhou (2016)
Gender	Cooner et al. (2004), Lathrop et al. (2010), Morena and Leix (2012), Finley et al. (2014), Ponnaluri (2016)
License state	Pour-Rouholamin et al. (2014), Pour-Rouholamin and Zhou (2016), Kemel (2015)
Impairment	Friebele et al. (1971), Copelan (1989), Braam (2006), Cooner et al. (2004), Scaramuzza and Cavegn (2007), SWOV (2009), Lathrop et al. (2010), Morena and Leix (2012), Finley et al. (2014), Zhou et al. (2015), Pour-Rouholamin et al. (2014), Pour-Rouholamin and Zhou (2016), FDOT (2015), Ponnaluri (2016)
Condition	Ponnaluri (2016)
Severity	Pour-Rouholamin and Zhou (2016)
Reason or contributing factor	Caltrans (2015), Scaramuzza and Cavegn (2007)
Violation type	Caltrans (2015), Scaramuzza and Cavegn (2007)
Vehicle	
Vehicle age	Kemel (2015)
Vehicle type	Lathrop et al. (2010), Zhou et al. (2015), Kemel (2015)
Headlight	Finley et al. (2014), FDOT (2015), Caltrans (2015)
Temporal	
Season	Braam (2006), Lathrop et al. (2010), Pour-Rouholamin et al. (2014), Pour-Rouholamin and Zhou (2016), FDOT (2015)
Day of the week	Zhou et al. (2015), FDOT (2015)
Geometric/Environmental	
Urban/Rural areas	Copelan (1989), Braam (2006), Zhou et al. (2015), Ponnaluri (2016)
Roadway type	FDOT (2015), Caltrans (2015)
Darkness	Copelan (1989), Scaramuzza and Cavegn (2007), Lathrop et al. (2010), Morena and Leix (2012), Zhou et al. (2015), FDOT (2015), Ponnaluri (2016)
Access control	Simpson and Bruggeman (2015)
Posted speed	Caltrans (2015), FDOT (2015), Finley et al. (2014), Xing (2014), Lathrop et al. (2010)
Roadway feature or countermeasure	Simpson and Bruggeman (2015), FDOT (2015), Finley et al. (2014)
Locality	Caltrans (2015), Cooner et al. (2004), FDOT (2015), Finley et al. (2014)
Weather	Ponnaluri (2016), Finley et al. (2014), FDOT (2015), Simpson and Bruggeman (2015)
Traffic	
Traffic volume	Ponnaluri (2016), NTSB (2012)
Traffic control	Simpson and Bruggeman (2015), FDOT (2015), Finley et al. (2014)
Crash	
Collision Types	Pour-Rouholamin and Zhou (2016), FDOT (2015),SWOV (2009)

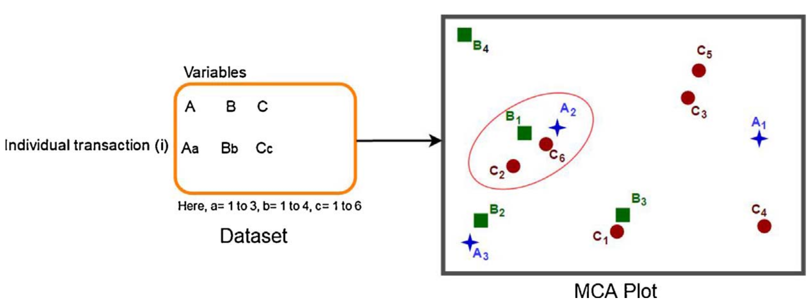

Fig. 1. MCA plot different categories.

Fig. 1 shows an easy representation of the cloud generation for combination of categories. In this figure, three variables are considered. Variable A has 3 categories (A_1, A_2, A_3), variable B has 4 categories (B_1, B_2, B_3, and B_4), and variable C has 6 categories (C_1, C_2, C_3, C_4, C_5, and C_6). The categories are plotted in the MCA plot representing their relative proximity on the two dimensional space. The plot shows a distinct cloud (red ellipse) associated with A_2, B_1, C_2, and C_6. This cloud is created based on the proximity of the coordinates of these 4 categories.

3.1. Limitations and comparison with other methods

MCA is a powerful tool to recognize patterns and associations in a dataset with multiple categorical variables. In the context of crash analysis, this approach focuses on the associations between the covariates of crashes rather than the associations between each covariate with the frequency or the odds of crashes, which is the emphasis of regression modeling tools. Rather than answering the question in regression analysis "How strong is the association of variable X with the frequency (or odds) of crashes, after accounting for other significant associations?" MCA focuses on answering the questions "Given a set of crashes that have occurred, how strongly group of variables do concentrate within this dataset?" The question is answered based on the relationships observed in a plot of orthogonal principal axes, as shown in Fig. 1. As with every other method based on cross-sectional data, MCA can only find patterns or correlations that must be interpreted carefully as statistical associations and avoiding unwarranted statements of causation.

4. Data

4.1. Data collection

To achieve the research objectives, this study used state maintained traffic crash database compiled from 2010 through 2014 in Louisiana.

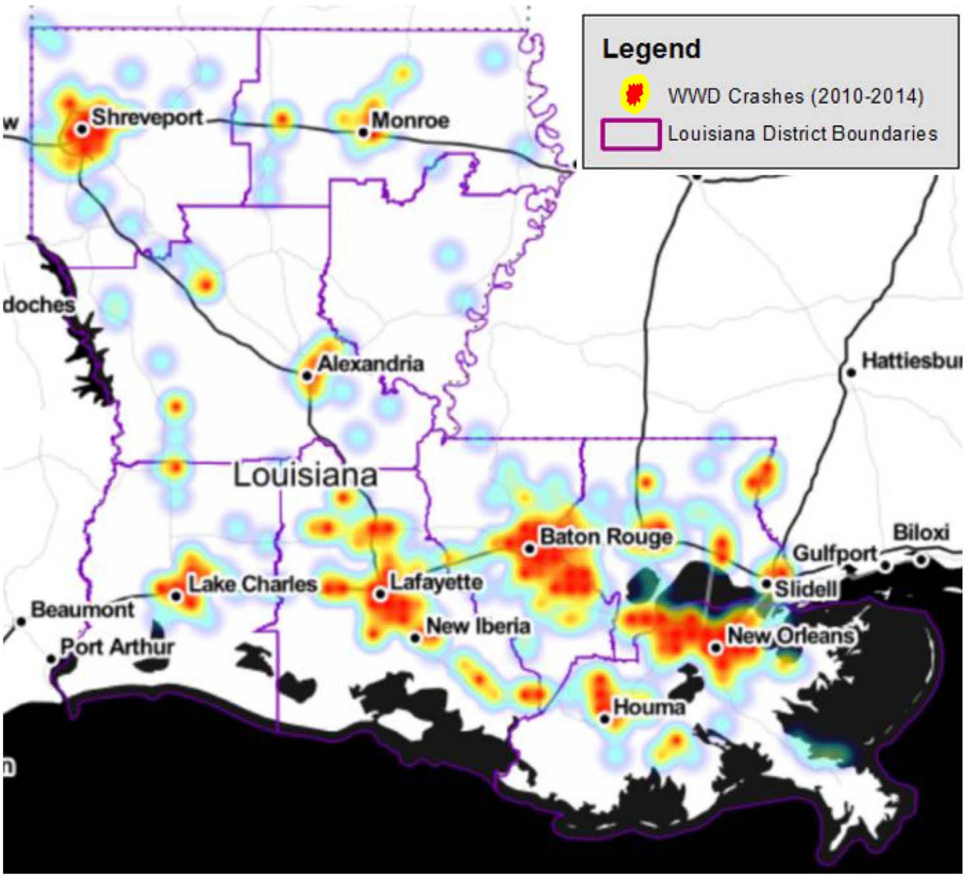

Fig. 2. WWD Crashes in Louisiana.

The primary dataset was prepared by merging information from three different databases (crash, roadway geometry, and vehicle). Altogether, 1873 possible WWD crashes were identified from the total 800,000 crashes from 2010 to 2014 in Louisiana. After performing an extensive review on these crashes, 1203 crashes were finally confirmed as WWD crashes. These crashes involved 1947 individuals. As the current research focus was to identify pattern of association factors involved with police reported at-fault drivers, the final dataset was prepared based on 1203 at-fault crash incidents. The police reported crash data has an identifier (VEH_NUM) in determining at-fault drivers. Value of 1 in this identifier (VEH_NUM) indicates at-fault drivers. The reporting police office provided this number to the vehicle or person based on his investigation on the crash occurrence site. Fig. 2 shows the spatial distribution of the WWD crashes in Louisiana. It clearly shows that WWD crashes mostly occur in populous cities and urban areas (highlighted by red center with yellow surroundings).

After removing non-pertinent information, a more precise database was prepared based on key contributing factors. The variable section method used the research findings from the past studies, as shown in Table 2.

4.2. Descriptive statistics

Table 3a and b shows the descriptive statistics of the final selected variables. Heat maps used in these tables will help to visualize the relative weightage of different categories.

4.2.1. Driver characteristics

In Louisiana WWD crashes, around 1.75% crashes are fatal in nature. When compared with non-WWD crashes, this rate is around 3 times higher. This finding is supported by other studies. For example, Pour-Rouholamin and Zhou (2016) found 1.34 fatalities per fatal WWD crashes in the U.S. from 2004 to 2013, while for other crashes the fatalities per fatal crash rate is 1.10 during the same time period. Driver impairment is found significant (around 19.31%) in Louisiana WWD crashes. Many studies concluded that driver impairment is significantly associated with WWD crashes (Friebele et al., 1971; Copelan, 1989; Braam, 2006; Cooner et al., 2004; Scaramuzza and Cavegn, 2007; SWOV, 2009; Lathrop et al., 2010; Morena and Leix, 2012; Finley et al., 2014; Zhou et al., 2015; Pour-Rouholamin et al., 2014; Pour-Rouholamin and Zhou, 2016; FDOT, 2015; Ponnaluri, 2016). Younger and older drivers are usually more prone towards WWD crashes. Louisiana crash data showed similar trends like other studies showed for younger drivers (Cooner et al., 2004; Braam, 2006; Scaramuzza and Cavegn, 2007; SWOV, 2009; Morena and Leix, 2012; Finley et al., 2014; Xing, 2014; Zhou et al., 2015; FDOT, 2015; Ponnaluri, 2016) and older drivers (Braam, 2006; Scaramuzza and Cavegn, 2007; SWOV, 2009; Lathrop et al., 2010; Morena and Leix, 2012; Xing, 2014; Zhou et al., 2015; Pour-Rouholamin et al., 2014; Pour-Rouholamin and Zhou, 2016; Ponnaluri, 2016). Distracted drivers showed higher trends in WWD crashes in Louisiana. Pour-Rouholamin and Zhou (2016) also associated inattention in older ages with higher number of WWD crashes. Table 1 also shows that male drivers were nearly twice in numbers when compared with female drivers. Similar findings are found in other studies (Cooner et al., 2004; Lathrop et al., 2010; Morena and Leix, 2012; Finley et al., 2014; Ponnaluri, 2016). Driver error (driver condition and driver violation) contributed significantly in WWD crashes. Similar results were found in Caltrans (2015) and Scaramuzza and Cavegn (2007). Driver's license state plays a key role in WWD crashes. Non local drivers usually end in WWD crashes due to the non-familiarity of the surroundings. In non-WWD crashes, around 18% of the crashes involved with out-of-state licensed drivers. On the other hand, this percentage is nearly double (31.42%) for WWD crashes. This finding is also in line with other studies (Pour-Rouholamin et al., 2014;

Table 3
Descriptive statistics.

a

Attribute	Frequency	%	Attribute	Frequency	%
Driver Severity			**Reason**		
Fatal	21	1.75%	Driver Violation	599	49.79%
Severe	34	2.83%	Driver Condition	236	19.62%
Moderate	148	12.30%	Normal Movement	92	7.65%
Complaint	303	25.19%	Vehicle Condition/Avoid things	48	3.99%
No Injury	697	57.94%	Other	228	18.95%
Driver Impairment			**Vehicle Headlight**		
Yes	260	21.61%	Headlights On	516	42.89%
No	694	57.69%	Headlights Off	325	27.02%
Not Recorded	249	20.70%	Daytime Running Lights	44	3.66%
Driver Age			Other	318	26.43%
15–24	229	19.04%	**Vehicle Type**		
25–34	272	22.61%	Passenger Car	587	48.79%
35–44	179	14.88%	Bus/Truck	284	23.61%
45–54	132	10.97%	Suv/Van	243	20.20%
55–64	96	7.98%	Motorcycle	21	1.75%
65–74	54	4.49%	Other	68	5.65%
74 plus	241	20.03%	**Vehicle Year**		
Driver Condition			1996–2005	554	46.05%
Inattentive/Distracted	422	35.08%	2005–2015	393	32.67%
Normal	224	18.62%	Older than 20 years	103	8.56%
Impaired	234	19.45%	Not Recorded	153	12.72%
Fatigue/Illness	44	3.66%	**Urban-Rural Type**		
Other	279	23.19%	Urban	539	44.80%
Driver Gender			Rural	96	7.98%
Male	667	55.44%	Not Recorded	568	47.22%
Female	375	31.17%	**Locality**		
Not Recorded	161	13.38%	Business, Mixed Residential	401	33.33%
Driver Violation			Business Continuous	317	26.35%
Improper Driving	418	34.75%	Residential District	226	18.79%
Fail to Control	211	17.54%	Open Country	111	9.23%
Driver Condition	132	10.97%	Residential Scattered	67	5.57%
Following Too Closely	11	0.91%	Manufacturing Or Industrial	44	3.66%
No Violations	27	2.24%	Other	37	3.08%
Other	404	33.58%			
Driver License State					
Louisiana	825	68.58%			
Other	378	31.42%			

b

Attribute	Frequency	%	Attribute	Frequency	%
Road Type			**Lighting Condition**		
Two-Way Road With No Physical Separation	416	34.58%	Daylight	599	49.79%
Two-Way Road With A Physical Separation	372	30.92%	Dark – Continuous Street Light	377	31.34%
One-Way Road	338	28.10%	Dark – No Street Lights	123	10.22%
Two-Way Road With A Physical Barrier	65	5.40%	Dark – Street Light At Intersection Only	65	5.40%
Other	12	1.00%	Dawn/Dusk	24	2.00%
Posted Speed			Other	15	1.25%
20 mph or less	177	14.71%	**Collision Type**		
21–30 mph	320	26.60%	Head-On	293	24.36%
31–40 mph	284	23.61%	Right Angle	206	17.12%
41–50 mph	190	15.79%	Sideswipe	210	17.46%
51–60 mph	142	11.80%	Non-Collision With Motor Vehicle	161	13.38%
61–70 mph	90	7.48%	Left Turn/Right Turn	90	7.48%
Countermeasure			Rear End	67	5.57%
White Dashed Line	337	28.01%	Other	176	14.63%
No Control	203	16.87%	**Day of the Week**		
Yellow No Passing Line	118	9.81%	Weekday	829	68.91%
Sign	88	7.32%	Weekend	374	31.09%
Signal	82	6.82%	**Season**		
Yellow Dashed Line	62	5.15%	Fall	277	23.03%
Other	313	26.02%	Spring	176	14.63%
Access Control			Summer	307	25.52%
Full Control	135	11.22%	Winter	443	36.82%
No Control	941	78.22%	**Weather**		
Partial Control	113	9.39%	Clear	943	78.39%
Other	14	1.16%	Cloudy	138	11.47%
Traffic Control			Rain	91	7.56%
Controls Functioning	936	77.81%	Fog/Smoke	11	0.91%

(*continued on next page*)

Table 3 (*continued*)

b

Attribute	Frequency	%	Attribute	Frequency	%
No Controls	207	17.21%	Other	20	1.66%
Other	60	4.99%			
Traffic Volume (vpd*)					
Less than 20,000	265	22.03%			
20,000–40,000	197	16.38%			
Over 40,000	173	14.38%			
Not Recorded	568	47.22%			

Note: * vpd = Vehicle per day.

Pour-Rouholamin and Zhou, 2016; Kemel, 2015).

4.2.2. Vehicle characteristics

Roadway sign and markings are keenly associated with vehicle headlights at night. State DOT WWD studies in Texas and California included this factor (Finley et al., 2014; FDOT, 2015; Caltrans, 2015). In Louisiana, vehicles with headlights off contributed largely in WWD crashes. Large vehicles (Bus/Truck, Suv/Van) contributed nearly 44% of WWD crashes in Louisiana. Other studies also showed similar trends (Lathrop et al., 2010; Zhou et al., 2015; Kemel, 2015). Kemel (2015) showed that vehicles more than 15 years old (respectively less than 5 years old) were over-represented among WWD crashes. Louisiana WWD crashes showed similar trends. Vehicles older than 10 years showed the higher (55%) involvement in WWD crashes.

4.2.3. Geometric/Environmental properties

Urban roadways contained the larger number of WWD crashes in Louisiana. Other studies showed similar result (Copelan, 1989; Braam, 2006; Zhou et al., 2015; Ponnaluri, 2016). Absence of adequate lighting at night is a key contributing factor in WWD crashes. Other studies supported this trend (Copelan, 1989; Scaramuzza and Cavegn, 2007; Lathrop et al., 2010; Morena and Leix, 2012; Zhou et al., 2015; FDOT, 2015; Ponnaluri, 2016). Two way roads were associated with higher crashes in Louisiana study. This finding is not in line with the findings of other studies probably because most of past research focused on freeway crashes. Around 28% of Louisiana WWD crashes happened on one way roadways. The majority of these one-way crashes were on exit ramp. Few studies associated impact of posted speed on WWD crashes (Caltrans, 2015; FDOT, 2015; Finley et al., 2014; Xing, 2014; Lathrop et al., 2010). In the Louisiana study, higher percentage of crashes is seen in lower speed roadways. Access control is a less studied variable in WWD crashes. Roadways with no access control were associated with 78% of WWD crashes. Very few studies (Simpson and Bruggeman, 2015; FDOT, 2015; Finley et al., 2014) considered presence of countermeasure as a significant factor. Finley et al. (2014) found that higher number of crashes happened on roadways either with white lines or with no controls. This study also used locality characteristics as a factor. Few studies (Simpson and Bruggeman, 2015; FDOT, 2015; Finley et al., 2014) considered such characteristics as a significant factor. Finley et al. (2014) found that locality with business entities was more prone to WWD crashes. Additionally, it was found that impact of inclement weather is significant for WWD crashes. Similar findings are mentioned in other studies (Ponnaluri, 2016; Finley et al., 2014; FDOT, 2015; Simpson and Bruggeman, 2015).

4.2.4. Traffic characteristics

Traffic control feature in Louisiana indicates the status of the traffic control devices in the surrounding roadways. Table 2 shows that higher number of WWD crashes occurred on low volume roadways.

4.2.5. Crash characteristics

Head-on collisions were higher in number in WWD crashes. Other studies found similar findings (Pour-Rouholamin and Zhou, 2016; FDOT, 2015; SWOV, 2009).

5. Analysis

In MCA, the approach is to analyze the rows and columns of a dataset while treating them as high-dimension geometry elements. The target is to show the co-occurrence of the categories in a lower dimensional space where proximity in the space potentially indicates meaningful associations among the categories. Graphical representations in MCA help to interpret data in a convenient way as they effectively summarize large, complex datasets by simplifying the structure of the associations between variable categories with a relatively simple view of the data (Greenacre and Blasius, 2006). A larger distance indicates a distant association. If the distance for a particular category is very far away from the centroid, it indicates that such category is different from the average profile.

Each of the categories is independent in principal and a co-occurrence based on weight proximity (by associating certain categories together in a cloud) tends to form a complete picture of certain scenarios. A percentage distribution for a single category conveys little meaning in many cases, but when combined with other categories due to the proximity in the space, a variety of implications can be potentially interpreted.

Table 4 shows the percentages of variance explained by the top 10 dimensions. The first principal axis explained 6.43% of the principal inertia, the second principal axis explained 5.32%. These percentages are calculated based on the eigenvalue (a value in between 0 and 1). Dimension with larger variance has a higher eigen-value magnitude. The first two dimensions explained 11.75% of variance and none of the remaining major dimensions explained more than 3.17%. As this percentage indicates a lower representation of the data, this study used another method (joint correspondence analysis, JCA) to identify the overall inertia of the first two planes. This method identifies the coverage around 78%. The comparison between MCA and JCA shows similar positions of the attributes up to the MCA axis limits. It indicates the MCA used in this study can identify significant associations even

Table 4
Percent variance explained in top 10 dimensions.

Dimensions	Eigen Value	% of Variance	Cumulative% of Variance
Dimension 1 (Dim 1)	0.238	6.428	6.428
Dimension 2 (Dim 2)	0.197	5.321	11.749
Dimension 3 (Dim 3)	0.117	3.169	14.918
Dimension 4 (Dim 4)	0.105	2.845	17.763
Dimension 5 (Dim 5)	0.091	2.445	20.207
Dimension 6 (Dim 6)	0.083	2.239	22.446
Dimension 7 (Dim 7)	0.079	2.118	24.564
Dimension 8 (Dim 8)	0.069	1.858	26.421
Dimension 9 (Dim 9)	0.066	1.779	28.200
Dimension 10 (Dim 10)	0.065	1.761	29.961

Table 5
Significance of Key Variables on the First Plane.

Dimension 1			Dimension 2		
	R^2	p-value		R^2	p-value
Dr_Impairment	0.535	< 0.001	Dr_Condition	0.631	< 0.001
Dr_Condition	0.531	< 0.001	Dr_Impairment	0.589	< 0.001
Reason	0.444	< 0.001	Dr_Gender	0.375	< 0.001
Posted_Speed	0.397	< 0.001	Reason	0.383	< 0.001
Dr_Gender	0.346	< 0.001	Dr_Age	0.318	< 0.001
Dr_Age	0.342	< 0.001	Lighting	0.308	< 0.001
Urban_Rural	0.314	< 0.001	Veh_Headlight	0.213	< 0.001
Traffic_Volume	0.290	< 0.001	Posted_Speed	0.212	< 0.001
Veh_Headlight	0.248	< 0.001	Locality	0.192	< 0.001
Locality	0.252	< 0.001	Collision	0.189	< 0.001
Veh_Yr	0.241	< 0.001	Dr_Lic_State	0.163	< 0.001
Countermeasure	0.206	< 0.001	Access_Control	0.158	< 0.001
Dr_Lic_State	0.173	< 0.001	Veh_Type	0.118	< 0.001
Traffic_Control	0.151	< 0.001	Urban_Rural	0.099	< 0.001
Dr_Violation	0.159	< 0.001	Traffic_Volume	0.088	< 0.001
Road_Type	0.153	< 0.001	Road_Type	0.087	< 0.001
Lighting	0.148	< 0.001	Dr_Violation	0.072	< 0.001
Access_Control	0.113	< 0.001	Veh_Yr	0.066	< 0.001
Veh_Type	0.108	< 0.001	Traffic_Control	0.052	< 0.001
Dr_Severity	0.089	< 0.001	Dr_Severity	0.055	< 0.001
Collision	0.090	< 0.001	Weather	0.050	< 0.001
Weather	0.066	< 0.001	Countermeasure	0.050	< 0.001
Day	0.006	0.01	Day	0.037	< 0.001
Season	0.006	0.03	Season	0.004	< 0.001

maintaining lower inertia by the two major planes.

Table 5 lists the significance of each of the variables in both dimensions. The higher R^2 values indicates higher association between the dimension and the variable. The most dominant variables in dimension 1 are: driver impairment, driver condition, reason, posted speed, and driver gender. For dimension 2, the most dominant variables are: driver condition, driver impairment, driver gender, reason, and driver age. The values from Table 4 indicate that dimension 1 and 2 both were governed by driver related variables. Posted speed is found as a top contributing variable in dimension 1.

Fig. 3 reveals the significance of the contributing categories. For both dimensions, driver related categories are most dominant. The other dominant categories are: vehicle lighting categories, street lighting categories, speed categories, locality categories, urban-rural categories, and countermeasure categories.

To verify the outcomes of MCA analysis, bootstrap resampling techniques were used to produce confidence regions on the first plane. In practice, around 100 replications were done to produce ellipses having similar shapes and sizes with similar categories. The bootstrap replication scheme shows 81.32% similarity of the clusters as shown in Fig. 4.

6. Results and discussion

It is important to note that the contribution of the variables depends on its number of categories, whereas the contribution of a category depends on the number of incidents coded under its categories. Fig. 4 shows the proximity distribution of all listed categories. Category coordinates are the weighted average of the individual incidents occurring in that category. Most of the variables show certain level of differences (exceptions are seen in driver gender and season). This study used open-source R software package 'FactoMineR' to perform MCA (R Core Team, 2016; Husson et al., 2014). This package also produces related graphics from the analysis. The research team used two other visualization packages 'ggplot2' and 'ggrepel' to develop MCA plots (Wickham, 2009; Slowikowski, 2016). The flexibility of functions in these two packages makes the graphics tidier. The three clusters shown in Fig. 4 are separately shown in three other figures (Fig. 5a–c) based on their coordinate values.

MCA is a powerful tool to distinct non-trivial categories in the MCA plot (shown in rectangles with dotted lines). One non-trivial cloud in Fig. 4 is shown in ellipses with dotted lines. Clouds 1–16 are shown in ellipses with solid lines. These are group of categories in the dataset that are associated in direct proportion to their relative distance in the plane.

In order to facilitate the interpretation of the clouds, two additional patterns of interest have been coded in Fig. 4: Triangular markers indicating the severity of crashes (most severe in red, least severe in green) and pentagonal markers indicating impaired drivers (blue) or non-impaired drivers (purple). From these two patters it is clear that dimension 1 is an inverse indicator of severity and driver impairment (over this dimension's negative range), while dimension 2 is a direct indicator of these variables (severity over its positive range, and driver impairment over both its positive and negative range). Therefore, a cloud of categories with low value in dimension 1 and a high in dimension 2 tends to generally associate with higher severity and presence of impaired drivers (such as clouds 1 and 2). Additionally, both dimensions in their positive extremes are direct indicators of crashes with unspecified characteristics in driver condition, driver impairment, gender, reason, traffic control, access control, vehicle year, road type, lighting, and weather (per the categories in cloud 4). Finally, it should be mentioned that the dimensions are associated with each and every category, so similar patterns as those described above can be inferred for each category. However, severity and driver impairment are features of great interest to this research and therefore will help to interpret the associations found in the first plane.

Finally, the plane associations are in a continuum rather than a discrete space. Even though clouds of categories can be identified by their proximity, proximity between clouds is also meaningful in the plane. Given this general pattern, the following are further interpretations of the additional clusters recognized in the first plane.

6.1. Cloud 1 (driver severity = fatal, posted speed = 60–70 mph, locality = open country)

The first cloud is associated with three significant factors. Many studies showed that higher posted speed is a significant factor in traffic fatalities (for example, Cooner et al., 2004; Finley et al., 2014). Interestingly, this cloud also specifies particular locality characteristics closely associating 'open country'. This association suggests a group of unintentional WWD crashes in rural, isolated conditions. Drivers approaching a freeway entry point in open country are more likely to make a wrong decision and enter the facility the wrong way. Once a wrong-way entry has occurred, it is more likely that a fatality would occur at locations with higher posted speeds. The location of this cloud also indicates a closer association with driver impairment than with not-impaired drivers, an association that has also been studied before (Finley et al., 2014). Countermeasures like reflective and raised wrong-way pavement arrow markers or proper signage would be helpful in mitigating much crashes.

6.2. Cloud 2 (lighting = dark with no street light, access control = full control, road type = two way road with physical barrier, urban_rural = rural)

The second cloud is associated with four factors. It suggests that full-controlled rural two lane roadways with a physical separation are likely to confluence in recorded WWD crashes. The potential explanation, again in the context of unintentional WW entry, is that once such WW entry has occurred on full-access controlled locations with physical separation in no-lit condition, the chances of a WW crash occurrence are increased, compared to locations with other access control features and without physical barriers in well-lit condition. This cloud is also associated with rural locations, which resonates with the first cloud

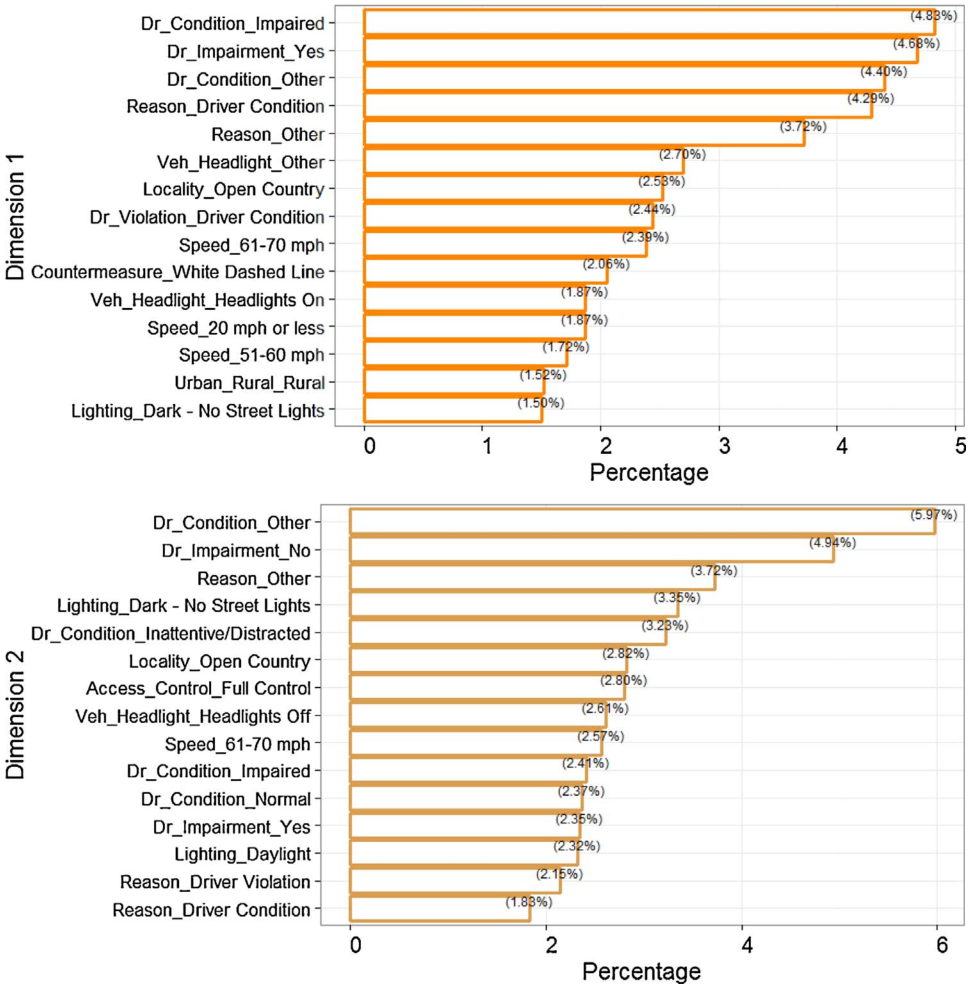

Fig. 3. Significance of top fifteen categories in Dimension 1 and Dimension 2.

described earlier. It should be kept in mind that, by its location in the plane, this cloud of conditions tends also to associate with crashes of high severity and with driver impairment.

Finally, the condition "no street light" is in the rough vicinity of both Cloud 1 and Cloud 2 in Fig. 4. Naturally, no street light relates to dark condition, but it is interesting that this analysis found dark conditions to closely correlate with fatal crashes, with driver impairment, rural environments, and full control and divided facilities. Countermeasures like adequate lighting at night in these sites would be beneficial.

6.3. Cloud 3 (traffic volume = over 40,000, driver severity = severe, countermeasure = white dashed line)

The second cloud is associated with three factors. It indicates that full-controlled rural two lane roadways with a physical separation are prone to WWD crashes. This cloud indicates that white dashed line is not enough to give visual guidance on avoiding wrong way driving. Proper wrong way sign and reflective and raised wrong-way pavement arrow markers would help in decreasing such crashes.

6.4. Cloud 4 (driver age = 74 plus, driver condition = other, reason = other, traffic control = other, vehicle type = other, driver impairment = not recorded)

1. This cloud suggests that crashes involving drivers older than 75 years of age tend to differentiate from other contributing factors (i.e., no driver impairment, vehicle types not classified, and reason, traffic control and driver condition not typified either). However, its distance to impaired or not impaired driver condition is roughly equal, thus indicating no clear association. The relationship to severity is not as strong as crashes with categories in cloud 1 and 2 and the three clusters in Fig. 5. However, by their position in the plot, these crashes tend to associate with low to moderate severity. Low mount oversized wrong way signs and oversized pavement arrows are particularly helpful in assisting older drivers to avoid WWD.

6.5. Cloud 5 (lighting = dark with intersection lighting, collision = head-on, road type = two way road with a physical separation)

This cloud suggests that head-on collisions may occur if a vehicle enters into a two way road with physical separation wrongfully in dark with intersection lighting. This scenario is possible due to lack of critical wrong way signage and signalization. This type of crash is also associated with moderate severity and more likely to involve impaired drivers than not. Geometric improvement and in-pavement warning lights would help mitigating these crashes.

6.6. Cloud 6 (urban_rural = urban, traffic volume = 40,000 or less, locality = residential, scattered, headlight = headlight on)

WWD crashes mostly happen on urban roadways due to their complex environment. This cloud indicates that WWD crashes mostly occur on medium volume roadway in urban residential areas. This cloud also includes that the vehicle headlights are turned on in these crashes, a circumstance that potentially indicate the dark lighting condition WWD crashes associate with a lack of other critical signage and signalization. Similar to the above type, this type of crash is also

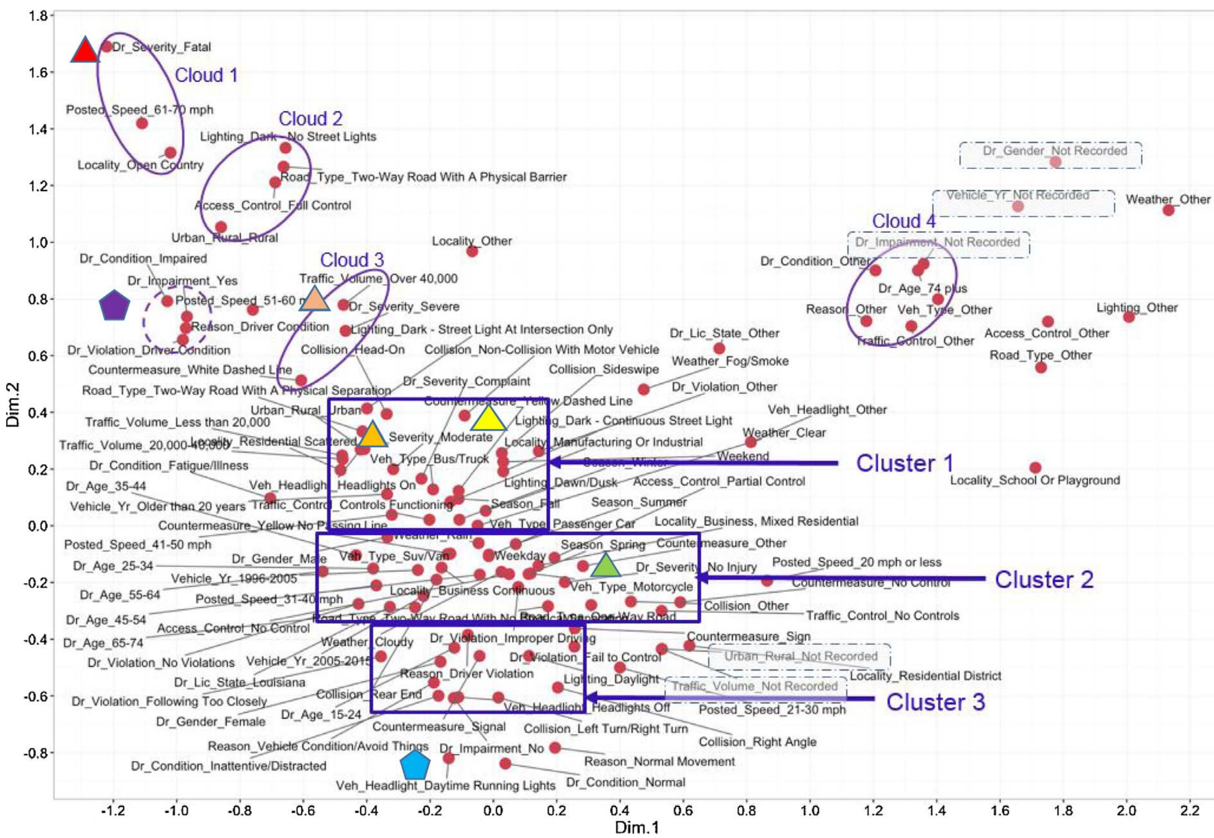

Fig. 4. Principal MCA plot for the variable categories.

associated with moderate severity and more likely to involve impaired drivers than not. Traffic calming countermeasures would be beneficial in avoiding these crashes.

6.7. Cloud 7 (driver severity = moderate, vehicle = older than 20 years, posted speed = 41–50 mph, driver gender = male)

This particular cloud indicates male drivers' involvement in WWD crashes while using older vehicles. These WWD crashes occurred mainly in roadways with moderate posted speed and these crashes involved moderate driver injuries. Driver education and strict state policies on vehicle maintenance would be helpful.

Other significant clouds (Cloud 8–Cloud 16) indicate various key association groups. This study identifies several countermeasures to mitigate such scenarios:

Cloud 8: This cloud indicates that large vehicle WWD crashes mostly occurred in manufacturing or industrial locations during dawn or dusk. The main feature of the roadway is that the locations were on roadways with yellow dashed lines. The temporal factor (Season = Fall/Winter) may indicate higher proportion of large truck movements compared to other seasons. Improvement of roadway signage and geometric features are needed to minimize large vehicle WWD crashes in the industrial areas.

Cloud 9: This cloud indicates that weekend collisions under dark with continuous lighting are mostly associated with sideswipe WWD crashes. The confluence of this set of factors suggests, perhaps, that sideswipe WW crashes seem to associate with moderately limited visibility and reduced traffic (i.e., weekend). Additionally, WWD crashes at night during weekend indicate possible association with intoxication or violation (non-identified violation is in the proximity of this cloud). Targeted enforcement and regulations would be helpful in reducing such crashes.

Cloud 10: Car and SUV/Vans are more likely to be involved in crashes during weekday. The crashes happened on locations with no passing lane. An interpretation of this association is that the drivers were involved in improper passing related crashes. It should be noted that compared to Fig. 5a, the scales in Fig. 5b and c are significantly differ, and thus, even though clouds 10 through 16 can be seen as separate groups of associated factors, each of these groups is more similar to its adjacent clouds than the average distance between categories in other clouds. Therefore, this study considers that when interpreting these clouds, it should be kept in mind that these clouds are subsets of bigger clusters, as shown in Fig. 4. However, the reduction of the scale may also result in individual clouds whose interpretation may not be practical, perhaps due to just randomness within the clusters. Adaptation of advanced warning systems (both infrastructure and in-vehicle) and cautionary signs/signals would be beneficial to reduced = these crashes.

Cloud 11: Locality specific WWD crashes require special attentions. Countermeasures like signage improvement, reflective raised pavement marker installation, and geometric feature improvement can be adopted to minimize these crashes.

Cloud 12: Motorcycle drivers are more likely to be involved in WWD crashes on business locations. This pattern associates with two harmful conditions: cloudy weather and no access control. However, severities of these crashes tend to be the lowest and slightly more likely to involve non-impaired drivers, rather than involve impaired drivers. Similar to the above category, these types of crashes could be intentional violations. In busy constraint environments, motorcycles may decide to travel in the opposite lane for shorter distances to gain on or avoid queues. Law enforcement presence may help alleviate these situations.

Cloud 13: Roadways with no controls are always more likely to be contributing in WWD crashes. This cloud associates with two major roadway types: one-way roadways or two-way roadways with no physical separation.

Cloud 14: It indicates that local older drivers (apparently involved in

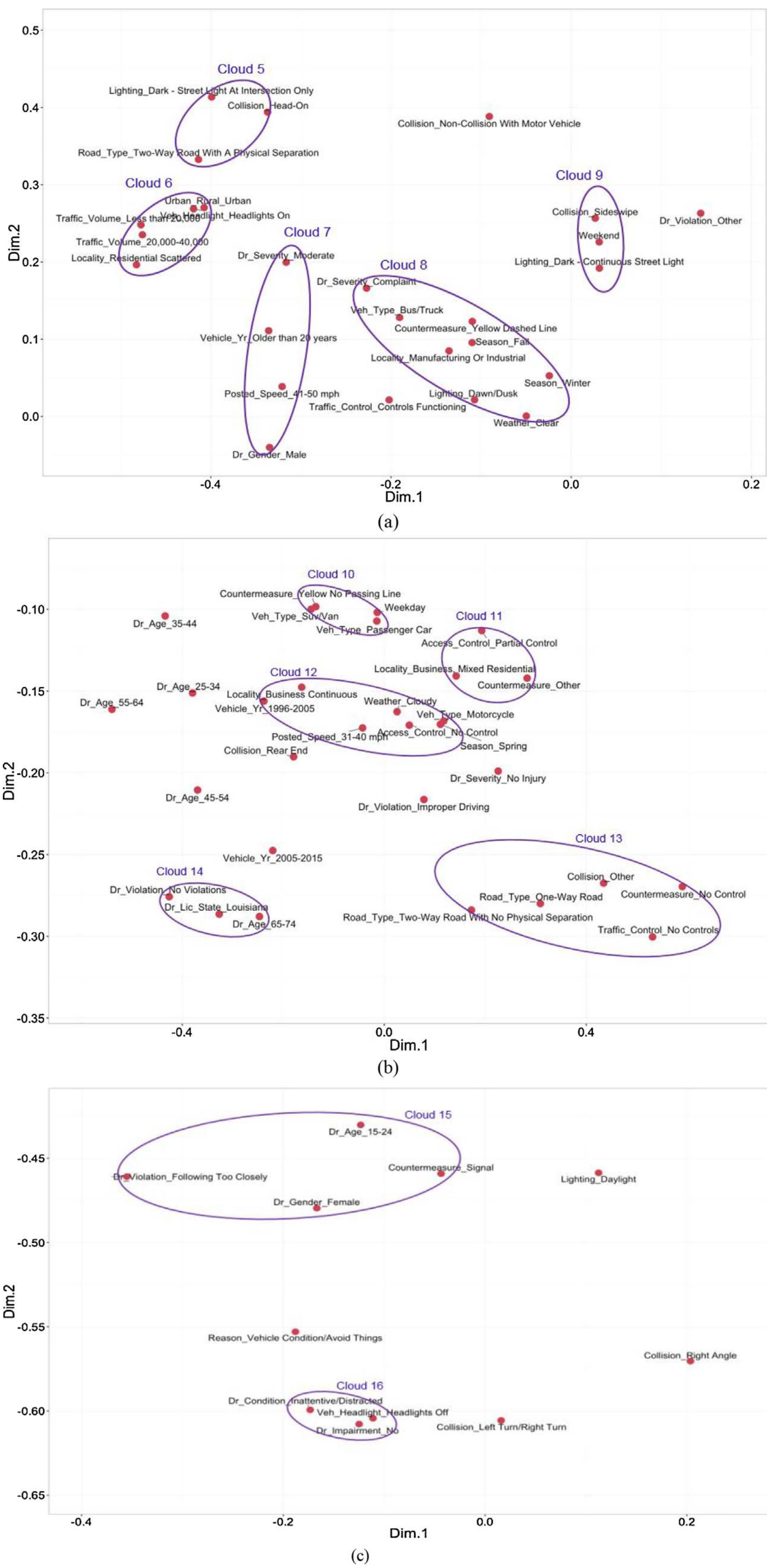

Fig. 5. MCA Plot for Cluster 1, 2, and 3.

no violations) are more likely to be involved in WWD crashes. Better education and policy guidance are beneficial in minimizing young driver related WWD crashes. Oversized 'do not enter' and 'wrong way' signs or lower mounter signs can help older drivers in avoiding WWD crashes.

Cloud 15: At signalized intersections, young female drivers were associated with driver violations (following too closely). Maneuvering through intersections requires better attention on the part of the drivers. Except for the gender specificity, this finding is consistent with previous research indicating an increased risk for teens of involvement in any type of crashes.

Cloud 16: This cloud indicates that distracted drivers (but not impaired) tended to be involved in WW crashes when their vehicle headlights were turned off. To decrease distraction related WWD crashes, better education, training, and safety campaigns are necessary.

7. Recommended countermeasures, policy implications, and study limitations

This study applied MCA on five years of WWD crashes in Louisiana. Variable selection was performed by the systemic literature review of past studies. As WWD crashes were associated with significant numbers of categorical factors, MCA is considered an extremely valuable tool to explore the association between these large numbers of variable categories. MCA helps in presenting proximity map of the variable categories in a low dimensional plane by revealing the main features from a multi-dimensional dataset. The key groups of the confluence of factors in WWD crashes found in this research (with addition of related countermeasures and policy implications) are discussed below:

- Locations with higher posted speed are associated with a specific locality condition where an open country condition exists. These types of crashes have a clear tendency to be most severe as well. These types of crashes should be the focus of safety programs intended to reduce WWD crashes. Possible countermeasures include improved signage and traffic control devices at such locations. Recent research has shown that WWD detection systems show promise in deterring drivers at entry points in freeway ramps (see Finley et al., 2014). From a policy perspective, it is desirable that regional or local agencies at risk move to implement pilot projects using technology alternatives and advanced signing to deter these most severity WWD instances.
- Rural areas with no lighting at night at full access control, divided facilities are associated with higher number of WWD crashes. These crashes tend to also have involvement of impaired drivers, rather than not. Similar to the recommendations above, possible countermeasures include improved signage and adequate lighting at such locations. Policymakers can use this finding in prioritizing initiatives or projects to reduce WWD crashes, such as pilot projects with WWD detection technology.
- Roadways with no control and roadways with no physical separation are more likely to be associated with WWD crashes. Adequate traffic control devices and geometric improvements can help reduce these incidents. However, expensive countermeasures such as physical separations require long term planning and justifications for the policymakers to proceed. However, in the case of no control, this finding may bear important implications for current policies in the U.S., where the Manual on Uniform Traffic Control Devices (FHWA, 2009) is the reference document for most control countermeasures. According to this document, no traffic control is recommended within a median opening at locations with medians narrower than 30 ft (9.14 m), but the finding of this study suggest that the use of traffic control countermeasures, such as Stop, Yield, or 'Do Not Enter' signs, may prove cost effective ways to prevent WWD instances at certain locations.
- Types of locality play a dominant role in WWD crashes. Open country, urban residential areas, and industrial zones show different types of key associations. For some locations, WWD crashes are more likely to be intentional. Targeted law enforcement at problem areas should help discourage intentional violations, and traffic calming countermeasures may help alleviate the severity of these crashes. This finding might have important policy implications. In the U.S., the guidelines in the MUTCD for WWD countermeasures do not distinguish between high speed and low speed facilities, nor between rural vs. urban environments. This research is in agreement with previous work that crash risk factors and specific characteristics vary by locality and might require differentiated guidelines for practitioners when selecting appropriate countermeasures.
- This study showed that countermeasures like white dashed edge line, yellow dashed line, and yellow no passing zone are not adequate enough in reducing WWD crashes. Innovative signing and pavement marking (low mounted wrong way sign, reflective wrong way arrows, pavement arrows, and flashing wrong way indicators) can be considered as targeted countermeasures.
- Crashes involving drivers older than 75 years of age seem to appear as a meaningful factor by itself, without a strong relationship with severity (though more likely to result in less severe crashes). Young female drivers involved in WWD crashes were associated with traffic violations. Possible countermeasures include low mounted visible signs and indicators for older drivers, and targeted enforcement for younger drivers. The knowledge gained from this study will help the policy makers in advocating WWD crash awareness to appropriate populations under risk of WWD involvement.

It is important to note that the groups of confluence factors are developed based on the first plane, which represents 12% variance of the complete database. However, the JCA method, which explains 78% variance of the data, has similar groups of confluence factors. Thus, the MCA exploration, as articulated in this WWD study, hopes to serve as a guide to safety professionals, transportation planners, and policy makers. The findings from this study will help the policy makers to introduce or develop new standards to deploy appropriate countermeasures based on the need from region or zone specific WWD safety facts.

A graphical display like MCA plot is useful for the general audience as it shows complex associations involving many factors into a lower dimensional space. One limitation of this study is that it fails to include any significance test for the cloud groups. It is important to note that other cluster techniques like K-means clustering and Principal Component Analysis (PCA) have similar strengths and disadvantages like as the MCA. Future research can improve the findings of MCA by connecting it statistical models like log-linear and risk models. Another limitation is that the study findings are based on the first plane that explained only 12% of the data inertia. The inclusion of more dimensions can potentially increase the number of association patterns underlying the dataset. This study compared the results by using joint correspondence analysis (JCA), which indicates high percentage (78%) of explained variance. Exploration on JCA could be considered as a future research scope. The conventional MCA was improved by using Kohonen algorithm and different variants of self-organizing map (SOM) algorithms were developed for better predication (Du and Swamy, 2014). Future research can consider using SOM to develop predictive models for WWD crashes.

8. Conclusions

Many research efforts on conventional crash data have been conducted to understand better the factors that influence the frequency and severity of WWD crashes and to provide more effective safety-related countermeasures. However, the number of WWD crashes is still at an unacceptable level, which is evident by the recent statistics. It shows that, in addition to current efforts, research needs to be conducted with additional resources and in newer directions. Moreover, the conventional way of associating effect of a single factor on the response variable is not sufficient to characterize the complex nature of a crash occurrence. The findings from this study demonstrate that MCA would be a viable tool in analyzing complex categorical data in search of meaningful associations between categorical factors. Conventional statistical modeling requires supervised data (clear definition of explanatory and response variables) and prior assumptions. Therefore, these models are more appropriate at a later stage of analysis, when the focus is on answering specific and clearly delimited questions. MCA, a distribution free method, does not require any assumptions and it can work on unsupervised data, and is most valuable when the dataset is approached without a specific hypothesis or set of hypotheses in mind. In MCA, the target is to show the co-occurrence of the categories in a low dimensional space where proximity in the space indicates the similarity of the categories. This method helps in understanding diverse variable categories and produces visual results from the key associations.

As the WWD crash dataset has limited number of cases, removing entries with noise would make a small dataset smaller. Applying MCA offers the advantage of removing noise (by representing the data in low dimensional spaces) without reducing the dataset. This feature helps to describe the significant associations between the categories of the complex dataset like WWD crashes. Prioritization of certain key association groups as well as target countermeasures listed in this study would help authorities in reducing WWD crashes.

References

Braam, A.C., 2006. Wrong-way Crashes: Statewide Study of Wrong-way Crashes on Freeways in North Carolina. North Carolina Department of Transportation, North Carolina, USA.
Caltrans Division of Research, Innovation and System Information, 2015. Wrong-Way Driving Prevention Methods. Caltrans.
Cooner, S.A., Cothron, A.S., Ranft, S.E., 2004. Countermeasures for Wrong-Way Movement on Freeway: Overview of Project Activities and Findings. Texas Transportation Institute, Texas, TX Publication FHWA/TX-04/4128-1.
Copelan, J.E., 1989. Prevention of Wrong Way Accidents on Freeways. California Department of Transportation, California, USA Publication FHWA/CA-TE-89-2.
Das, S., Sun, X., 2015. Factor association with multiple correspondence analysis in vehicle–pedestrian crashes. Transportation Research Record: Journal of the Transportation Research Board, No. 2519. Transportation Research Board, Washington, D.C, pp. 95–103.
Das, S., Sun, X., 2016. Factor Association knowledge for fatal run-off-road crashes by Multiple Correspondence Analysis. IATSS Res. 39 (2), 146–155.
Das, S., Brimley, B., Lindheimer, T., Pant, A., 2017. Safety Impacts of Reduced Visibility in Inclement Weather. Report: ATLAS-2017-19, Ann Arbor, MI.
Das, S. Statistical and Algorithmic Models on Crash Analysis: Research Papers and Abstract. Accessed November 2016. http://subasish.github.io/pages/TRB2016/crash.html.
De Leeuw, J., 1973. Canonical Analysis of Categorical Data. PhD Thesis. Republished in 1985 by DSWO- Press. University of Leiden, Leiden.
Du, K., Swamy, M., 2014. Neural Networks and Statistical Learning. Springer-Verlag, London.
Florida Department of Transportation, 2015. Statewide Wrong Way Crash Study. FDOT Financial Project ID: 190258-1-32-37.
Federal Highway Administration (FHWA), 2009. Manual on Uniform Traffic Control Devices (MUTCD).
Factor, R., Yair, G., Mahalel, D., 2010. Who by accident? The social morphology of car accidents. Risk Anal. 30 (9).
Finley, M., Venglar, S., Iragavarapu, V., Miles, J., Park, E., Cooner, S., Ranft, S., 2014. Assessment of the Effectiveness of Wrong Way Driving Countermeasures and Mitigation Methods. Report No. FHWA/TX-15/0-6769-1.
Fontaine, H., 1995. A typological analysis of pedestrian accidents. In: Presented at the 7th Workshop of ICTCT. Paris, 26–27 October.
Friebele, J., Messer, C., Dudek, C., 1971. State-of-the-art of Wrong-Way Driving on Freeways and Expressways. TxDOT Research Report Number 139-7.
Greenacre, M., Blasius, J., 2006. Multiple Correspondence Analysis and Related Methods. Chapman & Hall/CRC, FL.
Hoffman, D., De Leeuw, J., 1992. Interpreting multiple correspondence analysis as a multidimensional scaling method. Mark. Lett. 3, 259–272.
Husson, F., Josse, J., Le, S., Mazet, J. FactoMineR: Multivariate Exploratory Data Analysis and Data Mining with R. R package version 1.25. http://CRAN.R project.org/package=FactoMineR. Accessed July 21, 2014.
Kemel, E., 2015. Wrong-way driving crashes on French divided roads. Accid. Anal. Prev. 75, 69–76.
Lathrop, S.L., Dick, T.B., Nolte, K.B., 2010. Fatal wrong–way collisions on New Mexico's interstate highways, 1990–2004. J. Forensic Sci. 55 (2), 432–437.
Lord, D., Mannering, F., 2010. The statistical analysis of crash-frequency data: a review and assessment of methodological alternatives. Transp. Res. A: Policy Pract. 44 (5), 291–305.
Mannering, F., Bhat, C., 2014. Analytic methods in accident research: methodological frontier and future directions. Anal. Methods Accid. Res. 1, 1–22.
Morena, D.A., Leix, T.J., 2012. Where these drivers went wrong. Public Roads 75 (6).
Ponnaluri, R.V., 2012. Highway Special Investigation Report: Wrong-way Driving. Publication NTSB/SIR-12/01.
National Highway Traffic Safety Administration (NHTSA), 2013. Fatality Analysis Reporting System 2013 Data: Web-Based Encyclopedia. United States Department of Transportation. Available: https://www-fars.nhtsa.dot.gov/Main/index.aspx (Accessed: November, 2016).
Ponnaluri, R.V., 2016. The odds of wrong-way crashes and resulting fatalities: a comprehensive analysis. Accid. Anal. Prev. 88, 105–116.
Pour-Rouholamin, M., Zhou, H., 2016. Investigating the risk factors associated with pedestrian injury severity in Illinois. J. Saf. Res. 57, 9–17.
Pour-Rouholamin, M., Zhou, H., Shaw, J., 2014. Overview of safety countermeasures for wrong-way driving crashes. ITE J. 84 (12), 31–38.
R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Accessed November 2016 http://www.R-project.org.
Roux, B., Rouanet, H., 2010. Multiple Correspondence Analysis. Sage Publications, Washington D.C.
SWOV, 2009. Fact Sheet: Wrong-way Driving. Leidschendam, The Netherlands, http://www.swov.nl/rapport/Factsheets/UK/FSWrong way driving.pdf (Accessed November, 2016).
Savolainena, P., Mannering, F., Lord, D., Quddus, M., 2011. The statistical analysis of highway crash-injury severities: a review and assessment of methodological alternatives. Accid. Anal. Prev. 43 (5), 1666–1676.
Scaramuzza, G., Cavegn, M., 2007. Wrong-way drivers: extent-interventions. In: The European Transport Conference. Leeuwenhorst Conference Centre, The Netherlands.
Simpson, S., Bruggeman, D., 2015. Detection and Warning Systems for Wrong-Way Driving. Arizona Department of Transportation Report No. FHWA-AZ-15-741.
Slowikowski, K. ggrepel: Repulsive Text and Label Geoms for 'ggplot2'. R package version 0.6.3. Accessed in: November, 2016. https://CRAN.R-project.org/package=ggrepel.
Wickham, H., 2009. ggplot2: Elegant Graphics for Data Analysis. Springer, New York.
Xing, J., 2014. Characteristics of wrong-way driving on motorways in Japan. IET Intel. Transp. Syst. 9 (1), 3–11.
Xu, X., Wong, S.C., Choi, K., 2016. Quasi-induced exposure method for pedestrian safety at signalized intersections. J. Transp. Saf. Secur. 8 (2), 129–147.
Zhou, H., Zhao, J., Pour-Rouholamin, M., Tobias, P.A., 2015. Statistical characteristics of wrong-way driving crashes on Illinois freeways. Traffic Inj. Prev. 16 (6), 599–604.