

ANALYSIS OF CRASH LOCATION AND CRASH SEVERITY RELATED TO
WORK ZONES IN OHIO

Thesis

Submitted to

The School of Engineering of the

UNIVERSITY OF DAYTON

In Partial Fulfillment of the Requirements for

The Degree of

Master of Science in Civil Engineering

By

Ibrahim Alfallaj

Dayton, Ohio

August, 2014



ANALYSIS OF CRASH LOCATION AND CRASH SEVERITY RELATED TO
WORK ZONES IN OHIO

Name: Alfallaj, Ibrahim

APPROVED BY:

Deogratias Eustace, Ph.D., P.E., PTOE
Advisory Committee Chairperson
Associate Professor, Department of
Civil and Environmental Engineering
and Engineering Mechanics

Peter Hovey, Ph.D.
Committee Member
Associate Professor, Department
of Mathematics

Paul Goodhue P.E.
Committee Member, Senior Traffic
Engineer & Project Manager IBI Group

John G. Weber, Ph.D.
Associate Dean,
School of Engineering

Eddy M. Rojas, Ph.D., M.A., P.E.
Dean, School of Engineering

ABSTRACT

ANALYSIS OF CRASH LOCATION AND CRASH SEVERITY RELATED TO WORK ZONES IN OHIO

Name: Alfallaj, Ibrahim
University of Dayton

Advisor: Dr. Deogratias Eustace

Due to growth of vehicle travel using streets and highway systems in the United States, pavement repair and rehabilitation projects have increased. As a result, the presence of work zones has created traffic congestion and has increased the crash risk. The main object of this study was to identify significant factors that contribute to an increase in crash severity in the state of Ohio and recognize the most risk segment within the work zone locations. The work zone segment area is made of : (a) termination area (TA), (b) before the first work zone warning sign area (BWS), (c) advance warning area (AWA), (d) transition area (TSA), (e) activity area (AA). This study used a 5-year crash data from Ohio Department of Public Safety (ODPS) database from 2008 to 2012. In this study, classification tree modeling was used to investigate significant predictor variables of crash severity of work zone related crashes and recognize the most significant crash location within work zone areas in the state of Ohio.

Classification tree modeling identified ten important variables (factors) that explain a large amount of the variation in the response variable, crash severity. These predictor variables of work zone crash severity identified include collision type, motorcycle related, work zone crashes type, posted speed limit, vehicle type, speed related, alcohol related, semi-truck related, youth related and road condition. In case of work zone location analysis results, this study identified six significant factors, which include collision type, work zone crash type, posted speed limit, vehicle type, workers present, and age of driver. Collision type is the most significant factor that affects crash severity in a work zone. Likewise, for work zone location, the work-zone crash type was the most significant factor that contributed in increasing the probability of work zone location crashes.

ACKNOWLEDGEMENTS

First, I would like to thank Almighty God, who has helped and guided me successfully complete my program of study. I am very grateful to my principal advisor, Dr. Deogratias Eustace for his patience, encouragement, guidance and providing necessary support to achieve this research. Special thanks extended to Dr. Peter Hovey, statistics consultant for his review and advice in statistical analysis in this research. In addition, I would like to thank Mr. Paul Goodhue, who read and provided technical support to my thesis study. Thanks extended to my friends and colleagues at the University of Dayton who contributed in different ways to the successful completion of this thesis. Finally, my sincere thanks go to my parents for their continued support, prayers and contributions to completion of this research.

TABLE OF CONTENTS

ABSTRACT.....	iii
ACKNOWLEDGEMENTS.....	v
TABLE OF CONTENTS.....	vi
LIST OF FIGURES	x
LIST OF TABLES.....	xiii
CHAPTER ONE - INTRODUCTION.....	1
1.1 Introduction	1
1.2 Problem Statement	3
1.3 Research Objectives	5
1.4 Thesis Organization.....	5
CHAPTER TWO - LITERATURE REVIEW	7
2.1 Introduction	7
2.2 Work Zone Crash Characteristics	7
2.2.1 Crash Severity.....	8

2.2.2 Crash Location.....	9
2.2.3 Collision Types.....	10
2.2.4 Crash Time	11
2.2.5 Trucks and Safety in Work Zones	11
2.2.6 Crash Rate.....	12
2.2.7 Speed Related and Posted Speed Limit in Work Zones	13
2.2.8 Causal Factors.....	14
2.3 Traffic Control Devices.....	15
CHAPTER THREE - METHODOLOGY AND DATA COLLECTION	17
3.1 Data Collection.....	17
3.1.1 Data Collection Methodology	17
3.1.2 Merging Data Files	19
3.2 OH-1 Crash Data.....	20
3.2.1 Report Information	20
3.2.2 Vehicle Information.....	21
3.2.3 Roadway and Traffic Information	21
3.2.4 Crash Relationships	21
3.2.5 Crash Details.....	22
3.2.6 Driver/Passengers Information.....	22
3.3 Creating of Work Zone Related Traffic Crash Database	22

3.4 Methodology	23
3.4.1 Introduction	23
3.4.2 Classification Trees Modeling.....	24
3.4.2.1 Introduction.....	24
3.4.2.2 Classification Trees Formed	25
3.4.2.3 Pruning and Validation	26
3.4.2.4 Statistical Analysis.....	27
3.4.2.4.1 Node Splitting Criteria	27
3.4.2.4.2 Receiver Operating Characteristic (ROC) Curve.....	29
CHAPTER FOUR - RESULTS	31
4.1 Introduction	31
4.2 Description of Selected Variables	32
4.2.1 Description of Crash and Vehicle Characteristics	32
4.2.2 Description of Human and Driver Characteristics	34
4.2.3 Description of Roadway and Environmental Characteristics.....	35
4.3 Descriptive Results.....	37
4.3.1 Descriptive Results of Crash Severity	37
4.3.1.1 Distribution of Crash Severity by Crash Types	37
4.3.1.2 Distribution of Crash Severity by Work Zone Types	38
4.3.1.3 Distribution of Crash Severity by Posted Speed Limit.....	39

4.3.1.4 Distribution of Crash Severity by Crash Time	40
4.3.2 Descriptive Results of Crashes Relative to Work Zone Location	41
4.3.2.1 Distribution of Crashes by Work Zone Location.....	42
4.3.2.2 Distribution of Work Zone Location Crashes by Workers Presence.....	42
4.4 Results of Classification Tree Models	43
4.4.1 Introduction	43
4.4.2 Results of Crash Severity Model	44
4.4.2.1 Variables Important to Crash Severity Model	45
4.4.2.2 ROC Curve Results of Crashes Severity	46
4.4.3 Results of Work Zone Location Crashes	49
4.4.3.1 Variables Important to Work Zone Location Crashes Model.....	50
4.4.3.2 ROC Curve Results of Work Zone Location Crashes Model.....	51
4.4.4 Results Discussion	53
CHAPTER FIVE - CONCLUSIONS AND RECOMMENDATIONS.....	63
REFERENCES	67

LIST OF FIGURES

Figure 1.1: Component Areas of a Highway Work Zone (MUTCD; FHWA, 2003)	3
Figure 1.2: Work Zone Fatalities in the US 1982-2006 (FARS, 2006)	4
Figure 3.1: Four Files with Their Related Variable	18
Figure 3.2: The Process of Creating a Joint Crash-Unit-People File	20
Figure 3.3: An Example of a Classification Tree Form	26
Figure 3.4: An Example of a ROC Curve	30
Figure 4.1: Distribution of Crash Severity by Crash Types	38
Figure 4.2: Distribution of Crash Severity by Work Zone Types	39
Figure 4.3: Distribution of Crash Severity by Posted Speed Limit	40
Figure 4.4: Distribution of Crash Severity by Crash Time	41
Figure 4.5: Distribution of Crashes by Work Zone Location	42
Figure 4.6: Distribution of Work Zone Location Crashes by Worker Presence	43
Figure 4.7: Split History Before Pruning for Both Training and Validation Datasets of Crash Severity Model	45
Figure 4.8: Important Variables of Crash Severity Model	46
Figure 4.9: ROC Curve for Training Dataset of Crash Severity Model	47

Figure 4.10: ROC Curve for Validation Dataset of Crash Severity Model	47
Figure 4.11: Classification Tree Results for Crash Severity Model	48
Figure 4.12: Split History Before Pruning for Training and Validation Datasets of WZ Location Crashes Model	49
Figure 4.13: Important Variables of WZ Location Crashes Model	50
Figures 4.14: ROC Curve for Training Dataset of WZ Location Crashes Model	51
Figures 4.15: ROC Curve for Validation Dataset of WZ Location Crashes Model	52
Figure 4.16: Classification Tree Result for WZ Location Crashes Model	53
Figure 4.17: The First Split of Crashes Severity Model	55
Figure 4.18: Collision Types Splitting Nodes in the Crash Severity Model	56
Figure 4.19: Work Zone Crashes Types Splitting Nodes in the Crash Severity Model ...	57
Figure 4.20. Motorcycle Not Related Splitting Nodes in the Crash Severity Model	57
Figure 4.21. Collision Types (Rear-End, Angle or Head on) Splitting Nodes in the Crash Severity Model.....	58
Figure 4.22. The First Split of Work Zone Location Crashes Model	59
Figure 4.23. Work Zone Types Splitting Nodes in the Work Zone Location Crashes Model	60
Figure 4.24: Speed Limit (55-70 mph) Splitting Nodes in the Work Zone Location Crashes Model	60

Figure 4.25: Speed Limit (Less Than or Equal 50 mph) Splitting Nodes in the Work Zone Location Crashes Model	61
Figure 4.26. Vehicle Types Splitting Nodes in the Work Zone Location Crashes Model	62

LIST OF TABLES

Table 2.1 Distribution of total and truck crashes in different work zone locations (Khattak and Targa, 2004)	12
Table 3.1 Work Zone Types as Defined in the OH-1 Crash Documentation	22
Table 4.1 Description of Crash and Vehicle Characteristics	33
Table 4.2 Description of Human and Driver Characteristics.....	35
Table 4.3 Description of Roadway and Environmental Characteristics	36

CHAPTER ONE

INTRODUCTION

1.1 Introduction

A work zone is an area within a traffic zone with highway construction, maintenance or utility-work. A work zone may include motionless or mobility activities with short or long durations of length. In general, the range of work zone can be from the first warning sign or flashing lights until the last traffic control device. Since work zones increase likelihood of traffic crashes, they have become issues of concern to traffic engineers as well as to the Federal Highway Administration (FHWA) and state departments of transportation (DOT's). Studies by several researchers have found that severe crashes in work zone areas are greater than that in pre-work zone area and severe crashes would lead to fatalities or injuries among the drivers, passengers and workers as well (Ha and Nemeth, 1995; Rouphail et al, 1988).

In 2004, work zone related crashes in the United States accounted for about 1,068 fatal crashes and 49,620 injury crashes (FHWA, 2006). From 1995 to 1997, it is estimated that the cost of highway work zone crashes was about \$6.2 billion per year and at an average cost of \$3,687 per crash in the United States (Mohan and Gautam, 2002). In particular, (Nemeth and Migletz, 1978) performed a before, during, and after-work zone study on Ohio's rural interstate system and found that there is an increase in the accident

rate due to work zones. Work zones are likely to increase in number, duration, and length for several causes, such as an increase in traffic volumes, roadways deteriorating quicker and future highway maintenance needs. Thus, work zone improvement projects were extended to an average of 23,745 miles of road per year in year 1997 through 2001 (FHWA, 2001). As a result, work zone crashes have been increasing at an alarming rate.

Several risk factors affect work zone safety and many of them are not completely understood. Crash data can help to determine some risk factors that help in devising countermeasures, which can reduce risk levels and thus avoid severe crashes (mainly fatal and injury) in work zones. In this study, an extensive range of crash variables, such as work zone settings, environmental conditions, driver characteristics and crash information obtained from the Ohio Department of Public Safety (ODPS) work zone crash dataset was analyzed.

A work zone area is a section of road where construction or maintenance projects are carried out. As shown in Figure 1.1, the work zone area is divided into four major areas: advance warning, transition, activity and termination areas (Manual on Uniform Traffic Control Devices (MUTCD), FHWA, 2003). Drivers traveling through a work zone are usually cautioned in the advanced warning area and then immediately transfer to the transition area. The transition section is designed as a bottleneck, which could reduce the traffic speed through active area. Finally, the termination area is that section where drivers return to their normal path.

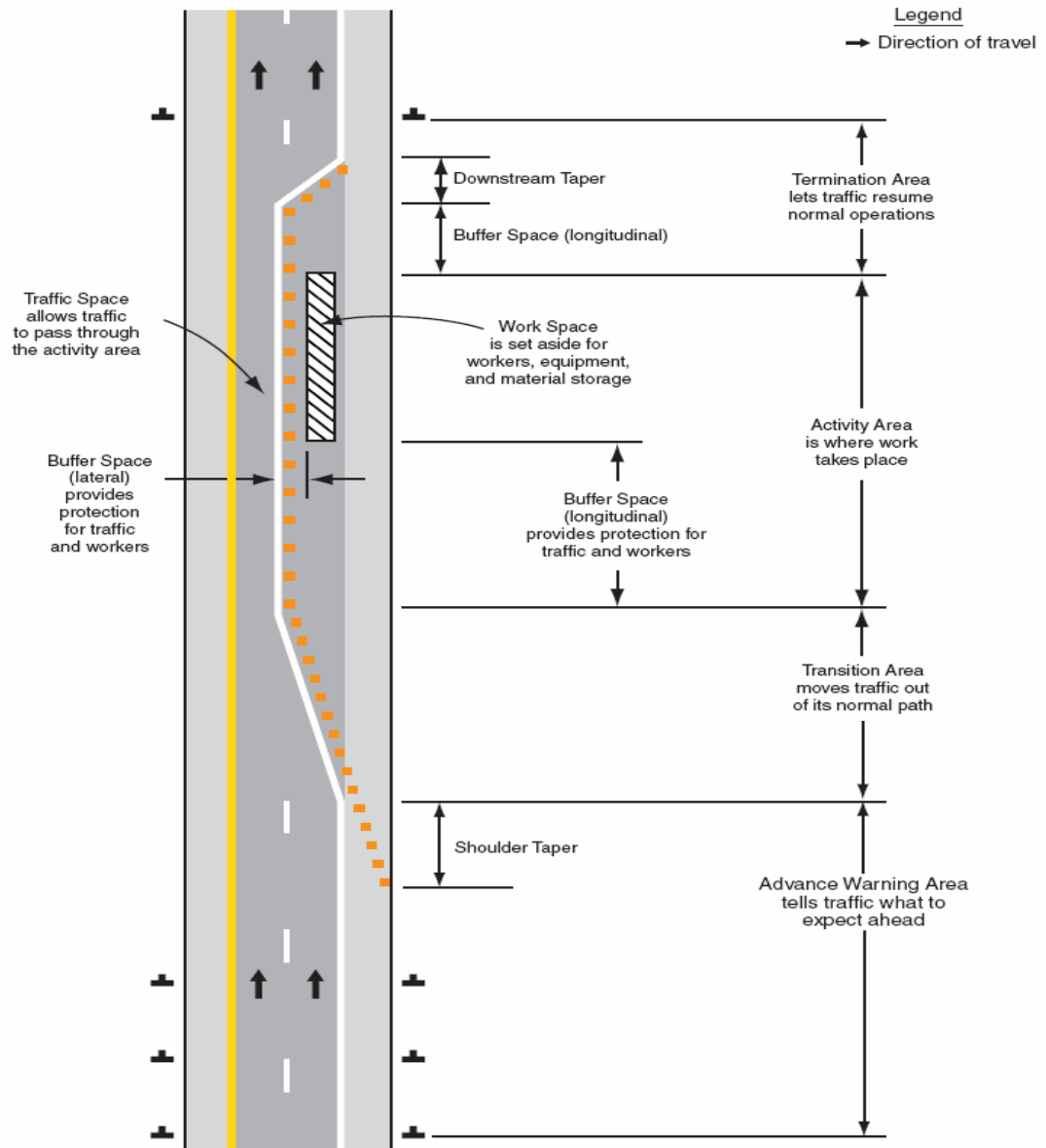


Figure 1.1: Component Areas of a Highway Work Zone (MUTCD; FHWA, 2003)

1.2 Problem Statement

Work zone crashes have been a major area of concern for engineers, government agencies, the highway industry, and the public. In 2008, there were over 40,000 injuries resulting from work zone crashes in the United States (FHWA, 2009). There was a 45 percent increase in work zone fatalities from 1996 to 2006, which reached to 1,010

fatalities in 2006. Also of concern, 235 of the 1,010 fatalities involved large trucks. On average, a work zone injury occurs once every 13 minutes and a fatality once every ten hours (FHWA, 2009). Figure 1.2 shows the work zone fatality rates from 1982 to 2006 in the United States.

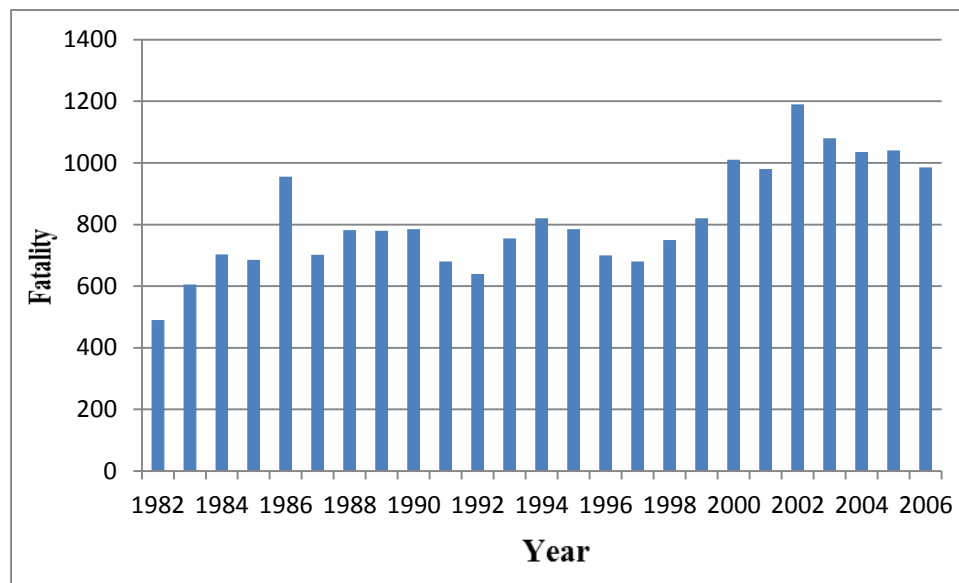


Figure 1.2: Work Zone Fatalities in the US 1982-2006 (FARS, 2006)

The Ohio Department of Transportation (ODOT) continues to expand its efforts to reduce work zone crashes in the state of Ohio. In 2004, 6,389 work zone crashes resulted in 14 fatalities and 2,250 injuries. In 2005, there were 5,854 work zone crashes with 20 fatalities and 2,076 injuries in the state of Ohio. While overall the work zone-related crashes and injuries have gone down, fatalities have gone up. ODOT (2013) reports in years 2003-2012, 56,945 vehicle crashes occurred in Ohio's work zones. Furthermore, about 169 fatal crashes occurred in Ohio work zones in the period of 2002-2011, and the highest number of fatal crashes occurred from 2009 to 2011 (National Highway Traffic Safety Administration (NHTSA), 2012). As a result, we need to identify

major factors that contribute to the increase of the likelihood of occurrences of work zone crashes in the state of Ohio. We also need to discover how to reduce the threat to the lives of Ohio drivers, passengers and highway workers.

1.3 Research Objectives

The main objective of this thesis study is to focus on crash severity and location of crashes within work zone areas in the state of Ohio using classification tree modeling. In addition, the current study assessed main explanatory factors (speed limit, weather, time-of-day, driver's age, road conditions, etc.), which cause an increase in work zone crashes and provided traffic control safety recommendations. The models in this study were estimated using crash data obtained from the Ohio Department of Public Safety (ODPS) database.

1.4 Thesis Organization

This thesis report is organized into five chapters as following:

- Chapter One: Introduction. This chapter includes a general introduction about work zone crashes, problem statement, research objectives and thesis organization.
- Chapter Two: Literature review. This chapter presents the previous findings of the literature review on work zone crashes.

- Chapter Three: Methodology and data collection. This chapter outlines the methodology and data collection of this study on work zone crashes. Also, in this chapter explains the methodology used in this study and describes the variables used in the data.
- Chapter Four: Results. This chapter includes the results of data analysis and discussion of the results.
- Chapter Five: Conclusions and recommendations.

CHAPTER TWO

LITERATURE REVIEW

2.1 Introduction

A review of the literature were corroborate that work zones create a significantly higher rate of crashes in the United States under certain conditions when compared to non-work-zone locations (Hall and Lorenz, 1989; Ha and Nemeth, 1995; Garber and Woo, 1990; Rouphail et al, 1988; Wang et al, 1996). (Hall and Lorenz .1989) stated that vehicle crashes increase about 26% in during work zone maintenance. In addition, (Rouphail et al, 1988) found that crash rates during construction activities increase by 88% when compared with non-work zones. In Virginia in period from 1996 to1999, work zones involved a higher proportion of fatal crashes than non-work-zone locations (Zhao, 2001). According to these facts need to understand work zone crash characteristics and risk factors and identify temporary traffic control to develop work zone safety bases on previous studies.

2.2 Work Zone Crash Characteristics

There are several studies, which have focused for many years in the investigation of the characteristics of work zone crashes. These research efforts are based on multi-state work

zone crash data records, so major characteristics documented from these studies vary with different data scope. In this section, crash characteristics from a number of previous studies were reviewed and are discussed in terms of the following major themes:

- Crash severity
- Crash rate
- Collision type
- Crash time
- Crash location
- Trucks and safety in work zones
- Speed related and posted speed limit in work zone
- Causal factors

2.2.1 Crash Severity

Work zone crashes have been related with more severe crashes as compared to non-work zone crashes. According to several states, such as Virginia (Garber and Zhao, 2002), Texas (Ullman and Krammes, 1990), Kentucky (Pigman and Agent, 1990), and Ohio (Nemeth and Migletz, 1978), discovered significant increases of severe crashes in work zones. Also, some studies found that both frequency of crash severity and average of fatalities per crash were higher in work zones across the nation (AASHTO, 1987). However, other studies (Chemblett et al, 2002; Ha and Nemeth, 1995; Hall and Lorenz, 1989) did not find any changes in crash severity rate in work zone compared with non-work zone crashes. Few studies found less severe work zone crashes than non-work zone

crashes (Wang et al, 1996; Garber and Woo, 1990; Rouphail et al, 1988; Hargroves, 1981).

2.2.2 Crash Location

The locations of work zones are a critical factor that affects an increase in work-zone accident rates. In previous investigations, studies did not reach a conclusion on which area is the most dangerous along the work zone. However, they agreed on the unbalanced crash distribution along the work zones. Some studies (Garber and Zhao, 2002; Schrock et al, 2004) argue that the activity area is the most dangerous area in terms of severe crash frequency. Another investigation study points out about 39.1 percent of crashes occurred in longitudinal buffer area and 16.6 percent of crashes occurred in the activity area (Nemeth and Migletz, 1978). However, the advanced warning area (Pigman and Agent, 1990), the transition area, and the termination area (Nemeth and Migletz, 1978; Hargroves, 1981) were highlighted as the most hazardous areas along work zones in different literatures. Furthermore, (Raub et al, 2001) when investigating crash locations within work zone areas in the state of Illinois they found that:

- The advance warning and transition areas recorded about 40 percent of all the work zone crashes and that greater than 30 percent of these crashes involved injury and mostly involved two vehicle collisions.
- In the activity area, usually involved more than two vehicle crashes and the most common crash severity was property damage only (PDO).

According to a national study, (AASHTO, 1987) the work zones on rural interstate systems accounted for roughly 69 percent of all fatal crashes. In addition, other

studies identified the rural highways (Pigman and Agent, 1990; Chembless et al, 2002) and two-lane highways (Roupail et al, 1988) as the most risky areas in which work zone crashes occur. However, in Virginia a study by Garber and Zhao (2002) found the urban highways system had much higher percentage of work zone crashes than the rural highways system. Moreover, Garber and Woo (1990) found that work zone crash rates increased by approximately 57% on two-lane urban highways.

2.2.3 Collision Types

In general, collision types of work zone are dependent on different locations and times. However, most studies agree that rear-end collisions are one of the most frequent work zone crash types (Mohan and Gautam, 2002; Garber and Zhao, 2002; Pigman and Agent, 1990; Nemeth and Migletz, 1978; Chembless et al, 2002; Hall and Lorenz, 1989; Garber and Woo, 1990; Roupail et al, 1988; Hargroves, 1981). One study also indicate that percentages of rear-end and sideswipe collisions during work zone were higher than the percentage of rear-end and sideswipe collisions in non-work zone crashes (Wang et al, 1996). Furthermore, some studies found that same-direction sideswipe collisions were the major type in work zone crashes (Pigman and Agent, 1990; Garber and Woo, 1990). Another common types of work zone crashes involved angle collisions (Pigman and Agent, 1990) and hitting-fixed-objects (Mohan and Gautam, 2002; Nemeth and Migletz, 1978; Hargroves, 1981). For a study conducted in Georgia (Daniel et al, 2000) single-vehicle crashes, angle, and head-on collisions were the main types of fatal work zone crashes.

2.2.4 Crash Time

According to several studies, the most common work zone crash time is at dusk (Garber and Zhao, 2002 and AASHTO, 1987). The Pigman and Agent (1990) study compared the relationship between crash severity and time of day. They indicate that overall nighttime crashes are more severe in work zones than daytime. Nemeth and Migletz (1978) found that transition area crashes were more likely to occur in nighttime. Additionally, a study by Nemeth and Migletz (1978) found that the percentage of large trucks and buses that caused severe crashes during nighttime was higher than the percentage of other types of vehicles. However, some other studies found work zone crashes frequently occurring in the daytime (Mohan and Gautam, 2002; Chembless et al, 2002; Hill, 2003 and Li and Bai, 2006) and during the most common season of construction, which is between June and October (Pigman and Agent, 1990).

2.2.5 Trucks and Safety in Work Zones

One of the major work-zone safety worries is regarding large trucks, because they are frequently involved in work zone crashes. Some previous studies had found that the percentage of crashes involving trucks are much higher in work zones (AASHTO, 1987; Pigman and Agent, 1990). In addition, some studies have found that crashes related to heavy trucks were more likely to involve multiple vehicles and frequently resulted in fatalities and large economic losses (Pigman and Agent, 1990; Schrock et al, 2004). According to truck drivers who were surveyed on, they considered about 90% of their

driving through work zones to be more hazardous than in other areas (Benekohal et al, 1995). Furthermore, according to The Federal Motor Carrier Safety Administration (FMCSA, 2002) the probability of fatal large truck crashes in work zones were about 24 percent of all crash severity types; that these were mainly due to their large sizes, limited maneuverability, and narrow lanes in work zone areas. An interesting study whose results are shown in Table 2.1 compared total crashes in work zones and work zone truck crashes in different work zone area locations (Khattak. and Targa, 2004). They did not find a significant difference in parentages of crash rates between total crashes in work zones and truck work zone crashes.

Table 2.1 Distribution of total and truck crashes in different work zone locations (Khattak and Targa, 2004)

Location	Parentage of work zone crashes	Parentage of work zone truck crashes
Advance warning area	21.2%	21.7%
In activity area	33.9%	33.7%
Adjacent to actual work area	44.9%	44.6%

2.2.6 Crash Rate

Several studies indicate that crash rates increase in highway work zones because work zones interrupt traffic flows in highways than in other places (Garber and Zhao, 2002; Pigman and Agent, 1990; AASHTO, 1987; Pal and Sinha, 1996; Graham et al, 1977).

Hall and Lorenz (1989) point out that crashes during a work zone period increased about 26 percent compared with the same period in the last year when there was no construction. Other studies (Ullman and Krammes, 1990; Rouphail et al. 1988) compared crash rates during work zone and non-work zone periods at long-term and short-term work zones. They indicate that the parentage of work zone crashes at long-term increased by 88 percent compared to non-work zone period, while the crash rates at short-term work zones were not affected by the roadwork. Likewise, the study by Garber and Woo (1990) found that crash rates increased by 57 percent on multilane highways work zone areas and on two-lane urban highways work zone areas increased about 168 percent when compared with the previous work zone studies.

2.2.7 Speed Related and Posted Speed Limit in Work Zones

Higher driving speeds in work zones are one of the leading causes of traffic crashes because posted speed limits in work zone segments are usually reduced from normal speed limits (Paulsen et al., 1978; Garber and Gadiraju, 1988). A study by Garber and Woo (1990) argue that changes in speed variance through work zone areas contributed to increased crash rates. In addition, Garber and Gadiraju (1988) report that most drivers did not comply with a posted speed limit in a work zone because it was much lower than the original roadway posted speed limit. However, a study by Nemeth and Rathi (1983) conclude that higher speeds and an introduction of trucks and probability of disturbance in the transition area were minimized when early merging drivers were assumed to respond to the lane closure signs immediately.

2.2.8 Causal Factors

Several factors can also contribute to an increase in work zone crashes. According to a number of previous studies reviewed, human errors are the major contributing factors in work zone crashes, such as following too close, inattentive driving, misjudging and lane change (Mohan and Gautam, 2002; Pigman and Agent, 1990; Chembless et al., 2002; Hargroves, 1981; Daniel et al., 2000; Hall and Lorenz, 1989). Other studies (Garber and Zhao, 2002; Ha and Nemeth, 1995; Nemeth and Migletz, 1978) found speed and inefficient traffic control as other significant factors causing crashes in work zones. In addition, according to Pigman and Agent (1990), traffic congestion, construction equipment and materials on roadways delay travels as well as increase work zone crashes. However, Hill (2003) points out that types of driver errors were dependent on time of day as daytime crashes and nighttime crashes. Studies by Nemeth and Migletz (1978) and Garber and Woo (1990) prove that environmental and road surface conditions did not contribute more to work zone related crashes than to crashes at other locations. Moreover, in Ohio from 2006 to 2010, fatal and bodily injury work zone crashes involving motorcycles increased from 166 to 175 per year. While the fatal and bodily injury crashes that occurred in work zones involving motor vehicles decreased from 1526 to 1296 over the same period (FARS, 2012).

2.3 Traffic Control Devices

Work zones' temporary traffic control (TTC) devices are used to provide warnings and alerting drivers in speed reduction zones, efficient traffic flows during roadwork and guiding and directing traffic safely through work zones. The common temporary traffic control (TTC) devices that are used in work zones include; flaggers, traffic signs, arrow panels and portable changeable message signs, channelization devices, pavement markings, lighting devices, temporary traffic control signals, and rumble strips (MUTCD; FHWA, 2003).

Flagger controls are devices such as STOP/SLOW paddles, lights, and red flags to direct road users through work zones. The flaggers should stand in appropriate locations or preceded by an advance warning sign that approaching road users have sufficient distance to stop at an intended stopping point (MUTCD, 2003). A number of studies indicate that flaggers are commonly used on two-lane, two-way rural highways and multilane urban highways; and appropriate for short-duration applications “less than one day” and mobility work zones (Richards and Dudek, 1986; Garber and Woo, 1990). Hill (2003) indicates that flaggers contributed in reducing fatal work zone crashes. Furthermore, Li and Bai (2008) found that flaggers in work zones could reduce the probabilities of causing fatalities when a severe crash occurred by 56%. However, a study by Benekohal et al., (1995) suggests improving flagging for heavy-trucks traffic, since truck drivers are hard in seeing flaggers, which may assist in increasing crash severity rates.

Changeable message signs are flexible message signs that display a diversity of warning messages. Several studies (e.g., Garber and Patel, 1994; Garber and Srinivasan,

1998; Dixon and Wang, 2002; Brewer et al., 2006) comment that a changeable message sign tended to be more effective in reducing vehicle speeds than traditional work zone message panels. However, another evaluation Richards and Dudek (1986) found that changeable message signs could lose their effectiveness when used alone or operated continuously for long periods without changing warning messages. Other common TTC devices are channelizing devices, which are used to alert drivers of changed traffic conditions in work zones and to direct road users to drive safely and smoothly through work zones (Pain et al., 1983). In addition, certain temporary traffic control signals, such as STOP/GO signals contributed in reducing fatal crashes in work zones (Hill, 2003).

The most common devices that attract drivers' attentions are lighting devices that illuminate work zones and alert road users of the complicated road conditions at both daytime and nighttime. A number of studies (e.g., Huebschman et al, 2003; Arnold, 2003) suggest that using flash warning lights such police vehicles with flashing lights, contributed in reducing speeds in work zones. In addition, Nemeth and Rathi (1983) points that using flash lights on four signs in the advance warning area on both sides of Ohio roadways assist in reducing motor vehicle speeds.

CHAPTER THREE

METHODOLOGY AND DATA COLLECTION

3.1 Data Collection

Crash data for the current study were obtained from the Ohio Department of Public Safety (ODPS) and the data covered five years from 2008 to 2012. ODPS crash database contains police-reported crashes that occurred on Ohio's public roads and streets.

3.1.1 Data Collection Methodology

This study used crash reports spanning five years, from 2008 to 2012. ODPS datasets include four files of crash records, unit records, people records, and ODOT records, which are organized in a relational format with crash records compiled together by calendar year. Two important variables are used to combine these files together into one file. These two variables are DOCNO and UNITNO; the DOCNO is a variable contained in all four files, which is very important as it relates all data cases to their respective crash incidents. In other words, this is a unique number that identifies each crash incident (record). The unit and people record files include a UNITNO variable, which is a variable that links all people involved in traffic crashes to their specific type of unit (i.e., vehicle) they were traveling in and their correct crash incidents they were involved. Figure 3.1

shows the four record files with their related variables that are used to properly and systematically join related cases together.

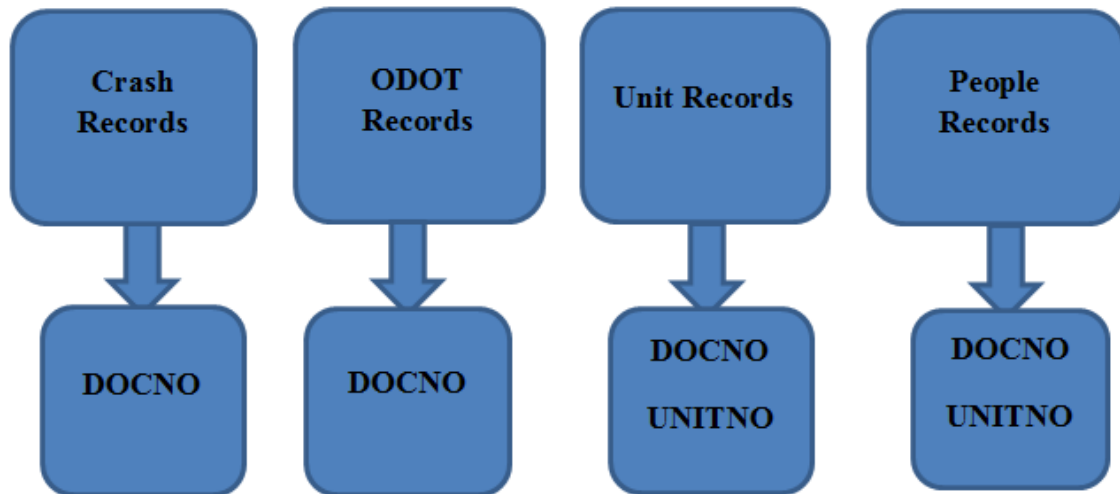


Figure 3.1: Four Files with Their Related Variable

These four related files can be briefly described as follows:

- A. The “crash records” file includes individual information for each occurring crash such as crash severity, vehicle in error, date of crash, time of crash, name of city, village or township where the crash occurred, FIPS place code, crash location, type of road, if alcohol or drug was involved, if speeding was involved, etc.
- B. The “ODOT records” file includes information such as county code, route type, latitude, longitude, construction plans, daily construction diaries, a database of the construction dates, crash type, etc.
- C. The “unit records” file includes specific information for each unit (vehicle), such as unit type (passenger vehicles, medium/heavy trucks, motorcycle, pedestrian, etc.), point of impact, number of occupants in the unit, etc.

- D. The “people records” file includes information for each person who was involved in each crash, such as person type (e.g., driver, occupant, or pedestrian), age, gender, severity of injury, seating position, use of safety equipment, and use of an air bag, etc.

3.1.2 Merging Data Files

As mentioned previously there are four files in ODPS database, but in this study only records from three files were needed, that is, crash records, unit records, and the people records files. These three files were merged together into a single file by using a one-to-many merging technique in SPSS (Version 22.0) software. First, DOCNO was used to join the “crash” and “unit” files together in one file known as “crash-unit” file. Then, the “people” file was merged with the “crash-unit” file by using both DOCNO and UNITNO to create a joint “crash-unit-people” file. Finally, the five created joint “crash-unit-people” files for each year were again joined together to create a new file that contained a five-year traffic crash data for years 2008 through 2012. A flowchart showing the files merging process is schematically illustrated in Figure 3.2.

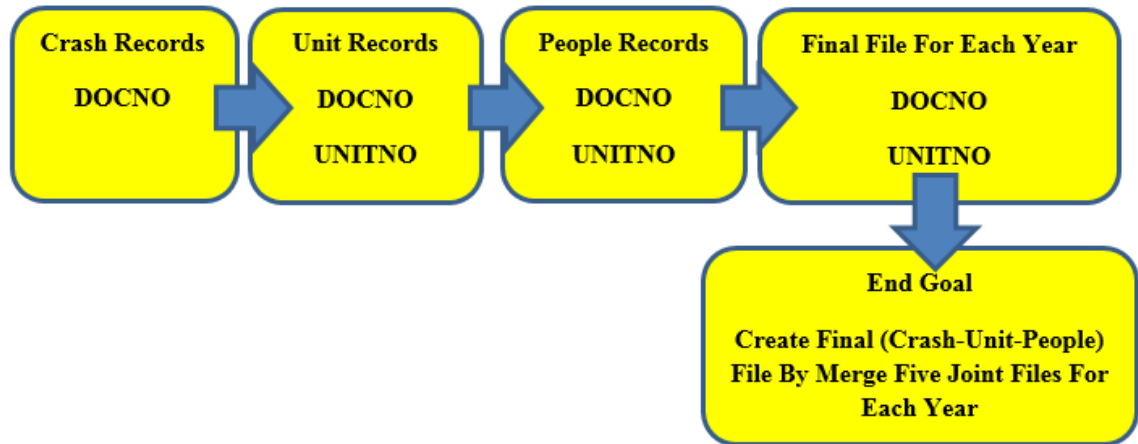


Figure 3.2: The Process of Creating a Joint Crash-Unit-People File

3.2 OH-1 Crash Data

In the current study, OH-1 Crash Data Documentation as a guide for crash variables was used. This OH-1 crash report includes several categories, which are essentially crash report records such as location conditions, roadway information, crash details, crash relationships, vehicle information, and driver/passenger information. The following sections discuss some categories of variables that are of interest to the current study.

3.2.1 Report Information

The reported information is individual information for each crash in the OH-1 documentation. The very important variable in this category is the document, which allows for a specific crash to be recognized and found when comparing crashes or locating other crash related documents.

3.2.2 Vehicle Information

This category recognizes and describes each vehicle involved in a work zone crash by using a unit number provided for each vehicle. Additional information includes unit type, which is a classification of each vehicle involved in a crash, such as passenger vehicles, medium/heavy trucks or combination units and bus/van, motorcycle, bicycle, etc.

3.2.3 Roadway and Traffic Information

The roadway and traffic information category recognizes the characteristics of roadway (e.g., average daily traffic, intersection type, road contour, road condition, road surface type, speed limit, number of lanes, and traffic control).

3.2.4 Crash Relationships

There are several factors believed to have a relationship with work zone crashes. These include alcohol, drugs, bicycles, motorcycles, speed, and other factors. In addition, there are factors in this category, which are important for this study, including work zone related, work zone type, work zone location and workers present. Work zone type is an important factor affecting the occurrence of work zone related crashes because it creates unfamiliar, confusing, or unexpected situation to the driver. As result, there are different precautions need to be taken when approaching each work zone types. These precautions are related to the hazards of merging traffic, altered lane movements; close quarters with work, and approaching work in traffic at large speed differences. Table 3.1 depicts different work zone types as defined in the OH-1 crash documentation.

Table 3.1 Work Zone Types as Defined in the OH-1 Crash Documentation

Work Zone Type	Description
Lane closure	One or more lanes are closed, consequential in a transfer of traffic from the closed lane to the residual open lane(s).
Lane shift or crossover	One or more lanes move laterally, without any of the lanes closing.
Work on the shoulder or median	The location of the work zone is happening on either the shoulder or the median of the roadway.
Intermittent or moving work	The work either is making frequent short-term stops or is moving at a slow speed.

3.2.5 Crash Details

This category explains the events leading up to the crash (unit speed, pre-crash actions, and sequence of events), the act of the crash (number of units, occurrence, and collision type), and the after-effects of the crash (crash severity, total injured, and total killed).

3.2.6 Driver/Passengers Information

This category describes how the driver/passengers may have affected the crash, and how they may have been affected by the crash. The OH-1 crash data describe driver and passenger's age, their gender and if the use alcohol or drugs were substantial factors in the occurrence of a work zone-related crash.

3.3 Creating of Work Zone Related Traffic Crash Database

Using the joined file, WORK ZONE RELATED variables were isolated by querying and sorting records to a new file that contained work zone crash records only. That is, this new file included crash information for crashes that occurred in work zones only. After

sorting out work zone crash data into a new file, this final file ended up with 58,368 records of work zone-related crashes that occurred since January 2008 through December 2012.

3.4 Methodology

3.4.1 Introduction

This study focuses on analyzing crash severity and crash location related to work zone areas by using crash data from traffic crashes that occurred in Ohio for the period of 5 years from 2008 to 2012. Classification tree modeling is a procedure that was used to identify significant factors that could increase crash risk in work zones. In this study, both crash severity and work zone crash locations were developed by using the classification tree modelling.

Breiman et al. (1984) used classification tree method to identify predictor variables that would make significant split of the data by dividing original group of data into pairs of subgroups in the response variable. The decision tree is a common modeling tool for both exploratory data analysis and predictive modeling applications because it is useful in identifying features and extracting patterns from large databases that are required in predictive and discrimination modeling (Myles et al, 2004).

3.4.2 Classification Trees Modeling

3.4.2.1 Introduction

Classification tree modelling is useful in dividing the database into smaller and more homogeneous subgroups by using a set of “if-statements” and some statistical measurements. The classification tree model divides the database according to the most predictive independent variables for the dependent variable. Moreover, it chooses an appropriate measurement according to the type of the dependent variable. Consequently, the main purpose of using classification tree method was to identify the main factors that affect crash severity and to recognize the significant areas within work zone locations.

The dependent and independent variables can be either continuous or categorical variables. The variable is split into two partitions according to cutting value of the variable, if the independent variable is continuous. However, if the independent variable is categorical (nominal or ordinal), the model splits the variable levels into two groups of levels. On the other side, when the dependent variable is continuous, the quantities of the difference of the two fitting groups are calculated as the sum of squares of the differences between means. But in case of a categorical dependent variable the fittings are calculated based on maximizing a *LogWorth* statistic, which is associated to the G^2 (likelihood ratio chi-square statistic). For the continuous dependent variable, the fitted values are the means within the two split groups, while in categorical variable; the fitted values are the assessed proportions within the groups.

3.4.2.2 Classification Trees Formed

The classification tree contains a number of nodes that represent attribute variables. These nodes are divided into three types, including a root node, branch nodes and terminal nodes. The root node is the most important node because it is the top most node of the decision tree and it contains the entire sample dataset. The root node is then split into two branches (nodes). It is noteworthy to mention that the classification tree (or decision tree) model generates binary trees to create the best homogeneity for a predictor variable (splitter). While, the branch nodes represent one of the cases of the one variable, the terminal node (leafs) can be used to identify the expected value of the variable class or target variable. Figure 3.3 illustrates the classification tree flow chart.

The principle of growing the classification tree is applied recursively (by a descending strategy) by partitioning the data variable to maximize “purity” in the child node. Each branch node, also known as parent node, is connected with two children nodes, which have more homogenous datasets than those in the upper parent node. Therefore, this process is repeated on each child node by a split recursively until all of them are pure (when all the subsets are of the same value) or their "purity" cannot be improved and this node becomes a terminal node also known as leaf node (e.g., nodes 2, 4, and 5 in Figure 3.3).

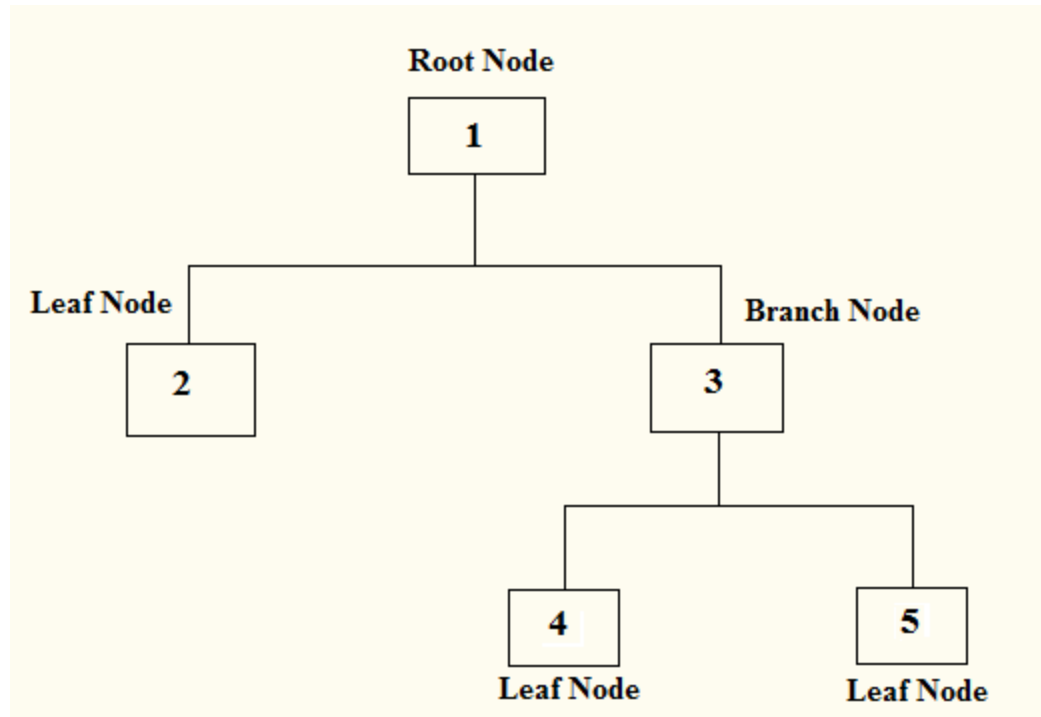


Figure 3.3: An Example of a Classification Tree Form

3.4.2.3 Pruning and Validation

The classification tree stops growing when these criteria are met, and then the tree is pruned to the smallest subtree according to the significant risk variables. The pruning technique is set up after the tree built based on split history chart (R^2 vs. number of splits). The pruning tree technique starts when splitting history for training and validation datasets have the same R^2 values. When growing the classification tree model usually there is a danger of overfitting the model, which is a result of a model being overly complex and having many less important variables. In addition, an overfitting issue tends to minimize the predictive power in the model, because it results into small fluctuations in the data. As a result, tree pruning is used to decrease overfitting of the model and thus improving its predictive accuracy. Furthermore, pruning is used to reduce the tree size,

complexity of the final model and to remove the less important splits of a classification tree, which may be based on noisy or erroneous data.

Validation process is used to assess how well the predictor variables predict the dependent variable. Validation is the technique of splitting the dataset into two parts (training set and validation set). The training set is used in the growth of the classification tree model and to estimate the model parameters. The second part is the validation set and is used to evaluate the predictive ability of the model developed with the first set of data. This is done by testing the performance of the model built on a set of data not used for training (or building it).

3.4.2.4 Statistical Analysis

3.4.2.4.1 Node Splitting Criteria

The main purpose of developing the classification tree model is to achieve the maximum purity in the nodes, so a *LogWorth* is used to identify the degree of purity of the node as well as in growing the tree. The *LogWorth* statistic parameter is used to decrease complexity and to facilitate pruning of the decision tree by using *p-value* measure that calculates the accuracy of criteria as opposed to the complexity in the number of nodes. Furthermore, *LogWorth* is used to identify the most significant independent (predictor) variables based on the larger *LogWorth* value. In general, the model splits the node based on the maximization of the *LogWorth* statistic, which is computed as illustrated by Equation 3.1.

$$LogWorth = -\log_{10}(p-value) \dots\dots\dots (3.1)$$

For a continuous dependent variable, the sum of squares (SS) is another parameter used to split the nodes with the *LogWorth* parameter. Essentially, this is the change in the error sum-of-squares due to the split. A candidate the sum of squares (SS) that has been chosen for splitting is computed as shown in Equation 3.2:

$$SS_{test} = SS_{Parent} - (SS_{Right} + SS_{Left}) \dots\dots\dots (3.2)$$

Where:

SS_{Parent} = Parent node total squares error

SS_{Right} = Child node total squares error in the right- hand side

SS_{Left} = Child node total squares error in the lift- hand side

SS in a node is just $S^2 (n - 1)$

Furthermore, the difference in the statistic for continuous dependent variable is also calculated. This is the difference between the predicted values for the two child nodes of a parent node.

For the case of a categorical dependent variable, the G^2 (log-likelihood-ratio chi-square), is computed to split the nodes with the *LogWorth* parameter. This is essentially twice the change in the entropy. This entropy is computed as shown in Equation 3.3:

$$G^2 = 2 \sum \left\{ f_0 \log \left(\frac{f_0}{f_e} \right) \right\} \dots\dots\dots (3.3)$$

Where:

f_0 = observation node frequency

f_e = expectation node frequency

Candidate G^2 chosen for splitting is computed as shown by Equation 3.4:

$$G_{test}^2 = G_{Parent}^2 - (G_{Right}^2 + G_{Left}^2) \dots\dots\dots (3.4)$$

3.4.2.4.2 Receiver Operating Characteristic (ROC) Curve

The receiver operating characteristic (ROC) curve is a statistical tool and a graphical plot of the sensitivity (true positive rate) vs. 1 – specificity (false positive rate) that is used to examine categorical response variables. An ROC curve is retrospectively assessing the accuracy (power) of predictions by using the area between the curve and the diagonal line to summarize the accuracy of the analysis data (Agresti, 2007). Additionally, this area can be used to identify the risk factors of the crash severity and the work zone area crashes in the current study. The ideal prediction performance has 100 % sensitivity and 100 % specificity, but this scarcely occurs in reality. Practically speaking, an ideal classification tree should depict a ROC curve inclined towards the left top of the graph. On the other hand, values of ROC curve range from 0 to 1, where a value of 0 shows a perfectly inaccurate test and a value of 1 reflects a perfectly accurate test. In addition, the ROC curve usually takes a shape of a concave curve which links the points (0, 0) on the bottom left corner and (1, 1) at the upper right (SAS, 2003) as shown in Figure 3.4.

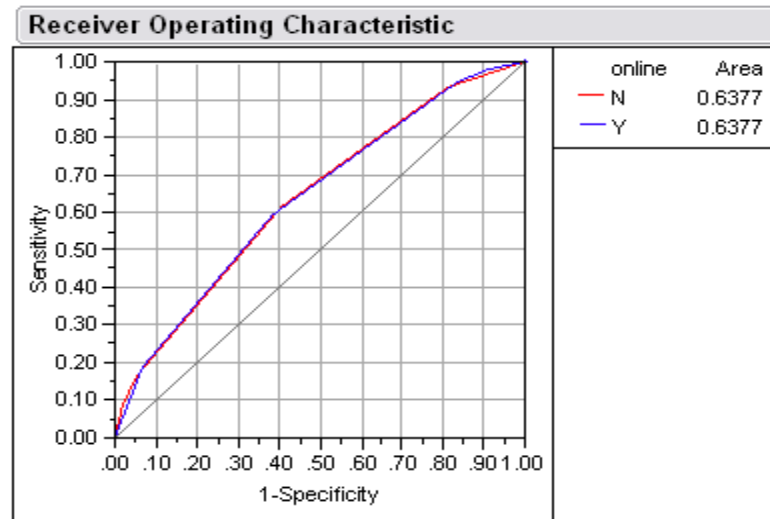


Figure 3.4: An Example of a ROC Curve

CHAPTER FOUR

RESULTS

4.1 Introduction

Several factors contribute to the increase of work zone crashes as related to crash severity and to specific location in the work zone areas. Such factors may include geometric, environmental, traffic, and drivers' behavioral factors. The current study analyzed these factors by using some powerful statistical modeling techniques in order to determine the most significant ones. In addition, this study selected twenty-two variables for exploratory analysis to investigate characteristics of dependent variables of work-zone crash severity and work zone location. A classification tree model was used to estimate statistically the effects of these variables in contributing to crash severity of work zone crashes. In addition, the classification tree model was used to identify the significant area in work zone location, which increases the frequency of crashes in work zone areas. A classification tree procedure in JMP software (version 11) was used for developing these estimates. This chapter discusses the descriptive results of the data analyzed in this study and the results of the two classification tree models developed to analyze significant predictor variables of the crash severity model and the WZ location crashes model.

4.2 Description of Selected Variables

This section discussed the characteristics of Ohio's work zone (WZ) crashes that occurred between 2008 and 2012. These characteristics were divided into three categories as depicted in Tables 4.1 through 4.4; which are crash/traffic characteristics, human/driver characteristics, and roadway/environmental characteristics, respectively. Additionally, the three tables include the number of observations for the two crash severity types considered in the current study, i.e., fatal/injury and PDO.

4.2.1 Description of Crash and Vehicle Characteristics

According to Table 4.1, WZ crash type was divided into five categories (levels), which include non-collision, collision with work zone equipment, collision with person/vehicle/object not fixed, collision with a fixed object, and unknown crash type., and their frequencies of occurrences were analyzed and the data reveal that they were 9.41%, 1.70%, 84.63%, 3.00% and 1.26%, respectively. In terms of vehicle type, it was recorded that passenger vehicles (cars, SUVs, minivans, and pickup trucks) made up 85.41%, trucks and buses 13.83%, motorcycles and bicycles 0.21% and emergency vehicles (police cars, ambulances, and fire trucks) 0.54% of WZ crashes. Posted speed limit was coded into three category levels, with 33.54 percent of WZ crashes occurring on roads with posted speed limits less than 40 mph, 31.51 percent occurring on roads with posted speed limits of 40-50 mph 34.95 percent occurring on roads with posted speed limits above 50 mph. Collision type was coded with six category levels. The largest recorded category was rear-end collisions, which constituted 45.54% of all WZ crashes, and the other five categories include no collision between two vehicles, head on, rear to

rear/backing, angle and sideswipe same direction/sideswipe opposite direction in the work zone and their percent contributions were 17.97%, 1.51%, 2.75%, 15.68% and 16.55%, respectively.

Table 4.1 Description of Crash and Vehicle Characteristics

Variable	Description (Code)	Fatal/ injury	PDO	%Total
WZ crash type	Not Collision Between Two Vehicles in Transport (NC)	1750	3741	9.41%
	Work Zone Equipment (WZE)	169	825	1.70%
	Collision with Person/Vehicle/Object not Fixed (PVUF)	14488	34910	84.63%
	Collision with Fixed Object (FO)	510	1239	3.00%
	Unknown (UN)	113	623	1.26%
Vehicle type	Passenger vehicles (PC)	14893	34962	85.41%
	Trucks/Buses (TB)	1945	6125	13.83%
	Motorcycles/motorized bicycles (MOB)	24	101	0.21%
	Emergency vehicles/Unknown (EMU)	168	150	0.54%
Posted speed limit	< 40 mph	4908	14671	33.54%
	40-50 mph	5449	12941	31.51%
	55-70 mph	6673	13726	34.95%
Collision Types	No collision between two vehicles in transport/unknown (NCU)	2537	7954	17.97%
	Rear-end (RE)	9521	17057	45.54%
	Head on (HO)	540	339	1.51%
	Rear to Rear/ Backing (RRB)	216	1388	2.75%
	Angle (AG)	2901	6253	15.68%
	Sideswipe same direction/ Sideswipe opposite direction (SSO)	1315	8347	16.55%
Work Zone Location	Termination Area (TA)	845	1990	4.86%
	Before The First Work Zone Warning Sign(BWS)	844	2249	5.30%
	Advance Warning Area (AWA)	1976	4091	10.39%
	Transition Area (TSA)	3316	9008	21.11%
	Activity Area (AA)	10049	24000	58.34%
Work Zone Types	Lane Closure (LC)	7500	18094	43.85%
	Lane Shift/Crossover (LSC)	2346	6374	14.94%
	Work on Shoulder or Median (WOSM)	4080	9346	23.00%
	Intermittent or Moving	3104	7524	18.21%

	Work/Unknown/Other (IMUO)			
Workers Present	No	10797	26791	64.40%
	Yes	6233	14547	35.60%
Motorcycle Related	No	16548	41187	98.92%
	Yes	482	151	1.08%
Semi-Truck Related	No	15247	36556	88.75%
	Yes	1783	4782	11.25%

The five segments that make up the work zone areas are shown in Table 4.1. The most significant area is the activity area, where 58.34% of all WZ crashes occurred, the second one is the transition area with 21.11% of all WZ crashes, and 20% of the observed WZ crashes occurred in the other three areas combined (i.e., the termination area, before the first work zone warning sign and advance warning area). As shown in Table 4.1 the highest number of crashes in work zones occurred when there was a lane closure with 43.85% of the work zone-related traffic crashes, while 14.94% occurred when there was a lane shift and a crossover, 23.00% occurred when work was on shoulder or median and 18.21% occurred during intermittent or moving work. Crashes that occurred when workers were present were about 35.60% of all WZ crashes. Motorcycle involvement was reported in only 1.08% of the total WZ crashes. Semi-trucks were estimated to be involved in work zone crashes by 11.25%.

4.2.2 Description of Human and Driver Characteristics

Table 4.2 summarizes the description of human and driver characteristics. People who were less than 20 years old made up 17.11% of all the people who were involved in the work zone crashes, people aged 20-25 years old made up 14.23%, 59.77% involved people aged 26-64) years old and people older than 64 years old were only 8.89% of the

total people involved. In term of gender, females were involved in 42.21% of the total WZ crashes, while males made up 57.79%. Alcohol and drugs were involved in only 2.81% and 0.77%, respectively of the work zone related crashes. Speeding was reported in 12.52% of the work zone related crashes.

Table 4.2 Description of Human and Driver Characteristics

Variable	Description (Code)	Fatal/ injury	PDO	% Total
Age of person	<20	3250	6736	17.11%
	20-25	2529	5779	14.23%
	26-64	9869	25018	59.77%
	65+	1382	3805	8.89%
Youth-related (16-25 years old)	No	9877	26422	62.19%
	Yes	7153	14916	37.81%
Senior-related (65+ years old)	No	14119	34642	83.54%
	Yes	2911	6696	16.46%
Gender of person	Female (F)	7712	16926	42.21%
	Male (M)	9318	24412	57.79%
Alcohol-related	No	16220	40507	97.19%
	Yes	810	831	2.81%
Drug-related	No	16785	41135	99.23%
	Yes	245	203	0.77%
Speed Related	No	14397	36665	87.48%
	Yes	2633	4673	12.52%

4.2.3 Description of Roadway and Environmental Characteristics

Roadway and environmental characteristics are displayed in Table 4.3. In terms of the roadway contour, the data reveal that 71.45% of all WZ-related crashes occurred on straight level segments, 20.27% of the total work zone-related crashes took place on straight graded segments, curved level segments were involved in only 3.87% of total work zone crashes, and 4.42% took place on curved graded segments. In addition, data in Table 4.3 show the road conditions when work zone crashes took place, where 81.50% of

the crashes happened when the road conditions were dry, 14.85% occurred when the roads were wet, 3.67 percent occurred when the roads were covered with snow, ice, mud, gravel, or slush. In term of light condition, the data show that the highest observations of all work zone-related crashes occurred during daylight, dawn or dusk with 81.62%, 10.24% occurred during nighttime on lit roadways and only 8.13% occurred during nighttime on unlit (dark) roadways.

In terms of weather condition, Table 4.3 shows that 62.42% of work zone -related crashes took place when the weather was clear, 25.39% of work zone crashes occurred when the weather was cloudy, and 12.20% of all work zone crashes occurred when there was rain, fog, sleet, snow, or winds. In addition, data reveal the time of the day when the work zone crashes occurred, where 82.34% occurred during early morning or daylight times (05:00-18:59), 11.04% took place during early nights (19:00-22:59) and 6.62% took place during late nights (23:00-04:59). In term of the day of week (categorized as either weekend or weekday), only 6.62% of total work zone-related crashes occurred on weekends and 81.94% % of work zone crashes occurred during weekdays.

Table 4.3 Description of Roadway and Environmental Characteristics

Variable	Description (Code)	Fatal/ injury	PDO	% Total
Road contour	Straight Level (SL)	11992	29710	71.45%
	Straight Grade (SG)	3579	8253	20.27%
	Curve Level (CL)	701	1555	3.87%
	Curve Grade (CG)	758	1820	4.42%
Road condition	Dry (DR)	14150	33419	81.50%
	Wet/water (WT)	2360	6299	14.84%
	Snow/ice/mud/oil/slush/gravel/other(OT)	520	1620	3.67%
Light condition	Daylight/Dawn/dusk (DLD)	13694	33948	81.62%
	Dark - lighted roadway (DL)	1926	4053	10.24%
	Dark - Unlighted roadway (DULU)	1410	3337	8.13%

Weather condition	Clear (CR)	10792	25639	62.42%
	Cloudy (CY)	4334	10483	25.39%
	Rain/fog/sleet/snow/wind/other (RSU)	1904	5216	12.20%
Time of crash	Early morning/daytime (EMD)	13736	34324	82.34%
	Early night (EN)	2055	4391	11.04%
	Late night (LN)	1239	2623	6.62%
Day of week	Weekends (WE)	3406	7135	18.06%
	Weekdays (WD)	13624	34203	81.94%

4.3 Descriptive Results

By testing the characteristics of crash data and identifying the main factors of the crash can assist the analysts in devising possible safety countermeasures. In this study, we used a total of 58,368 data of work zone crash that occurred on highways and public streets in the state of Ohio for the period of 2008- 2012.

4.3.1 Descriptive Results of Crash Severity

For crash severity, 17,030 observations involved fatal and injury crashes and 41,338 involved in property damage only (PDO) crashes. The major purpose of the descriptive results of work zone-related crash severity was to provide a better view of the characteristics of these crashes. In particular, this section pays more attention to traffic crashes that resulted into fatalities and injuries.

4.3.1.1 Distribution of Crash Severity by Crash Types

Figure 4.1 illustrates the relationship between crash severity and collision types. It clearly shows that head- on collisions have the highest rates of fatal and injury crashes in the work zones with probability of 61.4%. Therefore, head-on collisions were more

hazardous than any other collision type in overall, in terms of the rate of fatal and injury sustained. The next highest rates of fatal and injury crashes occurred to rear-end and angle collision types with the probability 35.8% for rear end and 31.7% for angle collisions.

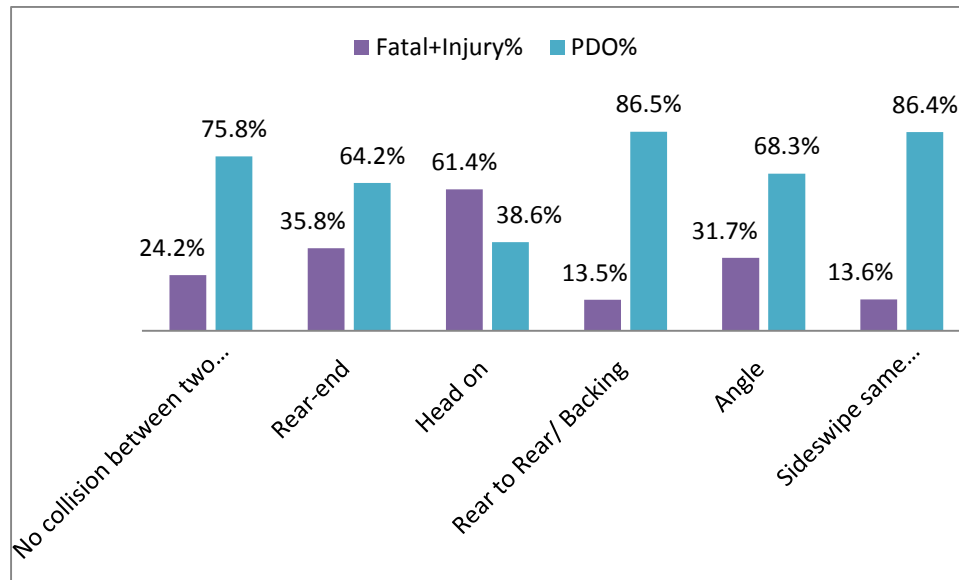


Figure 4.1: Distribution of Crash Severity by Crash Types

4.3.1.2 Distribution of Crash Severity by Work Zone Types

Figure 4.2 shows the relationship between probability of crash severity (fatal/injury and POD) and work zone type. The results show that work on shoulder or median, lane closure and intermittent or moving work/unknown/other have almost the same chances of their crashes becoming fatal or injury crashes with the probability of about 30%. However, the results show that a slightly lower chance of about 26.9% of the crashes that occurred in work zones with lane shift/crossover resulted into fatalities and injuries.

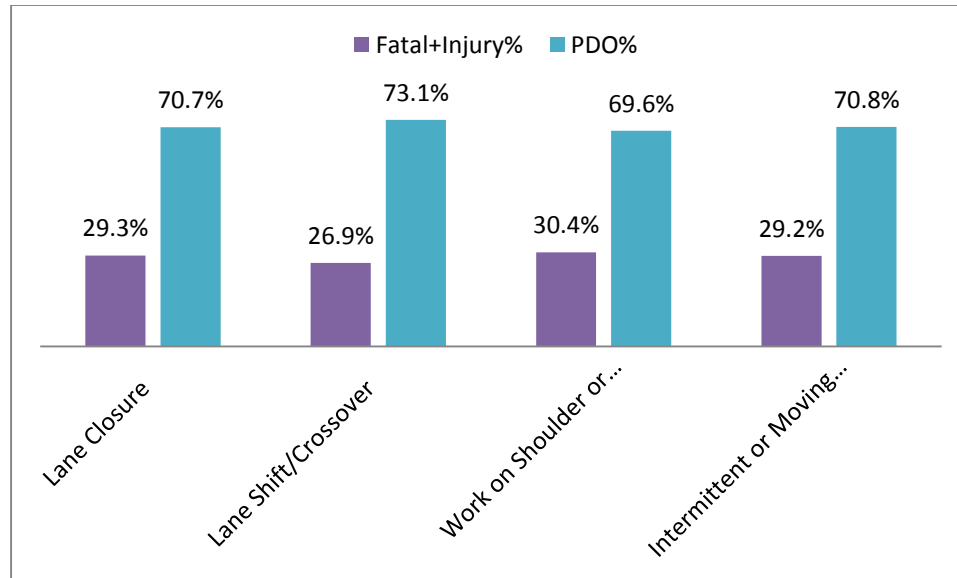


Figure 4.2: Distribution of Crash Severity by Work Zone Types

4.3.1.3 Distribution of Crash Severity by Posted Speed Limit

Figure 4.3 shows the relationship between crash severity and posted speed limit. It is noteworthy to mention that the posted speed limit here means the speed limit of the entire roadway, not the work zone posted speed limit (in the case that a temporary speed limit is posted pertaining to the WZ area only). The ODPS data does not include the work zone speed limit if it is different from the general highway posted speed limit. Figure 4.3 reveals that crashes that occurred on roads with posted speed limits between 55 mph and 70 mph had the highest likelihood of resulting into fatalities or injuries with a 32.7%. This fact indicates fatal and injury crashes more likely happened on roads with high-posted speed limits. For roads with posted speed limits between 40 and 50 mph fatal and injury work zone crashes accounted for about 29.6% of all crashes that occurred on these roads. However, it is not surprising fatal and injury crash rates were lowest for lower

posted speed limits roads with less than 40 mph where 74.9 percent of their crashes were just PDO crashes.

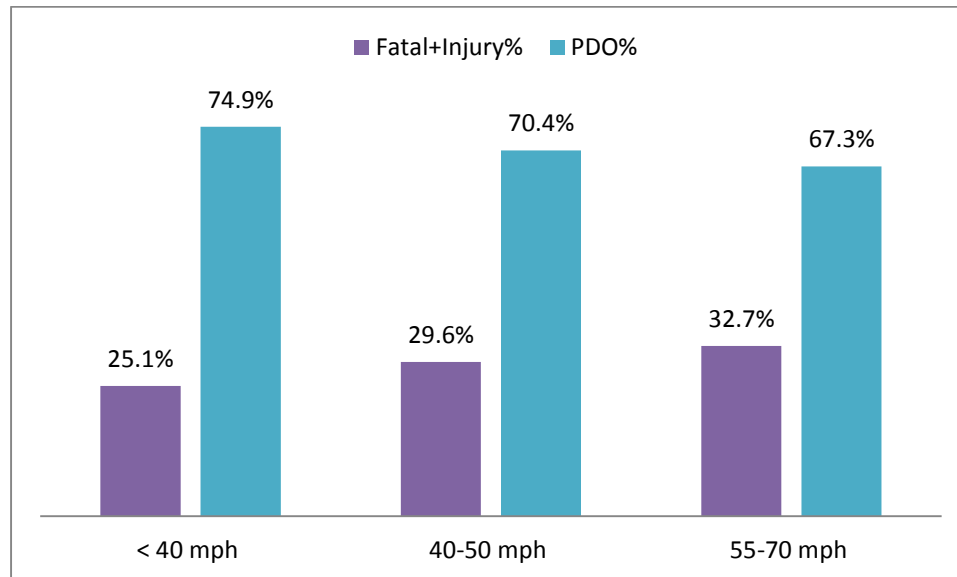


Figure 4.3: Distribution of Crash Severity by Posted Speed Limit

4.3.1.4 Distribution of Crash Severity by Crash Time

Figure 4.4 shows the relationship of WZ-related crash severity with crash time (categorized as early morning/daytime, early night and late night). These results show that crashes that occurred during early nights and late nights involved relatively significant percentages fatal and injury work zone crashes with probability about 32%. Early morning/daytime had bit lower rate of fatal and injury crashes was about 28.6%, about 3.5% less compared with the nighttime categories considered.

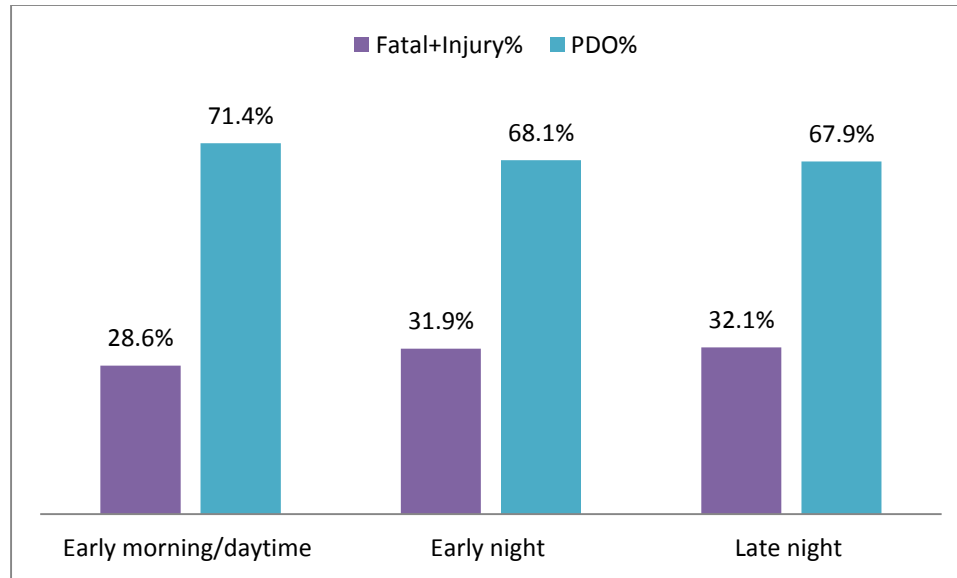


Figure 4.4: Distribution of Crash Severity by Crash Time

4.3.2 Descriptive Results of Crashes Relative to Work Zone Location

The work zone area is usually divided into five distinct sections as already mentioned earlier in this thesis report. The current study was also interested to know which area within the WZ is most prone to high severity crashes. Sections that make up the work zone area include termination area, before the first work zone warning sign area, advance warning area, transition area, and activity area and the total number of crashes in these segments were 2835, 3093, 6067, 12324, 34049, respectively. Consequently, the descriptive results of work zone location-specific crashes pay attention to these crashes within the work zone locations as they relate with other factors.

4.3.2.1 Distribution of Crashes by Work Zone Location

Figure 4.5 displays the likelihood of crash occurrences in different work zone locations (i.e., termination area, before the first work zone warning sign area, advance warning Area, transition area and activity area). These results clearly show that the majority of work zone crashes are more likely to occur in the activity area with a probability about 58.3%. The second location with the highest percent of work zone crashes is transition area with a probability of about 21.1%. However, the lowest rate of work zone crashes occurred in termination area with a percent of about 4.9% of all crashes.

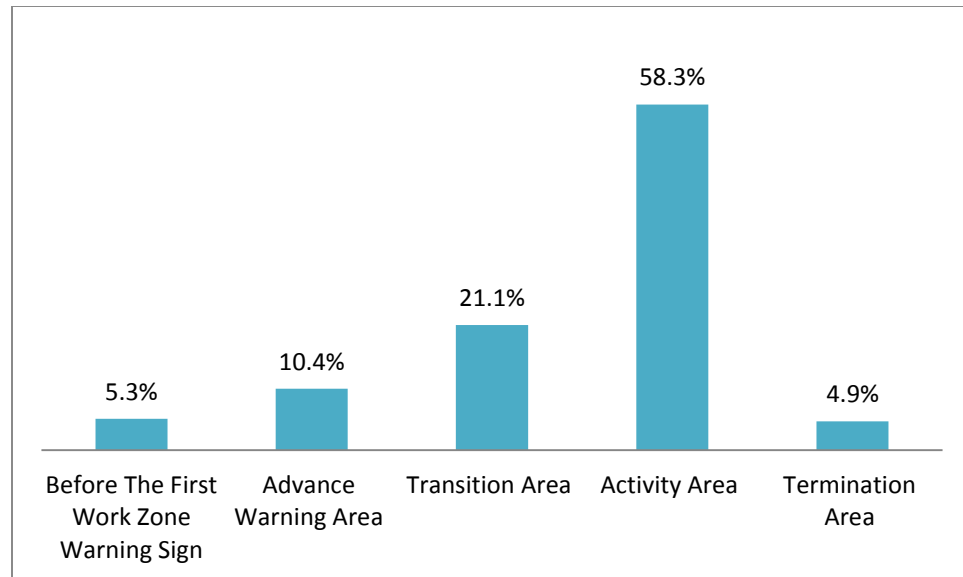


Figure 4.5: Distribution of Crashes by Work Zone Location

4.3.2.2 Distribution of Work Zone Location Crashes by Workers Presence

The relationship between work zone location crashes and whether workers were present when crash occurred is shown in Figure 4.6. About 64.4% of work zone crashes occurred when workers were not present at the construction sites. Therefore, about 35.6% of work

zone crashes occurred when workers were present. Consequently, these results agree with the Ohio Department of Transportation report for the period from 2003 to 2012 (ODOT, 2013).

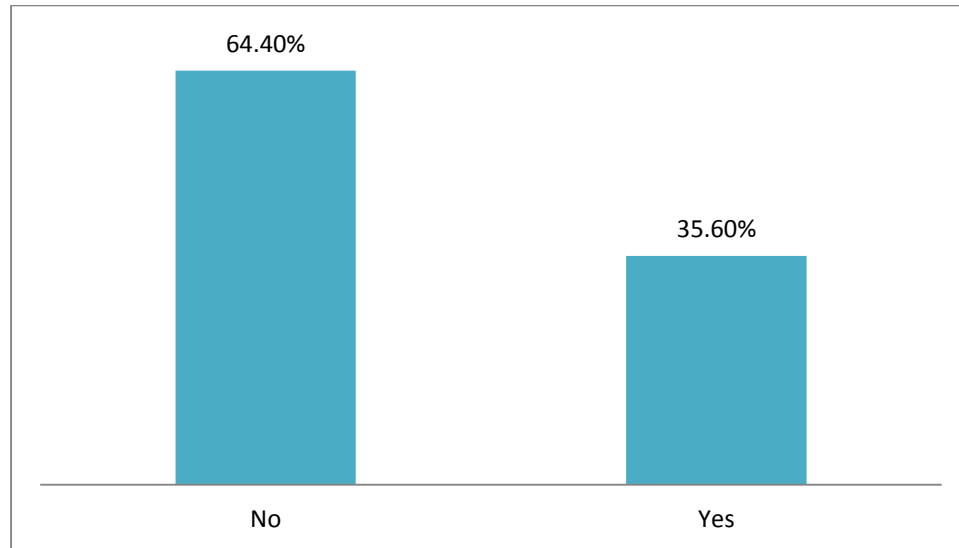


Figure 4.6: Distribution of Work Zone Location Crashes by Worker Presence

4.4 Results of Classification Tree Models

4.4.1 Introduction

A classification tree was used to investigate the complex relationships between the response variable and the predictor variables and recognized multilevel interactions. In this study, the crash severity was the dependent variable in the work zone crash severity analysis, which consisted of two levels of injury severity (fatal/injury and PDO) and was modeled as an ordinal variable. While for work zone location analysis, the dependent variable was work zone location, which consisted of five levels, i.e., termination area (TA) , before the first work zone warning sign area (BWS), advance warning area (AWA) ,transition area (TSA) and active are (AA) .Twenty two predictor variables were

selected among the variables recorded ODPS crash datasets. These predictor variables were grouped into two types of codes (multilevel nominal and binary), twelve variables were coded as multilevel nominal and the other ten were coded as binary nominal variables.

4.4.2 Results of Crash Severity Model

The total dataset inputted in the JMP program contained 58,368 observations of WZ crashes. For validation purposes, this dataset was split into two datasets (training sample set and validation sample set). The training sample set consisted of 52,667 observations (90% of the total observations) that were randomly selected by the program, and the validation sample set consisted of 5,701 observations (about 10% of the total observations). In the training set, which was used to grow the classification tree model, there were 15,366 fatal and injury crashes and 37,301 property damage only (PDO) crashes. The validation sample set is actually used to evaluate the predictive capability of the model developed using the training sample set.

As mentioned before, the classification tree is usually grown with overfitting, which is a result of a model being overly complex and having too many variables. In this study, the JMP software ran the model in full tree splitting where 118 splits were made. Based on the split history chart (R^2 vs. number of splits), as shown in Figure 4.7, it can be observed that split histories for training and validation datasets have almost a constant R^2 value after 18 splits. For that reason, 18 splits are the ideal tree size for this model and

node splitting has to be stopped here. By using these 18 splits, it is reasonable to identify the most powerful predictors of levels of crash severity.

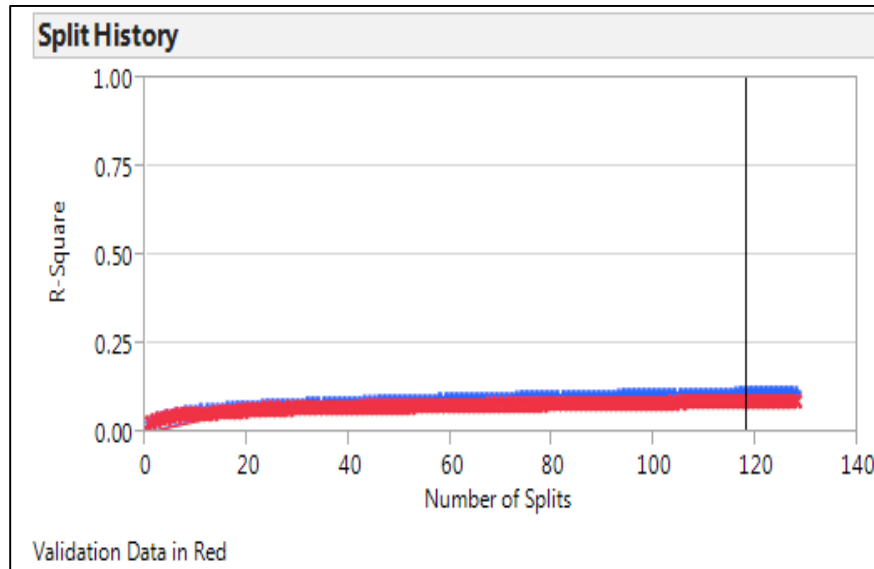


Figure 4.7: Split History Before Pruning for Both Training and Validation Datasets of Crash Severity Model

4.4.2.1 Variables Important to Crash Severity Model

The classification tree modeling process determines independent variables, which are of key importance in the prediction of the dependent variable (Kashani and Mohaymany, 2011). The column contributions report depicted in Figure 4.8 shows that the model selected ten independent variables of greatest influence on the crash severity work zone crashes. The ten independent variables identified under this analysis include collision type, motorcycle related, work zone crash type, posted speed limit, vehicle type, speed related, alcohol related, semi-truck related, youth related, and road condition. However, the most important variables with the highest G^2 statistic values include collision type,

motorcycle related, work zone crash type, posted speed limit, vehicle type and speed related.

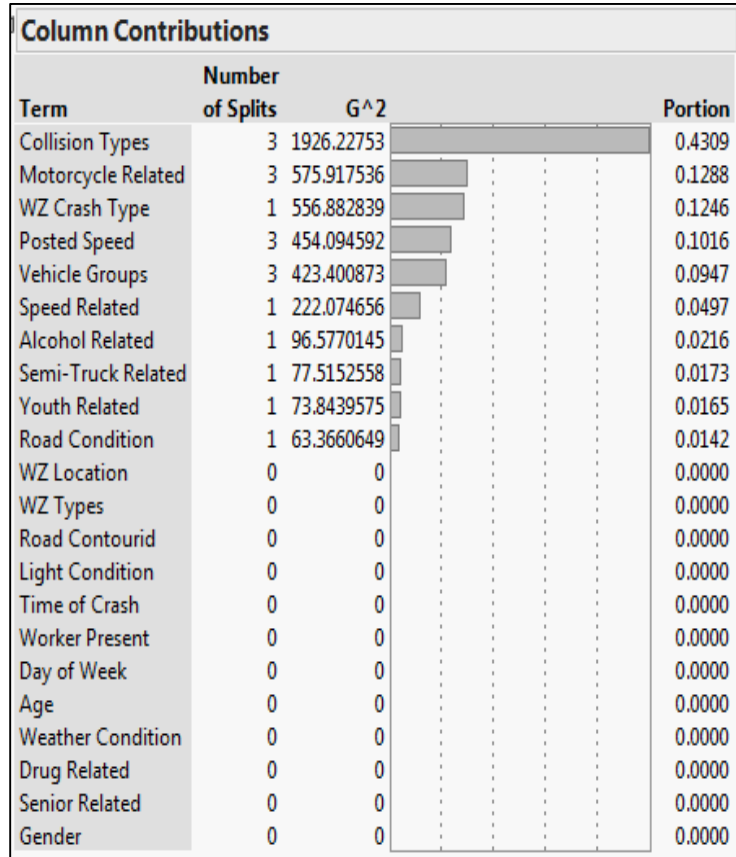


Figure 4.8: Important Variables of Crash Severity Model

4.4.2.2 ROC Curve Results of Crashes Severity

Figures 4.9 and 4.10 show the ROC curve results for training and validation models, respectively. Both curves show that the classification tree model developed in the current study is satisfactory for the levels of crash severity analyzed. Since both curves have high sensitivity and low 1-specificity, all the curves of the levels of crash severity in both plots are above a diagonal line “45-degree line” (i.e., each of them has a rate more than 0.5,

which is a point that specifies predictions based on a random guess). Also, the degree of accuracy of predicting fatal and injury crashes and POD crashes in both training and validation models are identical with about 67% for training dataset and about 66% for validation dataset.

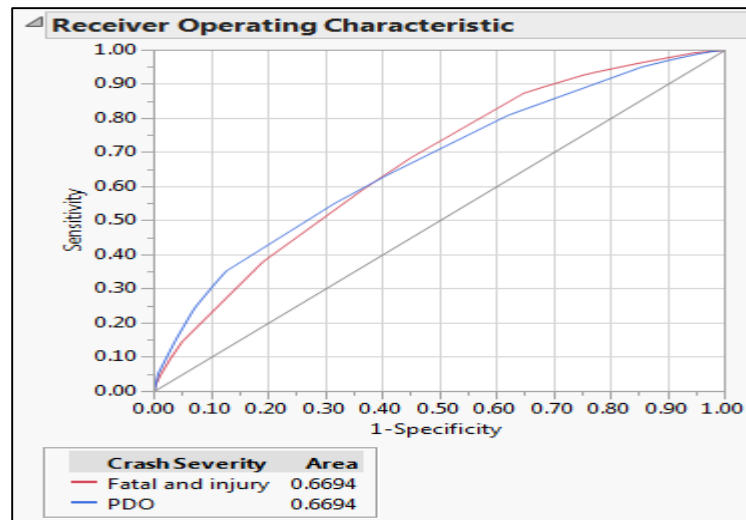


Figure 4.9: ROC Curve for Training Dataset of Crash Severity Model

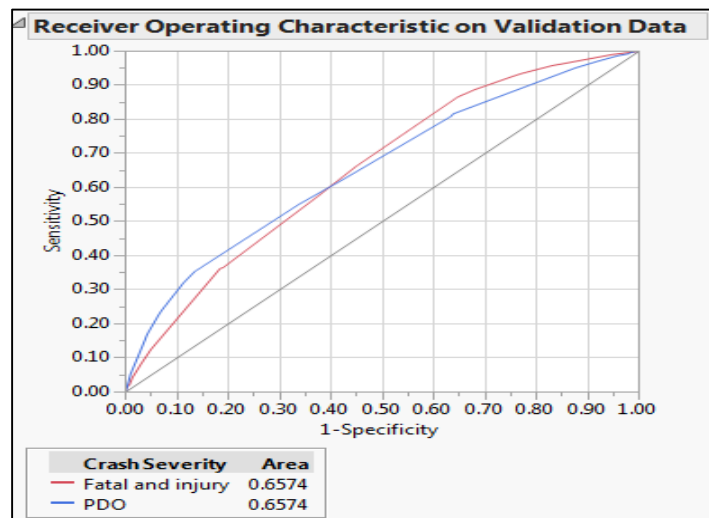


Figure 4.10: ROC Curve for Validation Dataset of Crash Severity Model

Figure 4.11 displays the final classification tree diagram of the crash severity model, which was specified to stop after performing 18 splits because this was assessed to

achieve the ideal tree size. All node boxes contain the two levels of crash severity (fatal/injury and PDO) and their probabilities. Furthermore, each parent node box shows the *LogWorth* value, which is the value for which the data split is based on.

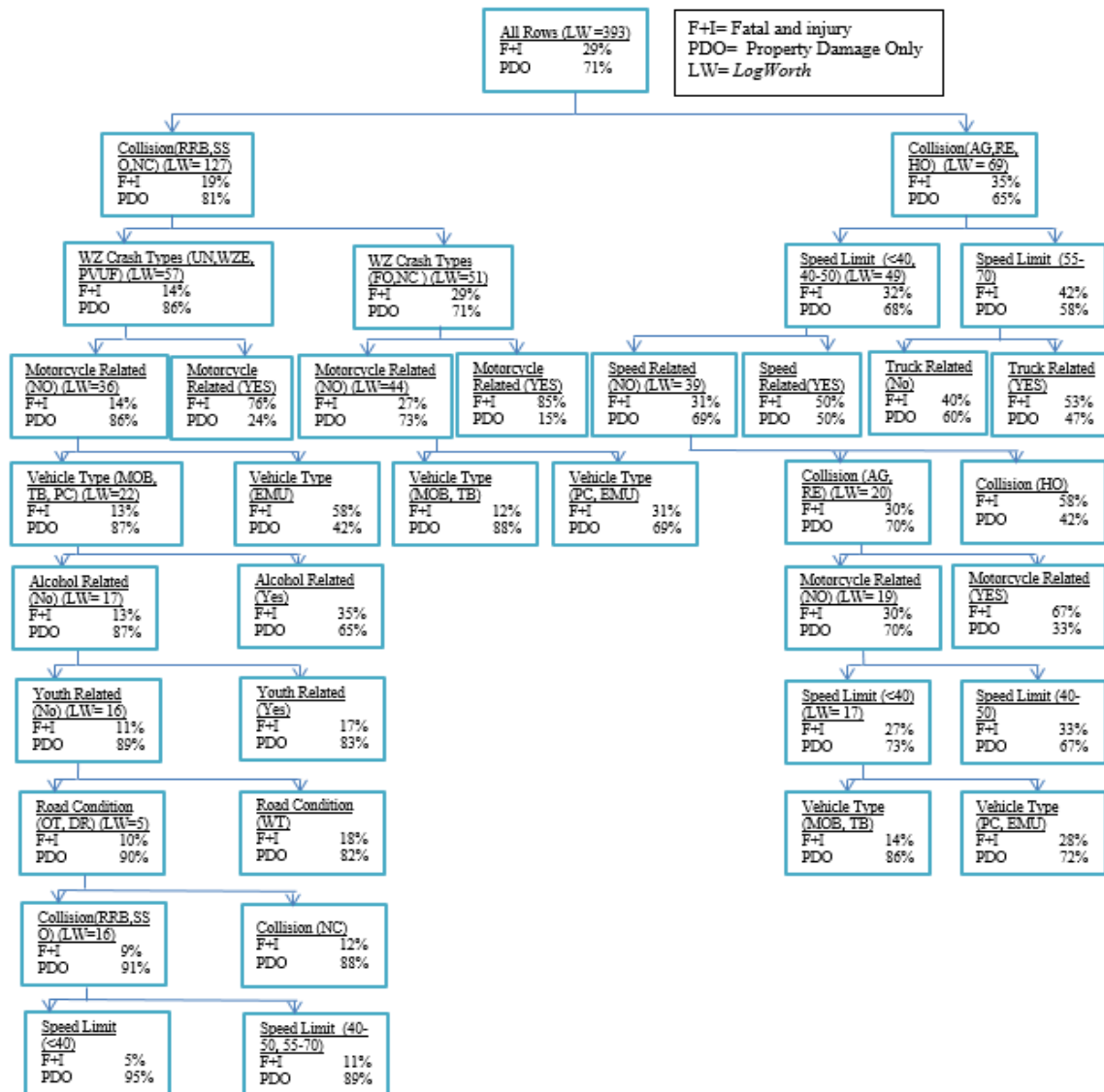


Figure 4.11: Classification Tree Results for Crash Severity Model

4.4.3 Results of Work Zone Location Crashes

The total dataset inputted in the JMP program for work zone location analysis contained 58,368 observations of work zone crashes. These observations were randomly divided into two datasets, the training sample set that contained 52,741 (90%) and a validation sample set that contained 5,627 (10%). Based on the dataset characteristics and the five segments that make up a work zone area, sample data were not equally distributed among the dependent variable's category levels. Activity area (AA) consisted the highest number of the training sample with 30,813 crash records, while the other segments termination area (TA), before the first work zone warning sign area (BWS), advance warning area (AWA) and transition area (TSA) contained 2556, 2769, 5474, 11129 crash records, respectively. Based on split history chart (R^2 vs. number of splits) as shown in Figure 4.12, the full tree made 66 splits. However, split history chart shows training and validation datasets the R^2 value remains the same after 10 splits. Consequently, by using the pruning feature, this model was pruned after 10 splits in order to achieve an ideal tree size and a better model.

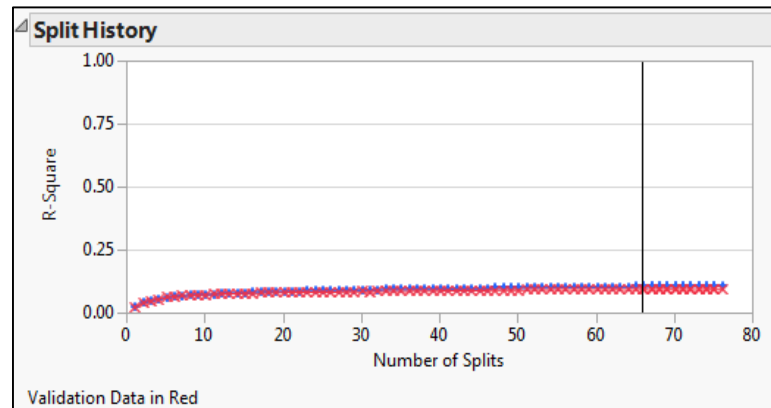


Figure 4.12: Split History Before Pruning for Training and Validation Datasets of WZ Location Crashes Model

4.4.3.1 Variables Important to Work Zone Location Crashes Model

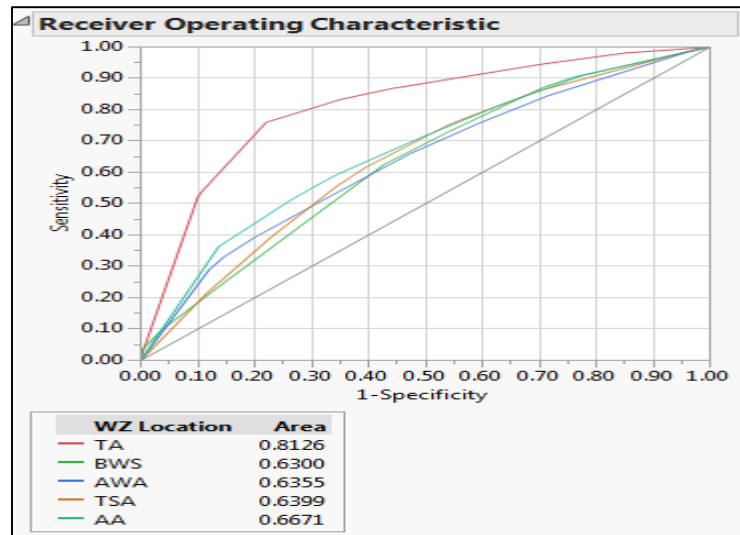
The column contributions report depicted in Figure 4.13 shows the six significant variables (factors), which have a large amount of variation in the response variable in terms of the statistic values identified by the classification tree model. Three out of these six predictor variables were also significant in the crashes severity model including collision type, posted speed limit and vehicle type. The other three significant variables, which were not significant in the crash severity model, are work zone crash type, workers present and age of driver.

Column Contributions			
Term	Number of Splits	G ²	Portion
WZ Types	2	4867.06208	0.5426
Worker Present	2	1698.08518	0.1893
Posted Speed	2	1314.43584	0.1465
Collision Types	2	686.037849	0.0765
Age	1	219.566131	0.0245
Vehicle Groups	1	184.541309	0.0206
Weather Condition	0	0	0.0000
WZ Crash Type	0	0	0.0000
Alcohol Related	0	0	0.0000
Light Condition	0	0	0.0000
Crash Severity	0	0	0.0000
Road Condition	0	0	0.0000
Time of Crash	0	0	0.0000
Youth Related	0	0	0.0000
Road Contourid	0	0	0.0000
Semi-Truck Related	0	0	0.0000
Senior Related	0	0	0.0000
Day of Week	0	0	0.0000
Motorcycle Related	0	0	0.0000
Speed Related	0	0	0.0000
Drug Related	0	0	0.0000
Gender	0	0	0.0000

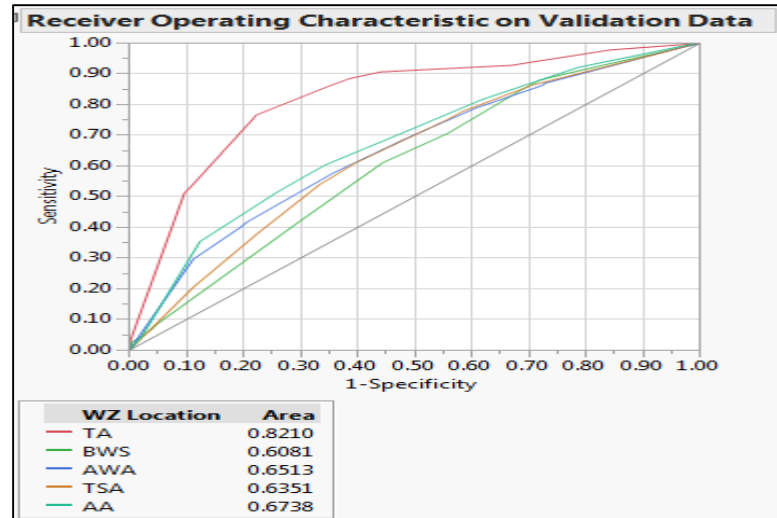
Figure 4.13: Important Variables of WZ Location Crashes Model

4.4.3.2 ROC Curve Results of Work Zone Location Crashes Model

Figures 4. 14 and 4.15 show the ROC curve results for training and validation datasets of the work-zone location crash model, respectively. Both curves show that the modelled classification tree is satisfactory for all levels of work zone locations considered. All curves of the levels of work zone location crashes in both plots are above a diagonal line “45-degree line”. The best degree of accuracy was obtained for predicting the expected number of crashes in the termination area (TA) in both training and validation samples with about 81% for training dataset and about 82% for validation dataset.



Figures 4.14: ROC Curve for Training Dataset of WZ Location Crashes Model



Figures 4.15: ROC Curve for Validation Dataset of WZ Location Crashes Model

The final classification tree diagram after 10 splits of the work zone location crashes is shown in Figure 4.16. All node boxes contain five levels of the work zone location crashes, i.e., termination area (TA), before the first work zone warning sign area (BWS), advance warning area (AWA), transition area (TSA) and activity area (AA) with their probabilities. The *LogWorth* value in each parent node box was also shown.

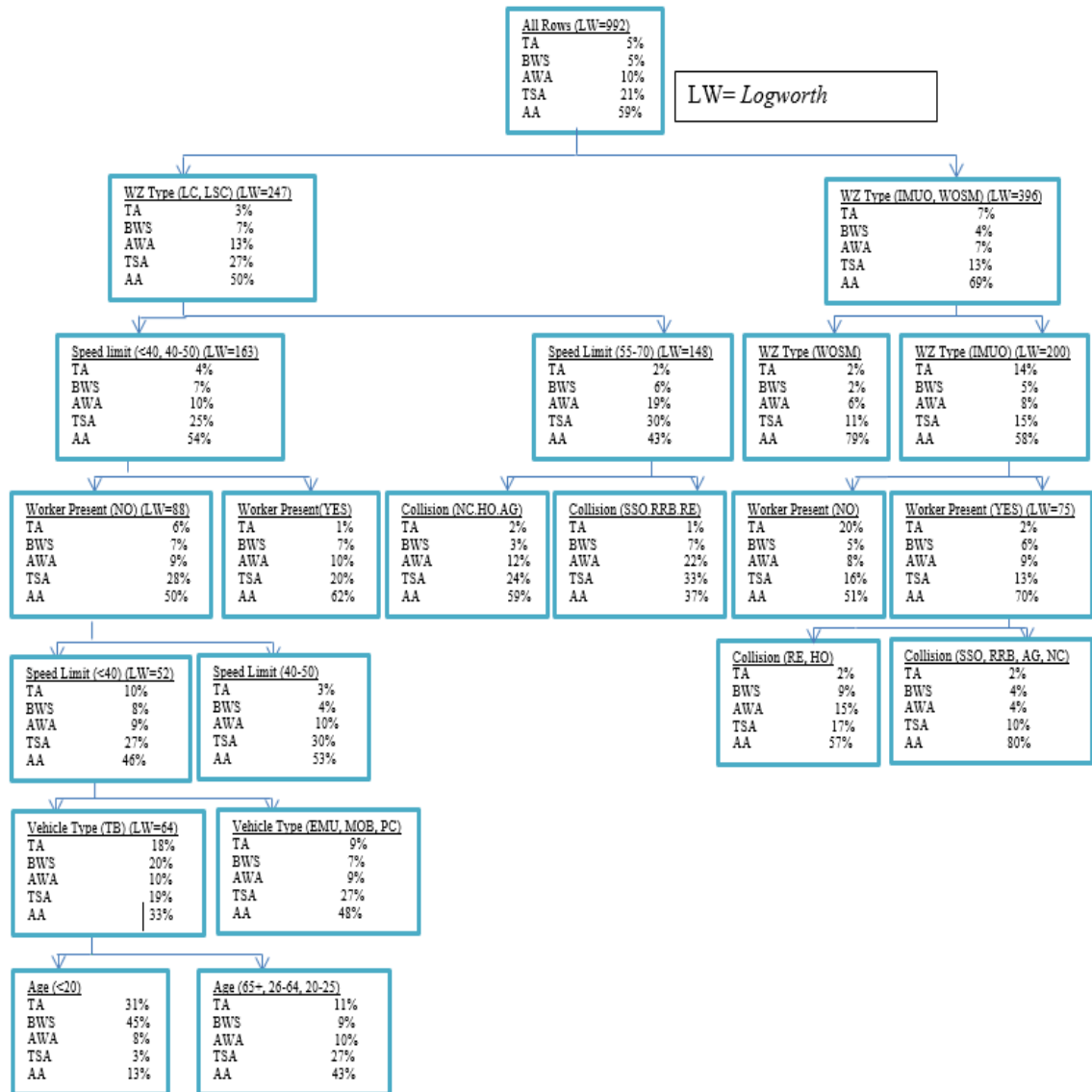


Figure 4.16: Classification Tree Result for WZ Location Crashes Model

4.4.4 Discussion of Results

The classification tree model identified ten independent variables that are important to the crash severity model and six important independent variables for the work zone location crashes model. There are three independent variables, which are important to both work zone location and crash severity models and these include collision type, posted speed

limit and vehicle type. Collision type was the most significant variable for crash severity model and the fourth significant variable for work zone location crashes model. It is a variable that is always highly significant in many similar studies (e.g., Mohan and Gautam, 2002; Garber and Zhao, 2002; Nemeth and Migletz, 1978; Hall and Lorenz, 1989; Pigman and Agent, 1990; Garber and Woo, 1990; Daniel et al, 2000; Salem et al, 2006). Whereas, work zone type was the most significant variable for work zone location model, but it was not significant in the crash severity model. Motorcycle related was the second most significant variable in the crash severity model, but it was not significant in the work zone location crashes model. Variables identified in the current study as contributing factors in increasing crash severity of work zone-related crashes are in agreement with other previous studies (e.g., FARS, 2012; Horswill et al., 2003; and Wong et al., 2010). Meanwhile, worker present variable was the second most significant variable in the work zone location crashes model but it was not significant in the crash severity model. Furthermore, work zone crash type was only significant for the crash severity model. Posted speed limit was a significant variable for both crash severity and work zone location crashes models, also this finding agrees with some previous studies (e.g., Graham et al., 1978, Garber and Gadiraju, 1988; Nemeth and Rathi 1983). Vehicle type was significant for crash severity model but it was a relative weakly significant variable for work zone location crashes model. In addition, semi-truck related was a relative weakly significant variable for crash severity model and this finding agrees with findings from other previous studies (e.g., Pigman and Agent, 1990; Schrock et al., 2004; Khattak and Targa, 2004), but it was not for work zone location crashes model.

Figure 4.17 shows the first split on the classification tree which splits the crash severity into two child nodes based on collision type variable. The two resulting child nodes are shown as follows: rear-end/head on/ angle, and no collision between two vehicles in transport/unknown/sideswipe same direction/sideswipe opposite direction/rear to rear/backing. In addition, Figure 4.17 shows if the collision type was rear-end/head on/angle, then the conditional probability of fatal and injury is 35 percent. Whereas, if the collision type was no collision between two vehicles in transport/unknown/ sideswipe same direction/sideswipe opposite direction/rear to rear/backing collision, the conditional probability of fatal and injury is only 19 percent. Therefore, crashes that involved collision type rear-end/ head on/ angle was relatively more hazardous than those that involved other collision type.

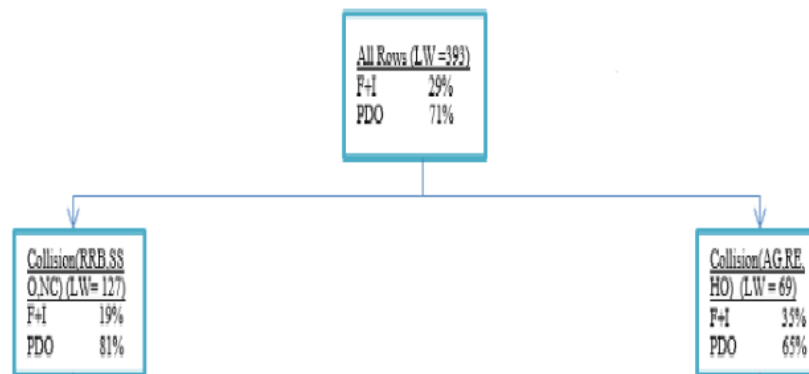


Figure 4.17: The First Split of Crashes Severity Model

When the collision types were no collision between two vehicles in transport/unknown/ sideswipe same direction/sideswipe opposite direction/rear to rear/backing, the tree continues to grow based on work zone crash types. Figure 4.18 shows that 29 percent of collision with fixed object/non-collision that involved collision types no collision

between two vehicles in transport/unknown/ sideswipe same direction/sideswipe opposite direction/rear to rear/backing were fatal and injury crashes. On the other hand, 14 percent of work zone equipment/collision with person/vehicle/object not fixed/unknown work zone crashes that involved in same collision types resulted into fatal and injury crashes.

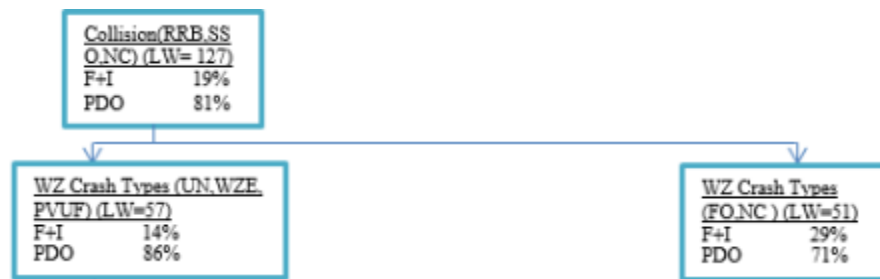


Figure 4.18: Collision Types Splitting Nodes in the Crash Severity Model

The classification tree shows that work zone crashes types (collision with fixed object/non-collision) node divides based on motorcycle related variable. Figure 4.19 shows that 85 percent of collision with fixed object/non-collision work zone crashes that involved collision types no collision between two vehicles in transport/unknown/ sideswipe same direction/sideswipe opposite direction/rear to rear/backing for which motorcycle was involved were fatal and injury crashes. Whereas, for the same types of work zone crashes and collisions for which motorcycle was not involved, only 27 percent of the crashes were fatal and injury crashes.

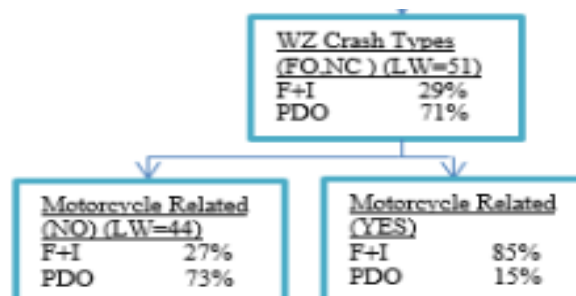


Figure 4.19: Work Zone Crashes Types Splitting Nodes in the Crash Severity Model

Illustrate 4.20 shows that motorcycle not related variable divides based on vehicle type variable. The model results show that 31 percent of crashes which occurred when vehicle types were emergency/unknown/passenger vehicles that involved collision types no collision between two vehicles in transport/unknown/ sideswipe same direction/sideswipe opposite direction/rear to rear/backing for which motorcycles not related were fatal and injury crashes. However, only 12 percent of crashes which occurred when vehicle types were trucks/buses/motorcycles/motorized bicycles that involved same collision types for which motorcycles not related were fatal and injury crashes.

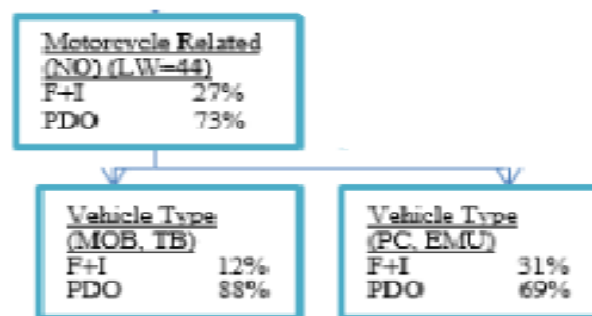


Figure 4.20. Motorcycle Not Related Splitting Nodes in the Crash Severity Model

Figure 4.21 shows that rear-end, angle or head on collision type node divides into two child nodes based on posted speed limit. Furthermore, Figure 4.21 shows that 42 percent of work zone crashes that involved collision types rear-end, angle or head on that occurred when the speed limit was more than or equal 55 mph were fatal and injury crashes. On the other hand, 32 percent of work zone crash that involved the same

collision types that occurred when the speed limit was 50 mph or less were fatal and injury crashes.

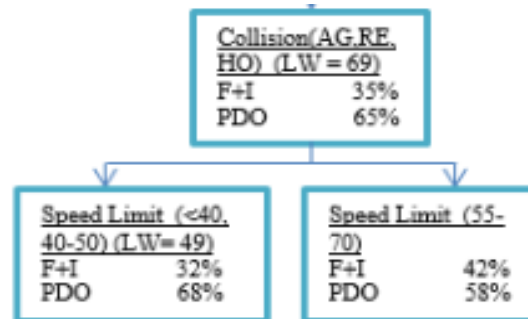


Figure 4.21. Collision Types (Rear-End, Angle or Head on) Splitting Nodes in the Crash Severity Model

Figure 4.22 shows the relationship between work zone location and work zone types for the work zone location crashes model. Figure 4.22 shows the first split on the classification tree which splits the work zone location crashes into two child nodes based on work zone type variable. It shows that 69 percent of work zone location crashes occurred on the activity area (AA) when work zone types were work on shoulder or median/intermittent or moving work/unknown/other, and 13 percent of the crashes occurred on the transition area (TSA). While, 50% of work zone location crashes occurred on the activity area (AA) when work zone were lane closure/lane shift/crossover, and 27 percent of the crashes occurred on the transition area (TSA).

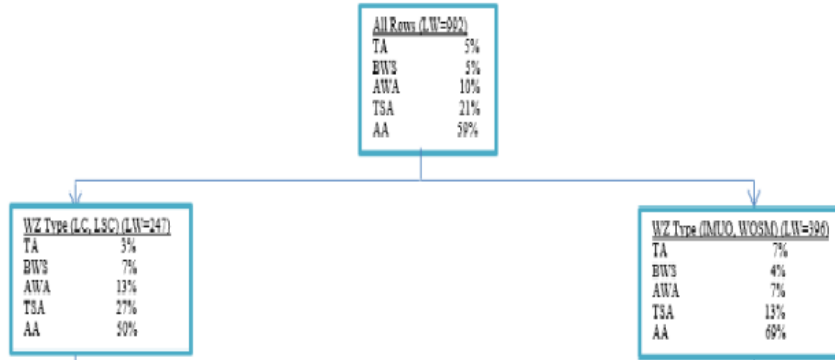


Figure 4.22. The First Split of Work Zone Location Crashes Model

Figure 4.23 shows that the work zone types (lane closure/lane shift/crossover) node splits into two nodes based on posted speed limit variable. Furthermore, it shows that 43% of work zone location crashes occurred on the activity area (AA) when work zone types were lane closure/lane shift/crossover that occurred on roadways with posted speed limits between 55-70 mph, and 30 percent of the crashes occurred on the transition area (TSA). However, 54 percent of work zone location crashes occurred on the activity area (AA) on same work zone types that occurred in roadways with posted speed limit 50 mph or less, and 25 percent of the crashes occurred on the transition area (TSA).

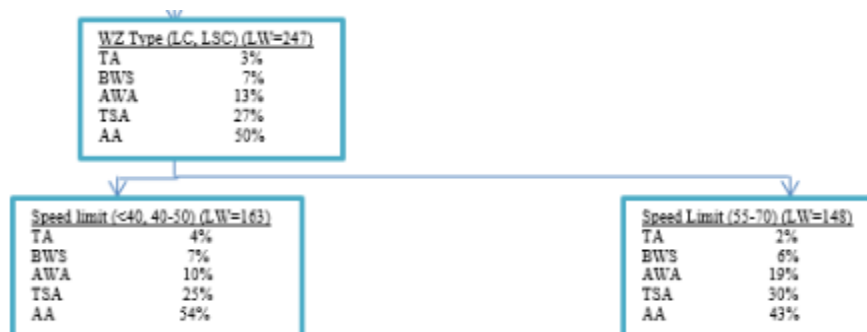


Figure 4.23. Work Zone Types Splitting Nodes in the Work Zone Location Crashes Model

Figure 4.24 shows that 37 percent of work zone location crashes occurred on the activity area (AA) when work zone types were lane closure/lane shift/crossover that occurred on roadways with posted speed limit between 55-70 mph which involved collision types rear-end/rear to rear/backing/sideswipe same direction/sideswipe opposite direction, 33 percent of the crashes occurred on the transition area and 22 percent of the crashes occurred on the advance warning area. Whereas, 59 percent of work zone location crashes occurred on the activity area (AA) that occurred on the same work zone types and same posted speed limit that involved collision types head on/angle/no collision between two vehicles in transport/unknown, 24 percent of the crashes occurred on the transition area and 12 percent of the crashes occurred on the advance warning area.

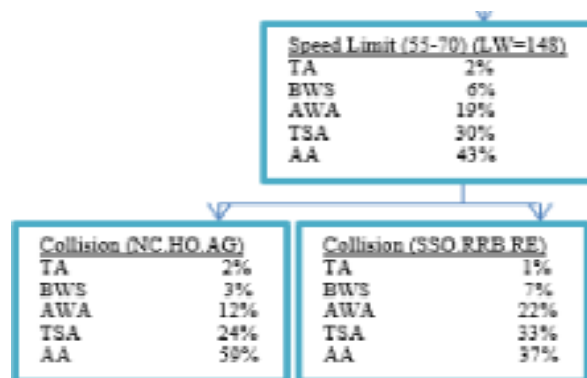


Figure 4.24: Speed Limit (55-70 mph) Splitting Nodes in the Work Zone Location Crashes Model

When posted speed limit was 50 mph or less, the tree continues to grow based on worker present related variable. Figure 4.25 shows that 62 percent of work zone location crashes occurred on the activity area when worker was present on work zone types were lane closure/lane shift/crossover that occurred in roadways with posted speed limit less than or equal 50 mph, and 20 percent of the crashes happened on the transition area. While, 50 percent of work zone location crashes occurred on the activity area when worker was not present in work zone types lane closure/lane shift/crossover that occurred in roadways with posted speed limit less than or equal 50 mph, and 28 percent of the crashes occurred on the transition area.

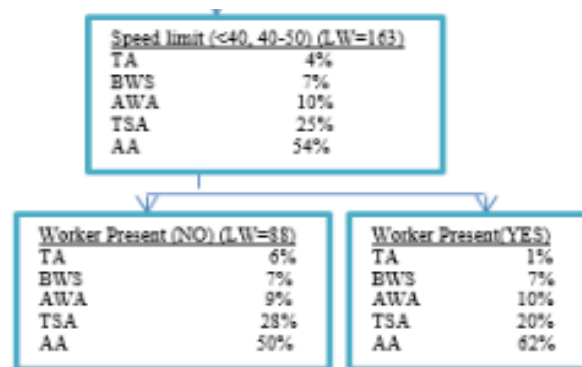


Figure 4.25: Speed Limit (Less Than or Equal 50 mph) Splitting Nodes in the Work Zone Location Crashes Model

Figure 4.26 shows that trucks and buses vehicle types node divides based on age of person. Furthermore, Figure 4.26 shows the most interesting result is that 45 percent of the work zone location crashes occurred on the before the first work zone warning sign (BWS) when work zone types were lane closure/lane shift/crossover that occurred on roadways with posted speed limit 40 mph or less which involved vehicle types trucks and buses and the age of person was less than 20 years, and 13 percent of work zone location

crashes occurred on the activity area (AA). However, when age of the person more than or equal 20 years, it shows that 9 percent of work zone location crashes occurred on the before the first work zone warning sign (BWS) and 43 percent of the crashes occurred on the activity area (AA).

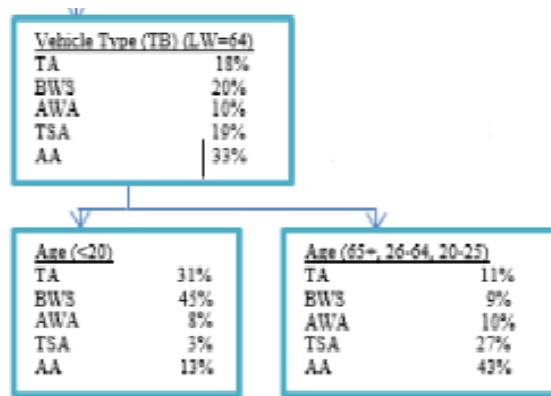


Figure 4.26. Vehicle Types Splitting Nodes in the Work Zone Location Crashes Model

CHAPTER FIVE

CONCLUSIONS AND RECOMMENDATIONS

A classification tree method was used to investigate the significant factors for crash severity and work zone location crashes models. Classification tree modelling procedure identified variables, which are important to the model and which have a large degree of variation in the response variable in terms of the G^2 statistic values. In terms of crash severity analysis, the model identified ten independent variables including collision type, motorcycles related, work zone crashes type, posted speed limit, vehicle type, speed related, alcohol related, semi-truck related, youth related and road condition. On the other, for the work-zone location crash analysis, the model identified only six significant independent variables including work zone type, workers present, posted speed limit, collision type, age of driver, and vehicle type. Three out of these six predictor variables were also significant in the crash severity model, which includes collision type, posted speed limit and vehicle type. Furthermore, the most significant variables for both crash severity and work zone location crashes models are collision types, work zone types, motorcycle related, workers present, work zone crashes types, posted speed limit and vehicle types.

Collision type was the most significant variable for the crash severity model and the most severe types of collision are angle and head-on when occurring on road with posted speed limits less than or equal 50 mph. Likewise, the proportion of work-zone crash in the activity area, transition areas and advance warning areas is found to be on the higher side as compared to other work zone locations. These types of crashes were mainly rear-end/rear to rear/backing/sideswipe same direction/ sideswipe opposite direction on high posted speed limit roads. Furthermore, the high-posted speed limit was a contributed factor that is likely to cause fatal and injury crashes in work zones and the most work zone crashes occurred in the activity area and transition areas.

The motorcycle involvement is likely to cause fatal and injury crashes in work zones. In addition, when emergency and passenger vehicles were involved in work zone crashes, the conditional probability of fatal and injury crashes increased comparing with other vehicles types.

The work-zone location results reveal that most work zone location crashes are more likely to occur in the activity area. Likewise, the second most likely location for a work zone crash was the transition area. Specifically, transition area crashes increase when work zone types are lane closure, lane shift, or crossover. In the case of posted speed limit, lower than or equal to 50 mph and no workers present in the work zone, the probability of work zone location crashes increase in the transition area. An interesting result was that if the driver is younger than 20 years old driving on a road with posted speed limit less than 40 mph, there was an elevated likelihood of crashes to occur before the first work zone warning sign was visible.

The results of crash severity model agree with some of the previous studies such as collision type (e.g., Mohan and Gautam, 2002; Garber and Zhao, 2002; Nemeth and Migletz, 1978; Hall and Lorenz, 1989; Pigman and Agent, 1990; Garber and Woo, 1990 and Daniel et al, 2000. Motorcycles related (e.g., FARS, 2012; Horswill et al., 2003; and Wong et al., 2010); posted speed limit (e.g., Paulsen, Glennon and Graham, 1978; Garber and Gadiraju, 1988); semi-truck related (e.g., Pigman , Agent, 1990; Schrock et al, 2004; Khattak, and Targa, 2004). Although crash time was not significant in this study, it was significant in some previous studies (e.g., Garber and Zhao, 2002; AASHTO, 1987; Pigman and Agent, 1990; Nemeth and Migletz, 1978; Chembless et al, 2002; Hill, 2003; Li and Bai, 2006).

The results of work zone location crashes analysis also agree with some of previous studies where activity area was the most prevalent crash location in a work zone (e.g., Garber and Zhao, 2002; Pigman and Agent, 1990; Nemeth and Rathi, 1983; Hargroves, 1981; Nemeth and Migletz, 1978). Additionally, Nemeth and Rathi (1983) study agrees that high posted speed limits contribute to increased likelihood of crashes in transition area. Another previous study conducted in Ohio agrees with the current study that rear-end crashes have higher probabilities of occurring in the advance warning area (Salem et al., 2006).

The results of the current study are helpful in identifying factors, which are significant factors for crashes in the work zones. Since most crashes occur in the activity area of the work zones, additional warning signs encouraging drivers to merge early before entering the activity area are recommended. Furthermore, for roads with posted speed limits higher than 50 mph, it is recommended to achieve work zone safety by

improving the implementation of additional countermeasures such as rumble strips, which cause vibrations in the vehicle to alert drivers to reduce vehicle speeds before entry into work zone areas. It is noteworthy to mention that Ohio has the fifth largest number of registered motorcycles in the United States (FHWA, 2011) and the majority of fatal and injury work zone crashes occurred when a motorcycle is involved. Therefore, it is suggested to increase enforcement on major roadways and encourage motorcyclists to be more careful when driving through work zones.

REFERENCES

- AASHTO. "Summary Report on Work Zone Crashes." Standing Committee on Highway Traffic Safety, American Association of State Highway and Transportation Officials, Washington, D.C., 1987.
- Agresti, A. "An Introduction to Categorical Data Analysis." Second Edition, John Wiley and Sons, Inc., New York, 2007.
- Arnold, E. D. "Use of Police in Work Zones on Highways in Virginia." Final Report VTRC 04-R9, Virginia Transportation Research Council, Charlottesville, Virginia, 2003.
- Benekohal, R. F., E. Shim, and P. T. V. Resende. "Truck Drivers' Concerns in Work Zones: Travel Characteristics and Accident Experiences." In *Transportation Research Record 1509: Journal of the Transportation Research Board*, 1995, pp. 55–64.
- Brewer, M. A., Pesti, G., and Schneider, W. "Improving Compliance with Work Zone Speed Limits: Effectiveness of Selected Devices." In *Transportation Research Record: Journal of the Transportation Research Board*, 2006, pp. 67–76.

Chambless, J., Chadiali, A. M., Lindly, J. K, and McFadden, J. “Multistate Work Zone Crash Characteristics” In *ITE Journal*, 2002, pp. 46–50.

Daniel J., K. Dixon, and D. Jared. “Analysis of Fatal Crashes in Georgia Work Zones.” In *Transportation Research Record: Journal of the Transportation Research Board*, 2000, pp. 18-23.

Dixon, K. K. and Wang, C. “Development of Speed Reduction Strategies for Highway Work Zones.” Georgia Transportation Institute, Georgia Institute of Technology, Atlanta, Georgia, 2002.

Fatality Analysis Reporting System (FARS). “Analytic Reference Guide 1982 to 2006” NHTSA Report DOT HS 810 937, National Highway Traffic Safety Administration, U.S. Department of Transportation, 2006.

Federal Motor Carrier Safety Administration. “2000 Work Zone Traffic Crash Facts.” U.S. Department of Transportation, Washington, D.C., 2002.

<http://ai.volpe.dot.gov/CarrierResearchResults/HTML/WORKZONE2000finalve>.

Accessed on May 20, 2014.

Federal Highway Administration (FHWA). “Highway Statistics 2001.” Report No. FHWA-PL-02-020, U.S. DOT, Washington D.C, 2001.

FHWA. “Manual on Uniform Traffic Control Devices (MUTCD) for Streets and Highway 2003 Ed.” U.S. Department of Transportation, Washington, D.C, 2003.

FHWA (2006). “Work Zone Safety Fact Sheet.”

http://www.pavements4life.com/qds/Safety_3FewerWorkzones.asp . Accessed on

May 11, 2014

- FHWA. "Work Zone Safety and Mobility Fact Sheet." FHWH Safety, 2009.
http://safety.fhwa.dot.gov/wz/wz_awareness/2009/factsht09.cfm . Accessed on May 18, 2014.
- FHWA. "State Motor-Vehicle Registrations – 2010"
<https://www.fhwa.dot.gov/policyinformation/statistics/2010/mv1.cfm> , 2011.
Accessed on April 28, 2014.
- Garber, N. J. and Woo, T. H. "Accident Characteristics at Construction and Maintenance Zones in Urban Areas." Report No. VTRC 90-R12, Transportation Research Council, Charlottesville, Virginia, 1990.
- Garber, N. J. and Ming Zhao. "Crash Characteristics at Work Zones." Virginia Transportation Research Council, Charlottesville, Virginia, 2002.
- Garber, N. and Gadiraju, R. "Speed Variance and Its Influence on Accidents." Foundation for Traffic Safety, Washington, DC. , 1988.
- Garber, N. J. and Srinivasan, S. "Effectiveness of Changeable Message Signs in Work Zones: Phase II." Report No. VTRC 98-R 10, Virginia Transportation Research Council, Charlottesville, Virginia, 1998.
- Garber, N. J. and Patel, S. T. "Effectiveness of Changeable Message Signs in Controlling Vehicles Speeds in Work Zones." Report No. VTRC 95-R4, Virginia Transportation Research Council, Charlottesville, Virginia, 1994.
- Graham, J., Paulsen, R., and Glennon, J. "Accident and Speed Studies in Construction Zones." Report FHWA-RD-77-80, FHWA, U.S. Department of Transportation, 1977.

- Ha, T. J. and Nemeth Z. A. "Detailed Study of Accident Experience in Construction and Maintenance Zones." In *Transportation Research Record: Journal of the Transportation Research Board*, No. 1509, 1995, pp. 38-45.
- Hall, J. W. and V. M. Lorenz. "Characteristics of Construction-Zone Accidents." In *Transportation Research Record: Journal of the Transportation Research Board*, No. 1230, 1989, pp. 20-27.
- Hargroves, B. T. "Vehicle Crashes in Highway Work Zones". In *Journal of Transportation Engineering*, Vol. 107, 1981, pp. 525 – 539.
- Hill, R. W. "Statistical Analysis of Fatal Traffic Accident Data." Master's Thesis, Texas Tech University, 2003.
- Horswill, M. S. & Helman, S. "A Behavioral Comparison between Motorcyclists and a Matched Group of Non-motorcycling Car Drivers: Factors Influencing Accident Risk." In *Accident Analysis and Prevention*, 2003, pp. 589- 597.
- Huebschman, C. R., Garcia, C., Bullock, D. M., and Abraham, D. M. "Construction Work Zone Safety." Report No. FHWA/IN/JTRP-2002/34, Transportation Research Program, Purdue University, West Lafayette, Indiana, 2003.
- Kashai, A., Mohaymany, A., Ranjbari, A. "A Data Mining Approach to Identify Key Factors of Traffic Injury Severity." *Promet-Traffic and Transportation*, 2011.
- Khattak, A., Targa, F. "Injury severity and total harm in truck-involved work zone crashes." In *Proceedings of 83rd Annual Meeting of the Transportation Research Board*, Washington, DC, January 11-15, 2004.
- Li, Y. and Bai, Y. "Investigating the Characteristics of Fatal Crashes in the Highway Construction Zones" In *proceedings of CIB W99 International Conference on*

- Global Unity for Safety & Health in Construction*, Beijing China, 2006, pp. 301 – 309.
- Li, Y. and Bai, Yong. “Effectiveness of Temporary Traffic Control Measures in Highway Work Zones.” In *Proceedings of the 87th TRB Annual Meeting*, Washington D.C., January 13-17, 2008.
- Mohan, S. B. and Gautam, P. “Cost of Highway Work Zone Injuries” *Practical Periodical on Structural Design and Construction*, ASCE, 2002, pp. 68 – 73.
- Myles, A.J., Feudale, R.N., Liu, Y., Woody, N.A., and Brown, S.D. “An Introduction to Decision Tree Modeling.” In *Journal of Chemometrics*, Vol. 18, 2004, pp. 275-285.
- National Highway Traffic Safety Administration. “The National Work Zone Safety Information” Published: 2012. <http://www.workzonesafety.org/> . Accessed on July 2, 2014.
- Nemeth Z. A., and D. J. Migletz. “Accident Characteristics Before, During, and After Safety Upgrading projects on Ohio’s Rural Interstate System.” In *Transportation Research Record 672: Journal of the Transportation Research Board*, 1978, pp. 19-23.
- Nemeth, Z. A., and Rathi, A. “Freeway Work Zone Accident Characteristics.” *Transportation Quarterly*, 371, 1983, pp. 145–159.
- Ohio Department of Transportation (ODOT). “Work Zone Safety Awareness Week” <http://www.dot.state.oh.us/Services/Pages/ODOT-2013-Work-Zone-Safety-Fact-Sheet.aspx> . Accessed on June 26, 2014.

- Pain, R. F., McGee, H. W., and Knapp, B. G. "Evaluation of Traffic Controls for Highway Work Zones." NCHRP Report 236, Transportation Research Board, Washington D. C, 1983.
- Pal, R. and Sinha, K.C. "Analysis of Crash Rates at Interstate Work Zones in Indiana." In *Transportation Research Record 1529: Journal of the Transportation Research Board*, 1996, pp. 43-53.
- Pigman, J. G. and Agent, K. R. "Highway Crashes in Construction and Maintenance Work Zones". In *Transportation Research Record 1270: Journal of the Transportation Research Board*, 1990, pp. 12–21.
- Raub, R.A., Sawaya, O.B., Schofer, J.L. and Ziliaskopoulos, A., "Enhanced Crash Reporting to Explore Work Zone Crash Patterns." Transportation Research Board, Washington D.C., 2001.
- Richards. S. H. and Dudek, C. L. "Implementation of Work-Zone Speed Control Measures." In *Transportation Research Record 1086: Journal of the Transportation Research Board*, 1986, pp. 36-42.
- Roupail, N. M., Zhao, S. Y., and Fazio, J. "Comparative Study of Short- and Long-Term Urban Freeway Work Zone." In *Transportation Research Record: Journal of the Transportation Research Board*, No. 1163, 1988, pp. 4-14.
- Salem, O. M., Genaidy, A.M, Wei, H. and Deshpande, H. "Spatial Distribution and Characteristics of Accident Crashes at Work Zones of Interstate Freeways in Ohio." In *Intelligent Transportation Systems Conference, 2006*. pp. 1642-1647.
- SAS Institute Inc. "SAS Help and Documentation," Version 9.1, Cary, NC, 2003.

- Schrock, D. S., Ullman, G. L., Cothron, A. S., Kraus, E., and Voigt, A. P. "An Analysis of Fatal Work Zone Crashes in Texas". Report FHWA/TX-05/0- 4028-1, FHWA, U.S. Department of Transportation, 2004.
- Ullman, G. L. and Krammes, R. A. "Analysis of Crashes at Long-Term Construction Projects in Texas". Report FHWA/TX-90/1108-2, FHWA, U.S. Department of Transportation, 1990.
- Wang, J., Hughes, W. E., Council, F.M., & Paniati, J.E. "Investigation of Highway Work Zone Crashes: What We Know and What We Don't Know" In *Transportation Research Record 1529: Journal of the Transportation Research Board*, 1996, pp. 54-64.
- Wong, J. T., Chung, Y. S., & Huang S. H. "Determinants Behind Young Motorcyclists' Risky Riding Behavior." In *Accident Analysis and Prevention*, 2010, pp. 275-281.