Summer 2018

Data driven decision making tools for transportation work zone planning

Samareh Moradpour

DATA DRIVEN DECISION TOOLS FOR TRANSPORTATION WORK ZONE PLANINNG

by

SAMAREH MORADPOUR

A DISSERTATION

Presented to the Faculty of the Graduate School of the

MISSOURI UNIVERSITY OF SCIENCE AND TECHNOLOGY

In Partial Fulfillment of the Requirements for the Degree

DOCTOR OF PHILOSOPHY

in

ENGINEERING MANAGEMENT

2018

Approved by:

Suzanna Long, PhD, Advisor
Steven Corns, PhD
Ruwen Qin, PhD
Dincer Konur, PhD
Ming Leu, PhD

PUBLICATION DISSERTATION OPTION

This dissertation consists of the following articles that have been submitted for publication or are in the process of preparation for submission:

Paper I: Pages 9-38 has been submitted to the Engineering Management Journal.

Paper II: Pages 39-63 will be submitted to the Case Studies on Transport policy Journal.

Paper III: Pages 64-82 will be submitted to the Transport and Health Journal.

ABSTRACT

This research provides tools and methods for integrating stakeholder input and crash data analytics to better guide transportation engineers in effective work zone design and management. Three key contributions are presented: the importance of stakeholder input in traffic management strategies, application of data mining and pattern recognition to identify high-risk drivers in work zones, and the use of multinomial logistic regression (MLR) as a tool to understand key findings from historic crash data. Work zone signage is mandated by the Manual on Uniform Traffic Control Devices (MUTCD), but the current configurations are often criticized by the driving public and state departments of transportation have questioned whether alternate signage would provide more cost-effective, equally safe options. A driving simulator study funded by the Missouri Department of Transportation (MoDOT) evaluated one such alternate sign configuration and determined that it received higher levels of driver satisfaction with no statistical impact on safety. Findings of driver preference for the alternate configuration are considered high value by MoDOT with respect to both mobility and safety. A second contribution focused on risk mitigation through data analytics. Pattern recognition and data mining techniques were applied to driving simulator data as part of a multi-criteria decision making tool to identify drivers with high risk potential. Findings related to age and gender suggest opportunities for driver education and training to increase safety. The third contribution identifies a method for analyzing historic crash data to determine key risk factors in fatality and serious injury accidents in work zones. Multinomial logistic regression (MLR) is used. Findings outline patterns and scenarios that should be integrated into work zone design to enhance safety and improve mobility with respect to work zone lighting, impact of weather, and the like.

ACKNOWLEDGEMENTS

I would like to express my sincere gratitude for the advice, guidance, and support provided by my dear advisor, Professor Suzanna Long. What I have learned from her was far more than research and the art of engineering.

The insight, advice, and the support provided by my committee members, Dr. Corns, Dr. Qin, Dr. Konur, and Dr. Leu is highly appreciated. The support provided by the faculty and staff at Engineering Management and System Engineering department is highly appreciated. This research could not have been done without the support provided by Missouri Department of Transportation.

I would like to appreciate the great support provided by my dear friends and colleagues at Missouri S&T. None of these could have been accomplished without their help and contributions. It has been a great pleasure to work alongside these great colleagues.

This dissertation could never be written without the unconditional love and support from my beloved parents. They have been my role models in life and supported me during all moments of difficulty. I also would like to thank my sister and brother who have always helped and encouraged me.

And last but not least, I offer my deepest gratitude for the unconditional love, support, and dedication of my beloved spouse, Hamed, who has been my greatest support. I could never accomplish this without him. Being with him in this journey, made it possible to go through the impossible. I can never thank him enough for his understanding and support during the difficult moments of life.

TABLE OF CONTENTS

Page

LIST OF ILLUSTRATIONS

Figure

Section Page

LIST OF TABLES

NOMENCLATURE

Symbol	Description
NHTSA	National Highway Traffic Safety Administration
NSC	National Safety Council statistics
PDO	Property Damage Only
OR	Operation Research
RF	Risk Factor
SDLMS	Simplified Dynamic Lane Merging Systems
TMS	Transportation Management System
DS	Driving Simulator
DM	Data Mining
MCDM	Multi Criteria Decision Making
MoDOT	Missouri Department of Transportation
FHWA	Federal Highway Administration
FARS	Fatality Analysis Reporting System
DOT	Departments of Transportation
MUTCD	Manual on Uniform Traffic Control Devices
DBQ	Driver Behavior Questionnaires
MOT	Maintenance of Traffic

1. INTRODUCTION

This section offers a summary of the work conducted throughout the dissertation.

1.1. BACKGROUND

Despite the research done to demonstrate the risk factors (RFs) in work zones, the rates of crashes and fatalities are still high. Regarding the Federal Highway Administration (FHWA) data, 1.8 fatalities per day were recorded in work zones in 2014. Fatality Analysis Reporting System (FARS) statistics revealed work zone fatalities had increased by 13% from 2013 to 2014 (FHWA, 2017). Mandatory lane changing and merging in work zones with lane closures can increase drivers' dangerous maneuvers, which increase crashes (Fei et al., 2016). The likelihood and severity of crashes in work zones are higher than on normal roads. The results of a survey indicated that the rates of no injury and injury accidents in work zones are 23.8% and 17.3% higher than those of the normal roads, respectively (Khattak et al., 2002).

Studies on roadway work zone safety cover a wide range of research topics. These include studies to identify the common factors in roadway work zone accidents, evaluating the effectiveness of various traffic control methods, studying the effects of work zone configurations on drivers' behavior, evaluating the safety apparel of roadway work zone workers, evaluating the cognitive processes and behavior of drivers around work zones, and performing risk modelling and risk assessment on roadway work zones (Ean Harn et al., 2013, Long et al., 2014).

Based on available literature there are two major types of research on work zone safety. These two types are as follow:

1.1.1. Evaluation of Stakeholders Perception Regarding Implementing New Temporary Traffic Control (TTC) Sign Configuration. Departments of Transportation (DOTs) use a variety of methods to inform drivers of upcoming work zones, including work zone signage, flaggers, arrow panels (Moradpour et al., 2015). Efficiency and satisfaction of stakeholders about work zone sign configuration is a major concern of DOT managers. However, the public drivers often criticize the current configurations. So DOTs implementing new sign configurations in work zone areas and evaluate the efficiency of alternate signs. The reaction of drivers to alternate sign configurations should be explored in addition to their driving patterns through the work zones where such new signage is incorporated in order to measure safe implementation (Thind et al., 2017).

Much research has been conducted regarding the safety benefits of implementing new/alternative signs to foster traffic safety in work zones (Reyes et al., 2008). A dynamic late-merge scenario was evaluated in Tappahannock, Virginia and the usefulness of before and after those scenarios were examined. The research findings revealed that the number of vehicles in the closed lane increased when compared to the late merge with the Manual on Uniform Traffic Control Devices (MUTCD) scenario. There were not any statistically significant differences in throughput volumes and delay time between the MUTCD scenario and the late merge (Beacher et al., 2004). The Simplified Dynamic Lane Merging Systems (SDLMS) for early- and late-merging scenarios were used in Florida’s Maintenance of Traffic (MOT) plans. The study demonstrated the highest queue discharge values (or capacity) of the work zone in the early merging scenarios (Harb et al., 2009).

Effect of joint lane merge (JLM) on traffic in a controlled work zone was investigated by Idewu and Wolshon (2010) through a field study in Louisiana. The merging speed was compared between the JLM and the Conventional Lane Merge (CLM). Results of this research revealed no significant difference at volumes ranging from 600 to 1200 vehicles per hour. However, it was suggested that when going through the JLM scenario, the drivers were more cautious and experienced a smoother lane merge. In the case of the JLM, the drivers had a lower number of lane changes and entered the transition zones with lower speeds during congested periods. In other research conducted by Shakouri et al. (2014), JLM results were compared to those obtained from CLM. Based on the results mean maximum braking forces are lower in the JLM configuration compared to the CLM configuration (Shakouri et al., 2014).

In a simulation-based study, the Missouri Department of Transportation's (MoDOT) alternate signage was compared to MUTCD lane shift signs. The results did not reveal significant differences between the two signs with respect to drivers' performance (Long et al., 2017; Thind et al., 2017). Despite several studies to make alterations to the work zone configurations and to improve work zone safety, the accident rate throughout work zone areas is still alarmingly high.

1.1.2. Identifying Risk Factors (RFs). The second type of literature related to work zone safety is about identifying risk factors in work zones. Vehicle crashes as a system consists of independent variables such as the driver, vehicle characteristics, environmental and geographical conditions, occupants and other road users, and the roadway (Bayam et al., 2005).

Each of these variables consists of different characteristics. The driver variable consists of the driver's age, gender, and driving experience. The vehicle variable is composed of vehicle characteristics such as the vehicle type and year. The roadway variable consists of road attributes, such as road type and road surface. The environmental and geographical conditions consist of weather conditions, light conditions, and the date and time. The occupants and other road users include pedestrians, other occupants of the car, drivers of other vehicles, and occupants of other vehicles (Bayam et al., 2005).

It is not easy for researchers to evaluate the contribution of these variables (Bedard et al., 2002), so matrix was developed by Haddon (1980) to help investigators categorize accident factors (Shankar et al., 2004). Based on Haddon index, three time frames (pre-crash, crash, and post-crash) and factors such as the human (driver), vehicle, and environment should be considered to analyze vehicle crashes (Haddon, 1980). Table 1.1 summarizes the Haddon matrix that may helpful for identifying countermeasures to vehicle crashes.

These variables interact with each other, and one of these interactions can cause an accident on the road. These interactions can consist of speeding, alcohol and drug use, rapid lane changing, failure to wear a seatbelt, improper weather, road conditions, inattentive or negligent driving convictions, engaging in distracting behaviors, following other cars too closely, improper turn convictions, and road light, etc. (Zamorski & Kelley, 2011). Identifying these risk factors consist of determining risky drivers and also evaluating historical data of work zone crashes.

Table 1.1. The Haddon matrix template (Bayam et al., 2005)

		Factors		
		Human(Driver)	Vehicle	Environment
Phases	Pre-crash			
	Crash			
	Post-crash			

Risky driving behavior is a reason for the high likelihood of severe crashes in work zones. Identifying high-risk drivers is significant to reducing RFs due to the increasing rate of fatalities and the high impact of driver errors on work zone crashes.

Based on statistics, driver errors can cause 75% to 95% of work zone crashes (Stanton & Salmon, 2009). Regarding research conducted at Kansas State, 92% of work zone crashes in Kansas are caused by risky drivers (Li & Bai, 2006). Even though only 6% of total drivers are considered risky drivers, these drivers cause 65% of crashes (Guo & Fang, 2012).

These risky behaviors include aggressive lane changing, speeding, careless driving, not paying attention to pedestrians, and not considering the traffic control signs (Weng & Meng, 2012; Luke & Heyns, 2014). This highlights the fact that the effect of drivers on work zone safety is a significant factor that needs to be considered. In addition to drivers, environmental conditions, road geometry, and road condition have a significant effect on the severity of crashes in work zone.

Many studies focus on the effect of these factors on work zone safety. The data from fields, driving simulators, and driver behavior questionnaires (DBQs) were used to evaluate the effect of drivers' characteristics on their driving patterns.

Driver casualty risk in the construction, maintenance, and utility work zones was investigated by using data from the FARS. Based on the multiple t-test results, the work zone type has an effect on driver casualty risk. Moreover, the rate of driver casualty risk is highest in construction work zones, followed by maintenance and utility work zones. Based on the results, traffic control devices and restraint use are related to reduce driver casualty risk (Weng & Meng, 2011).

Driving attitudes and self-reported behavior of drivers were compared in a study. Participants filled out two questionnaires regarding risky driving. The multivariate analysis of variance (MANOVA) and univariate regression analysis were carried out to determine risky drivers. Gender was an important factor in demonstrating risky attitudes, and male drivers had riskier responses (Harré et al., 2000).

The effects of personality traits and gender on risky driving behavior and accident involvement were investigated by using a questionnaire survey. Results indicated that over 37% of the variance in risky driving was explained by personal behaviors and gender. In the case of young drivers, it was observed that both gender and certain personality traits affected the risky driving behaviors (Oltedal & Rundmo, 2006).

A survey was conducted in the State of Alabama to determine correlation between risk perception, positive affect, and risky driving. The results of a regression analysis of gender revealed that male drivers are engaging in risky driving behaviors more than female drivers (Rhodes & Pivik, 2011).

The results of different states such as Southeast Michigan, Florida, and Tennessee crash records revealed the importance of roadway geometry, weather conditions, driver characteristics such as age and gender, lighting conditions, and driving under the influence of alcohol and/or drugs in work zone crashes (Harb et al., 2008; Wei et al., 2017; Meng et al., 2010; Weng & Meng, 2012).

The decision tree method was employed to determine the effects of environmental, vehicle, and driver characteristics on drivers' behavior in work zones. Data from Michigan highway work zones were used for the analysis. The results revealed that gender was a significant factor in drivers' driving behavior. Middle-age drivers are more likely to engage in risky behavior at the lower work zone speed limit (Weng & Meng, 2012).

1.2. RESEARCH OBJECTIVES AND CONTRIBUTIONS

The goal of this dissertation is to propose an analytic tool for work zone safety. These tools help transportation managers reach a better understanding of crucial factors in the work zone. Figure 1.1 includes a framework of the research. Three contributions of this dissertation consists of:

Research I: Because dissatisfaction of public drivers regarding work zone sign configuration, the evaluation of new sign configurations is necessary to compare the efficiency of alternate sign configurations with the MUTCD sign configuration. This paper evaluates MoDOT alternate sign configuration based on stakeholders' reaction.

Paper II: Exploring the driving patterns as one of the significant risk factors is very helpful for researchers to determine drivers with a high-risk potential. This research proposes a hybrid of DM and MCDM methods for identifying drivers' pattern. The goal of this research is to develop an analytic tool to identify high-risk drivers in work zone.

Paper III: Road accidents and crashes are unpredictable and knowledge of the relevant factors are necessary for analysis. The historical data from Missouri state work zone crashes will be used to identify, evaluate, and model trends that are related to severe crashes The results of this study will help transportation managers to understand significant RFs. Effective safety countermeasures may be designed at the work zone planning to prevent safety deficiencies.

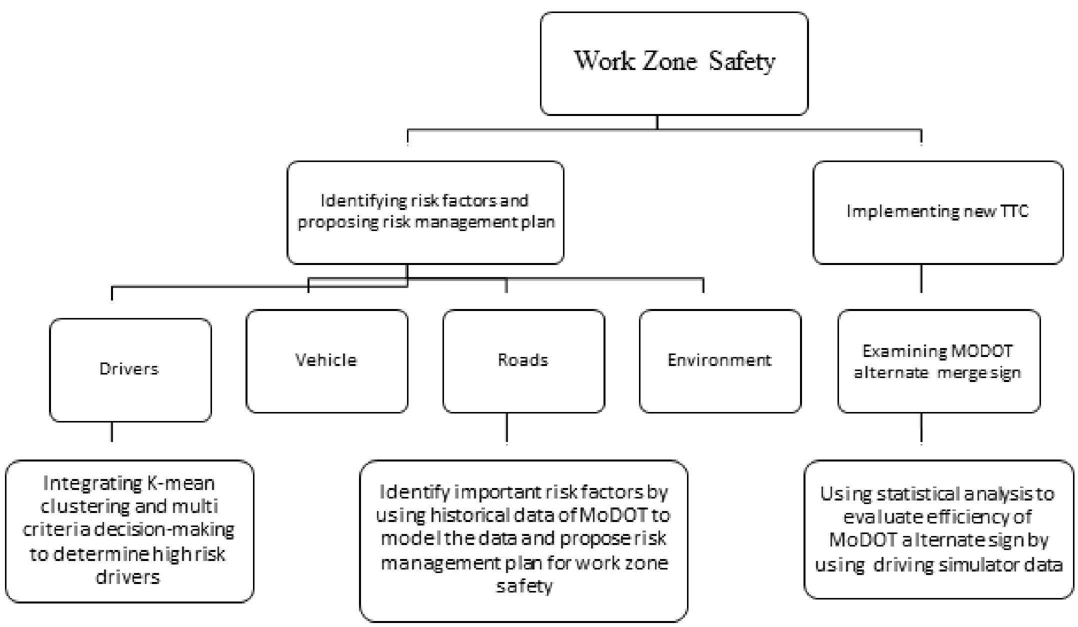

Figure 1.1. Dissertation framework

PAPER

I. EVALUATING WORK ZONES SIGN CONFIGURATIONS USING A DRIVING SIMULATOR

Samareh Moradpour, Suzanna Long

Department of Engineering Management and Systems Engineering, Missouri University of Science and Technology 600 W. 14th St. Rolla, MO 65409-0370

ABSTRACT

This research presents a study designed to assess drivers' responses to work zone sign configurations utilizing statistical analysis. A driving simulator is used to compare the effectiveness of national standard work zone signage based on the Manual on Uniform Traffic Control Devices (MUTCD) with Missouri Department of Transportation (MoDOT) alternative signage. Seventy-five participants were selected to complete four driving scenarios. Statistical data analysis was used to investigate the effectiveness of the alternate configurations employed under different scenarios. The results of analysis of variance (ANOVA) suggest MoDOT alternate signs and MUTCD signs do not have any statistically significant effect on the merge location and mean speed while age and gender have significant effects on the merge location and mean speed for the simulated scenarios. In terms of safety, the number of drivers with late merge in the MoDOT scenarios was less than in the MUTCD scenarios. This suggests that the MoDOT alternative signage is a good alternative for the MUTCD signs.

Keywords: Data Analysis, Driving Simulator, Merging Behavior, Work Zone Signage

1. INTRODUCTION

Due to ongoing efforts to improve aging transportation infrastructure throughout the U.S., work zones are frequently encountered and necessary for reconstruction and maintenance of roads. It is estimated that 20% of U.S. highways, approximately 3,000 work zones, are under repair during the peak construction season (FHWA, 2009). Often, when roadwork is being done, it is necessary to close one or more lanes of traffic, causing lane changes and merges. Due to changeable traffic conditions, work zones pose a significant threat to drivers and their passengers, as well as the workers present in the construction zones and the likelihood and severity of crashes in work zones are higher than on normal roads (FHWA, 2017). The National Work Zone Safety Information reported 669 fatalities in work zone crashes in 2014 (NHTSA, 2014). Based on the statistics released by the National Safety Council (NSC), the rate of fatalities occurring in work zones has increased from 576 to 1,074 between 2005 and 2014 (NSC, 2016). Therefore, safety and mobility are great concerns of transportation policy makers (Hurwitz, Heaslip, & Moore, 2012). Specifically, effective traffic management through work zones is crucial for increasing safety for all (Grillo, Datta & Hartner, 2008).

Driving is a complex and potentially dangerous task and can be affected by factors such as the signs, road conditions, and individual driving behaviors. In recent years, several researchers have studied the efficiency of different work zone sign configurations, risk characteristics of roads (surface type, road light, etc.), and the effect of drivers' characteristics (age, gender, etc.) for increasing work zone safety (Harb, Radwan, Yan, Pande, & Abdel-Aty, 2008; Meng & Weng, 2010; Oltedal & Rundmo, 2006; Edara, Sun, & Zhu, 2013; Zhu, Edara, & Sun, 2015; Blinded for Review, et al., 2017). The signage used in the advance warning area of a work zone provides critical information to drivers such as information regarding the closed lane, when to merge, when to reduce speed limits, etc. These types of information are critical to the overall safety of the work zone (Zhu et al., 2015).

Although there are existing work zone sign configurations approved by MUTCD, other sign configurations are possible and may be evaluated against traffic management goals for traffic flow, driver behavior, driver satisfaction, and the like. The reaction of drivers to alternate sign configurations, in addition to their driving patterns through the wok zones where such new signage is incorporated, must be evaluated in order to assure safe implementation (Blinded for Review, 2017) and before traffic management agencies can request their use.

Transportation professionals have used a variety of traffic control methods over the past two decades for work zone traffic management. Such methodologies include Conventional Lane Merge (CLM) proposed by the United States Department of Transportation, as well as the early merge (EM), and late merge (LM). The EM and LM strategies are divided into two categories of static and dynamic. Each of these approaches has some specific characteristics that limit their usage in congested and uncongested traffic flow conditions. The differences between these methods refer to the location where drivers merge to the open lane. In other words, the objective of the late merge is to use the maximum available roadway space by using the whole available traffic lanes up to the merge point. The early merge strategy encourages drivers to merge early before work zone lane closures to reduce the potential for merging friction near the merge point adjacent to the lane closure spot.

These scenarios have been investigated in several previous studies. For instance, Beacher (2004) evaluated the efficiency of the dynamic late merge scenario during a case study in Tappahannock, Virginia. The research findings revealed that the number of vehicles in the closed lane increased when compared to the Manual on Uniform Traffic Control Devices (MUTCD) scenario. There were no statistically significant differences in throughput volumes and delay time of the MUTCD scenario with those of the late merge scenario (Beacher, 2004). The Simplified Dynamic Lane Merging Systems (SDLMS) for early and late merging scenarios were used in Florida's Maintenance of Traffic (MOT) plans. The study demonstrated the highest queue discharge values (or capacity) of the work zone in the early merging scenarios (Harb et al., 2009).

Zhu et al. (2015) presented the findings of a recent MoDOT field study that compared MUTCD merge sign with MoDOT alternate merge signs. Behavior characteristics of the drivers, including speeds and open lane occupancies were investigated as part of the study. Results indicated that the MoDOT alternate sign configuration led to 11% higher traffic upstream of the merge sign in the open lane. This is a positive finding for the MoDOT alternate sign from both a safety point of view and the ability to minimize the conflicts associated with lane drops. The authors found no statistically significant differences between the speed characteristics of the investigated sign configurations. The MoDOT alternative sign configuration was shown to be equal to that of the MUTCD sign configuration (Zhu et al., 2015; Edara et al., 2013). MoDOT continues to consider alternative sign configurations and several recent studies (Edara, Sun, & Brown, 2017; Long et al., 2017) explore lane shift sign configurations and another (Brown, Sun, & Cope, 2015) explored the addition of mobile alarm systems to work zone sign configurations.

Despite attempts to make alterations to the merge configurations and to improve work zone safety, the accident rate throughout work zone areas is still alarmingly high. This can be contributed to a deficient measures for reducing risky driving patterns (Aghazadeh, Ikuma, & Ishak, 2013). Moreover, most of the available literature on work zone safety is devoted to investigations that explore the static lane merge configurations. Therefore, more studies are needed on alternative signs, sign placement, and driver response to current and suggested signs.

This research addresses this gap in the literature and presents the results of a simulation-based study where the MoDOT alternate signage was compared to MUTCD signs for work zone management. The simulator study considered the factors from the previous MoDOT field study (Edara et al., 2013) and also added driver preference as a consideration. The results revealed no significant differences between the two sign configurations with respect to overall drivers' reaction, but drivers did report increased satisfaction with the MoDOT alternate configurations. Some differences were also identified with respect to age and gender. The findings of this study suggest opportunities for traffic managers to consider driver preference, as well other factors, in work zone sign configuration.

2. METHODOLOGY

The study compared the Conventional Lane Merge (CLM) configurations using MUTCD sign configurations against MoDOT's alternate sign configurations. The test scenarios simulated both right and left work zone lane closures for both the MUTCD and MoDOT sign alternatives (see Section 2.3). Statistical data analysis was used to investigate the efficiency of different configurations employed in the study.

The research sequence used to assess the effectiveness of work zone signs effectiveness consisted of four stages: simulation design/programming, participant selection, data collection, and data analysis. In the first step, the relevant scenarios were designed and programmed into the driving simulator using data inputs on work zone design from MoDOT traffic engineers. In the second step drivers were then selected by pre questionnaire for participation in the simulation. In the third step, participants drove four scenarios. During each driving simulation, the data acquisition board in the simulator recorded relevant data such as time, speed, position (x,y), acceleration, deceleration, and steering angle. In addition, post-simulation questionnaires were used to determine user satisfaction and preference for the sign configurations. The final step of the study involved using statistical analysis on the collected data. The drivers' merging patterns and speed were analyzed against demographic characteristics of the participants. Statistical data analysis was used to investigate the effectiveness of the alternate configurations employed under different scenarios. An analysis of variance (ANOVA) was conducted to determine the effect of age and gender on merge location and speed. A flow-chart of the methodology is shown in Figure 1.

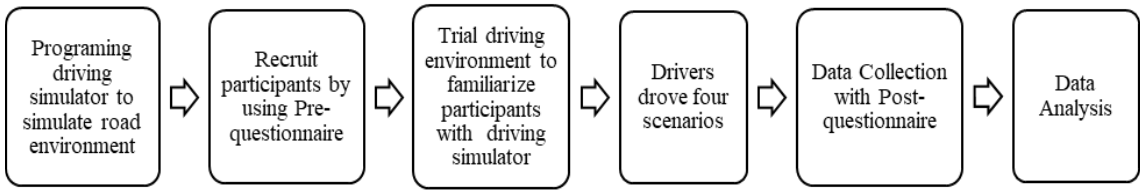

Figure 1. Research Sequence Flow-Chart

2.1. DRIVING SIMULATOR DESCRIPTION

This study utilized a driving simulator, as opposed to studying an actual work zone due to cost, the difficulty of manipulating the site for different scenarios, and factors such as environmental changes that come into play with real sites (Blinded for Review et al., 2016; Blinded for Review et al., 2015). Additionally, real life evaluations may introduce unnecessary risks for both test participants and investigators. Driving simulators provide a safe, virtual-reality environment to evaluate a wide range of interventions and have been used extensively in previous research (Reyes & Khan, 2008). They are useful for evaluating sign configuration and analyzing driver behavior. The driving simulator used in this study consisted of a Ford Ranger pickup cabin, held at a fixed base (Figure 2). This simulated cabin included a steering wheel, accelerator pedal, brake pedal, and speedometer. The simulated environment was created using three 3,000-lumen Liquid Crystal Display (LCD) projectors, a projection screen, and a master simulation computer. The projection screen had a projection angle of 52.5°, an arc width of 25 feet and a height of 6.6 feet from the ground to provide a realistic field of view of 115° (Figure 2). The Blender 3D graphics software and Python software were used to program and simulate roads and the driving environment. A data acquisition system was also used to record time, speed, position, acceleration, deceleration, and steering angle during the simulation.

Figure 2. Driving simulator (left) and the screen view (right)

Information about the operation of the simulator was provided prior to the start of the official test. Specifically, drivers were given information regarding the location of all the controls such as brake pedal, seat and steering wheel adjustments. Participants completed all four scenarios and the total average time to complete was 30 minutes or less.

The Driving Simulator used for this study was validated in terms of relative and absolute validity using both subjective and objective evaluations. The framework, methods and results can be found in Bham, et al., 2014. Summary results demonstrated the applicability of the simulator for work zone studies.

2.2. PARTICIPANT SELECTION

Selecting appropriate participants is one of the most important steps in the sequence. As part of the research design, a kick-off meeting was held with MoDOT and FHWA traffic engineers. The sample size was set at 75 participants during that meeting with a follow-on requirement that the sample approximate the demographic percentages from within the state in terms of age and gender, in addition native language, and years of driving experience were tracked (see Table 1). Income, education levels, and job categories were not specifically considered as part of the study. Participants in this research were separated into four age groups: 18-24, 25-44, 45-64, and over 65 years. Each participant chosen for the study completed the four driving scenarios using the Driving Simulator. During each simulation run, a member of the research team observed driver reactions and participant questions for each of the scenarios. This qualitative information was combined with the quantitative simulator data to generate data records for each participant. The pre questionnaire was given before participants entered the simulator. Participants also had to meet the following qualifications based on MoDOT/FHWA requirements: (a) valid driving license, (b) no prior information or knowledge of the study being conducted, and (d) alcohol and drug free for the past 24 hours. In order to recruit participants for this study, an email was sent to university faculty, staff and students and advertisements placed in the area community. All interested individuals were asked to complete a pre-experiment questionnaire to determine their eligibility. The participants were given the opportunity to become familiar with the driving simulator before the test began, including the completion of a trial driving experience. One volunteer experienced simulator sickness during the trial experience and was excused from participation in the formal study. Additionally, a $10 gift card was offered as an incentive to participants, awarded upon completion of the simulation.

Table1. Demographic information of participants

Age				Gender		Native Language		Driving Experience (Year)			
18-24	25-44	45-64	≥65	Female	Male	English	Non English	<1	1-5	5-10	>10
111	228	227	99	41	34	67	8	22	99	33	61

2.3. DRIVING SCENARIOS

Four merge scenarios were considered within this study (Figure 3). Right and left lane closures were simulated using MUTCD and MoDOT configurations respectively. In each of these scenarios, the MUTCD merge configuration was compared to the corresponding MoDOT alternative merge configuration. Each scenario consisted of two lanes, each lane was 6 meters wide and the roads were 6 km long. The start point was located at an approximate distance of 4 km before the work zone. The first sign, Road Work Ahead, was located 1,466 meters before the work zone. The second sign, Right/Left Lane Closed, was located 752 meters prior to the work zone. The third sign of the MUTCD scenario, Lane Closed, was installed 305 meters before the start of the work zone. The third sign in the MoDOT scenario consisted of two separate signs, Merge/arrow and Right/Left Lane Closed, and they installed 305 meters before the work zone. A STOP sign was placed at the end of the work zone, instructing drivers to come to a halt which is less than 1 km after the end of the work zone. To simplify the research, a straight highway road with no curves or traffic was used in this simulation.

3. DATA ANALYSIS METHODOLOGY

The methodologies incorporated for data analysis are elaborated in this section.

3.1. MERGE ANALYSIS METHODOLOGY

To measure the effectiveness of the alternate sign configuration against the MUTCD sign configuration, analysis of variance (ANOVA) was used to test performance differences between each pair of changing lane configurations according to the following hypotheses at a significance level of 0.05 (α=0.05):

H0: There was no significant difference between the mean locations of lane changes in the different scenarios.

Ha: At least one of the scenarios had a different mean location of lane change.

The ANOVA analysis is based on the fact that, for a P-value less than α, the factor(s) interaction is significant. Otherwise, for a P-value greater than α, the factor or interaction is not significant (Sadati, Arezoumandi, Khayat, & Volz, 2016; Elrod, Daughton, Murray, & Flachsbart, 2010).

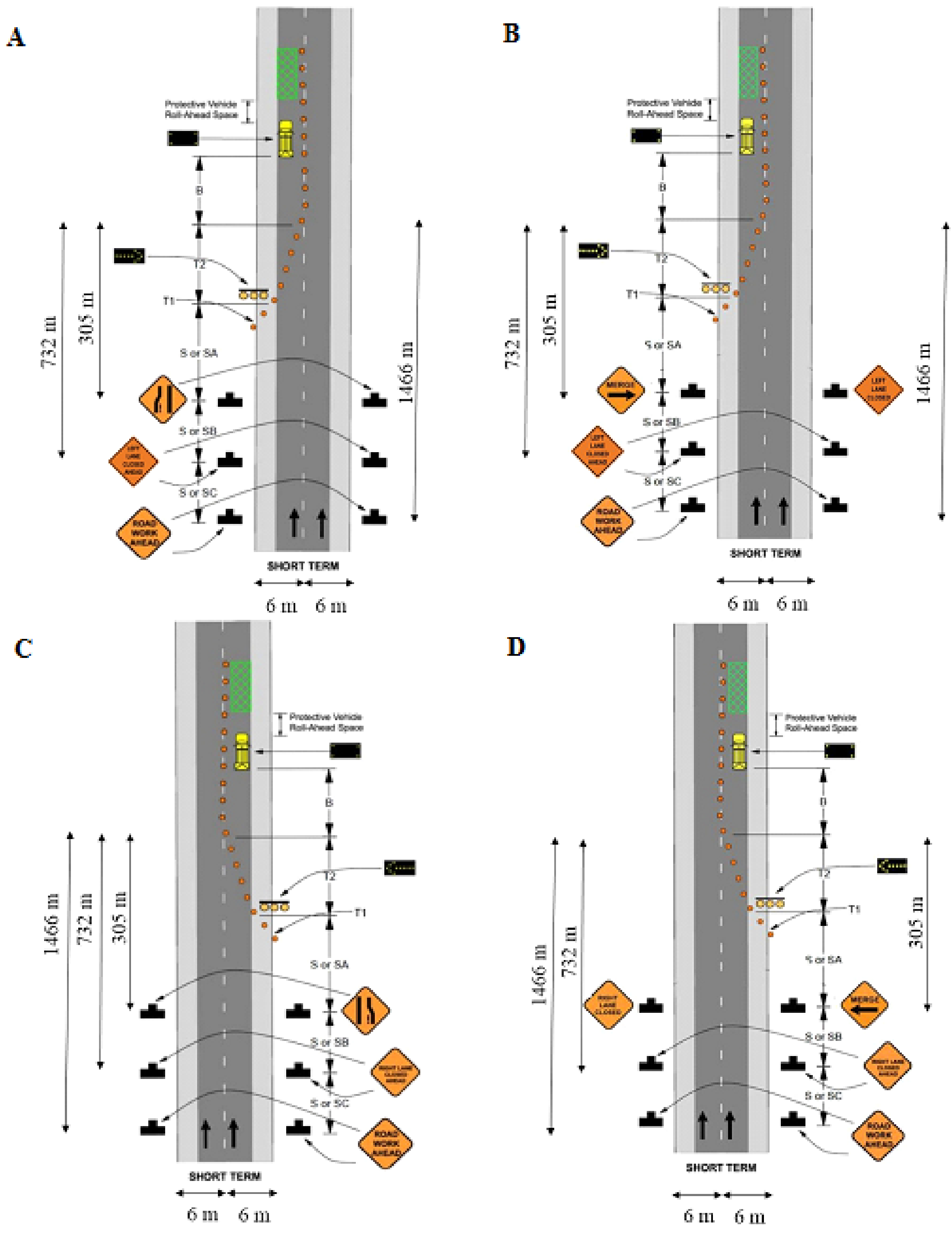

Figure 3. (A) MUTCD merge right, (B) MoDOT alternate merge right, (C) MUTCD merge left, (D) MoDOT alternate merge left

3.2. SPEED ANALYSIS METHODOLOGY

Speed is one of the most significant causes of crashes in work zones. It is important to encourage drivers to be cautious and observe the speed limits (Brewer, Pesti, & Schneider, 2006). Evaluating characteristics such as mean, standard deviation, and 85th percentile of speed are significant for safety in work zones. The test used in this research is presented as:

$$\text{Mean speed} = \frac{\sum_{i=1}^{n} X_i}{n} \tag{1}$$

$$\text{Standard deviation} = \sqrt{s_x^2} \tag{2}$$

$$\text{85th percentile} = \frac{X_{([n0.85]+1)} - Y_{([n0.85]+1)}}{1.530\sqrt{\frac{s_y^2}{n_y} + \frac{s_x^2}{n_x}}} \tag{3}$$

Where n = sample size for the two data sets, x and y; $\bar{y}$ and $\bar{x}$= sample means. The $X_{([n0.85]+1)}$ and $Y_{([n0.85]+1)}$ represent the 85th speed percentiles for two independent random samples; s_y^2 and s_x^2 = are variances for sample; $s_x^2 = 1/n_x - 1\sum_{i=1}^{n_x}(x_i - \bar{x})^2$; and n_x and n_y represent sample size for the two data sets, x and y, equal to 75 in this study (Hou , Sun, &Edara,2012).

4. RESULTS

The normality of data was tested by using the Kolmogorov-Smirnov and Anderson-Darling normality tests. The P-values of the MUTCD and MoDOT values were less than 0.01 for both scenarios on the Kolmogorov-Smirnov. Moreover, the P-values were 0.005 for both scenarios based on the Anderson-Darling method. Based on the research results, the merge location and speed data are not normally distributed. Since the sample size for this study is greater than 30 (sample size = 75), the ANOVA test could be used, although the normality assumption is not justified (Montgomery & Runger, 2008). All statistical analysis was done by using Minitab version 17.

4.1. MERGE PATTERN

Each participant completed four different scenarios in the simulation: MUTCD left lane merge, MoDOT left lane merge, MUTCD right lane merge, and MoDOT right lane merge. The driving path consisting of (x, y) coordinates recorded approximately each second the individual drove on the simulated road.

The individual driving paths obtained from the 75 participants were investigated. The data were incorporated to analyze and model the driving pattern for MUTCD and MoDOT configurations for right/left merge scenarios. Figures 4-7 show a plot of the 75 driving paths collected from the driving simulator of the merge scenarios for MUTCD and MoDOT right/left merge signs.

In order to gain a better understanding of drivers’ merging behavior, the road was divided into three parts. One is within $y = [-2400, -93]$, $y = [-94, 670]$ and y $= [1100, 1400]$, termed Z_1, Z_2, and Z_3, respectively. Based on the figures, the merging points where drivers preferred to join the other lane are within these three parts.

Figures 4 and 5 indicate some driving patterns that are easily observable from these plots. In both the MUTCD and MoDOT merge right scenarios, about half of the drivers started merging to the lane on the right after the simulation started. The other drivers, stayed in the left lane for more than 2,000 meters, then merged to the right. A few drivers merged to the right very late, around y=600 meters. Some of drivers merged back to the left lane during the simulation study, but most drivers were in the right hand lane when the simulation was over.

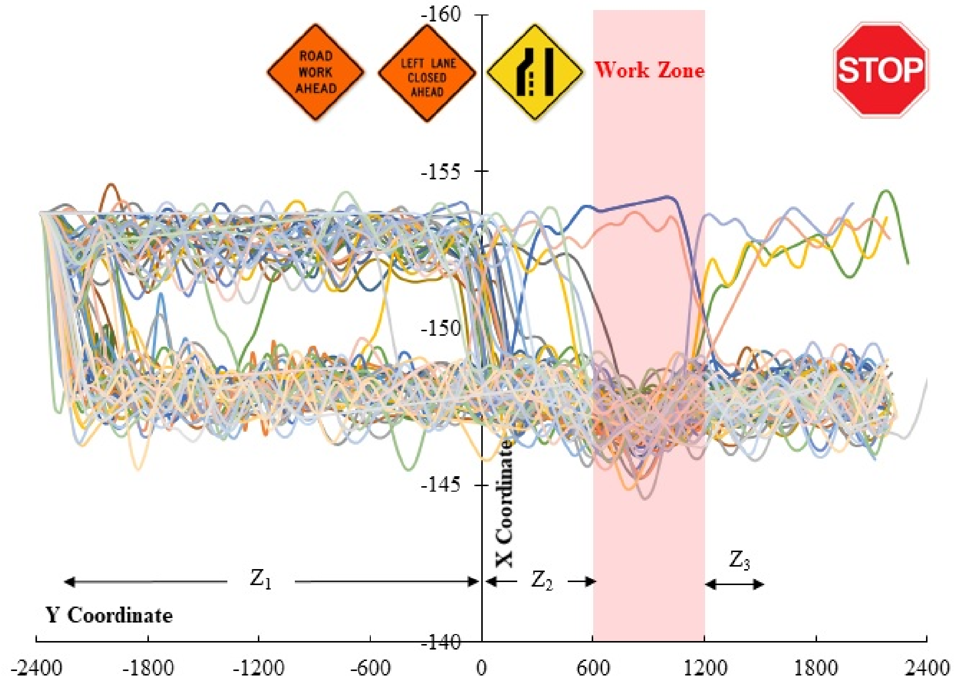

Figure 4. Plot of 75 driving paths - MUTCD right merge scenario

The results of the drivers' merging behavior in both of the MUTCD and MoDOT right merge scenarios revealed that 74% of the drivers merged into the right lane on Z_1. Further analysis indicated that 77% of the drivers who merged in Z_1 were between 45-64 years old. Approximately 68% and 65% of the drivers were female in the MUTCD and MoDOT right merge scenarios, respectively.

The results of drivers' merging behavior in the MUTCD and MoDOT right merge scenarios revealed that 25% of the drivers merged to the right lane in Z_2. For merging right in Z_2, 63% and 72% of the drivers were in the age range of 18-24 years old in the MUTCD and MoDOT scenarios, respectively. Of these drivers merging in Z_2, 52% and 55% were male in the MUTCD and MoDOT scenarios, respectively.

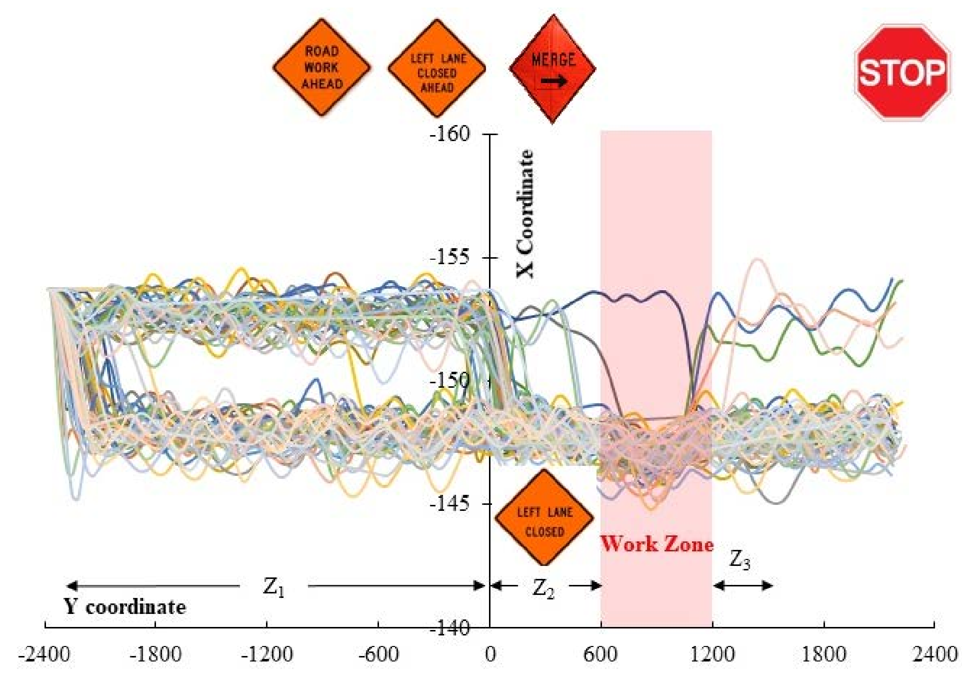

Figure 5. Plot of 75 driving paths - MoDOT right merge scenario

In both MUTCD and MoDOT right merge scenarios, around 1% of the drivers merged left again after their first merge to the right in Z_3. In the MUTCD right merge scenario, 50% of the drivers were in the age range of 18-24 years and remaining 50% were in the range of 25-44 years. In the MoDOT right merge scenario, 75% of the drivers were in the age range of 18-24 years and the others were 25-44. In both scenarios there was an equal distribution of male and female drivers (Table 2).

Table 2. Merging behavior of drivers

Zone		MUTCD right merge	MoDOT right merge
Z_1	Percentage of drivers	74	74
	Age group	77% drivers aged 45-64	77% drivers aged 45-64
	Gender	68% female	65% female
Z_2	Percentage of drivers	25	25
	Age group	63% drivers aged 18-24	72% drivers aged 18-24
	Gender	52% male	55% male
Z_3	Percentage of drivers	<1	<1
	Age group	50% drivers aged 18-24 and 50% aged 25-44	75% drivers aged 18-24 and 25% aged 25-44
	Gender	50% female	50% female

Based on the results, drivers in the range of 45-64 years, merged right immediately after they started driving, but younger drivers, between the age of 18-24 years preferred to merge to the right in the middle of the path before the work zone.

In both MUTCD right merge scenarios and MoDOT right merge scenarios two drivers missed the signs and drove throughout the work zone. Data corresponding to these drivers was eliminated from the analysis.

As with the right merge scenarios, some driving patterns are easily observable from left merge scenarios plots. Figures 6 and 7 indicate two zones where most of the drivers actively merged. For example, in the MUTCD merge left scenario, more than 90% of the drivers started merging to the right lane after the simulation study started.

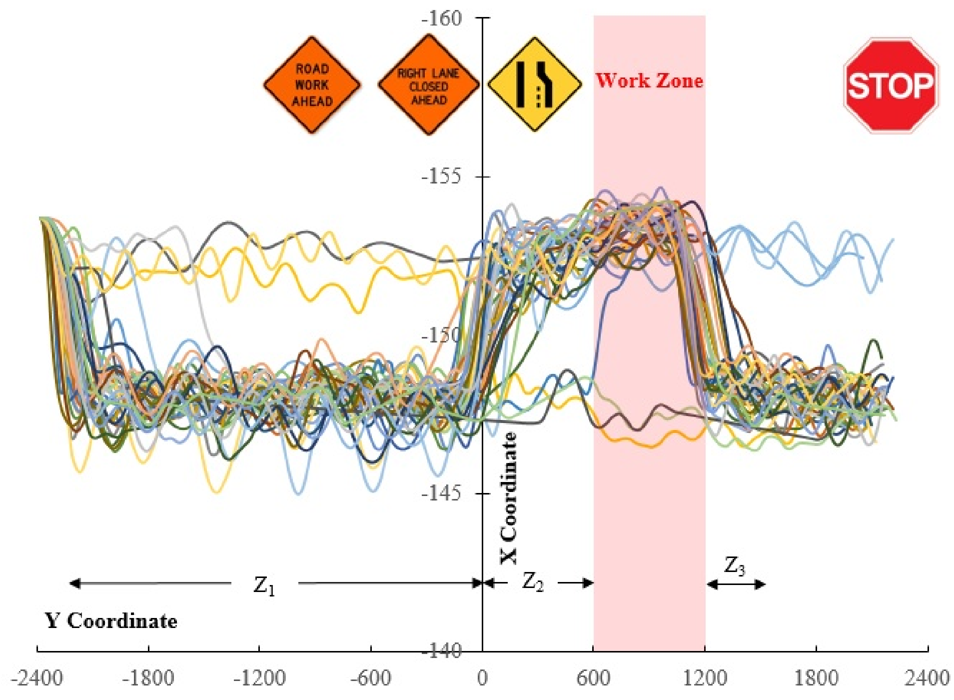

Figure 6. Plot of 75 driving paths - MUTCD left merge scenario

There were several drivers who remained in the left lane upon completion of the simulation, and therefore, failed to complete the test. The drivers who merged to the right tended to stay in the right lane for at least 2,000 meters and merge to the left soon after that. Most of the drivers merged back to the right lane during the simulation study, but a few of them were still in the left lane when the simulation was over. During the MoDOT left merge scenarios, most drivers did continue on the lane and merged to the left after about 2,000 meters.

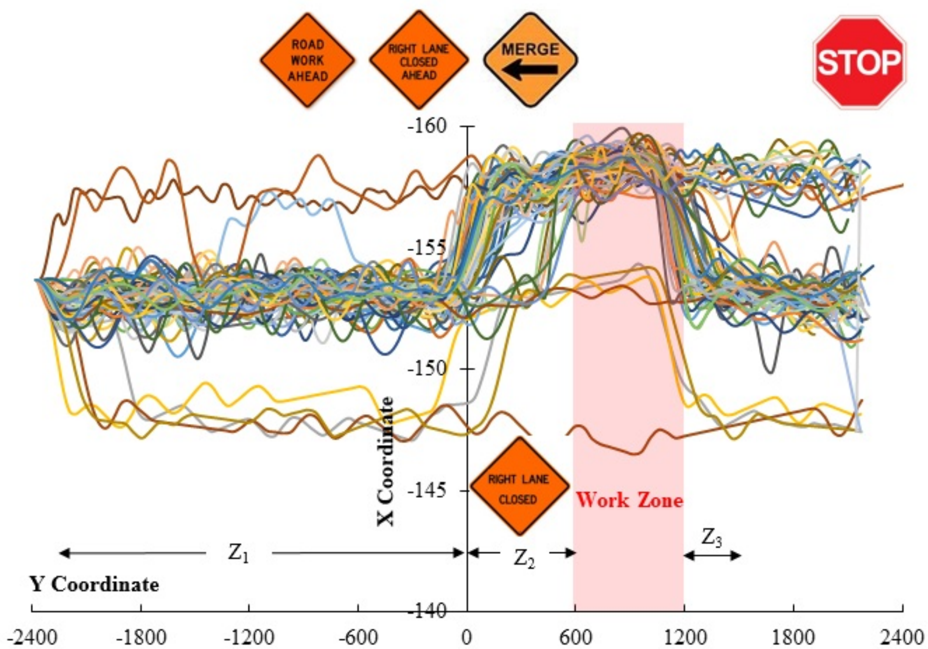

Figure 7. Plot of 75 driving paths - MoDOT left merge scenario

An analysis of drivers' merging behavior in the MUTCD and MoDOT left merge scenarios indicated that more than 90% of drivers merged in Z1. About 42% and 71% of the drivers who merged in Z_1 were between 45-64 years old in the MUTCD and MoDOT scenarios, respectively. There was no correlation with gender and merging in Z1 during each of the left merge configuration as gender was split between males and females.

Analysis of drivers' merging behavior in the MUTCD and MoDOT left merge scenarios indicates that about 9% of the participants merged in Z2 . For merging left in Z_2, 70% and 42% of drivers were 18-24 years old in the MUTCD and MoDOT scenarios, respectively. In regards to gender, about 70% and 45% are females for the MUTCD and MoDOT scenarios, respectively.

In both MUTCD and MoDOT left merge scenarios, about 1% of the drivers merged right again after their first merge to the left in Z_3. In the MUTCD scenario, 62% of drivers were in the range of 25-44 years. In the MoDOT scenario, 50% of the participants were in the range of 25-44 years. In the MUTCD scenario, 62% of the drivers were female, while 50% of drivers were female in the MoDOT scenario (Table 3).

Based on the results, drivers in the range of 45-64 years merged left immediately after they started driving, but younger drivers in the range of 18-24 years, preferred to merge left in the middle of the path before the work zone. Based on merging pattern of drivers of four scenarios, the middle age drivers (25-64 years) prefer to merge to the other lane immediately after they start driving while young drivers (18-24 years) tend to merge to the other lane in the middle of the path, before start of work zone. Regarding the variations in merging pattern of different genders, female drivers merge to other lane after they start driving while male drivers prefer to merge to open lane in the middle of path before work zone.

In MUTCD left merge scenarios and MoDOT left merge scenarios four and five drivers missed the signs respectively, and drove throughout the work zone. Data corresponding to these drivers was eliminated from the analysis.

Table 3. Merging behavior of drivers

Zone		MUTCD left merge	MoDOT left merge
Z_1	Percentage of drivers	92	91
	Age group	42% drivers aged 45-64	71% drivers aged 45-64
	Gender	52% female	58% female
Z_2	Percentage of drivers	8	9
	Age group	70% drivers aged 18-24	42% drivers aged 18-24
	Gender	70% female	45% female
Z_3	Percentage of drivers	<1	<1
	Age group	62% drivers aged 25-44	50% drivers aged 25-44
	Gender	62% female	50% female

The merging point is important for analyzing drivers' reactions to different merge signs. The majority of work zone crashes occur in lane closure areas due to driver merge driving behaviors and late lane merges are a significant cause of work zone crashes. Late lane merges occur when drivers decide to merge to the open lane at the very last moment before work zones, creating a safety threat for both drivers and workers in work zones (Datta, Schattler, Kar, & Guha, 2004). It is safest if the vehicles move into the open lane as far before the work zone as possible. The sooner the merge starts, the safer the travel through a work zone will be. Therefore, how the merge changes were determined with alternative signs on average using the patterns collected from the driving simulation. In the MUTCD right merge scenario, four drivers merged right late, near the taper. In the MoDOT right merge, two drivers merged late to the right lane. In the MUTCD left merge, one driver merged late to the left lane, while there were no drivers in the MoDOT left merge who merged late to the left lane.

An analysis of variance was conducted to find out the effect of sign configuration, age, and driver's gender on merge location; Table 4 presents the results. The scenario type (MUTCD vs. MoDOT) does not play a significant role in the location of merging, with P-values of 0.918. In other words, the null hypothesis will not be rejected. Given that the observed P-values are less than 0.05 for age and gender, it can be observed that these two factors play a significant role in the merging location. In other words, the null hypothesis (H0) will be rejected and these two factors have statistically significant effects on merging location.

Table 4. Analysis of Variance of Lane Change

Source	DF	Adj SS	Adj MS	F-Value	P-Value
Scenario	3	0.0025	0.00246	0.01	0.918
Gender	1	1.1897	1.18967	5.16	0.025
Age	3	2.4404	0.81346	3.53	0.017

4.2. SPEED ANALYSIS

In this section, driver mean speed, standard deviation, and the 85th percentile test at two locations for the four scenarios are evaluated. The driver’s speed was recorded at the sign “Merge” in MUTCD and “MERGE/arrow” (y= 370) in MoDOT, and at the start of the work zone (y=670). The results reveal that at both points of measurement during the right merge scenarios, speeds were the lowest for the MUTCD configuration. The differences of 0.005 kph and 0.004 kph in the 85th percentile and 0.06 kph and 0.006 kph in the standard deviation at the merge sign and the beginning of work zone in MUTCD and MoDOT right merge scenarios, respectively, are not significant (Table 5).

Table 5. Speed behavior of drivers in right merge scenarios

	Sign “Merge” in MUTCD and “Merge/arrow” in MoDOT (y=370)		Start of work zone (y=670)	
	MUTCD right merge	MoDOT right merge	MUTCD right merge	MoDOT right merge
Mean	56.09	57.57	54.49	55.35
Standard deviation	15.175	15.115	15.621	15.615
85th Percentile	70.138	70.143	70.135	70.139
P value (Mean speed)	0.174		0.44	

During left merge scenarios, somewhat lower speeds were recorded for simulations that featured the MoDOT merge sign. The difference of 0.89 kph in mean speed was not statistically significant. The speeds recorded at the beginning of the work zone were lower for the MUTCD sign configuration. The difference of 1.39 kph in mean speed was recorded at the beginning of the work zone. The differences of 0.006 kph and 0.002 kph for the 85th percentile and 0.46 kph and 0.445 kph for the standard deviation at merge sign and the beginning of work zone in MUTCD and MoDOT left merge scenarios, respectively, are not significant (Table 6).

Thus, based on the speed analysis, there were no significant differences between the MoDOT sign and the MUTCD sign. As expected, in both scenarios, mean speed decreases from the merge sign to the start of the work zone. Given the same results of speed for both signs, MoDOT alternative signs could be considered viable alternatives for MUTCD signs. The 85th percentile values presented a similar trend as the mean values.

Table 6. Speed behavior of drivers in left merge scenarios

	Sign "Merge" in MUTCD and "Merge/arrow" in MoDOT (y=370)		Start of work zone (y=670)	
	MUTCD left merge	MoDOT left merge	MUTCD left merge	MoDOT left merge
Mean	57.4	56.51	55.09	56.48
Standard deviation	15.165	14.705	14.986	14.541
85th Percentile	70.144	70.138	70.139	70.141
P value (Mean speed)	0.447		0.151	

The analysis of variance was done to measure the effectiveness of the MoDOT sign as compared to the MUTCD sign. This is presented in Table 7. The scenario type

(MUTCD vs. MoDOT) does not play a significant role in speed, with P-values of 0.649; so it could be concluded that the null hypothesis will not be rejected. Considering the observed P-values of 0.000 for age and gender, it can be concluded that these factors have a significant role in the driver's speed.

The results of mean speed are in agreement with observations reported by Zhu et al. (2015) and Edara et al. (2013), where the authors reported lower mean speeds for the MUTCD right merge in front of Merge arrow and start of work zone.

Table7. Analysis of Variance of Speed

	DF	Adj SS	Adj MS	F-Value	P-Value
Scenario	3	259	86.3	0.55	0.649
Gender	1	4449	4449.5	28.30	0.000
Age	3	49958	16652.6	105.92	0.000

4.3. DISCUSSION AND IMPLICATION FOR TRAFFIC MANAGEMENT PROFESSIONALS

It is evident in the literature that the sooner the merge starts; the safer it is to travel through a work zone. Therefore, data analysis and results presented for this study focused on how to determine the start -of-the-merge change behavior varied when comparing the MUTCD sign configuration against the MoDOT alternative signs. The start- of-the merge-points was determined individually for each driver and for each configuration. These individual points allowed the calculation of a more representative start of-the-merge point for each participant and for each scenario using graphical analysis as presented above. Further, this allowed the deletion of inconsistencies of behavior that would not be present in an actual work zone. The valid data was then used for the comparisons of the MUTCD sign configurations with the MoDOT alternate sign configurations.

Further, it was clear from post questionnaire results, that the first sign, "Work zone ahead," is the most critical to alert drivers that they are approaching a work zone. Also, participants noted that they preferred the MoDOT alternate sign configurations, including the positioning of signs on each side of the roadway, to the MUTCD-approved sign configurations. Traffic managers can use this information in sign placement and other warning strategies to alert drivers of upcoming work zones.

5. CONCLUSIONS

The frequency and severity of crashes in work zones are remarkably higher than those occurring on normal roads. This is most likely due to capacity reduction and lane changes throughout work zones. Improving safety throughout work zones is a major concern of traffic managers. A literature review shows that temporary traffic control signs are useful for the improvement of safety in work zones by guiding and directing in regards to upcoming work zones. This study demonstrates the importance of collecting and analyzing driving patterns with a driving simulator to evaluate the effectiveness of traffic management measures in work zones.

Based on the data analysis, there was not a noticeable, statistical difference in location of merging between the MUTCD and MoDOT alternative signs. The simulation results showed that the age of the drivers had a significant effect on the location of merging, which was expected. Similarly, the data showed that drivers' gender has a significant effect on the location of merging. In particular, based on the P-values, which are less than 0.05, hypothesis H0 is rejected; thus, there is sufficient evidence for one to conclude that both age and gender have significant effects on the location of merging.

In terms of safety, it is observed that fewer drivers that had late merge late in MoDOT scenarios compared to the MUTCD scenarios. Regarding the speed analysis, there is no difference between the average speeds of drivers in any of the scenarios. Based on statistical analysis, different scenarios did not have a significant effect on drivers' speed, but age and gender did seem to have a significant effect.

In future work, researchers should consider the impact of traffic, multiple lane closures, and day versus night hour to evaluate MUTCD and MoDOT sign configurations. Although outside the scope of this project, it would be interesting to gauge the reaction of professional drivers to the two sign configurations to determine the implications for roadway freight corridor design and management. The impact of distracted driving should also be considered.

ACKNOWLEDGEMENTS

We would like to thank the Missouri Department of Transportation (MoDOT TR201512) for partially funding this research. We would also like to acknowledge the

Missouri S&T Engineering Management and Systems Engineering Department and the Intelligent Systems Center for providing student support.

REFERENCES

Aghazadeh, F., Ikuma, L. H., & Ishak, S. (2013). Effect of changing driving conditions on driver behavior towards design of a safe and efficient traffic system (No. SWUTC/13/600451-00103-1).

Beacher, A. G. (2004). Final report evaluation of the late merge work zone traffic control strategy.

Bham, G. H., Leu, M. C., Vallati, M., and Mathur, D. R. (2014). Driving Simulator Validation of Driver Behavior with Limited Safe Vantage Points for Date Collection in Work Zones. Journal of Safety Research (49), 53-60.

Brewer, M., Pesti, G., & Schneider IV, W. (2006). Improving compliance with work zone speed limits effectiveness of selected devices. Transportation Research Record: Journal of the Transportation Research Board, (1948), 67-76.

Brown, H., Sun, C., Cope, T. (2015). Evaluation of Mobile Work Zone Alarm Systems (Report No. cmr 15-011). Missouri Department of Transportation.

Datta, T. K., Schattler, K. L., Kar, P., & Guha, A. (2004). Development and evaluation of an advanced dynamic lane merge traffic control system for 3 to 2 lane transition areas in work zones (No. Research Report RC-1451).

Edara, P. K., Sun, C., & Zhu, Z. (2013). Investigation of Alternative Work Zone Merging Sign Configurations (No. MATC-MU: 176). Mid-America Transportation Center.

Edara, P.K., Sun, C., and Brown, H. (2017). Evaluation of Work Zone Split Traffic Symbol Sign (Report No. cmr17-009). Missouri Department of Transportation.

Elrod, C. C., Daughton, W. J., Murray, S. L., & Flachsbart, B. B. (2010). Evaluating the engineering management brand meaning. Engineering Management Journal, 22(2), 16-26.

FHWA (2009). “Work Zone safety and mobility fact sheet - 10th Annual National Work Zone Awareness Week.” https://safety.fhwa.dot.gov/wz/wz_awareness/2009/factsht09.cfm (May 25, 2014).

FHWA (2017). Facts and Statistics –Work Zone Safety. http://www.ops.fhwa.dot.gov/wz/resources/facts stats/injuries fatalities.htm (July 10, 2017).

Grillo, L., Datta, T., & Hartner, C. (2008). Dynamic late lane merge system at freeway construction work zones. Transportation Research Record: Journal of the Transportation Research Board, (2055), 3-10.

Harb, R., Radwan, E., Ramasamy, S., Abdel-Aty, M., Pande, A., Shaaban, K., & Putcha, S. (2009). Two simplified dynamic lane merging system (SDLMS) for short term work zones.

Harb, R., Radwan, E., Yan, X., Pande, A., & Abdel-Aty, M. (2008). Freeway work-zone crash analysis and risk identification using multiple and conditional logistic regression. Journal of Transportation Engineering, 134(5), 203-214.

Hou, Y., Sun, C., & Edara, P. (2012). Statistical test for 85th and 15th percentile speeds with asymptotic distribution of sample quantiles. Transportation Research Record: Journal of the Transportation Research Board, (2279), 47-53.

Hurwitz, D., Heaslip, K., & Moore, D. (2012). Relating transportation systems management and operations strategies to policy goals: A framework for quantitative decision making. Engineering Management Journal, 24(3), 32-42.

Injuries and fatalities unfortunately do happen, 2016. National Safety Council, Retrieved from http://www.nsc.org/WorkplaceTrainingDocuments/Workzone%20Injury%20Facts.pdf.

Khattak, A. J., Khattak, A. J., & Council, F. M. (2002). Effects of work zone presence on injury and non-injury crashes. Accident Analysis & Prevention, 34(1), 19-29.

Meng, Q., & Weng, J. (2011). Evaluation of rear-end crash risk at work zone using work zone traffic data. Accident Analysis & Prevention, 43(4), 1291-1300.

Meng, Q., Weng, J., & Qu, X. (2010). A probabilistic quantitative risk assessment model for the long-term work zone crashes. Accident Analysis & Prevention, 42(6), 1866-1877.

Montgomery, D. C., & Runger, G. C. (2010). Applied statistics and probability for engineers. John Wiley & Sons.

NHTSA (2014). “A Compilation of Motor Vehicle Crash Data from the Fatality Analysis Reporting System and the General Estimates System.” https://crashstats.nhtsa.dot.gov/Api/Public/ViewPublication/812261 (2014).

Oltedal, S., & Rundmo, T. (2006). The effects of personality and gender on risky driving behaviour and accident involvement. Safety science, 44(7), 621-628.

Reyes, M. L., Khan, S. A., & Initiative, S. W. Z. D. (2008). Examining driver behavior in response to work zone interventions: A driving simulator study. Master’s Thesis, University of Iowa, USA.

Sadati, S., Arezoumandi, M., Khayat, K. H., & Volz, J. S. (2016). Shear performance of reinforced concrete beams incorporating recycled concrete aggregate and high-volume fly ash. Journal of Cleaner Production, 115, 284-293.

Zhu, Z., Edara, P., & Sun, C. (2015). Case study of an alternative merging sign design for temporary traffic control in work zones. Journal of transportation engineering, 142(1), 05015005.

II. USING A COMBINATION OF MULTI-CRITERIA DECISION-MAKING AND DATA MINING METHODS FOR WORK ZONE SAFETY: A CASE ANALYSIS

Samareh Moradpour, Suzanna Long

Department of Engineering Management and Systems Engineering, Missouri University of Science and Technology 600 W. 14th St. Rolla, MO 65409-0370

ABSTRACT

Work zone accidents are important concerns for transportation decision-makers. Therefore, knowledge of driving behaviors and traffic patterns are essential for identifying significant risk factors (RF) in work zones. Such knowledge can be difficult obtain in a field study without introducing new risks or driving hazards. This research uses integrated data mining and multi-criteria decision-making methods as part of a simulator-based case study of work zone logistics along a highway in Missouri. The research design incorporates *k*-mean clustering to cluster driving behavior trends, step-wise weight assessment ratio analysis (SWARA) to determine weights for criteria that are most likely to impact work zones, and the VIKOR method to rank the alternatives (clusters). Transportation engineers and decision makers can use results from this case study to identify driving populations most likely to engage in risky driving behaviors within work zones, and to provide guidance on effective work zone management.

Keywords: Case Study, Multi-criteria decision-making, Data mining, *k*-mean clustering, SWARA, VIKOR method

1. INTRODUCTION

Work zone safety and mobility are major concerns for traffic managers due to the high rate of accidents in work zones. Mandatory lane changing and consequent merging of traffic in work zones resulting from lane closures tend to increase drivers' dangerous maneuvers. This increases the likelihood and severity of crashes in work zones compared to unencumbered roads (Fei et al., 2016). According to the Federal Highway Administration (FHWA) statistics, there were 669 fatalities in work zone crashes in 2014, or about 1.8 work zone fatalities per day (FHWA, 2017).

To model work zone crashes as a system requires knowledge of several independent variables such as driver behavior, vehicle characteristics, environmental and geographical conditions, occupant behavior, other road user mechanics, and roadway conditions (Bayam et al., 2005). Of these, Stanton and Salmon (2009) showed statistical evidence that driver error was the more frequent cause (75 to 95% of the cases) of work zone crashes. Further, Guo and Fang (2012) suggest that only 6 % of total drivers exhibit risky driving behavior, but these drivers cause 65% of work zone crashes. Risky driving behavior includes aggressive lane changing, speeding, careless driving, not paying attention to pedestrians, and ignoring traffic control signs (Weng & Meng, 2012; Luke& Heyns, 2014; American Transportation Research Institute, 2011). Therefore, it is essential to evaluate driver behavior as a risk factor (RF) in any model designed to improve work zone safety and management.

This research presents a case study in which a driving simulator is used to identify risky drivers in work zones. Data analytic tools determine patterns and cluster behaviors within the simulations. For example, *k*-mean clustering detects trends between simulation runs based on the available similarities, step-wise weight assessment ratio analysis (SWARA) method weights the factors most likely to impact work zone safety and efficiency (henceforth call "criteria"), and the VIKOR method ranks the alternatives(clusters). The result of this research is the development of analytic data mining and Multi-Criteria Decision-Making (MCDM) methods that improve the safety and efficiency of work zones.

2. RELATED WORK

A significant means to reduce RFs in work zones (especially those due to the rate of fatalities and high impact of driver error) is to identify high-risk drivers. Some recent work on such identification has focused on demographics. Oltedal & Rundmo (2006) investigated the effects of personality traits and gender on risky driving behavior and accident involvement. Results indicate that over 37% of the variance in risky driving is explained by the personal behaviors and gender. Moreover, Rhodes & Pivik (2011) was conducted a survey in Alabama determined that male drivers engaged in risky driving behavior more frequently than female drivers did. They found that teen drivers are more frequently engaged in risky driving behavior.

Long et al. (2017) evaluated driver reaction to Manual on Uniform Traffic Control Devices (MUTCD) lane-shift sign configurations and alternative lane-shift sign configurations for work zones. Seventy-five participants tested two scenarios in a driving simulator, and found gender had no significant effect on driver lane-change patterns, but driver age could affect their lane-shift patterns. In addition, both age and gender had an effect on driver average speed (Long et al., 2017; Thind et al., 2017).

Weng and Meng (2011) used data from the Fatality Analysis Reporting System (FARS) to investigate driver casualty risk in work zones. Results indicated a 17% increase in risk of injury or fatal accidents in construction work zones for middle-aged drivers compared to those of the young ones. The rates were even higher for maintenance work zones, with a 24% increase for middle-aged drivers. Moreover, a higher casualty risk was observed for the female drivers in construction and utility work zones.

Harré et al., (2000) studied risky driving behavior using the Multivariate Analysis of Variance (MANOVA) and univariate regression analysis. Gender was identified as a key RF.

These studies use univariate statistical or multivariate regression methods to identify associated RFs and different groups such as age or gender. However, univariate statistical methods only consider a single factor at a time. Given the potential interactions that different contributing factors can have on risky behavior, the isolation of a single factor for analysis, while treating all else as fixed, can lead to bias. Alternatively, multivariate regression methods address independency between state variables (An increase in input value for one variable forces a reduction or increase in the values for other variables), however, such an assumption is not typically valid in driving behavior analysis. Therefore, the multivariate regression method may not accurately represent the relationship between the risky driving behavior and its governing factors (Weng & Meng, 2012).

In addition, road accidents and crashes are unpredictable and knowledge of the relevant factors are necessary for analysis. These accidents are associated with normally discrete variables, and therefore, heterogeneity within the data yields insight into any interdependence. Therefore, *k*-mean clustering analysis is useful to highlight this issue (Weng and Meng, 2011).

This research presents a case study that addresses these gaps by using a combination of data mining (DM) and MCDM methods. The proposed analytic method emphasizes that integrating DM and MCDM methods can provide a comprehensive assessment of driving behavior identification and allows transportation professionals a roadmap for better decision making to promote safety in work zones.

3. METHODOLOGY

The case study design separates the research method into four phases. The first phase focuses on data collection. Data collected in simulator runs included driver characteristics, merge locations, merge speed, mean speed, sign locations, and the like. More detail regarding data collection is described in part 4.1. The second phase of the framework uses *k*-mean clustering as part of data mining. Clustering analysis was helpful in extracting patterns from a large amount of data and in the identification of underlying patterns in driving behavior. The third phase was the use of SWARA methods. The SWARA method is used to calculate the relative weights for the criteria. Finally, the weight of the criteria is established using the VIKOR method to rank the clusters (alternatives).

3.1. K-MEAN CLUSTERING METHOD

A significant challenge in pattern recognition within transportation problems is how to process the huge amount of data. *K*-mean clustering methods are capable of extracting patterns from large amounts of transportation data (Jain, 2010). Indeed, *K*-mean clustering techniques were designed to identify hidden patterns by extracting information from the data to predict activities, determine trends among the data, and group (cluster) data based on similarities (Moradpour et al., 2017; Rygielski et al., 2002).

K-mean clustering method was developed over 50 years ago, and it is one of the more common clustering methods. It has been widely used in such diverse disciplines as psychology, biology, and marketing research (Jain, 2010; Zhu et al., 2018). *K*-mean clustering segments data into clusters (groups) based on similarities and characteristics between the data (Peng et al., 2011). The outputs of *k*-mean clustering (k clusters) are the inputs (alternatives) to be used in the VIKOR method which needs to be ranked (Saxena et al., 2017).

The following optimization model determines the cluster means of $\{\bar{y}_k\}$, by minimizing the sum of the squared error.

Minimize:

$$\underset{\{\bar{y}_k\}}{sse_K} = \sum_{i \in I_{ML}} \sum_{k=1}^{K} z_{ki} \left\| y_{ML,i} - \bar{y}_k \right\|^2 \tag{1}$$

Subject to

$\sum_{k=1}^{K} z_{ki} = 1$, for $i \in I_{ML}$

$z_{ki}'s$ are binary variables (Moradpour & Long, 2017)

The k-mean clustering procedure consists of four phases (Jain, 2010):

Select the initial number of cluster (k)

Assign patterns to the nearest cluster

Compute the sum of square error

Repeat phase 2 and 3 until the cluster sum of square error stabilizes.

3.2. SWARA METHOD

MCDM is an area of operation research (OR) that helps researchers evaluate, rank, and select under conflicting criteria based on the priorities of the decision-maker(s). MCDM methods are capable to consider experts' and decision-makers' opinion and ideas regarding criteria importance and weight in the decision-making process. In transportation research, decision-makers usually deal with complex and sometimes conflicting criteria related to the environment, safety, economic, sustainability, and pattern recognition. This ability makes MCDM methods applicable to transportation decision-making and policy regulation.

The SWARA method was used in this research to determine the weights of the incorporated criteria. The aim of the SWARA method is the opportunity to estimate experts' opinions about the ratio of criteria for determining weight. In this method, the most important criterion is given the top rank while the least important criterion is given the lowest rank. The process of determining the weight of the criteria that helps to estimate the differentiation of their importance is described below.

Step 1. Sort the criteria based on their expected importance in descending order.

Step 2. Compute the comparative importance of the average value, sj.

Step 3. Compute the coefficient kj as follows:

$$k_j = s_j + 1 \tag{2}$$

Step 4. Compute the recalculated weight, wj, as follows:

$$w_j = \frac{k_{j-1}}{k_j} \tag{3}$$

Step 5. Compute the relative weights of the evaluation criteria (Dehnvai et al., 2015)

$$q_j = \frac{w_j}{\sum w_j} \tag{4}$$

MCDM is an area of operation research (OR) that helps researchers evaluate, rank, and select alternatives under conflicting criteria based on the priorities of the decision-maker(s). MCDM methods are capable of integrating expert and decision-making ideas and opinions regarding criteria. In transportation research, decision-makers usually deal with complex and sometimes conflicting criteria related to the environment, safety, economy, sustainability, and pattern recognition. This ability makes MCDM methods applicable to transportation decision-making and policy regulation (Zopounidis & Doumpos, 2002).

The SWARA method is used in this research to determine weights for the selected criteria. The SWARA method quantifies expert determinations of the relative weights between a ratio (pair) of criteria. In this method, the most important criterion is ranked the highest while the less important criterion is ranked lower. This process of weighting the criteria is described below:

Step 1. Sort the criteria based on their expected importance in descending order.

Step 2. Compute the comparative importance of the average value, s_j.

Step 3. Compute the coefficient k_j as follows:

$$k_j = s_j + 1 \tag{5}$$

Step 4. Compute the recalculated weight, w_j, *as* follows:

$$w_j = \frac{x_{j-1}}{k_j} \tag{6}$$

Step 5. Compute the relative weights of the evaluation criteria (Dehnvai et al., 2015),

$$q_j = \frac{w_j}{\sum w_j} \tag{7}$$

3.3. VIKOR METHOD

The VIKOR method ranks a set of alternatives by measuring the closeness of the solution to an ideal solution (Opricovic & Tzeng, 2007; Moradpour et al., 2011). The VIKOR algorithm is as follows:

Step 1. Determine the best and the worst values of all criterion functions, for $i=1, 2, .., , n;$

$$f^* = max\ f_{ij} \tag{8}$$

$$f^- = min\ f_{ij} \tag{9}$$

Step 2. Compute values of S_j and R_j, $j=1, 2,\ldots, n$. The w_j is the weight of criteria which was calculated by the SWARA method;

$$S_j = \sum_{i=1}^{n} w_i \,(f_i^* - f_{ij})/(f_i^* - f_i^-) \tag{10}$$

$$R_j = \max[\, w_i \,(f_i^* \, - f_{ij})/(f_i^* - f_i^-)\ ] \tag{11}$$

Step 3. Compute Q_j , $J=1, 2, \ldots, n$,

$$Q_j = \frac{v\big(\,S_j - S^*\big)}{S^- - S^*} + (1 - v)(R_j - R^*)/(R^- - R^*) \tag{12}$$

where,

$$S^* = min\, S_j \text{ , } S^- = max\, S_j \tag{13}$$

$$R^* = min\, R_j \text{ , } R^- = max\, R_j \tag{14}$$

and *v* is the weight of the strategy of the majority of criteria. Normally, *v* was assumed as *v*=0.5. However, v can take any value from zero to one (San Cristóbal, 2011).

Step 4. Rank the alternatives based on *S*, *R*, and *Q* in descending order in three lists.

Step 5. Suggest a compromise solution of the alternative (a') that is on the top of the ranked list of Q if satisfy two conditions:

Condition 1: Acceptable advantage: if Q (a'') - Q (a') > DQ, where a'' is the second best alternative based on Q ranking, and DQ = $\frac{1}{(J-1)}$, while J is the number of alternatives.

Condition 2: Acceptable stability in decision-making. Alternative a' should be the best alternative based on S and R rankings. This compromise solution can be considered stable in a decision-making process, which could be ‘‘voting by majority rule’’ (when v > 0.5 is needed), ‘‘by consensus’’ v ~ 0.5, or ‘‘with veto’’ (v < 0.5), where v is the weight obtained for the strategy of decision-making ‘‘the majority of criteria’’ (or ‘‘the maximum group utility’’).

A set of compromising solutions is suggested for situations where one of the aforementioned conditions is not met:

• Alternatives a' and a'' if only Condition 2 is not satisfied.

• Alternativesa', a'',..., a^M if Condition 1 is not satisfied, and a^M is determined by the relation Q(a^M) - Q(a') < DQ for maximum M (the positions of these alternatives are ‘‘in closeness’’). The best alternative based on ranking of Q values has the minimum Q, but the main ranking result is the compromise ranking of alternatives (Opricovic & Tzeng, 2004).

4. CASE STUDY

The procedures used for data collection and analysis are elaborated in this section.

4.1. DATA COLLECTION

A case study is used to demonstrate the application of this research design by comparing safety and efficiency of traffic merging patterns associated with a short-term work zone along a Missouri highway. Driver driving data are recorded in these two different scenarios. The first scenario incorporated the Missouri Department of Transportation's (MoDOT) alternative merge sign, while the second scenario is based on Manual on Uniform Traffic Control Devices (MUTCD) Temporary Traffic Control's (TTC) merge signage for short-term work zones.

The driving simulator (DS) used for data collection was a fixed base DS with a Ford ranger pickup cabin, three 3,000-lumen Liquid Crystal Display (LCD) projectors, a projection screen, and a master simulation computer (Figure 1) (Moradpour et al., 2015). The simulated cabin included a steering wheel, accelerator pedal, brake pedal, speedometer that had sensors inputting data into a data acquisition system. The data acquisition system collected driver data such as *x* and *y* location coordinates steering wheel angel, braking amount, time, and speed.

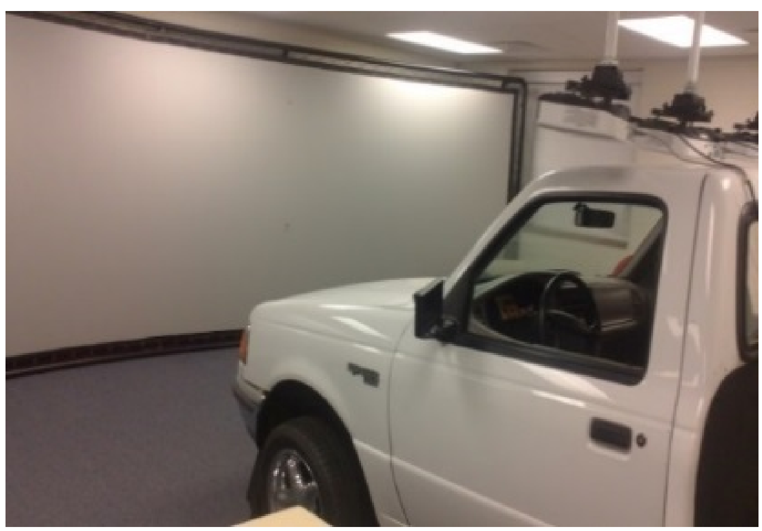

Figure 1**.** Driving Simulator

Each of the seventy-five participants drove the two work zone scenarios. The ages of the participants are grouped into four age bins: 18-24, 25-44, 45-64, and over 65 years. Each participant also completed a questionnaire before entering the simulator the results of which are shown in Table 1.

Table 1. Demographic information of participants

Age				Gender		Number of Mile Driven Annually (Mileage)				Driving Experience (Year)			
18-24	25-44	45-64	≥65	Female	Male	<=1000	1000-5000	5000-10000	>=10000	<1	1-5	5-10	>10
11	28	27	9	41	34		11	12	46	2	9	3	61

The participants became familiar with the DS before the test began. They were also able to stop the test at any time if they felt uncomfortable. One volunteer experienced simulator sickness during the trial experience and was excused from participation in the formal study. Table 2 extracts a part of the dataset and consists of drivers driving data such as average speed, merge location, and merge speed for the two scenarios.

Table 2. A part of drivers' dataset

	Average Speed (km/h)		Merge Location (m)		Merge Speed (km/h)	
Driver	MODOT	MUTCD	MODOT	MUTCD	MODOT	MUTCD
1	57.498	58.036	-2106.1	-2181.6	56.750	48.390
2	50.385	50.678	-2226.6	-2273.7	44.995	42.590
-	-	-	-	-	-	-
74	61.331	60.440	-2141.9	-1954.8	66.259	75.843
75	53.355	57.2691	117.5	120.0	66.661	68.688

4.2. K-MEAN CLUSTERING METHOD

The first step in the *k*-mean clustering method is selecting the number of clusters (k). In this study, the elbow method was used for selecting *k* (Bholowalia & Phagwara, 2014), which led to the selection of four clusters. The statistical software Minitab 17 is used for clustering data into four different clusters. In this method, drivers are clustered based on their average speed, merge location, and merge speed. Table 3 presents a summary of the clustering results.

4.3. SWARA METHOD

The main risky driver behaviors that influence work zone safety are identified based from the literature (Table 4). The criteria, average speed (C_1), merge location (C_2), and merge speed (C_3), were selected in evaluating and ranking the alternatives (clusters).

Speeding is a significant factor the cause of accidents. Speeding reduces the driver's ability to control the vehicle for braking and going around curves and increases the severity of work zone crashes. In addition, speeding reduces the time to react to a changing situation. Based on the statistics, the rate of crashes in a road with posted limit of 60 km/h are doubled as a result of a 5 km/h increase in speed compared to the posted limited. (Luke & Heyns, 2014; Li & Bai, 2009).

Table 3. *K*-mean clustering results

	Scenario	Cluster1	Cluster 2	Cluster 3	Cluster 4
Average speed (km/h)	MODOT	51.010	49.844	50.676	42.426
	MUTCD	48.249	48.977	49.781	44.293
Merge location (m)	MODOT	2132.617	20	2315.239	152.5445
	MUTCD	166.837	38.233	2347.271	2166.709
Merge speed (km/h)	MODOT	59.210	47.563	59.312	50.016
	MUTCD	48.658	48.637	56.955	53.812

The location of the point of merger driving behavior generally ensures that most crashes occur in the lane closure area of work zones. Therefore the location of the point of merging (merge location, (C_2)) is significant for risk mitigation in work zone. Late lane merges are a significant cause of work zone crashes. Late lane merges occur when drivers decide to merge to the open lane at the very last moment before entering work zones, which creates a safety threat for both drivers and workers in work zones (Datta et al., 2004). Meng and Weng (2011) suggested that merging early is the most effective method to reduce rear-end crashes in work zones.

The merge speed (C_3) may have an effect on safe merging. Higher speeds during congested merging leads to crashes in work zones (Ahmmed et al., 2008) and these can be quite catastrophic.

Table 4. List of experts used for criteria evaluation

Criteria	Expert
Average speed	Meng et al., (2010), Li and Bai (2009), Aarts and Schagen (2006), Li and Bai (2009), Debnath et al. (2015), Li and bai (2006)
Merge location	Meng and Weng (2011), Weng and Meng(2011)
Merge speed	Ahammed et al. (2008)

The results of the SWARA method calculations revealed that average speed had the highest weight between criteria $w_1 = 0.39$. The merge location had the second highest weight of 0.32 between the considered criteria by $w_2 = 0.32$. Finally, based on the SWARA analysis, merge speed weight is equal to 0.29 ($w_3 = 0.29$). The results of the SWARA method are shown in Table 5.

Table 5. SWARA analysis results

	Index	Comparative importance of average value	Coefficient	Recalculated weight	Weight
Average speed	w_1	0	1	1	0.39
Merge location	w_2	0.22	1.22	0.82	0.32
Merge speed	w_3	0.11	1.11	0.73	0.29

4.4. VIKOR METHOD

In the last step of the framework, the VIKOR method was used to rank the clusters (alternatives). This method includes different decision-maker perceptions in the process.

The results of ranking alternatives (clusters) are as follow: $S_2 < S_4 < S_1 < S_3$, $R_4 < R_1 < R_2 < R_3$, and $Q_4 < Q_1 < Q_2 < Q_3$. Based on Q ranking, alternatives 1 and 4 are the top ranked. These two results are compared with the required solution conditions:

Condition1: the DQ = $\frac{1}{(J-1)}$ = 0.33, and Q (1)-Q (4) = 0.96>= 0.33. As a result, condition 1 is satisfied.

Condition 2: based on the VIKOR ranking $Q_4 < Q_1$, $S_2 < S_4$, and $R_4 < R_1$. These results did not satisfy Condition 2.

Therefore, one of the two conditions of the VIKOR analysis was not satisfied and a compromise solution was the outcome of this problem. Based on this compromise solution, Alternatives 1 and 4 (Cluster 1 and Cluster 4) are best. In other words, drivers in these two clusters are the best drivers based on safety considerations (see Table 6). Driver characteristics for each cluster are presented in Table 7.

Table 6. VIKOR ranking results

Alternatives	S	R	Q
Cluster 1	1.301402	0.39	1.015848
Cluster 2	0.745881	0.337007	1.572052
Cluster 3	2.051735	0.456937	2
Cluster 4	0.813312	0.293501	0.051638

Table 7. Characteristics of drivers in each cluster (Percentage)

Drivers characteristics		Cluster1	Cluster 2	Cluster 3	Cluster 4
Number of Driver		0.1	0.48	0.33	0.09
Female		0.146	0.561	0.195	0.122
Male		0.059	0.382	0.500	0.029
Age (year)	18-24	0.273	0.182	0.455	0.091
	25-44	0.000	0.464	0.464	0.071
	45-64	0.185	0.593	0.185	0.037
	>=65	0.000	0.556	0.222	0.222
Driving experience (year)	<1	0.500	0.000	0.500	0.000
	1-5	0.333	0.111	0.556	0.000
	5-9	0.000	0.333	0.667	0.000
	>=10	0.066	0.557	0.279	0.098
Driving mileage (mile)	<1000	0.333	0.000	0.667	0.000
	1000-5000	0.091	0.273	0.546	0.091
	5000-10000	0.083	0.417	0.417	0.083
	>=10000	0.087	0.609	0.217	0.087

5. RESULTS AND DISCUSSIONS

Based on the analysis, Clusters (alternatives) 1 and 4 are the best alternatives regarding driver safety patterns. Cluster (alternative) 3 is the least desired alternative in the ranking. In other words, this group has the least safe driving pattern compared to other alternatives.

Cluster 1 consisted of 8 drivers, more females ranging from 18 to 24 years old. Most of the drivers in this cluster had less than 1 year driving experience. Cluster 4 consisted of 6 drivers, more female ranging older than 65 years old. Most of the drivers in this cluster had more than 10 year driving experience.

Cluster 3 contains 25 drivers, more middle age (25-44 years old) male drivers compared to the other clusters. The plurality of the drivers in this cluster have between 5 and 9 years of driving experience. Most of the drivers in this cluster drive less than 1000 miles per year. This result was in agreement with Weng & Meng (2012), which stated that middle-age male drivers engaged in risky driving more than other drivers drive.

These results are in general agreement with Kleisen (2011) and Ericsson (2000), which determined female drivers participate in fewer accidents than male drivers do. In other word, male drivers are characterized as more risky drivers and drive at high speed than female drivers (Kleisen, 2011; Ericsson, 2000).

The literature suggests that driver intervention strategies focused on driver education are beneficial and that drivers completing training program have safer records (Gregersen, 1994; Takeda et al., 2011); this is supported through the findings of this case study. Results from this study suggest that driver education should target select scenarios and select driver demographics. Driving scenarios should provide training on safe merge behavior, merge timing, and merge speed control to promote early merge behaviors. Participant strategies should focus on male drivers between the age of 25-44 and older drivers. These trainings can be offered as part of driving improvement programs for the public, but could also be part of mandated driver safety protocols for those who have driving violations. Results show the importance of integrating multiple analytic methods in order to develop robust traffic management and driver education programs.

6. CONCLUSIONS

Work zones have a significant effect on traffic flow and safety. It is essential to identify key risk factors and include effective countermeasures as part of a comprehensive traffic management design for work zones. This case study addressed a gap in the literature by considering these risk factors in combination, rather than in isolation. Key findings provide effective validation of prior work while also providing fresh directions for work zone management and driver education.

Driver patterns and behaviors must be included as a key risk factor. This extends the findings of previous research that focused on age and gender. By using a combination of data mining (DM) with multi-criteria decision making (MCDM) methods, this case study identified patterns and behaviors associated with work zone merge scenarios most likely to contribute to an accident.

Driver training safety programs can be updated to include modules on merge behavior. These modules can be part of general driver-education training programs or built into mandatory driver improvement training for traffic offenders.

Future work should consider historical crash data as part of integrated DM/MCDM strategies. These data are often not considered due to challenges with data format, terminology, and related data integration issues. Nevertheless, this data source contains a vital record of insights and findings from investigating officials that may prove useful in advancing work zone safety.

REFERENCES

Aarts, L., & Van Schagen, I. (2006). Driving speed and the risk of road crashes: A review. Accident Analysis & Prevention, 38(2), 215-224.

Ahammed, M. A., Hassan, Y., & Sayed, T. A. (2008). Modeling driver behavior and safety on freeway merging areas. Journal of Transportation Engineering, 134(9), 370-377.

American Transportation Research Institute, 2011, Predicting truck crash involvement: A 2011 update, viewed 22 February 2013, from http://www.atri-online.org/ research/results/ATRI_Crash_Predictor One_Pg_Summary_Apr_2011.pdf

Bayam, E., Liebowitz, J., & Agresti, W. (2005). Older drivers and accidents: A meta-analysis and data mining application on traffic accident data. Expert Systems with Applications, 29(3), 598-629.

Datta, T. K., Schattler, K. L., Kar, P., & Guha, A. (2004). Development and evaluation of an advanced dynamic lane merge traffic control system for 3 to 2 lane transition areas in work zones (No. Research Report RC-1451).

Debnath, A. K., Blackman, R., & Haworth, N. (2015). Common hazards and their mitigating measures in work zones: A qualitative study of worker perceptions. Safety science, 72, 293-301.

Dehnavi, A., Aghdam, I. N., Pradhan, B., & Varzandeh, M. H. M. (2015). A new hybrid model using step-wise weight assessment ratio analysis (SWARA) technique and adaptive neuro-fuzzy inference system (ANFIS) for regional landslide hazard assessment in Iran. Catena, 135, 122-148.

Ericsson, E. (2000). Variability in urban driving patterns. Transportation Research Part D: Transport and Environment, 5(5), 337-354.

Fei, L., Zhu, H. B., & Han, X. L. (2016). Analysis of traffic congestion induced by the work zone. Physica A: Statistical Mechanics and its Applications, 450, 497-505./

Federal Highway Administration (FHWA) (2017). Facts and statistics –Work zone safety. http://www.ops.fhwa.dot.gov/wz/resources/facts stats/injuries fatalities.htm (July 10, 2017).

Gregersen, N. P. (1994). Systematic cooperation between driving schools and parents in driver education, an experiment. Accident Analysis & Prevention, 26(4), 453-461.

Guo, F., & Fang, Y. (2013). Individual driver risk assessment using naturalistic driving data. Accident Analysis & Prevention, 61, 3-9.

Harré, N., Brandt, T., & Dawe, M. (2000). The development of risky driving in adolescence. Journal of Safety Research, 31(4), 185-194.

Jain, A. K. (2010). Data clustering: 50 years beyond K-means. Pattern recognition letters, 31(8), 651-666.

Juarez, P., Schlundt, D. G., Goldzweig, I., & Stinson, N. (2006). A conceptual framework for reducing risky teen driving behaviors among minority youth. Injury Prevention, 12(suppl 1), i49-i55.

Keršuliene, V., Zavadskas, E. K., & Turskis, Z. (2010). Selection of rational dispute resolution method by applying new step-wise weight assessment ratio analysis (SWARA). Journal of Business Economics and Management, 11(2), 243-258.

Kleisen, L. M. (2013). A positive view on road safety: Can ‘car karma’contribute to safe driving styles?. Accident Analysis & Prevention, 50, 705-712.

Li, Y., & Bai, Y. (2006). Fatal and injury crash characteristics in highway work zones. In Proc., Transportation Research Board 87 th Annual Meeting.

Li, Y., & Bai, Y. (2009). Effectiveness of temporary traffic control measures in highway work zones. Safety science, 47(3), 453-458.

Li, Y., & Bai, Y. (2009). Highway work zone RFs and their impact on crash severity. Journal of Transportation engineering, 135(10), 694-701.

Long, S., Qin, R., Konur, D., Leu, M. C., Moradpour, S., Thind, S., ... & Brown, H. (2017). Work zone simulator analysis: driver performance and acceptance of Missouri alternate lane shift configurations (No. cmr 17-002).

Luke, R., & Heyns, G. J. (2014). Reducing risky driver behaviour through the implementation of a driver risk management system. Journal of Transport and Supply Chain Management, 8(1), 1-10.

Mardani, A., Nilashi, M., Zakuan, N., Loganathan, N., Soheilirad, S., Saman, M. Z. M., & Ibrahim, O. (2017). A systematic review and meta-Analysis of SWARA and WASPAS methods: Theory and applications with recent fuzzy developments. Applied Soft Computing.

Meng, Q., & Weng, J. (2011). Evaluation of rear-end crash risk at work zone using work zone traffic data. Accident Analysis & Prevention, 43(4), 1291-1300.

Meng, Q., Weng, J., & Qu, X. (2010). A probabilistic quantitative risk assessment model for the long-term work zone crashes. Accident Analysis & Prevention, 42(6), 1866-1877.

Moradpour, S., Ebrahimnejad, S., Mehdizadeh, E., & Mohamadi, A. (2011). Using hybrid fuzzy PROMETHEE II and fuzzy binary goal programming for risk ranking: A case study of highway construction projects. Journal of Optimization in Industrial Engineering, (9), 47-55.

Moradpour, S., Wu, S., & Leu, M. C. (2015, January). Use of traffic simulators to determine driver response to work zone configurations. In Proceedings of the International Annual Conference of the American Society for Engineering Management. (p. 1). American Society for Engineering Management (ASEM).

Moradpour,S., Long, S.(2017). K-mean clustering method in transportation problems, a work zone simulator case study. In Proceedings of the International Annual Conference of the American Society for Engineering Management. American Society for Engineering Management (ASEM).

National Research Council. (2007). Preventing teen motor crashes: Contributions from the behavioral and social sciences: Workshop Report. National Academies Press.

Oltedal, S., & Rundmo, T. (2006). The effects of personality and gender on risky driving behaviour and accident involvement. Safety science, 44(7), 621-628.

Opricovic, S., & Tzeng, G. H. (2004). Compromise solution by MCDM methods: A comparative analysis of VIKOR and TOPSIS. European journal of operational research, 156(2), 445-455.

Opricovic, S., & Tzeng, G. H. (2007). Extended VIKOR method in comparison with outranking methods. European journal of operational research, 178(2), 514-529.

Peng, Y., Zhang, Y., Tang, Y., & Li, S. (2011). An incident information management framework based on data integration, data mining, and multi-criteria decision-making. Decision Support Systems, 51(2), 316-327.

Rhodes, N., & Pivik, K. (2011). Age and gender differences in risky driving: The roles of positive affect and risk perception. Accident Analysis & Prevention, 43(3), 923-931.

Rygielski, C., Wang, J. C., & Yen, D. C. (2002). Data mining techniques for customer relationship management. Technology in society, 24(4), 483-502.

Saxena, A., Prasad, M., Gupta, A., Bharill, N., Patel, O. P., Tiwari, A., ... & Lin, C. T. (2017). A review of clustering techniques and Developments. Neurocomputing.

Stanton, N. A., & Salmon, P. M. (2009). Human error taxonomies applied to driving: A generic driver error taxonomy and its implications for intelligent transport systems. Safety Science, 47(2), 227-237.

Stanujkic, D., Karabasevic, D., & Zavadskas, E. K. (2015). A Framework for the selection of a packaging design based on the SWARA method. Engineering Economics, 26(2), 181-187.

Takeda, K., Hansen, J. H., Boyraz, P., Malta, L., Miyajima, C., & Abut, H. (2011). International large-scale vehicle corpora for research on driver behavior on the road. IEEE Transactions on Intelligent Transportation Systems, 12(4), 1609-1623.

Thind, S.S., Moradpour,S., Konur,D., Long, S., Qin, R.(2017). Statistical comparison of lane shift sign configurations with driving simulation. Industrial and Systems Engineering Research Conference.

Weng, J., & Meng, Q. (2011). Analysis of driver casualty risk for different work zone types. Accident Analysis & Prevention, 43(5), 1811-1817.

Weng, J., & Meng, Q. (2011). Modeling speed-flow relationship and merging behavior in work zone merging areas. Transportation research part C: emerging technologies, 19(6), 985-996.

Weng, J., & Meng, Q. (2012). Effects of environment, vehicle and driver characteristics on risky driving behavior at work zones. Safety science, 50(4), 1034-1042.

Zhu, Q., Sarkis, J., & Lai, K. H. (2018). Regulatory policy awareness and environmental supply chain cooperation in China: A regulatory-exchange-theoretic perspective. IEEE Transactions on Engineering Management, 65(1), 46-58.

III. WORK ZONE SAFETY IN MISSOURI: A STATISTICAL ANALYSIS

Samareh Moradpour[1], Brian Smith[2], Suzanna Long[1]

[1]Department of Engineering Management and Systems Engineering, Missouri University of Science and Technology 600 W. 14th St. Rolla, MO 65409-0370

[2]Mississippi State University, Industrial and Systems Engineering, McCain Engineering Building 260G, Starkville, MS 39762

ABSTRACT

Although tremendous amounts of crash data is collected, little of it is analyzed to improve work zone safety. Transportation managers usually focus on reducing risk factors that lead to crashes in work zones and require robust tools and analysis processes to identify these risk factors. In this research, multinomial logistic regression (MLR) is used to model historical data of Missouri work zone crashes to identify patterns and categories of factors that statistically contribute to work zone crashes. Results confirm that road grade, road curvature, lighting, and weather have statistical impacts on the severity and impact of crashes. By sorting these factors into functional categories, the results will assist transportation decision makers in integrating signage and communication strategies into work zone design and management.

Keywords: Multinomial Logistic Regression, Work Zone Safety, Decision Makers, Risk Factor

1. INTRODUCTION

In recent years, the focus of many states has shifted from building new highways to maintenance and rehabilitation, which gives rise to scheduled construction activities on existing roadways as part of managed work zones. In peak construction season, about twenty percent of all U.S. highways are under construction, which involves over three thousand work zones (Yang et al., 2015).

Work zone safety and mobility is one of the main concerns of Department of Transportations (DOTs), the Federal Highway Administration (FHWA), drivers, and work zone workers. The FHWA statistics indicate that 669 fatalities in work zone crashes in 2014 equated to 1.8 work zone fatalities per day (FHWA, 2017) and that work zones increase the severity and probability of crashes. Based on work zone studies, the total crash rate in work zones was 21.5% higher than that found on general roadways (Khattask et al., 2012).

Studies on roadway work zone safety cover a wide range of research topics. These include studies to find the root cause and identify the common factors in roadway work zone accidents, evaluating the effectiveness of various traffic control methods, examining the effects of various physical features and barriers on roadway work zone accident rates, studying the effects of work zone configurations on drivers' behavior, evaluating the safety apparel of roadway work zone workers, evaluating the cognitive processes and behavior of drivers around work zones, and performing risk modelling and risk assessment on roadway work zones (Ng et al., 2013, Long et al., 2014).

Correlation between fatal crashes and risk factors were considered in research such as the Georgia and Kansas DoTs research. The results of these studies display the importance of light conditions, truck involvement, roadway functional classification pavement center/edge lines, and usage of flaggers and flashers in work zones (Daniel et al., 2000; Li et al., 2009). The results of Southeast Michigan, Florida, and Tennessee crash records revealed the importance of roadway geometry, weather conditions, driver characteristics such as age and gender, lighting conditions, and driving under the influence of alcohol and/or drugs in work zone crashes (Harb et al., 2008; Wei et al., 2017; Meng et al., 2010; Weng & Meng, 2012). These findings are evaluated as part of this study for their generalizability and are used to determine risk categories.

This research evaluates historical crash data as part of a case study in Missouri. In Missouri, 69 people were killed in work zone crashes between 2012 and 2015. In addition, nineteen Missouri Department of Transportation (MoDOT) employees have died in the line of duty since 2000, with thirteen of the fatalities taking place in work zones (MoDOT, 2017). The effect of risk factors on property damage only (PDO), Minor injury (MI), and Disability injury and Fatality accidents (DI/FA) are considered as different levels of crash severity. In addition, this article considers the effect of collision type of two vehicle crashes as independent variables and the relationship of these factors on crash severity. Results confirm that road grade, road curvature, lighting, and weather have statistical impact on the severity and impact of crashes. By characterizing these elements into functional categories, the results will assist transportation decision makers in developing countermeasures in work zone design and management to improve work zone safety.

2. METHODOLOGY

In this research, multinomial logistic regression (MLR) is used to model the raw data. MLR is used when the dependent variable is nominal and the number of categories is more than two. In situations when the dependent variable cannot be perfectly predicted by independent variables, MLR is useful. This method does not assume normality and linearity of variables (Chan, 2005). MLR method uses the maximum likelihood ratio to calculate the probability of the categorical membership of the dependent variable.

Several methodologies are proposed for modeling MLR. These methodologies are mainly based on construction of a linear predictor function that attributes a score from a set of weights that are linearly combined with independent or explanatory variables using a dot product. The MLR development is based on determining the relationship among the dependent and independent variables. One category of the dependent variables is selected as the reference category in this regression method.

The equation for the MLR model is:

$$g(x) = [\frac{\pi(x)}{1-\pi(x)}] = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_n x_n \quad (1)$$

where $\pi(x)$is conditional probability of a accident;

x_n are independent variables (environment, geometry of road, traffic, etc.);

β_n is model coefficient, which directly determines odd ratio

Odd ratios of an event are defined as the probability of the event not occurring (Yan et al., 2005).

3. MODELING THE DATA

The data for this study is compiled from the Missouri Transportation Management System (TMS); these data are modeled to extract relationships between dependent and independent variables. These historical data from Missouri work zone crashes are used to identify, evaluate, and model trends that are related to severe crashes. The findings are sorted into work zone risk categories that can be integrated into work zone design and management strategies.

The independent variables consist of variables such as accident type, environment, geometry, and traffic condition categories. The accident type includes motor vehicles (MV) in transport, MVs on other roadways, and parked MVs. Environmental data contained information related to light conditions (daylight, dark with streetlights on, dark with streetlights off, dark with no streetlights, and indeterminate), road condition (dry, wet, snow, ice, slush, mud, standing water, and other), vision obscurity (load on vehicle, tree/bush, building, embankment, signboards, hillcrest, parked cars, moving cars, glare, not-obscured, and other).

The geometry data included road alignment (straight, curve) and road profile (level, grade, hillcrest). The traffic conditions were reported as normal, accident ahead, and congestion ahead.

In this study, the independent variables are those which might have an effect on the dependent variable (i.e., severity of crash).The crash severity was categorized to PDO, MI, and DI/FA. The number (N) displays the total number of observations corresponding to a particular category. For instance, the values of PDO, MI, and DI/FA are 129032, 60646, and 9158, respectively. Table 1 presents common abbreviations used in the manuscript.

Table 1. Key Terms

Term	Description
PDO	Property Damage Only
MI	Minor Injury
DI	Disability Injury
FA	Fatality Accident
MV	Motor Vehicle

The marginal percentage determines the proportion of valid observations found in the variable's group. For example, the marginal percentage of the PDO, MI, and DI/FA were 64.9%, 30.5%, and 4.6%, respectively. Table 2 displays the percentage and frequency of crashes based on severity.

Table 2. Dependent Variables

Severity	Code	Number (N)	Marginal Percentage
Property Damage Only (PDO)	1	129,032	64.90%
Minor Injury (MI)	2	60,646	30.50%
Disabling Injury (DI) and Fatal	4	9158	4.60%

Table 3 offers a summary of the descriptive statistics of the data set comprised of 225,383 observations, including 198,836 valid observations and 26,547 missing or blank data. The descriptive statistics display the quantitative features of the subgroups in the sample. Valid observations are the ones with no missing dependent or independent variables. The missing observations are the ones with missing data from either the dependent or independent variables, or both.

4. DATA ANALYSIS

The results of the model outputs are presented in Table 4. PDO was considered as the reference category for the dependent variables and has the highest numeric value (among dependent variables). The model conducts comparisons between the PDO and MI, as well as comparisons between the PDO and DI/FA. Given that PDO is treated as the reference group, models are estimated for an MI relative to PDO and a model for DI/FA in reference to PDO. The beta coefficient represents the effect of the independent variables on the dependent variables.

For the case of the present analysis, a positive β value indicates that the investigated independent category is more likely to impact the category of a dependent variable with respect to the reference category, while for $\beta < 0$, it is less likely to impact the dependent variable. For values of $\beta = 0$, the particular category and the reference category are equally likely to impact the dependent variable.

Table 3. Independent Variables

Variable		Categories	Code	N	Marginal Percentage
Accident	Type	MV in transport	7	195936	98.50%
		MV on other roadway Parked MV	8	136	0.10%
		Parked MV	9	2764	1.40%
	Two vehicle analysis	Head on	60	2486	1.30%
		Rear-end	61	137725	69.30%
		Sideswipe-meeting	62	2928	1.50%
		Sideswipe-passing	63	20277	10.20%
		Angle	64	28080	14.10%
		Backed into	65	3954	2.00%
		Other	67	3266	1.60%
Road Geometry	align ment	Straight	1	180460	90.80%
		Curve	2	18376	9.20%
	profile	Level	1	123829	62.30%
		Grade	2	71276	35.80%
		Hillcrest	3	3731	1.90%
Environment	Light condition	Daylight	1	164568	82.80%
		Dark with streetlights on	2	16768	8.40%
		Dark with streetlights off	3	1037	0.50%
		Dark with no streetlights	4	15417	7.80%
		Indeterminate	5	1046	0.50%
		Cloudy	2	45072	22.70%
	Road condition	Rain	3	6169	3.10%
		Snow	4	455	0.20%
		Sleet	5	20	0.00%
		Freezing(temp)	6	392	0.20%
		Fog/mist	7	620	0.30%
		Indeterminate	8	148	0.10%
		Dry	1	178752	89.90%
		Wet	2	18274	9.20%
		Snow	3	571	0.30%
		Ice	4	382	0.20%
		Slush	5	76	0.00%
		Mud	6	24	0.00%

Table 3. Independent Variables (Cont.)

		Standing water	7	12	0.00%
		other	9	745	0.40%
	Vision obscurity	Windshield	1	301	0.20%
		Load on vehicle	2	400	0.20%
		Tree/bush	3	87	0.00%
		Building	4	40	0.00%
		Embankment	5	62	0.00%
		Signboards	6	52	0.00%
		Hillcrest	7	687	0.30%
		Parked cars	8	483	0.20%
		Moving cars	9	2372	1.20%
		Glare	10	919	0.50%
		Other	11	2288	1.20%
		Not-obscured	12	191145	96.10%
Traffic	Traffic control	Normal	1	90619	45.60%
		Accident ahead	2	7222	3.60%
		Congestion ahead	3	100995	50.8%

The exponential beta value shows the odds ratio obtained for the independent variables. This ratio represents the variations in likelihood of the dependent variable being in a particular category compared to the reference, corresponding to one unit change of the independent variable. An odds ratio of greater than 1.0 indicates that the risk of the outcome falling in the comparison group relative to the risk of the outcome falling in the reference group increases as the variable increases, so it is more likely for the outcome to fall in the comparison group. An odds ratio of lower than 1.0 indicates that the risk of the outcome falling in the comparison group relative to the risk of the outcome falling in the reference group decreases as the variable increases. In general, for the odds ratio lower than 1.0, the outcome is more likely to be in the reference group.

The P value is usually tested at a threshold value of 5% or 1%. If the P value is less than the threshold value, the null hypothesis is rejected and the test hypothesis is accepted as valid. In this study, a 5% significance level is used in the model. Therefore, if the P value is less than 0.05, it can be concluded that the effect of the independent variable is statistically valid. Since the last category of each independent variable is used as the reference category, its β value is denoted as 0b.

The results obtained from the MLR model indicated that 44 variables are significant for MI and 41 variables for DI within a 0.05 significance level. All these results are based on the p-value measurements, beta coefficients (B), and the exponential beta coefficients (odds ratio) (See equation 1). The high level of statistical significance for the evaluated data variables demonstrates both the efficacy of using crash data to determine countermeasures as well as the need to refine results into categories to maximize their usefulness.

4.1. ACCIDENT TYPE VARIABLES

Accident: MLR compares MVs in transport to a parked MV for an MI relative to PDO while the other variables in the model are held constant. MVs in transport with a B value of 0.642 are more likely to cause a MI than a parked MV. An MV in transport has an odds-ratio of 1.901, which is a relative risk ratio compared to a parked MV for an MI relative to a PDO. In other words, a MV in transport is more likely than a parked MV to be in an MI over a PDO. The P value for MV in other roadways (0.431) is higher than the significant level (0.05), which means it is not a statistically significant factor.

Results of two-vehicle analyses reveal that rear-end, sideswipe (meeting and passing), angle, and backed into collisions are all more likely to cause minor injury (MI) when compared to head-on collisions. The most likely factor was a head-on collision with a β value of 2.964 and an odds ratio of 19.379. Head-on categories of two-vehicle analyses are more likely to cause disabling injury/fatality accident (DI/FA) with a regression coefficient of 6.124. Rear-end collisions are often associated with lower travel speeds, while the sideswipe collisions are associated with lane changing/merging maneuvers (Bham et al., 2012; Daniel et al., 2000). These results are in agreement with Li et al. (2007a) and indicate that head-on collisions are the most common cause for fatal work zone accidents. These results, although common sense, clearly demonstrate the importance of controlling related work zone design elements that can allow head-on collisions. As an example, when roadways collapse from controlled-access highway to two-way traffic or on undivided highways, it is essential that signage, lighting, and speed are sufficiently controlled.

MLR analysis of road alignment reveals straight roads are more likely to cause both MI and DI/FA than a curved road. These results are in agreement with research studies conducted in Alabama and New Jersey that which stated that a small proportion of crashes occurred on curved roads (Sisiopiku et al.2015; Yang et al., 2013; Harb et al., 2008) as drivers are more cautious on curves than straight roads. The results of the road profile analysis indicate that grade roads are more likely to cause a DI/FA than a level road.

Table 4. Model Results

	SEVERITY a	ß	Df	Sig.	Exp(ß)	ß	Df	Sig.	Exp(ß)
		PDO compares to MI				PDO compares to DI and fatal			
	Intercept	-4.474	1	0.000		7.393	1	0.000	
Accident Type	MV in transport	0.642	1	0.000	1.901	0.621	1	0.000	1.861
	MV on other roadway	-1.289	1	0.431	0.276	5.698	1	0.000	298.406
	Parked MV	0b	0	.	.	0b	0	.	.
Two Vehicle Analysis	Head on	2.964	1	0.000	19.379	6.124	1	0.000	456.675
	Rear-end	1.624	1	0.000	5.073	0.853	1	0.000	2.348
	Sideswipe- meeting	1.146	1	0.000	3.144	2.281	1	0.000	9.789
	Sideswipe-passing	0.518	1	0.000	1.678	1.329	1	0.000	3.778
	Angle	1.498	1	0.000	4.474	0.781	1	0.000	2.183
	Backed into	0.490	1	0.000	1.632	-0.235	1	0.120	0.790
	Others	0b	0	.	.	0b	0	.	.
Road Alignment	Straight	0.377	1	0.000	1.458	0.647	1	0.000	1.910
	Curve	0b	0	.	.	0b	0	.	.
Road profile	Level	0.622	1	0.000	1.863	0.237	1	0.031	1.268
	Grade	0.565	1	0.000	1.759	1.183	1	0.000	3.263
	Hillcrest	0b	0			0b	0		
Light conditions	Daylight	-0.726	1	0.000	0.484	0.408	1	0.020	1.504
	Dark with streetlights on	-0.757	1	0.000	0.469	0.205	1	0.256	1.227
	Dark with streetlights off	-0.745	1	0.000	0.475	0.305	1	0.219	1.357
	Dark with no streetlights	-0.293	1	0.000	0.746	3.313	1	0.000	27.467
	Indeterminate	0b	0	.	.	0b	0	.	.

Table 4. Model Results (cont.)

Weather	Clear	0.629	1	0.005	1.875	0.426	1	0.522	1.531
	Cloudy	0.801	1	0.000	2.228	0.831	1	0.211	2.295
	Rain	0.795	1	0.000	2.214	-0.599	1	0.370	0.549
	Snow	0.984	1	0.000	2.676	1.241	1	0.071	3.460
	Sleet	4.664	1	0.000	106.051	2.523	1	0.472	12.472
	Freezing (temp)	0.397	1	0.135	1.487	-0.371	1	0.595	0.690
	Fog/mist	1.611	1	0.000	5.008	-2.522	1	0.000	0.080
	Indeterminate	0b	0	.	.	0b	0	.	.
Road Condition	Dry	0.828	1	0.000	2.288	0.836	1	0.003	2.307
	Wet	0.265	1	0.008	1.303	0.890	1	0.002	2.434
	Snow	0.420	1	0.016	1.522	1.809	1	0.000	6.107
	Ice	-0.610	1	0.002	0.543	0.475	1	0.149	1.608
	Slush	-0.600	1	0.214	0.549	-0.256	1	0.805	0.774
	Mud	-0.128	1	0.862	0.880	0.363	1	0.842	1.437
	Standing water	-0.727	1	0.491	0.484	3.983	1	0.126	53.693
	Other	0b	0	.	.	0b	0	.	.
Vision Obscurity	Windshield	0.022	1	0.881	1.022	-2.192	1	0.000	0.112
	Load on vehicle	0.004	1	0.976	1.004	-0.888	1	0.007	0.411
	Tree/bush	-0.460	1	0.132	0.631	-0.950	1	0.293	0.387
	Building	-53.407	1	.	.	-42.294	1	.	1.000E-013
	Embankment	-1.668	1	0.000	0.189	-1.669	1	0.134	0.189
	Signboards	-0.362	1	0.275	0.696	-0.449	1	0.658	0.638
	Hillcrest	0.052	1	0.587	1.053	-2.701	1	0.000	0.067
	Parked cars	-0.105	1	0.328	0.900	-0.144	1	0.571	0.866
	Moving cars	0.800	1	0.000	2.226	-0.758	1	0.000	0.469
	Glare	-0.551	1	0.000	0.577	-0.466	1	0.008	0.627
	Other	-0.405	1	0.000	0.667	-0.890	1	0.000	0.411
	Not-obscured	0b	0	.	.	0b	0	.	.

Table 4. Model Results (cont.)

Traffic Control	Normal	-0.138	1	0.000	0.871	0.194	1	0.000	1.214
	Accident ahead	0.139	1	0.000	1.149	0.788	1	0.000	2.199
	Congestion ahead	0b	0	.	.	0b	0	.	.

Table 5 summarizes the most likely factors that can contribute to work zone crash severity based on the results of the MLR.

Table 5. Most likely factors in work zone crash severity

Independent variables	Crash Severity	
	MI	DI/FA
Accident type	MV on arterial roadway	MV on major, undivided roadway
Two vehicle analysis	Head on	Head on
Road alignment	Straight	Straight
Road profile	Level/Grade	Grade
Light condition	Indeterminate	Dark with no streetlight
Weather	Sleet	Snow
Road condition	Dry	Snow
Vision obscurity	Moving car	-
Traffic control	Accident ahead	Accident ahead

Grade roads are more likely than level roads to cause DI/FA by a β value of 1.183 and an odds ratio of 3.263. These results are in agreement with the findings of Bham et al. (2012); grade profiles increase severe crashes on undivided highways. Grade (especially downhill grades) may have effect on vehicle speed and more failure in controlling the vehicle that increase risk of accident in work zone. For light conditions, dark roadways without adequate lighting is more strongly correlated to DI/FA and has the highest odds ratio of 27.467. These results are in agreement with the analysis of Li et al. (2009) and Wei et al. (2017) which also found that poor light conditions (dark with no streetlights) is a significant risk factor in fatal accidents. Sleet and snow are the most likely contributing weather factor to a DI/FA, with snow statistically significant.

Managing behavior in changing roadway conditions is key. The use of alternative sign configurations and lane shift sign configurations (Edara et al., 2017; Long et al., 2017) support the integration of communication techniques as part of roadway design. Similar to the findings of Brown et al. (2015) the results of this study suggest that mobile alarm systems may prove valuable in alerting drivers to changing conditions or roadway patterns.

5. CONCLUSIONS

Despite several attempts to change merge configurations and improve work zone safety, the accident rate throughout work zone areas is still alarmingly high. This can be attributed to insufficient policies and measures for reducing risk factors. Analysis of historical work zone data assists managers in identifying risk factors. These data enable managers to extract significant information which can be used in planning and designing the work zone. Results of this study demonstrate the correlation between head-on collision, road grade and curvature, roadway lighting, and weather impacts on roadway safety and mobility. The greatest opportunities for improving roadway and work zone safety are linked to roadway design and management using effective signs and light configurations.

These findings provide strong guidance for the installation of temporary traffic control (TTC) signs or variable message sign (VMS) before work zone to inform drivers about an upcoming work zone and any driving pattern changes as well as any weather scenarios. Messages must be short and succinct to provide maximum information at a glance. Signs positioned in tandem may be suitable solution to address time/message length constraints. The signage used in the advance warning area of a work zone provides critical information to drivers such as information regarding the closed lane, when to merge, when to reduce speed limits, etc. These types of information are critical to the overall safety of the work zone (Zhu et al., 2015). Although there are existing work zone sign configurations approved by MUTCD, other sign configurations are possible and may be evaluated against traffic management goals for traffic flow, driver behavior, driver satisfaction, and the like. The reaction of drivers to alternate sign configurations, in addition to their driving patterns through the work zones where such new signage is incorporated, must be evaluated in order to assure safe implementation (Thind et al., 2017; Long et al., 2016) and before traffic management agencies can request their use.

Findings outline patterns and scenarios that should be integrated into work zone design to enhance safety and improve mobility with respect to work zone lighting, impact of weather, and the like. In addition to work zone crash data analysis, MoDOT work zone survey, traffic control signs, education, and laws are different methods that help transportation decision makers eliminate or reduce risk factors.

Future work should more carefully consider driver behavior as one of the important risk factors in work zone crashes. The length of work zone, road type (rural/urban), speed limit, vehicle type, and crash location are among factors that were not considered in the current research.

ACKNOWLEDGEMENTS

This project was funded by the Missouri Department of Transportation under TR201312. The authors gratefully acknowledge the assistance provided by graduate research assistant Paul Robin for his assistance in the data analysis as well as contributions by colleagues Drs. Carlos Sun and Ean Ng for their work with other phases of the project.

REFERENCES

Chambless, J., Belou, L., Lindly, J. K., & McFadden, J. (2001). Identification of overrepresented parameters for work zone crashes in Alabama, Michigan, and Tennessee. In Transportation Research Board 2001 Annual Meeting.

Chan, Y. H. (2005). Biostatistics 305. Multinomial logistic regression. Singapore medical journal, 46(6), 259.

Daniel, J., Dixon, K., & Jared, D. (2000). Analysis of fatal crashes in Georgia work zones. Transportation Research Record: Journal of the Transportation Research Board, (1715), 18-23.

Dissanayake, S., and Akepati, S. R. (2009). Identification of Work Zone Crash Characteristics, Final Report. Smart Work Zone Deployment Initiative.

Harb, R., Radwan, E., Yan, X., Pande, A., & Abdel-Aty, M. (2008). Freeway work-zone crash analysis and risk identification using multiple and conditional logistic regression. Journal of Transportation Engineering, 134(5), 203-214.

Khattak, A. J., Khattak, A. J., & Council, F. M. (2002). Effects of work zone presence on injury and non-injury crashes. Accident Analysis & Prevention, 34(1), 19-29.

Li, Y., & Bai, Y. (2009). Highway work zone risk factors and their impact on crash severity. Journal of Transportation engineering, 135(10), 694-701.

Long, S., Smith, B. K., Ng, E. H., & Sun, C. (2014). Work zone safety: physical and behavioral barriers in accident prevention (No. cmr 14-013). Mid-America Transportation Center, University of Nebraska-Lincoln.

Meng, Q., Weng, J., & Qu, X. (2010). A probabilistic quantitative risk assessment model for the long-term work zone crashes. Accident Analysis & Prevention, 42(6), 1866-1877.

MoDOT. 2017. http://www.modot.org/workzones/index.htm

Ng, E. H., Long, S., Smith, B., & Sun, C. (2013, January). A Framework for Improving Roadway Work Zone Safety Using Action Research. In IIE Annual Conference. Proceedings (p. 746). Institute of Industrial and Systems Engineers (IISE).

Sisiopiku, V. P., Ramadan, O. E., Ismail, M. I. E., & Cavusoglu, O (2015). Analysis of Crash Causes, Costs, and Countermeasures in Alabama Work Zones 2.

Srinivasan, S., Carrick, G., Zhu, X., Heaslip, K., and Washburn, S. (2008). Analysis of Crashes in Freeway Work Zone Queues.

Wei, X., Shu, X., Huang, B., Taylor, E. L., & Chen, H. (2017). Analyzing traffic crash severity in work zones under different light conditions. Journal of Advanced Transportation, 2017.

Weng, J., & Meng, Q. (2011). Analysis of driver casualty risk for different work zone types. Accident Analysis & Prevention, 43(5), 1811-1817.

Weng, J., & Meng, Q. (2012). Effects of environment, vehicle and driver characteristics on risky driving behavior at work zones. Safety science, 50(4), 1034-1042.

Yan, X., Radwan, E., & Abdel-Aty, M. (2005). Characteristics of rear-end accidents at signalized intersections using multiple logistic regression model. Accident Analysis & Prevention, 37(6), 983-995.

Yang, H., Ozbay, K., Ozturk, O., & Xie, K. (2015). Work zone safety analysis and modeling: a state-of-the-art review. Traffic injury prevention, 16(4), 387-396.

Yang, H., Ozturk, O., Ozbay, K., & Bartin, B. Analysis of driver injury severity in single- vehicle work zone crashes.

SECTION

2. CONCLUSIONS OF DISSERTATION AND FUTURE WORK

This chapter overviews the conclusion of this dissertation and discusses potential future work. This research contributes to the body of knowledge in work zone management and design through the integration of data mining and decision analytics, along with qualitative stakeholder inputs, into conventional temporary traffic control scenarios. Although existing research provides strong results when considering individual scenarios, this research uses research tools and techniques to create a mixed methods multi-criteria decision research design. Results from this dissertation can help transportation managers to reach a better understanding of crucial factors in improving work zone safety and mobility.

The dissertation includes three key contributions designed to address critical gaps in the traffic engineering and engineering management work zone management literature. The results of three related case studies are presented that showcase strategies for work zone management and improvement with respect to safety and mobility.

The first contribution considers the importance of stakeholder input, or integrating the voice of the customer, into traffic engineering design. Work zone signage is mandated by the Manual on Uniform Traffic Control Devices (MUTCD), but the current configurations are often confusing to the driving public. State departments of transportation have questioned whether alternate signage would provide more cost-effective, equally safe options. The first article used a driving simulator to model four work zone merge scenarios. The scenarios were based on an actual work zone along a

Missouri interstate highway. Drivers' responses to these signs were analyzed by data analysis and statistical data method to investigate the efficiency of sign configurations employed in the study. The results revealed no significant differences between the two sign configurations with respect to overall drivers' reaction, but drivers did report increased satisfaction with the MoDOT alternate configurations. The findings of this study suggest opportunities for traffic managers to consider driver preference, as well as other factors, in work zone sign configuration.

Despite many attempts at countermeasures, accident rates in work zones remain high. Understanding driver patterns and behaviors is critical to improving safetly. Due to many related variables and the amount of data generated in this field, evaluation and analysis of drivers' behavior is complicated. The second research contribution uses a combination of data mining and multi-criteria decision making to uncover driver characteristics that contribute to risky driving behaviors. The k-mean clustering method is used to cluster large amounts of data, which makes it easier for decision makers to evaluate these clusters rather than all of the data. Additional data analytics are used to weight categories and provide additional guidance as part of a multi-criteria decision framework. The proposed analytic tool can provide a comprehensive assessment of driving behavior identification and allows transportation professionals a roadmap for better decision making to promote safety in work zones. Driver training safety programs can be updated to include modules on merge behavior. These modules can be part of general driver-education training programs or built into mandatory driver improvement training for traffic offenders.

The third contribution addresses a challenge of data integration from historical crash data into decision platforms. Although tremendous amounts of crash data are collected, little are analyzed to improve work zone safety. Transportation managers usually focus on reducing risk factors that lead to crashes in work zones and require robust tools and analysis processes to identify these risk factors. Multinomial logistic regression (MLR) is used to model historical data of Missouri work zone crashes to identify patterns and categories of factors that statistically contribute to work zone crashes. Results of this study demonstrate the correlation between head-on collision, road grade and curvature, roadway lighting, and weather impacts on roadway safety and mobility. The greatest opportunities for improving roadway and work zone safety are linked to roadway design and management using effective signs and light configurations.

Future work will try to use data from field database to evaluate drivers' behavior. For expanding the model, safety criteria other than the criteria considered in this study will be used for driver behavior analysis. By using more expert opinions, the questionnaire could be a good choice for identifying safety criteria. The fuzzy method is another choice to convert experts' linguistic opinions to numerical data, which could be useful for data analysis models.

BIBLIOGRAPHY

Arnau-Sabatés, L., Sala-Roca, J., & Jariot-Garcia, M. (2012). Emotional abilities as predictors of risky driving behavior among a cohort of middle aged drivers. Accident Analysis & Prevention, 45, 818-825.

ARTBA (2014). Fatalities in Motor Vehicle Traffic Crashes by State and Work Zone. The National Work Zone Safety Information Clearinghouse. http://www.workzonesafety.org/crash_data/workzone_fatalities (last accessed 3/21/14).

Bayam, E., Liebowitz, J., & Agresti, W. (2005). Older drivers and accidents: A meta-analysis and data mining application on traffic accident data. Expert Systems with Applications, 29(3), 598-629.

Beacher, A. G., Fontaine, M. D., & Garber, N. J. (2005). Part 2: Work Zone Traffic Control: Field Evaluation of Late Merge Traffic Control in Work Zones.Transportation Research Record: Journal of the Transportation Research Board, 1911(1), 32-41.

Federal Highway Administration, 2017, http://www.ops.fhwa.dot.gov/wz/resources/facts_stats/safety.htm

Fei, L., Zhu, H. B., & Han, X. L. (2016). Analysis of traffic congestion induced by the work zone. Physica A: Statistical Mechanics and its Applications, 450, 497-505.

Guo, F., & Fang, Y. (2013). Individual driver risk assessment using naturalistic driving data. Accident Analysis & Prevention, 61, 3-9.

Haddon Jr, W. (1980). Advances in the epidemiology of injuries as a basis for public policy. Public health reports, 95(5), 411.

Harb, R., Radwan, E., Ramasamy, S., Abdel-Aty, M., Pande, A., Shaaban, K., & Putcha, S. (2009). Two simplified dynamic lane merging system (SDLMS) for short term work zones.

Harré, N., Brandt, T., & Dawe, M. (2001). The development of risky driving in adolescence. Journal of Safety Research, 31(4), 185-194.

Idewu, W. I., & Wolshon, B. (2010). Joint Merge and Its Impact on Merging Speeds in Lane Reduction Areas of Construction Zone. Transportation Research Record: Journal of the Transportation Research Board, 2169(1), 31-39.

Khattak, A. J., Khattak, A. J., & Council, F. M. (2002). Effects of work zone presence on injury and non-injury crashes. Accident Analysis & Prevention, 34(1), 19-29.

Kumar, S., & Toshniwal, D. (2015). A data mining framework to analyze road accident data. Journal of Big Data, 2(1), 26.

Li, Y., & Bai, Y. (2006). Fatal and injury crash characteristics in highway work zones. In Proc., Transportation Research Board 87 th Annual Meeting.

Long, S., Qin, R., Konur, D., Leu, M. C., Moradpour, S., Thind, S., ... & Brown, H. (2017). Work zone simulator analysis: driver performance and acceptance of Missouri alternate lane shift configurations (No. cmr 17-002).

Long, S., Smith, B. K., Ng, E. H., & Sun, C. (2014). Work zone safety: physical and behavioral barriers in accident prevention(No. cmr 14-013). Mid-America Transportation Center, University of Nebraska-Lincoln.

Luke, R., & Heyns, G. J. (2014). Reducing risky driver behaviour through the implementation of a driver risk management system. Journal of Transport and Supply Chain Management, 8(1), 1-10.

Moradpour, S., Wu, S., & Leu, M. C. (2015, January). Use of Traffic Simulators to Determine Driver Response to Work Zone Configurations. In Proceedings of the International Annual Conference of the American Society for Engineering Management. (p. 1). American Society for Engineering Management (ASEM).

Ng, E. H., Long, S., Smith, B., & Sun, C. (2013, January). A Framework for Improving Roadway Work Zone Safety Using Action Research. In IIE Annual Conference. Proceedings (p. 746). Institute of Industrial and Systems Engineers (IISE).

Oltedal, S., & Rundmo, T. (2006). The effects of personality and gender on risky driving behaviour and accident involvement. Safety science, 44(7), 621-628.

Rakha, H., El-Shawarby, I., & Setti, J. R. (2007). Characterizing driver behavior on signalized intersection approaches at the onset of a yellow-phase trigger. IEEE Transactions on Intelligent Transportation Systems, 8(4), 630-640.

Reyes, M. L., Khan, S. A., & Initiative, S. W. Z. D. (2008). Examining driver behavior in response to work zone interventions: A driving simulator study.

Rhodes, N., & Pivik, K. (2011). Age and gender differences in risky driving: The roles of positive affect and risk perception. Accident Analysis & Prevention, 43(3), 923-931.

Shakouri, M., Ikuma, L. H., Aghazadeh, F., Punniaraj, K., & Ishak, S. (2014). Effects of work zone configurations and traffic density on performance variables and subjective workload. Accident Analysis & Prevention, 71, 166-176.

Stanton, N. A., & Salmon, P. M. (2009). Human error taxonomies applied to driving: A generic driver error taxonomy and its implications for intelligent transport systems. Safety Science, 47(2), 227-237.

Thind, S.S., Moradpour,S., Konur,D., Long, S., Qin, R.(2017). Statistical Comparison of Lane Shift Sign Configurations with Driving Simulation. Industrial and Systems Engineering Research Conference.

Weng, J., & Meng, Q. (2011). Analysis of driver casualty risk for different work zone types. Accident Analysis & Prevention, 43(5), 1811-1817.

Weng, J., & Meng, Q. (2011). Modeling speed-flow relationship and merging behavior in work zone merging areas. Transportation research part C: emerging technologies, 19(6), 985-996.

Weng, J., & Meng, Q. (2012). Effects of environment, vehicle and driver characteristics on risky driving behavior at work zones. Safety science, 50(4), 1034-1042.

Zamorski, M.A. & Kelley, A.M., 2011, Chapter 5 – Risky driving behavior, viewed 22 February 2013, from http://ftp.rta.nato.int/public//PubFullText/RTO/TR/RTO-TR-HFM-164///TR-HFM-164-05.pdf

VITA

Samareh Moradpour was born in 1984 in Iran. She earned her bachelor's degree in Industrial Engineering from Iran University of Science and Technology (IUST) in Behshahr, Iran, in 2006. In 2008, Samareh started her studies as a master's student of Industrial Engineering under the supervision of Dr. Ebrahimnezhad at Islamic Azad University in Qazvin, Iran, and graduated in 2012.

In August 2014, Samareh started her PhD program in the Department of Engineering Management and System Engineering at Missouri University of Science and Technology in Rolla, Missouri, under the supervision of Dr. Suzanna Long. Her research was devoted to the development of decision-making tools for work zone safety. She received a PhD in Engineering Management from Missouri University of Science and Technology in July 2018. Results of her research were published in several technical reports, journal papers, and conference proceedings and presentations, as well as the current dissertation.