Safety Assessment Tool for Construction Zone Work Phasing Plans

Final Report
May 2016

SWZDI
Smart Work Zone Deployment Initiative

Sponsored by
Smart Work Zone Deployment Initiative
Federal Highway Administration
(TPF-5(535))

About SWZDI

Iowa, Kansas, Missouri, and Nebraska created the Midwest States Smart Work Zone Deployment Initiative (SWZDI) in 1999 and Wisconsin joined in 2001. Through this pooled-fund study, researchers investigate better ways of controlling traffic through work zones. Their goal is to improve the safety and efficiency of traffic operations and highway work.

ISU Non-Discrimination Statement

Notice

Quality Assurance Statement

Iowa DOT Statements

The preparation of this report was financed in part through funds provided by the Iowa Department of Transportation through its "Second Revised Agreement for the Management of Research Conducted by Iowa State University for the Iowa Department of Transportation" and its amendments.

Technical Report Documentation Page

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
InTrans Project 15-535		

4. Title and Subtitle	5. Report Date
Safety Assessment Tool for Construction Zone Work Phasing Plans	May 2016
	6. Performing Organization Code
7. Author(s)	**8. Performing Organization Report No.**
Henry Brown, Carlos Sun, Praveen Edara, and Roozbeh Rahmani	
9. Performing Organization Name and Address	**10. Work Unit No. (TRAIS)**
University of Missouri-Columbia Department of Civil & Environmental Engineering E 2509 Lafferre Hall Columbia, MO 65211	**11. Contract or Grant No.**
12. Sponsoring Organization Name and Address	**13. Type of Report and Period Covered**
Midwest Smart Work Zone Deployment Initiative Iowa Department of Transportation 800 Lincoln Way Ames, Iowa 50010 Federal Highway Administration U.S. Department of Transportation 1200 New Jersey Avenue SE Washington, DC 20590	Final Report
	14. Sponsoring Agency Code
	TPF-5(295)

15. Supplementary Notes

Visit www.intrans.iastate.edu/smartwz/ for color pdfs of this and other Smart Work Zone Deployment Initiative research reports and to download the associated Work Zone Safety Assessment Tool from this project's webpage.

16. Abstract

The Highway Safety Manual (HSM) is the compilation of national safety research that provides quantitative methods for analyzing highway safety. The HSM presents crash modification functions related to freeway work zone characteristics such as work zone duration and length. These crash modification functions were based on freeway work zones with high traffic volumes in California.

When the HSM-referenced model was calibrated for Missouri, the value was 3.78, which is not ideal since it is significantly larger than 1. Therefore, new models were developed in this study using Missouri data to capture geographical, driver behavior, and other factors in the Midwest. Also, new models for expressway and rural two-lane work zones that barely were studied in the literature were developed.

A large sample of 20,837 freeway, 8,993 expressway, and 64,476 rural two-lane work zones in Missouri was analyzed to derive 15 work zone crash prediction models. The most appropriate samples of 1,546 freeway, 1,189 expressway, and 6,095 rural two-lane work zones longer than 0.1 mile and with a duration of greater than 10 days were used to make eight, four, and three models, respectively.

A challenging question for practitioners is always how to use crash prediction models to make the best estimation of work zone crash count. To solve this problem, a user-friendly software tool was developed in a spreadsheet format to predict work zone crashes based on work zone characteristics. This software selects the best model, estimates the work zone crashes by severity, and converts them to monetary values using standard crash estimates.

This study also included a survey of departments of transportation (DOTs), Federal Highway Administration (FHWA) representatives, and contractors to assess the current state of the practice regarding work zone safety. The survey results indicate that many agencies look at work zone safety informally using engineering judgment. Respondents indicated that they would like a tool that could help them to balance work zone safety across projects by looking at crashes and user costs.

17. Key Words	18. Distribution Statement
crash mitigation—highway safety—safety models—work-zone safety	No restrictions.

19. Security Classification (of this report)	20. Security Classification (of this page)	21. No. of Pages	22. Price
Unclassified.	Unclassified.	134	NA

SAFETY ASSESSMENT TOOL FOR CONSTRUCTION ZONE WORK PHASING PLANS

Final Report
May 2016

Principal Investigator
Henry Brown, Research Engineer

Co-Principal Investigators
Carlos Sun, Associate Professor
Praveen Edara, Associate Professor

Research Assistant
Roozbeh Rahmani

Authors
Henry Brown, Carlos Sun, Praveen Edara, and Roozbeh Rahmani
University of Missouri-Columbia
Department of Civil & Environmental Engineering

Sponsored by the Midwest Smart Work Zone Deployment Initiative and
the Federal Highway Administration (FHWA) Pooled Fund Study TPF-5(295):
Iowa (lead state), Kansas, Missouri, Nebraska, Wisconsin

Preparation of this report was financed in part
through funds provided by the Iowa Department of Transportation
through its Research Management Agreement
with the Institute for Transportation
(InTrans Project 15-535)

A report from
Institute for Transportation
Iowa State University
2711 South Loop Drive, Suite 4700
Ames, IA 50010-8664
Phone: 515-294-8103 / Fax: 515-294-0467
www.intrans.iastate.edu

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

This research was conducted under the Midwest Smart Work Zone Deployment Initiative (SWZDI) and Federal Highway Administration (FHWA) Pooled Fund Study TPF-5(295), involving the following state departments of transportation:

- Iowa (lead state)
- Kansas
- Missouri
- Nebraska
- Wisconsin

The authors would like to thank the FHWA, the Iowa Department of Transportation (DOT), and the other pooled fund state partners for their financial support and technical assistance.

The authors are thankful for the assistance provided by the members of the technical advisory committee (TAC) for the project: Dan Smith from the Missouri DOT (MoDOT), John Miller from MoDOT, and Kristi Ericksen from the Kansas DOT (KDOT). The authors would also like to thank all who completed the survey and participated in phone interviews. The authors would like to acknowledge the contributions of John W. Perlik, Xiaochuan Wang, and Ryan A. Rands who helped with compiling and processing work zone related data.

EXECUTIVE SUMMARY

The Highway Safety Manual (HSM) (AASHTO 2010) introduced quantitative methods to be used by transportation engineers and practitioners for safety and capacity assessment. Although the HSM includes methods to predict crashes for many different road facility types, it only gives two Crash Modification Factors (CMFs) to calculate the effect of increase or decrease of freeway work zone length and duration on the crash count. The HSM methodology for work zones is based on 36 freeway work zones with high traffic volumes in California. HSM models were calibrated in a recent research by Rahmani et al. (2016) using data from Missouri and the study determined a calibration factor of 3.78 which creates concerns since it is significantly larger than 1. This report describes the research conducted to make 15 different models to predict crashes for work zones on three facility types (freeway, expressway and rural two-lane highways) using Missouri data.

For work zone safety studies, different databases such as work zone characteristics, crash database and road network information need to be linked together. The tremendous amount of effort required for data collection and checking process makes work zone safety studies challenging. Of the 20,837 Missouri freeway, 8,993 expressway and 64,467 rural two-lane work zones that were analyzed in this report, samples of 1,546 freeway, 1,189 expressway, and 6,095 rural two-lane work zones were used to make eight, four and three models respectively. The samples were extracted using work zones longer than 0.1 mile and with a duration of greater than 10 days. The thresholds for minimum work zone length and duration were developed using a theoretical method devised by the authors. Most work zones in database were small work zones with short durations and no crashes. Using all of these work zones in the sample is possible but increases the uncertainty of the resulting model's predictions. However, by increasing the minimum length and duration threshold the sample size decreases. Thus, there is a tradeoff between dropping more small work zones and the sample size. This study tested different length and duration thresholds to extract the sample, and made work zone crash prediction models. By comparing the accuracy of the developed models, the optimum thresholds for minimum length and duration were found. Table ES-1 presents the characteristics of the work zones such as length, duration, AADT and number of crashes for all three facility types. The table shows that the work zone data represented a wide variety of work zones.

In work zone databases, the footprint of a work zone is typically recorded as the beginning and end of the work area. To account for the crashes that occur in the advance warning area, transition area, buffer area and termination area of work zones, most studies in the literature considered a constant threshold before the start and after the end of each work zone. The model used by the HSM (similar to most studies in the literature) classified all crashes within 0.5 mile (0.8 km) of the beginning and 0.5 mile (0.8 km) after the end of the work zone as work zone crashes. In contrast, as a new contribution this study used more accurate variable MUTCD (FHWA 2009a) recommended temporary traffic control plans' thresholds for freeway, expressway and rural two-lane work zones.

Table ES-1. Summary of work zone data characteristics

Freeway Work Zones			
Length, Duration and AADT			
Variables	Average	Min	Max
Length of work zone, mi (km)	5.048 (8.125)	0.101 (0.163)	29.920 (48.151)
AADT (vehicles per day)	30,531	757	128,756
Work Zone Duration (days)	43.4	10	290
Urban/rural percent	69% / 31%		
Number of observations	1,546		
Crashes			
Number of Crashes	All crashes	PDO	Fatal-Injury
Sum	9,199	6,975	2,224
Average	5.950	4.152	1.439
Min/max	0/175	0/136	0/39
Expressway Work Zones			
Length, Duration and AADT			
Variables	Average	Min	Max
Length of work zone, mi (km)	4.074 (6.557)	0.107 (0.172)	29.606 (47.474)
AADT (vehicles per day)	8,767	713	34,744
Work Zone duration (days)	51.3	10.3	298.3
Urban/rural percent	51% / 49%		
Number of observations	1,189		
Crashes			
Number of Crashes	All crashes	PDO	Fatal-Injury
Sum	3,047	1,624	591
Average	2.563	2.707	0.985
Min/max	0/74	0/42	0/32
Rural Two-Lane Work Zones			
Length, Duration and AADT			
Variables	Average	Min	Max
Length of work zone, mi (km)	5.803 (9.339)	0.1 (0.161)	29.897 (48.114)
AADT (vehicles per day)	778.6	50	10,325
Work Zone duration (days)	30.9	10	300
Number of observations	6,095		
Crashes			
Number of Crashes	All crashes	PDO	Fatal-Injury
Sum	1,077	1,077	552
Average	0.267	0.177	0.091
Min/max	0/32	0/23	0/9

All 15 of the models developed in this study were programmed in a user-friendly spreadsheet tool for practitioners. An illustrative example is presented to show how this software can be used for assessing the safety of different work zone plans. Figure ES-1 and ES-2 show the software graphical user interface and an example of output respectively.

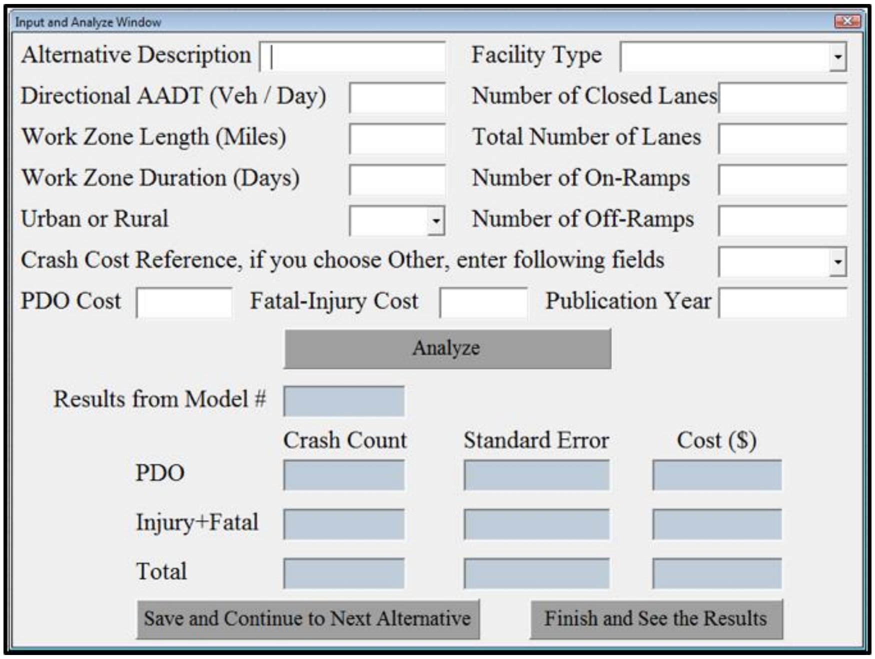


Figure ES-1. User input window of the work zone crash costs software

Alternatives Comparison			
	Output		
	Freeway Shoulder	Rural Two-Lane Resurfacing	Project 1
Expected Number of PDO Crashes	7.8	0.17	7.49
Standard Error of PDO Estimation	4.085	0.412	6.819
Expected Number of Fatal and Injury Crashes	2.51	0.06	2.49
Standard Error of Fatal and Injury Estimation	1.852	0.245	2.608
Total Crash Cost; value in 2016	$928,800	$21,953	$917,653
Model Used: #	6	14 & 15	12
	Input		
	Freeway Shoulder	Rural Two-Lane Resurfacing	Project 1
AADT	50000	1950	30000
Duration	47	15	54
Length	3	2	3
Urban/Rural	Urban	Rural	Urban
Number of Closed Lanes	1		
Total Number of Lanes	3		
Number of On-ramps	1		
Number of Off-ramps	2		
Number of Signalized Intersections		1	1
Crash Cost Reference; Publication Year	HSM (2010)	HSM (2010)	HSM (2010)
PDO Crash Cost	$7,400	$7,400	$7,400
Fatal and Injury Crash Cost	$158,200	$158,200	$158,200
Facility Type	Freeway	Rural Two-Lane	Expressway
			Developed by University of Missouri-Columbia; TransZou

Figure ES-2. Sample output of the software

This study also included a survey of DOTs, FHWA representatives, and contractors to assess the current state of the practice regarding work zone safety. Two separate online surveys were developed. One survey was for contractors and the other survey was for both DOT and FHWA representatives. There were seven respondents to the contractor online survey and 29 respondents (27 DOT respondents and 2 FHWA respondents) to the DOT and FHWA online survey. In addition, follow-up phone interviews were conducted with one contractor, eight DOT representatives, and one FHWA representative. Speed reduction was the most important factor identified by the contractors for freeway work zone safety while the number of intersections was the most important factor identified by the contractor respondents for facilities with at-grade intersections. The factors that more than half of the contractors took into account for work zone safety evaluations included traffic volumes, crash history, site characteristics and experience. Work zone traffic volume was the most significant factor affecting freeway work zone safety identified by the DOT and FHWA respondents, and the number of intersections in the work area was the most significant factor identified by these respondents for facilities with at-grade intersections. The factors that more than half of DOT and FHWA respondents considered for work zone safety evaluation include traffic volume, crash history, site characteristics and experience. The survey results indicate that many agencies look at work zone safety informally using engineering judgment. Respondents indicated that they would like a tool that could help them to balance work zone safety across projects by looking at crashes and user costs.

1 INTRODUCTION

In addition to connectivity and efficiency, safety is an important factor of any road network. According to FHWA (2009b) during the peak construction season, there are more than 3000 work zones on the National Highway System (NHS), and there are almost 12 billion vehicle miles traveled through work zones each year in the United States. More than 40,000 injuries happen at work zones which is equivalent to an injury each 13 minutes (FHWA 2009b). Work zones include many components that increase the crash occurrence risk, such as lane closures, lane width reductions, changes in road geometry, and the presence of construction workers.

Work zones have both traffic and safety impacts. Transportation agencies are in charge of assessing these effects. There are many tools available for analyzing the traffic impacts of work zones such as analytical tools (i.e. QuickZone, QUEWZ-98, and CA4PRS) based on the Highway Capacity Manual (TRB 2010). In addition, simulation tools can be used to analyze traffic impacts of complex work zones. A discussion of the traffic impact analysis tools can be found in Edara (2009) and Edara et al. (2013). As shown in Figure 1-1, practitioners need to estimate road user costs resulting from the traffic and safety impacts of a work zone.

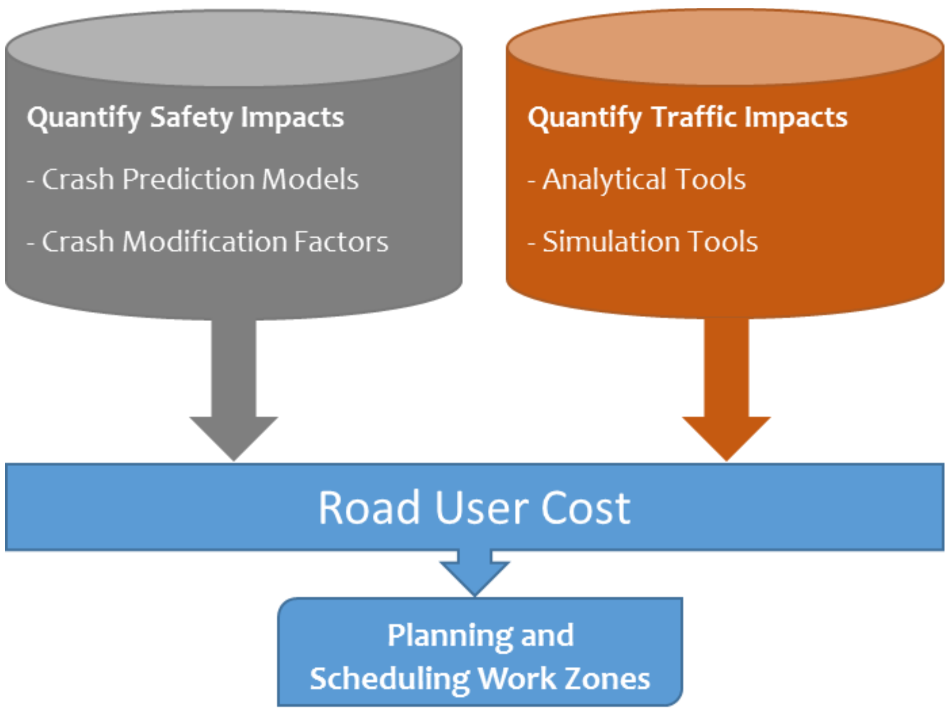


Figure 1-1. Assessing traffic and safety impacts for planning work zones

There are not many studies on quantifying work zone safety in the literature. The Highway Safety Manual (HSM) (AASHTO 2010) introduced two crash modification factors (CMFs) for freeway work zone length and duration derived from Khattak et al.'s (2002) Safety Performance Function (SPF). This SPF was made by using 36 high impact work zones in California. CMFs and SPFs will be explained and defined clearly in sections 1.1 and 1.2.

The main goal of this study is to develop a structured safety assessment tool to help decision makers to evaluate the safety impacts of different construction work zone phasing plans. To accomplish this goal, models were developed to predict work zone crashes for freeways, expressways and rural two lane highways. Different models were made using a large database of Missouri work zones between 2009 and 2014. The models were incorporated into a user-friendly spreadsheet tool for practitioners.

Based on the input data provided by the user, the software finds the proper and most accurate model to quantify the work zone safety and shows the results. The output includes the number of Property Damage Only (PDO crashes), number of Fatal-Injury crashes, total number of crashes, and their standard error. A smaller standard error means higher accuracy. Further, these predicted number of crashes are converted to monetary values using standard crash costs (Sun et al. 2014a) to facilitate comparison of alternative plans and schedules.

This report presents the process and tasks that were executed to accomplish the goals of this project. In the next sections of this Chapter, SPFs and CMFs are defined. In Chapter 2, a thorough literature review on the quantifying work safety studies is described. Two surveys were designed for Department of Transportation (DOT) and Federal Highway Administration (FHWA) employees and contractors to learn about existing practices for work zone safety. The surveys and results are described in Chapter 3. Chapter 4 explains the data collection, fusion, and sampling as the essential components of any data-driven study. Chapter 5 focuses on the necessary statistical background for modeling. Chapter 6 describes the modeling process for this project and presents the models developed for freeways, expressways and rural two-lane highways. Sample applications and the software tutorial are presented in Chapter 7.

1.1 Safety Performance Function (SPF) and Crash Modification Function (CMF)

"Safety performance functions (SPFs) are regression equations that estimate the average crash frequency for a specific site type (with specified base conditions) as a function of annual average daily traffic (AADT) and, in the case of roadway segments, the segment length (L)." (HSM) (AASHTO 2010). Any SPF is made based on a set of specific geometric and geographic characteristics called base conditions. To use a SPF for a condition different from the base condition, it needs to be multiplied by Crash Modification Factors (CMFs). The following equation shows the general form of a crash prediction model for a site.

$$\boldsymbol{N_{predicted} = N_{SPF} \times CMF_1 \times CMF_2 \ldots \times C}$$

Where:

$N_{predicted}$ is the predicted crash frequency for a site,
N_{SPF} is the predicted crash frequency for specified base conditions,
CMF_i is the crash modification factor *i* reflecting a prevailing site condition that differs from the base condition,

C is the calibration factor which accounts for differences (jurisdictional and time period) between the sample used for SPF development and the one for which the crash frequency is currently being estimated.

"The relative change in crash frequency due to a change in one specific condition (when all other conditions and site characteristics remain constant)" is represented by the CMF (HSM) (AASHTO 2010). The HSM provides CMF values for several facility types derived by synthesizing previous research. The crash data used to develop the SPFs usually comes from several states. Chapter 16 of HSM introduces two CMFs for work zone duration and length that were developed using California data (AASHTO 2010):

$$\boldsymbol{CMF_{d,all} = 1.0 + \frac{(\%\ \text{increase in duration}\ x\ 1.11)}{100}}$$

The crash modification factor for work zone length for all crash severities is presented as (AASHTO 2010):

$$\boldsymbol{CMF_{l,all} = 1.0 + \frac{(\%\ \text{increase in length}\ x\ 0.67)}{100}}$$

Both CMFs specify a linear relationship between the CMF value and work zone duration or length.

Work zones SPFs could be modeled in two ways. The first method considers the work zone as a base condition and makes an SPF for work zone situation. The second method is a before-after study with data for both a work zone and non-work zone situation. The non-work zone situation could be achieved from the same location in a period before implementing the work zone or finding another segment/site with similar characteristics such as duration (cross-sectional before-after study, which is not as reliable as the other one). In before-after studies, the most important factor is the similarity of the before and after conditions in all aspects, except for the target treatment (implementing a work zone). However, even if the samples are chosen properly, as crash frequency is a random variable there is no guarantee that before/after durations are similar to each other. This phenomenon is called regression to mean (RTM). In before-after studies to overcome RTM, Empirical Bayes (EB) is the best solution. EB method considers the difference between predicted and observed crashes and tries to reduce the regression to mean phenomenon.

2 LITERATURE REVIEW

While existing research is somewhat inconsistent on the impact of work zone presence on crash severity, most studies show that work zone presence has a negative impact on crash frequency. According to a recent review from Yang et al. (2015), 48% of previous studies on work zone crash severity indicate no clear evidence that there is an increase in crash severity during work zone conditions. On the other hand, the majority of previous studies regarding work zone crash frequency show an obvious increase in crash frequencies during work zone operations. Crash frequency is usually used as a safety evaluation measure for work zones and is expressed in the total number of crashes in a given time period.

Although there are many studies on crash frequency modeling, only a few of them focus on work zone presence. Pal and Sinha (1996) conducted a study on Indiana highway work zones and found that crash rates in work zones were significantly higher than non-work zone conditions. They developed two normal regression models to compare the predicted crash rate for different types of lane closures. Although the normal regression model seemed to have better prediction power over the negative binomial and Poisson models, it produced negative crash rates in several cases. To ensure non-negative predicting results, researchers started using and fine-tuning negative binomial models and Poisson models. Venugopal and Tarko (2000) developed two negative binomial models with duration of work, type of work, AADT and work zone length as the main variables. The two models were calibrated for two regions: the region approaching the work zone and the region containing the work zone. They also added cost of work to the model as an indicator of the intensity of work and showed AADT, work zone length and duration to be major safety related factors. Khattak et al. (2002) developed a negative binomial model using before-and-after data with crash rates of 0.65 crashes per million vehicle kilometers without work zones and 0.79 crashes per million vehicle kilometers with work zones. Thus the models they developed showed higher crash tendency for work zones. Their findings were consistent with previous studies which suggested that a higher AADT along with a longer work zone duration and work zone length led to a higher crash rate. The current HSM CMF for work zones is derived from the aforementioned model by Khattak et al. (2002). To account for zero-crash work zones, researchers have suggested using zero-inflated negative binomial models (a zero-inflated model is based on a statistical distribution that allows for frequent zero-valued observations.). Although there were studies comparing zero-inflated negative binomial models with negative binomial models for crash frequency prediction modeling (Lord et. al 2005; Lord and Mannering 2010), no one has tested and compared zero-inflated negative binomial models with other models using work zone data. Qi et al. (2005) built a zero-inflated negative binomial model but did not compare it to the truncated negative binomial model in their study. Srinivasan et al. (2011) developed negative binomial SPFs for all crashes, injury crashes, and PDOs, and then used the empirical Bayes method to estimate different CMFs for daytime and nighttime work zones. Recently, Ozturk et. al (2013) developed a negative binomial-based model with further temporal adjusted daytime and nighttime traffic volumes and found that “work zone duration”, “length of work zone” and “traffic volumes” had the most impact on work zone safety. Chen and Tarko (2014) proposed a new fixed-parameter negative binomial model with random effects as an alternative to random parameters model, and obtained similar crash frequency prediction accuracies. Since previous studies have shown reliable results by using the negative binomial model in work zone crash frequency modeling, this study also developed negative binomial models.

3 WORK ZONE SURVEY

Two online surveys were prepared for contractors and DOT and FHWA representatives and were sent to respondents by email. Thirty-six survey responses were received, seven from contractors and 29 from DOT and FHWA representatives. The survey summary results are explained in sections 3-1 and 3-2 respectively. The surveys were designed using a web tool (i.e. Survey Monkey) and were sent to candidate respondents.

3.1 Work Zone Survey of Contractors

This section summarizes the responses from the 15-question contractor survey. Each question form the survey is repeated below along with a summary of results:

Q1: What agency do you represent?

Seven different respondents answered the questions of this survey and due to privacy issues, their information are kept confidential.

Q2: Do you believe, generally, that the presence of work zones increases the crash frequency?

As Table 3-1 shows, among the contractor respondents, 57.1% of them believed that work zone presence increases the crash frequency, while 28.6% of them thought it does not increase crash frequency. One of the seven respondents was unsure.

Table 3-1. Summary of results for Question 2 of contractor survey

Answer Options	Response Percent	Response Count
Yes	57.1%	4
No	28.6%	2
Unsure	14.3%	1

Q3: To what degree do you believe that the following factors impact work zone safety on freeways? Please rate the factors on a scale from Not Important to Highly Important.

Figure 3-1 shows the answer options for this question.

	Not Important	Somewhat Important	Important	Highly Important
Work zone length	○	○	○	○
Work zone duration	○	○	○	○
Work zone traffic volume: work zone AADT	○	○	○	○
Terrain (flat, rolling, etc.)	○	○	○	○
Urban versus rural roadway	○	○	○	○
Speed decrease	○	○	○	○
Type of work: lane closure (drop in number of lanes)	○	○	○	○
Type of work: lane shift/crossover	○	○	○	○
Type of work: work on shoulder	○	○	○	○
Type of work: moving WZ	○	○	○	○
Work zone warning signs	○	○	○	○
Number of on-off ramps	○	○	○	○
Contracting elements such as liquidated damages, incentive/disincentives, cost+time	○	○	○	○
Contract cost per mile per duration	○	○	○	○

Add any other factors that impact work zone safety

Figure 3-1. Answer options for Question 3 of contractor survey

The importance levels were ranked numerically from zero to three as: 0 for not important to 3 for highly important. Figure 3-2 summarizes the average importance of each factor in freeway work zones' safety based on the responses.

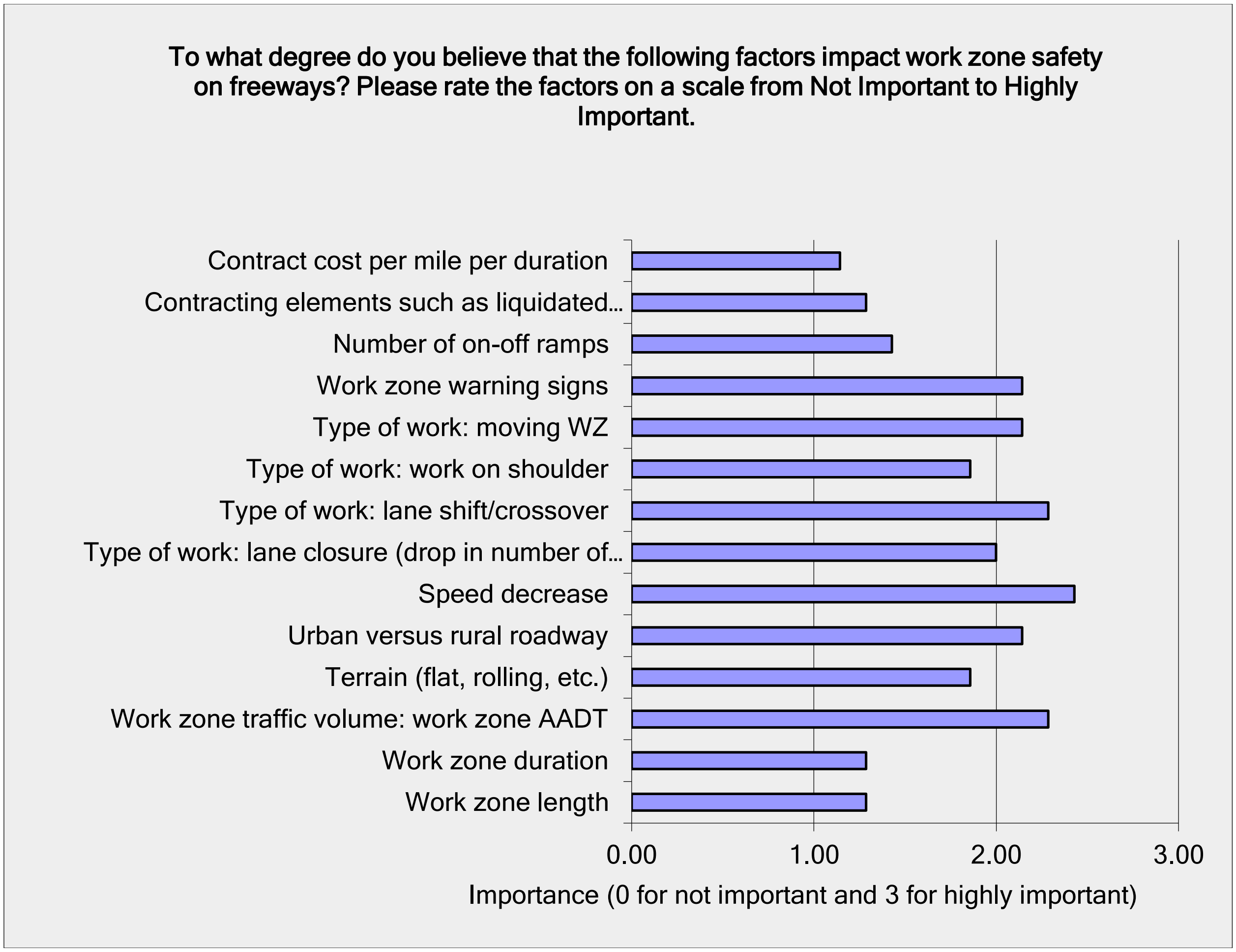
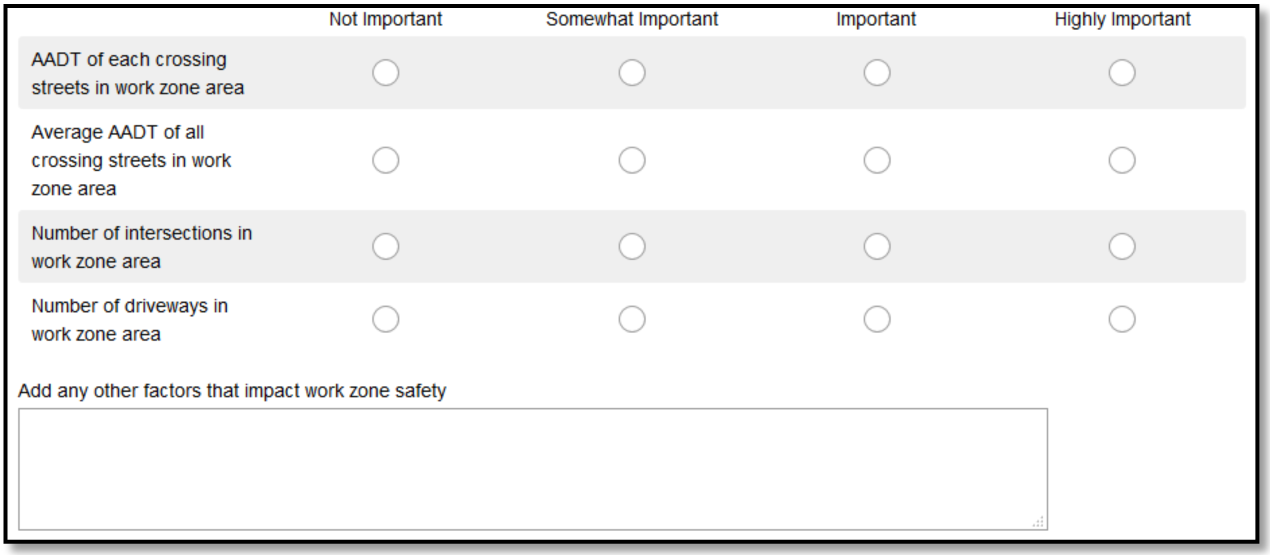


Figure 3-2. Average importance of work zone safety factors for Question 3 of contractor survey

The factors with an average importance of 2 and more were: AADT, urban-rural classification, speed reduction, type of work zone (lane shift, crossover, lane closure and moving work zone) and work zone warning signs. Speed reduction was the most important factor identified by the contractors. A respondent mentioned “depending on if there's night work, high visibility is a key factor with speed, road design, and volume.”

Q4: To what degree do you believe that the following additional factors impact the safety of work zones on facilities with at-grade intersections? Please rate the factors on a scale from Not Important to Highly Important.

Figure 3-3 shows the answer options for this question.



	Not Important	Somewhat Important	Important	Highly Important
AADT of each crossing streets in work zone area				
Average AADT of all crossing streets in work zone area				
Number of intersections in work zone area				
Number of driveways in work zone area				

Add any other factors that impact work zone safety

Figure 3-3. Answer options for Question 4 of contractor survey

The importance levels were numerically ranked from zero to three as: 0 for not important to 3 for highly important. Figure 3-4 summarizes the average importance of each factor in arterial work zones' safety based on the responses.

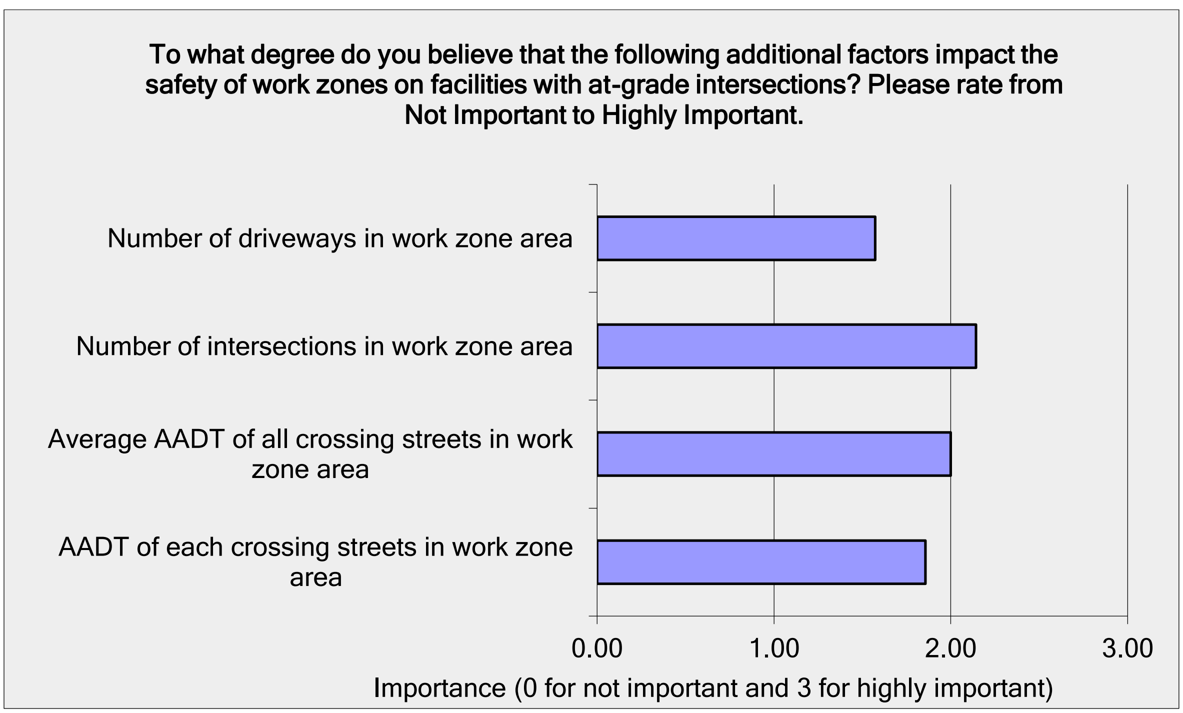


Figure 3-4. Average importance of different factors for Question 4 of contractor survey

The factors with an average importance of 2 and more were: number of intersections in work zone segment and average AADT of crossing roads. The number of intersections was the most important factor identified by the respondents.

A respondent also had the option to add any other factors not included in the list of answers. A written answer from a respondent was: “Depending on if there's night work, high visibility is a key factor with speed, road design, and volume.”

Q5: For freeway work zones, how far upstream before the transition area and downstream after the activity area would you consider a crash to be most likely influenced by the work zone?

As shown in Table 3-2, one respondent did not answer this question and the rest of respondents believed that work zones influence area length is less than a mile upstream and downstream of work zones. The upstream length of the influence area was longer than the downstream length.

Table 3-2. Summary of results for Question 5 of contractor survey

Answer Options	0-0.25 mi	0.25-0.5 mi	0.5-1 mi	1-1.5 mi	1.5-2 mi	>2mi	Response Count
Peak-Hour in Urban Areas							
Upstream	1	3	2	0	0	0	6
Downstream	2	1	1	0	1	0	5
Off-Peak in Urban Areas							
Upstream	3	3	0	0	0	0	6
Downstream	2	1	1	0	0	1	5
Rural							
Upstream	2	4	0	0	0	0	6
Downstream	3	0	1	0	0	1	5

Q6: Based on your experience, rank the following work zone characteristics in order of safety risk (1 for the most probability of having an incident and 4 for the least):

The answer options for this question are shown in Table 3-3, and the results are shown in Table 3-4.

Table 3-3. Answer options for Question 6 of contractor survey

	Freeways	Arterials	Nighttime	Peak-Hour
For Workers	1-4	1-4	1-4	1-4
For Driving Public	1-4	1-4	1-4	1-4

Table 3-4. Summary of results for Question 6 of contractor survey

Answer Options	1	2	3	4	Response Count
Freeways					
For Workers	3	2	2	0	7
For Driving Public	2	2	2	1	7
Arterials					
For Workers	1	3	0	3	7
For Driving Public	0	1	3	3	7
Nighttime					
For Workers	5	1	1	0	7
For Driving Public	3	2	2	0	7
Peak-Hour					
For Workers	3	2	1	1	7
For Driving Public	3	4	0	0	7

The results indicate that the contractors viewed nighttime and peak-hour work zones as a safety concern. Contractors also viewed freeways as more of a safety concern than arterials. It should be noted that some of the contractors did rate the characteristics 1 through 4, but the results still provide some insights into contractors' views of work zone safety risks.

Q7: How serious do you think work zone motor vehicle crashes are compared to other work zone safety hazards (e.g. equipment misuse)?

As shown in Table 3-5, 57.2% of respondents believed that work zone motor vehicle crashes are more serious compared to other work zone safety hazards. One respondent thought it is less serious.

Table 3-5. Summary of results for Question 7 of contractor survey

Answer Options	Response Percent	Response Count
Much less serious	0.0%	0
Less serious	14.3%	1
Same	14.3%	1
More serious	28.6%	2
Much more serious	28.6%	2
Unsure	14.3%	1

Q8: Do you think a greater police presence at work zone construction sites help to reduce the number of work zone crashes, injuries and/or fatalities?

Six respondents answered 'yes' to this question while one believed police enforcement does not improve work zone safety.

Q9: Which section of a typical work zone do most crashes occur?

As shown in Table 3-6, all respondents believed that the transition area is the most probable section of work zone for crashes to happen. Interestingly, the empirical data in Chapter 4 indicate a different result.

Table 3-6. Summary of results for Question 8 of contractor survey

Answer Options	Response Percent	Response Count
Advance Warning Area	0.0%	0
Transition Area	100.0%	7
Activity Area	0.0%	0
Termination Area	0.0%	0

Q10: To what extent do work zone incidents and/or crashes delay your construction schedule?

Table 3-7 shows that respondents mentioned that work zone incidents and crashes delayed construction from somewhat to significantly.

Table 3-7. Summary of results for Question 9 of contractor survey

Answer Options	Response Percent	Response Count
Not at all	0.0%	0
Somewhat	57.1%	4
Significantly	42.9%	3
Very Significantly	0.0%	0

Q11: To what extent is your firm's Experience Modification Rate (EMR) affected by highway work zone safety considerations?

Six respondents answered this question, Table 3-8, and four of them mentioned that their firm's EMR was somewhat affected by work zone safety considerations.

Table 3-8. Summary of results for Question 11 of contractor survey

Answer Options	Response Percent	Response Count
Not at all	16.7%	1
Somewhat	66.7%	4
Significantly	0.0%	0
Very Significantly	16.7%	1

Q12: To what extent does worker and public safety play a role in winning a construction/rehabilitation bid?

Four of seven respondents mentioned that worker and public safety plays an important role in their chance of winning a construction/rehabilitation bid (Table 3-9).

Table 3-9. Summary of results for Question 12 of contractor survey

Answer Options	Response Percent	Response Count
Not Important	14.3%	1
Important	57.1%	4
Somewhat Important	14.3%	1
Highly Important	14.3%	1

Q13: What factors do you take into account for evaluating work zone safety? (check all that apply)

Six respondents answered the question and a summary of the results is shown in Table 3-10.

Table 3-10. Summary of results for Question 13 of contractor survey

Answer Options	Response Percent	Response Count
Traffic volumes	100.0%	6
Crash history	66.7%	4
Site characteristics	100.0%	6
Knowledge/experience	83.3%	5
Highway Safety Manual (HSM)	33.3%	2
Others	16.7%	1
Answered question		6
Skipped question		1

The factors that more than half of these contractors took into account for work zone safety evaluations were traffic volumes, crash history, site characteristics and experience. This result underscores the need for a quantitative work zone safety assessment tool.

Q14: Do you use any tools or quantitative measures to compare the safety of different alternative work zone phasing plans? If so, which one(s)?

As shown in Table 3-11, six of the respondents mentioned that they do not have any quantitative measure to compare different work zone alternatives.

Table 3-11. Summary of results for Question 14 of contractor survey

Answer Options	Response Percent	Response Count
No or Engineering Judgment only	100.0%	6
HSM work zone CMF	0.0%	0
Other published sources	33.3%	2

A respondent mentioned MUTCD as a source, which does not have any quantitative work zone safety evaluation equation. The results show the importance and need for a quantitative work zone safety assessment tool for practitioners.

Q15: If there were a tool for quantifying the safety of different work zone configurations, how frequently would you use it on the following types of work zones?

Q16: Would you like to receive a copy of the final project report and work zone safety assessment tool when they are completed?

The answers to questions 15 and 16, Table 3-12 and Table 3-13, show the practitioners' eagerness to have access to an analytical work zone safety assessment tool, which is the goal of this project.

Table 3-12. Summary of results for Question 15 of contractor survey

Answer Options	Always	Often	Sometimes	Rarely	Never	Response Count
Full closure work zones	3	2	0	1	0	6
Crossover work zones	2	4	0	0	0	6
Lane closure work zones	2	4	0	0	0	6
Work zones on shoulder	2	2	1	1	0	6
Short term work zones	2	3	1	0	0	6
Moving work zones	3	2	1	0	0	6
Answered question						6
Skipped question						1

Table 3-13. Summary of results for Question 16 of contractor survey

Answer Options	Response Percent	Response Count
Yes	83.3%	5
No	16.7%	1
Answered question		6
Skipped question		1

Q17: Are there any additional comments that you have on work zone safety?

None of the respondents had any additional comments on work zone safety.

3.2 Work Zone Survey of DOT and FHWA Representatives

This section summarizes the answers of 10-question survey of DOT and FHWA representatives, answered by 27 DOT representatives and 2 FHWA representatives.

Q1: What agency do you represent?

The survey was sent to different DOT and FHWA representatives and 29 different respondents completed it. Due to privacy issues, their information is kept confidential.

Q2: Do you believe, generally, that the presence of work zones increases the crash frequency?

The results of this question are shown in Table 3-14.

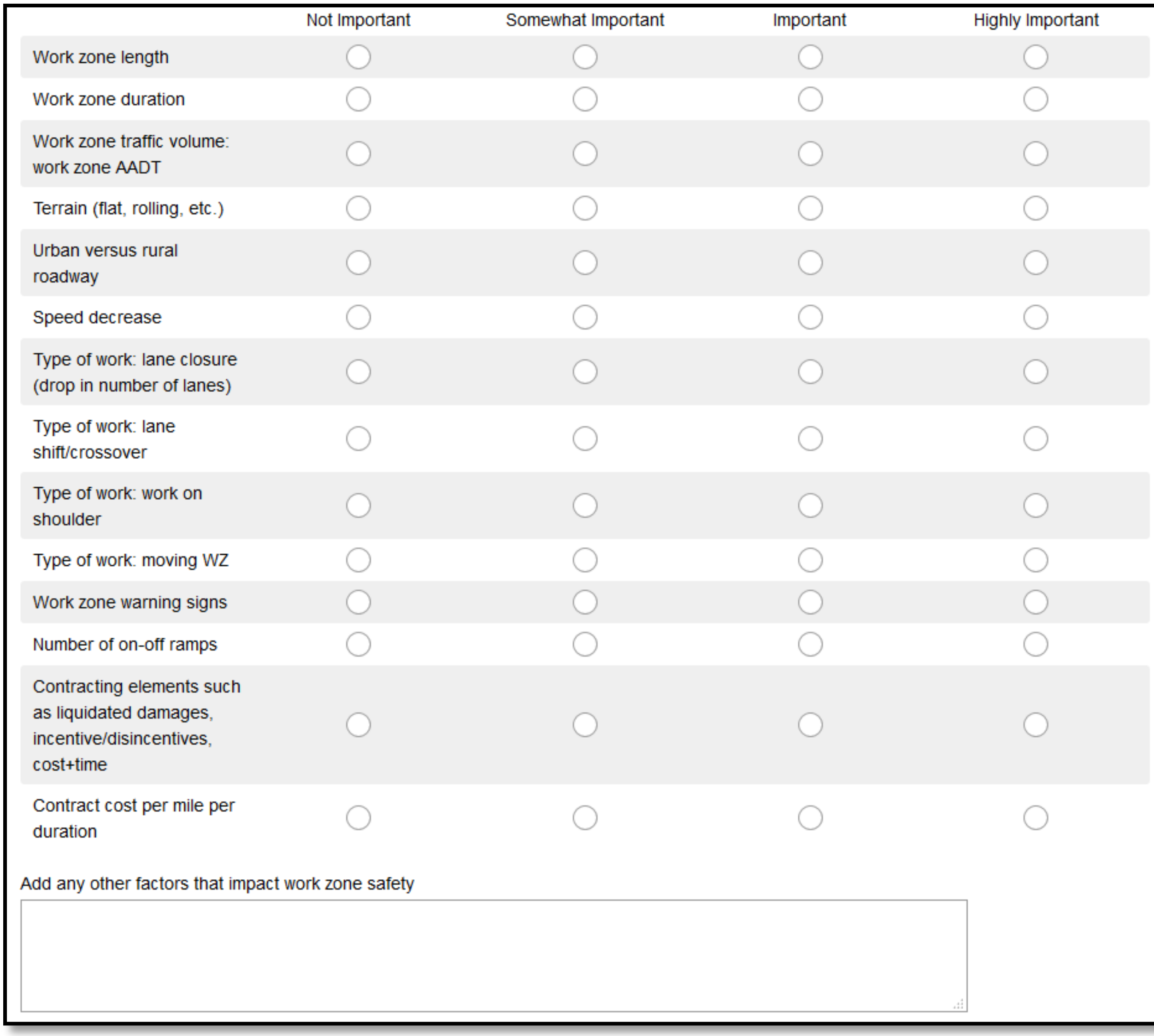
Table 3-14. Summary of results for Question 2 of DOT/FHWA survey

Answer Options	Response Percent	Response Count
Yes	48.3%	14
No	34.5%	10
Unsure	17.2%	5

Among the DOT and FHWA respondents, 48.3% of them believed that work zone presence increases the crash frequency, while 34.5% thought it does not increase crash frequency. Five of the 29 respondents were unsure about this question.

Q3: To what degree do you believe that the following factors impact work zone safety on freeways? Please rate the factors on a scale from Not Important to Highly Important.

Figure 3-5 shows the answer options for this question.



	Not Important	Somewhat Important	Important	Highly Important
Work zone length	◯	◯	◯	◯
Work zone duration	◯	◯	◯	◯
Work zone traffic volume: work zone AADT	◯	◯	◯	◯
Terrain (flat, rolling, etc.)	◯	◯	◯	◯
Urban versus rural roadway	◯	◯	◯	◯
Speed decrease	◯	◯	◯	◯
Type of work: lane closure (drop in number of lanes)	◯	◯	◯	◯
Type of work: lane shift/crossover	◯	◯	◯	◯
Type of work: work on shoulder	◯	◯	◯	◯
Type of work: moving WZ	◯	◯	◯	◯
Work zone warning signs	◯	◯	◯	◯
Number of on-off ramps	◯	◯	◯	◯
Contracting elements such as liquidated damages, incentive/disincentives, cost+time	◯	◯	◯	◯
Contract cost per mile per duration	◯	◯	◯	◯

Add any other factors that impact work zone safety

Figure 3-5. Answer options for Question 3 of DOT/FHWA survey

The importance levels were ranked numerically from zero to three as follows: 0 for not important to 3 for highly important. Figure 3-6 summarizes the average importance of each factor in freeway work zone safety based on the responses.

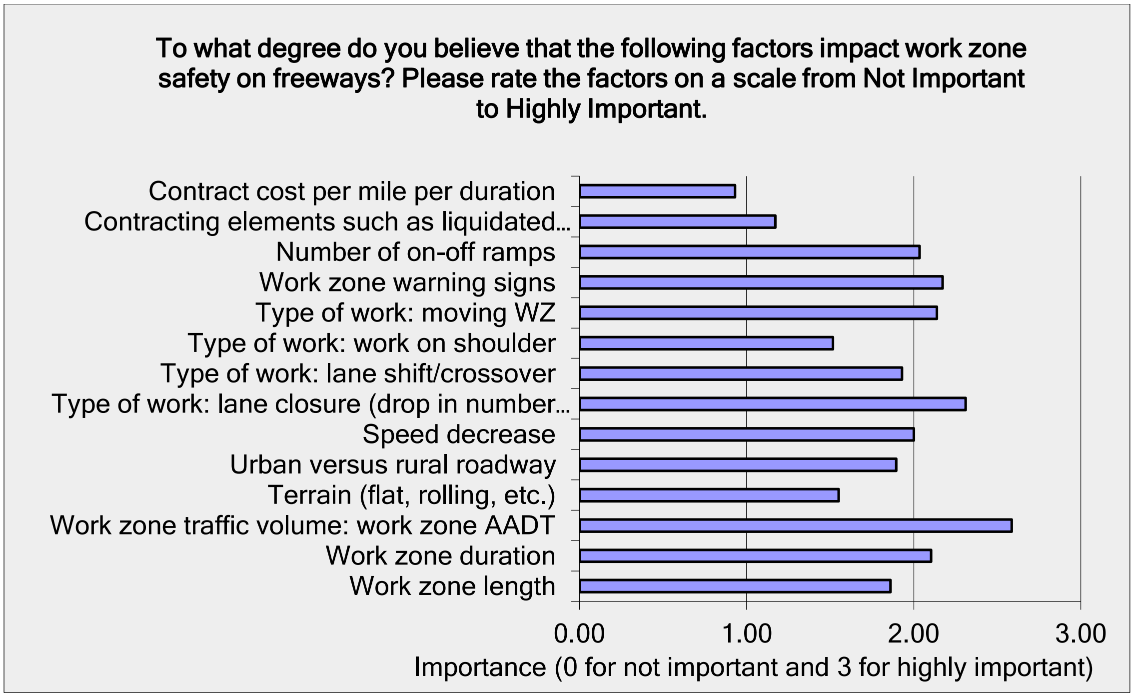


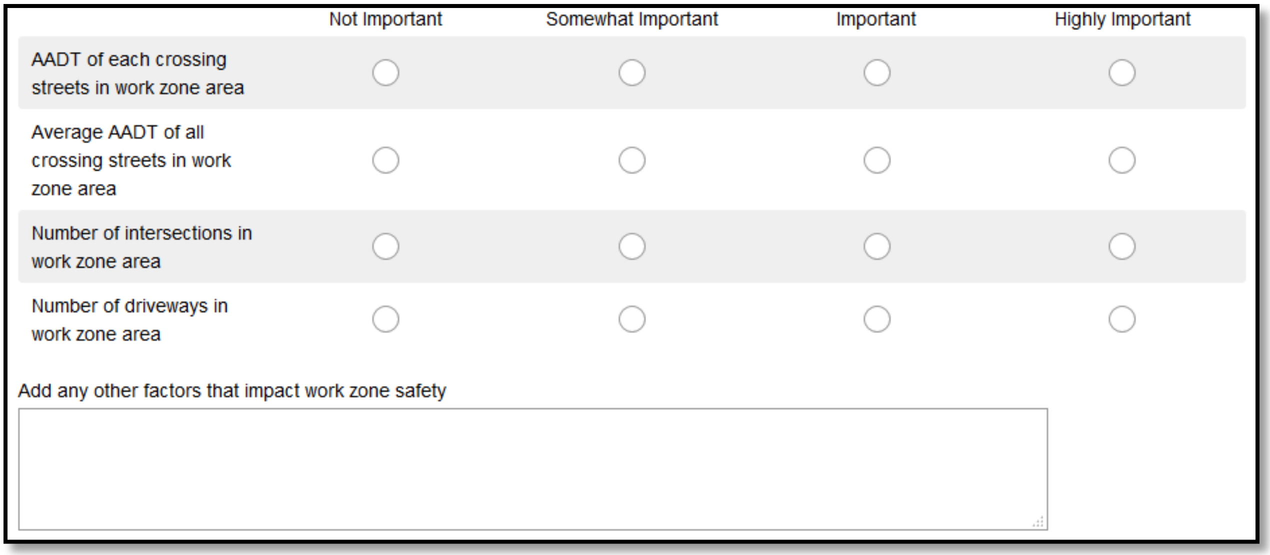
Figure 3-6. Average importance of different factors for Question 3 of DOT/FHWA survey

The factors with average importance of 2 and more were AADT, duration, speed reduction, type of work zone (lane shift, crossover, lane closure and moving work zone), number of on /off-ramps and work zone warning signs. Work zone traffic volume was the most significant factor identified by the respondents.

Some respondents mentioned a few different factors other than the ones in the list such as number of traffic phases, presence or absence of shoulders (especially for moving or short duration work), sight distance along the highway night construction, lane width, daytime versus nighttime seasons (e.g. summer versus winter), large truck volumes or percentage of total ADT, state construction versus maintenance operations versus utility contractors under permit, predictability of traffic delays (i.e. expected end of queue versus unexpected end of queue), types of positive protection used (i.e. temporary barrier, mobile barrier, barricades, traffic cones, impact attenuation devices), and advance notification with the use of ITS to alert of work zone areas (cameras, message boards, and sensors).

Q4: To what degree do you believe that the following additional factors impact the safety of work zones on facilities with at-grade intersections? Please rate from Not Important to Highly Important.

Figure 3-7 shows the answer options for this question.



	Not Important	Somewhat Important	Important	Highly Important
AADT of each crossing streets in work zone area				
Average AADT of all crossing streets in work zone area				
Number of intersections in work zone area				
Number of driveways in work zone area				

Add any other factors that impact work zone safety

Figure 3-7. Answer options for Question 4 of DOT/FHWA survey

The importance levels were ranked numerically from zero to three as follows: 0 for not important to 3 for highly important. Figure 3-8 summarizes the average importance of each factor in arterial work zones' safety based on the responses.

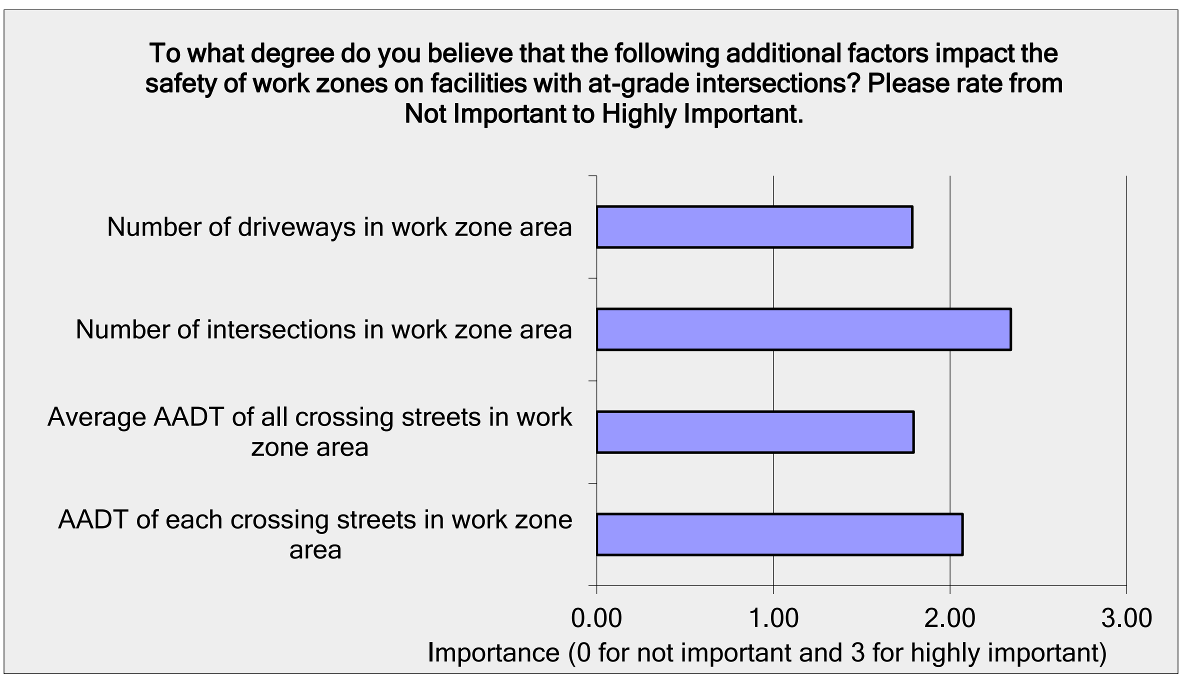


Figure 3-8. Average importance of different factors for Question 4 of DOT/FHWA survey

The factors with average importance of 2 and more were number of intersections in work zone segment and AADT of each crossing road. The number of intersections in the work area was the most significant factor identified by the respondents.

Some DOT respondents mentioned a few different factors other than the ones in the list such as maintenance of work zone devices, police enforcement of speeds and other traffic regulations, number of businesses within the work zone, spacing of intersections/accesses, available alternate routes for intersecting local/surface streets, access types (e.g. public parks, private businesses, large businesses, schools and factories, private residence or apartment complex), number of pedestrians and bicyclists that travel in and around work zone area, and quality of devices and layout (effectiveness) especially in intersections.

Q5: For freeway work zones, how far upstream before the transition area and downstream after the activity area would you consider a crash to be most likely influenced by the work zone?

As shown in Table 3-15, 27 respondents answered to this question and more than 80% of them believed that work zones influence area length is less than a mile upstream and downstream of work zones. The upstream influence length was longer than downstream.

Table 3-15. Summary of results for Question 5 of DOT/FHWA survey

Answer Options	0-0.25 mi	0.25-0.5 mi	0.5-1 mi	1-1.5 mi	1.5-2mi	>2mi	Response Count
Peak-Hour in Urban Areas							
Upstream	2	7	8	3	1	6	27
Downstream	15	9	2	1	0	0	27
Off-Peak in Urban Areas							
Upstream	8	9	5	4	1	0	27
Downstream	19	5	1	0	1	0	26
Rural							
Upstream	6	5	10	3	1	2	27
Downstream	18	5	1	1	1	0	26

Q6: How do you account for safety in work zone planning/design? (check all that apply)

Results concerning safety factors are shown in Table 3-16.

Table 3-16. Summary of results for Question 6 of DOT/FHWA survey

Answer Options	Response Percent	Response Count
Traffic volumes	96.6%	28
Crash history	62.1%	18
Site characteristics	96.6%	28
Knowledge/experience	96.6%	28
Highway Safety Manual (HSM)	34.5%	10
Other	17.2%	5
Others (please specify)		8

The factors that more than half of respondents considered for work zone safety evaluation include traffic volume, crash history, site characteristics and experience. The results were similar to the results from the contractor survey. The need for a quantitative work zone safety assessment tool is apparent.

Some respondents mentioned other sources they use to account for work zone safety planning and design such as the Virginia Work Area Protection manual (Virginia's version of Part 6 to the MUTCD), standardized design of TCPs Pre-construction operating speeds safety and operational review of detour routes, MUTCD, and the Montana Department of Transportation Standard Specifications and Detailed Drawings. There were two other interesting comments. The first comment was regarding ITS usage: “account for ITS architecture to assist while work zones in place. i.e. cameras, message boards, sensors to be used for diversion and notification”. The second comment concerned driver behavior: “driver behavior considerations if the info is available - usually falls under experience.”

Q7: Does your agency utilize innovative contracting techniques (e.g. incentive/disincentive contracts, A+B bidding, etc.) to improve work zone safety?

As shown in Table 3-17, among the 28 respondents that answered this question, 79% of them indicated that their agency uses innovative contracting techniques (e.g. incentive/disincentive contracts, A+B bidding, etc.).

Table 3-17. Summary of results for Question 7 of DOT/FHWA survey

Answer Options	Response Percent	Response Count
Yes	78.6%	22
No	10.7%	3
Unsure	10.7%	3
Explain, if necessary		8
Answered question		28
Skipped question		1

The methods some respondents mentioned include incentive/disincentive, I/D lane rental, A+B (Cost + Time) and A+C (Cost + Pre-Qualifications), and accelerated bridge construction techniques.

Q8: If there were a tool for quantifying the safety of different work zone configurations, how frequently would you use it?

Q9: Would you like to receive a copy of the final project report and work zone safety assessment tool when they are completed?

The answers to questions 8 and 9, in Tables 3-18 and 3-19, show the DOT and FHWA representatives' eagerness to have access to an analytical work zone safety assessment tool, which is the goal of this project.

Table 3-18. Summary of results for Question 8 of DOT/FHWA survey

Answer Options	Always	Often	Sometimes	Rarely	Never	Response Count
Choose the most applicable	2	13	11	3	0	29

Table 3-19. Summary of results for Question 9 of DOT/FHWA survey

Answer Options	Response Percent	Response Count
Yes	96.4%	27
No	3.6%	1
Answered question		28
Skipped question		1

Q10: Are there any additional comments that you have on work zone safety?

Several respondents provided additional comments as summarized below:

- Having experienced work zone contractors and project inspectors is important.
- A good sequence of work and constructability review are beneficial.
- Proper deployment of maintenance of traffic in the work zone is helpful.
- Speeding in the work zone is a concern. One possible countermeasure is the use of radar speed feedback signs in the work zone.
- A safety assessment tool should be kept updated with current technologies.
- Each work zone is unique and requires engineering judgement for the design of the temporary traffic control.
- A variety of factors should be taken into consideration when evaluating work zone safety.

3.3 Phone Interviews

In addition to the online survey, ten follow-up phone interviews were conducted with some of the participants from the online survey to learn more about their perspectives on work zone safety. The interviewees were from the following states: Oregon, Iowa, District of Columbia, Kansas, Missouri, Virginia, and Nebraska. The interviews included eight representatives from state DOTs, one representative from the Federal Highway Administration (FHWA), and one contractor. Some of the questions asked of the interviewees included:

- Do you currently incorporate safety analysis into your evaluation of work zone phasing alternatives? If so, how do you perform this analysis?

- What features would you like to see incorporated into a safety assessment tool to evaluate work zone phasing alternatives?

Many interviewees responded that they do not perform a formal safety analysis of work zone alternatives but look at safety informally using engineering judgment. Some agencies indicated that it is difficult to obtain sufficient data for a formal analysis. Other agencies look at crash rates as an indicator of work zone safety. Anticipation of queuing is also used by some agencies to help reduce rear end crashes in work zones.

In response to the question about features in safety assessment tool for work zones, interviewees indicated that they would like a tool that could help them to balance safety across projects by looking at crashes and user costs. Work zone duration and traffic counts were mentioned as important exposure variables that should be investigated. The tool should give some guidance to practitioners regarding factors to consider when evaluating work zone safety and should help to provide recommendations to practitioners while recognizing that each project is different.

Interviewees were also asked about other aspects of work zone safety such as nighttime work zones, quality assurance, and work zone speed limits. Some of the other key findings from the phone interviews include:

- Work zone phasing that provides for separation between traffic and construction workers is ideal. A complete closure with a detour is preferable from a safety standpoint but often difficult to implement due to resistance from stakeholders and the general public. Other options to be considered in order of preference include the use of a temporary bypass to divert traffic from the construction area, staged construction with a horizontal offset or concrete barrier, and the use of channelizers to separate the traffic from construction workers.

- Consistency in work zone implementation within an agency can be a challenge, especially since there can be differences in goals between jurisdictions (such as rural versus urban). Training can help to improve consistency in work zone implementation.

- Agencies use various methods to help improve visibility and safety for nighttime work zones such as the development of specifications, use of brighter signs, use of drums instead of cones for channelizers, and lighting requirements such as lighting for flagger stations.

- Short duration work zones can also be challenging. Some agencies use additional measures for short duration work zones such as the use of a special handbook for operations in short duration work zones or trying to make construction vehicles more visible.

- Communicating work zone information to the public through public outreach and Dynamic Message Signs (DMS) is an important aspect of work zone safety.

- Agencies have different policies for setting work zone speed limits. For example, Missouri typically uses a 10 mph speed limit reduction on freeways. Iowa reduces work zone speed limits from 70 mph to 55 mph for two-lane two-way operations on multi-lane highways. Nebraska has implemented a policy that sets work zone speed limits based on the type of facility and type of work. Virginia uses a spreadsheet to analyze work zone speed limits but tries not to lower speed limits if possible. The District of Columbia typically uses a 5 mph speed reduction for work zones. In Kansas, work zone speed limits are set in the field and are typically 10 mph to 15 mph less than the posted speed limit.

- Most interviewees generally thought that speed enforcement helps to improve work zone safety, but it helps if the enforcement is visible and announced in advance. In some cases, enforcement can impact traffic and safety negatively if it causes traffic backup.

- Many agencies perform work zone audits to help evaluate the safety and operations of work zones.

- The use of Personal Protective Equipment by construction workers is an important component of work zone safety.

4 DATA

The dependent variable in work zone crash prediction models is most often the crash count based on crash severity. The crash temporal-spatial characteristics are available in some DOT crash archives. The Missouri DOT databases contained all the necessary information; thus they were used for model development in this study. By knowing the exact time and place of work zones, crashes that occurred in the exact temporal-spatial footprint of work zones could be found. The number of crashes that occurred in a work zone is the only necessary dependent variable in developing work zone crash prediction models. However, the locations of crashes in work zones could be interesting to DOTs; so this report also discusses crash location analysis. Typically, a work zone is divided into five independent segments: advance warning area, transition area, buffer area, activity area, and termination area (MUTCD 2009). These segments are described in greater detail in subsequent sections of this report.

4.1 Databases

To develop a work zone SPF, three categories of data are needed: work zone characteristics, crash characteristics, and road and traffic characteristics. The challenging part is fusing the data from these different databases. Some of the necessary data in each category are as follows:

Work zone characteristics

- Freeway identifier (e.g. I-70, I-63, etc.)
- Work zone dates and location (mile post)
- Cost of the project
- Lane closure
- Duration of the work zone
- Length of the work zone

Crash characteristics

- Exact location of the crash
- Date and time of the crash
- Number of injuries, fatalities, etc.
- Number of vehicles involved
- Type of collision

Road and traffic characteristics

- Average daily traffic (ADT) or AADT with seasonal adjustment factor
- Number of lanes
- Number of intersections
- Percent of heavy vehicles

Figure 4-1 shows the schematic data collection process.

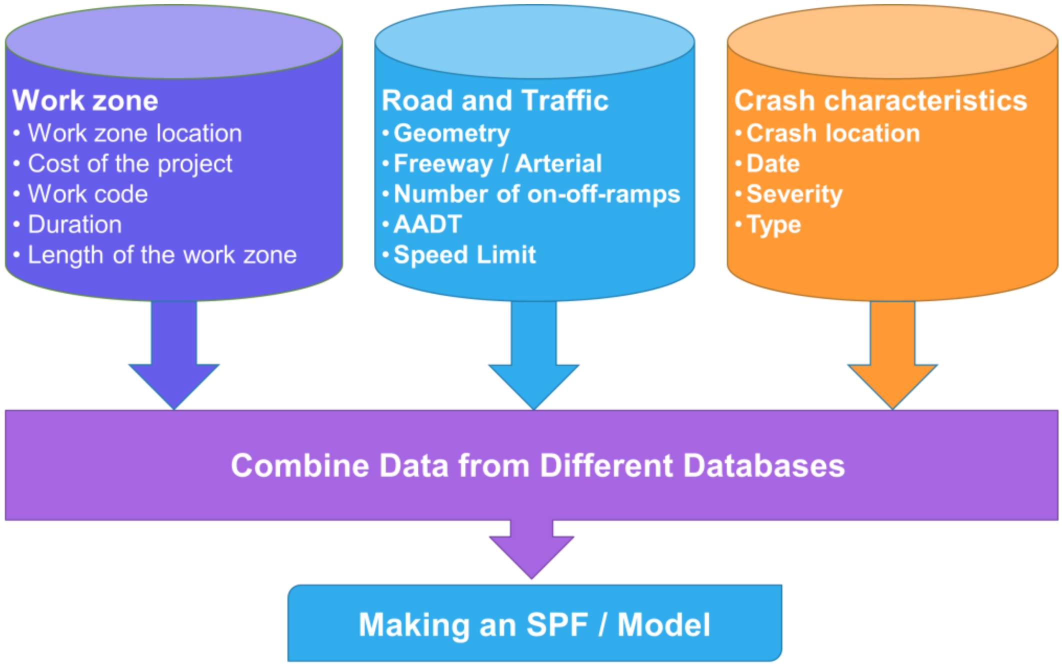


Figure 4-1. Schematic work zone studies data collection process

There are several variables that can be used in making a crash prediction model. At first glance, the inclusion of more variables leads to a more accurate model. However, a large number of variables has several disadvantages. First, it increases the data requirements and may require the collection of data that are not readily available. Second, it can lead to over fitting of the model (especially in big samples). An over fitted model has many parameters and describes random error instead of an underlying relationship between the variables. Such a model has a poor predictive performance and can exaggerate results of a small change in exposure variables. Third, the use of too many variables can make the model usage by practitioners more difficult due to the extensive data requirements, as some data are not readily available to practitioners.

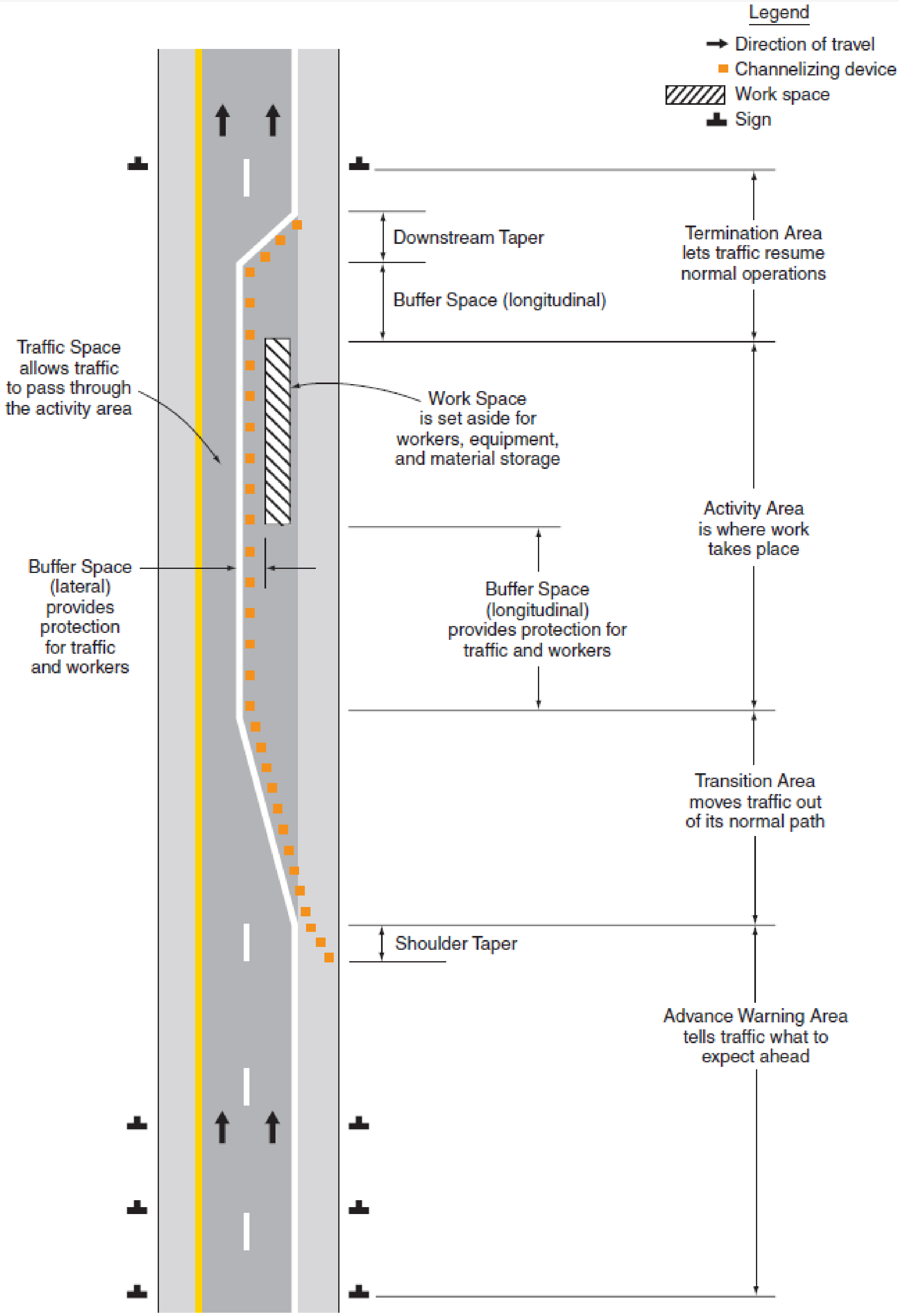
Data fusion was used to merge the information contained in the work zone, crash, and road segment databases. Databases need to be organized in a way that makes linking them possible. Missouri databases are indexed in a way that crashes, road segments and work zone databases could be linked in a proper way. Because the data fusion process is complex and time consuming, prior research has typically relied on small sample sizes for building work zone crash models. For example, Khattak et al. (2002) used a sample size of 36 work zones in building the model used in HSM. One contribution of this research is the use of a large sample size for model development. The sample sizes used in this study are significantly larger than those used in any of the published literature in work zone safety. This study used 1,546 freeway, 1,189 expressway and 6,095 rural two-way work zones to develop 15 different work zone safety models.

The work zone database included a unique work zone ID, a roadway segment ID, start and end date, time of work, and start and end location. The crash database contained archived highway patrol reports. Even though there is a column in crash reports indicating work zone presence, it was not relied upon, because it was based upon a police officer's judgment at the scene which could be inaccurate. Instead, tempo-spatial matching was used to match the crashes with the time and location of each work zone. Thus there are crashes that occurred in work zones that are not reported as work zone-related crashes in crash reports. A FHWA study (FHWA 1996) tested four work zones and found that as many as 77 percent of the crashes that occurring in these work zones were not coded as work zone-related crashes by police officers.

In work zone databases, the footprint of a work zone is recorded as the beginning and end of the work area. To account for the work zone signage areas, including the advance warning areas, this study used MUTCD recommended temporary traffic control plans' thresholds for freeway, expressway and rural two lane work zones. To this end, the road functional type, speed limit, lane width and area designation (urban-rural) were also collected. The process is explained further in the next section 4.2, Crash Assignment to Work Zones. In contrast, the model used by the HSM classified all crashes within 0.5 mile (0.8 km) of the beginning and 0.5 mile (0.8 km) after the end of the work zone as work zone crashes.

4.2 Crash Assignment to Work Zones (based on MUTCD)

As mentioned in the previous section, the crashes were matched to work zones based on MUTCD recommended distances. Work zones have five different parts: advance warning area, transition area, buffer area, activity area, and termination area (FHWA 2009a). This study considered activity and buffer areas together, and the remaining areas separate. Figure 4-2 shows the schematic plan of work zone parts.



FHWA 2009a

Figure 4-2. Work zone components based on MUTCD

Table 4-1 shows the information from the MUTCD that was used to compute the advanced warning area minimum distance. This minimum distance depends on the facility functional type, speed and work zone area urban-rural designation.

Table 4-1. Advanced warning area distances, MUTCD recommendations

Road Type	Distance Between Signs**		
	A	B	C
Urban (low speed)*	100 feet	100 feet	100 feet
Urban (high speed)*	350 feet	350 feet	350 feet
Rural	500 feet	500 feet	500 feet
Expressway / Freeway	1,000 feet	1,500 feet	2,640 feet

* Speed category to be determined by the highway agency

** The column headings A, B, and C are the dimensions shown in Figures 6H-1 through 6H-46. The A dimension is the distance from the transition or point of restriction to the first sign. The B dimension is the distance between the first and second signs. The C dimension is the distance between the second and third signs. (The "first sign" is the sign in a three-sign series that is closest to the TTC zone. The "third sign" is the sign that is furthest upstream from the TTC zone.)

Source: FHWA 2009a

The buffer distance only depends on the road speed limit as shown in Table 4-2. This space could be included in both the before and after work area of the work zone.

Table 4-2. Buffer area, MUTCD recommendations

Stopping Sight Distance as a Function of Speed	
Speed*	Distance
20 mph	115 feet
25 mph	155 feet
30 mph	200 feet
35 mph	250 feet
40 mph	305 feet
45 mph	360 feet
50 mph	425 feet
55 mph	495 feet
60 mph	570 feet
65 mph	645 feet
70 mph	730 feet
75 mph	820 feet

* Posted speed, off-peak 85th-percentile speed prior to work starting, or the anticipated operating speed

Source: FHWA 2009a

The transition area is based on the lane closure, speed, and the lane width. Based on the MUTCD work zone schematic plan, the shoulder taper is not in the transition area and is included in the advanced warning area. So there is no need to compute this distance. Table 4-3 shows the equation for computing transition areas. The distance needed to add to the start of the work area is the summation of the buffer area, transition area, and advance warning area.

Table 4-3. Transition and termination area, MUTCD recommendations

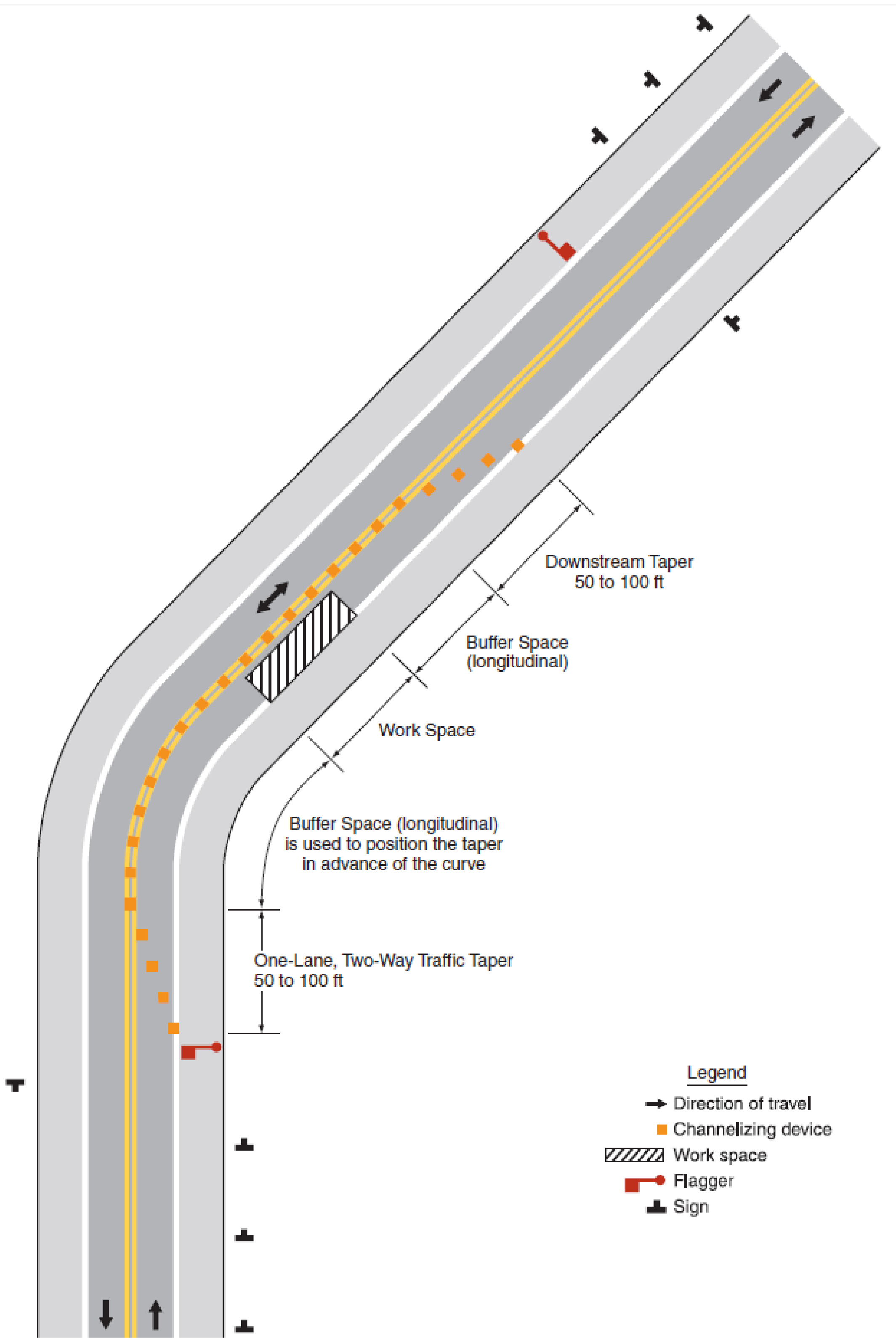
Taper Length Criteria for Temporary Traffic Control Zones	
Type of Taper	**Taper Length**
Merging Taper	at least L
Shifting Taper	at least 0.5 L
Shoulder Taper	at least 0.33 L
One-Lane, Two-Way Traffic Taper	50 feet minimum, 100 feet maximum
Downstream Taper	50 feet minimum, 100 feet maximum

Formulas for Determining Taper Length	
Speed (S)	**Taper Length (L) in feet**
40 mph or less	$L = \frac{WS^2}{60}$
45 mph or more	$L = WS$

L = taper length in feet
W = width of offset in feet
S = posted speed limit, or off-peak 85th-percentile speed prior to work starting, or the anticipated operating speed in mph

Source: FHWA 2009a

Downstream from each work zone work area are two different parts: buffer space and termination area. The buffer area is considered the same as the upstream buffer of the work zone, and the termination area is 50-100 ft for each closed lane. Figure 4-3 shows a schematic plan of a two-lane work zone.



FHWA 2009a

Figure 4-3. Rural two-lane schematic work zone parts, MUTCD

4.3 Sampling and Data Descriptive Statistics

The Missouri work zone database had 110,287 work zones between January of 2009 and December of 2014. Data for years before 2009 were available, but the crash rate of the years before 2009 was different than the years after 2009. The years between 2009 and 2014 seem not to be significantly different. The concern with using pre-2010 data, was that the difference in the crash rate was due to factors not captured in the available variables, e.g. the Great Recession.

Table 4-4 shows the number of work zones in each facility type.

Table 4-4. Number of work zones by facility type, 2009 to 2014

Operation Type	Number of WZs
3 LANE SECTION	474
5 LANE SECTION	3922
EXPRESSWAY	8993
FREEWAY	20873
MULTI-LANE	2300
ONE-WAY	216
RAMP	4083
SHARED FOUR LANE	226
SUPER 2-LANE	2191
TWO-LANE	64476
(blank)	2533
Grand Total	**110287**

The facilities having a large number (>5000) of work zones were freeway, expressway, and two lane roads. Freeway and expressway work zones were divided almost equally between urban and rural roads, while the two lane roads were mostly rural. So, three categories of models were developed for freeway, expressway and rural two lanes.

Most of work zones in the database were short length and short duration work zones, with no crashes. Table 4-5 shows the minimum, maximum and average duration, length and AADT of these work zones.

Table 4-5. Descriptive statistics of 110,287 Missouri work zones between 2009 and 2014

	Duration (day)	Length (mile)	AADT (veh/day)
Min	0.02	0.01	4
Max	1096.42	282.89	241418
Average	6.9	3.28	16990

Table 4-6 shows the length and duration of these 110,287 work zones.

Table 4-6. Distribution of 110,287 work zones by length and duration

	Duration (days)					
Length (miles)	**<30**	**30-120**	**120-210**	**210-300**	**>300**	**Grand Total**
0-2	61153	2867	303	118	104	64545
2-4	16083	495	44	35	20	16677
4-6	9284	305	20	6	13	9628
6-8	5956	230	7	13	6	6212
8-10	4031	133	14	7	17	4202
>10	8402	522	68	18	9	9019
Grand Total	**104913**	**4552**	**456**	**197**	**169**	**110287**

As can be seen in Table 4-6, more than half of the work zones (61153) were shorter than 2 miles with duration less than 30 days.

Among these 110,287 work zones, based on police officers' judgement, only 2,618 of 110,287 work zones contained at least 1 work zone related crash. Tables 4-7 and 4-8 show the distribution of work zones with no crashes and the work zones with at least 1 crash, respectively. Table 4-7 shows that most of "No-Crash-Work zones" are the short ones with duration less than a month.

Table 4-7. Distribution of work zones with no crashes by length and duration

	Duration (days)					
Length (miles)	**<30**	**30-120**	**120-210**	**210-300**	**>300**	**Grand Total**
0-2	60568	2707	255	94	85	63709
2-4	15841	424	22	17	5	16309
4-6	9085	251	10	4	5	9355
6-8	5824	180	4	7	3	6018
8-10	3904	95	7	1	9	4016
>10	7874	335	29	8	6	8252
Grand Total	**103100**	**3992**	**327**	**131**	**113**	**107663**

Table 4-8. Distribution of work zones with at least one crash by length and duration

Length (miles)	Duration (days)					
	<30	**30-120**	**120-210**	**210-300**	**>300**	**Grand Total**
0-2	579	160	48	24	19	830
2-4	242	71	22	18	15	368
4-6	199	54	10	2	8	273
6-8	132	50	3	6	3	194
8-10	127	38	7	6	8	186
>10	528	187	39	10	3	767
Grand Total	**1807**	**560**	**129**	**66**	**56**	**2618**

Table 4-9 shows the average crash count of the work zones with at least 1 crash (crash per work zone) indicated for different combinations of work zone duration and length.

Table 4-9. Average crash count per work zone (for work zones with at last one crash)

Length (miles)	Duration (days)					
	<30	**30-120**	**120-210**	**210-300**	**>300**	**Grand Total**
0-2	1.17	2.24	4.02	10.17	5.89	1.91
2-4	1.38	2.49	4.32	6.17	5.73	2.18
4-6	1.35	2.43	2.70	41.00	7.13	2.07
6-8	1.48	3.28	6.00	26.67	11.00	2.94
8-10	1.49	3.34	6.43	15.33	19.88	3.29
>10	1.66	4.63	16.79	22.80	37.33	3.56
Grand Total	**1.40**	**3.26**	**8.01**	**13.89**	**9.98**	**2.62**

The table shows that by increasing the duration and length of the work zone, the probability of having a higher crash frequency, increases too.

As discussed previously, most of the work zones do not have any crashes. Table 4-10 shows the number of work zones with different crash frequency for the years between 2009 and 2014.

Table 4-10. Number of work zones with different crash counts

Number of Crashes (n)	Work Zones with n crashes
0	103418
1	1719
2	384
3	156
4	91
5	64
6	35
7	35
8	18
9	15
>=10	101

Table 4-11 shows the average crash count (i.e. crashes per work zone) by severity for rural, urban, and urbanized areas for work zones with at least one crash.

Table 4-11. Average crash count based on severity and area designation

Area Designation	Number of Work Zones	Number of Crashes	Crashes per Work Zone			
			PDO	Minor Injury	Disabling Injury	Fatal
RURAL	1214	2254	1.461	0.451	0.138	0.029
URBAN	234	657	1.863	0.598	0.085	0.030
URBANIZED	1170	3958	2.405	0.721	0.068	0.003

An area is classified as rural if it has fewer than 5,000 people, urban if it has between 5,000 and 50,000 people, urbanized if it has between 50,000 and 250,000 people, and metropolitan if it has more than 250,000 people. in this study urban, urbanized and metropolitan work zones were categorized as one group named urban area. PDO and Minor Injury crashes increase from rural to urbanized area, while Disabling Injury and Fatal crashes decrease. Table 4-11 shows that there is significant difference between urban and rural work zones. So, the work zones were separated based on their urban-rural designation.

Another way to analyze the work zone crashes is to study them by time of occurrence. Table 4-12 indicates that most of the crashes occurred during the day. “(F+DI)/Grand Total ratio” indicates the percentage of severe crashes, which is more than 7% for 6 AM, 7 and 8 PM. It means that during these hours it is more probable to see a severe crash.

Table 4-12. Timely distribution of work zone crashes by severity

Time	FATAL	DISABLING INJURY	MINOR INJURY	PROPERTY DAMAGE ONLY	(F+DI)/Total crashes ×100	Total	Percentage
12 AM				10	0.0	10	0.15
1 AM				2	0.0	2	0.03
2 AM			1	1	0.0	2	0.03
3 AM			1	5	0.0	6	0.09
4 AM				1	0.0	1	0.01
5 AM			1	4	0.0	5	0.07
6 AM		5	5	36	10.9	46	0.67
7 AM	2	18	66	285	5.4	371	5.40
8 AM	8	20	86	334	6.3	448	6.52
9 AM	3	19	118	428	3.9	568	8.27
10 AM	6	23	129	496	4.4	654	9.52
11 AM	3	30	177	571	4.2	781	11.37
12 PM	6	26	220	582	3.8	834	12.14
1 PM	5	35	184	574	5.0	798	11.62
2 PM	3	22	200	615	3.0	840	12.23
3 PM	4	25	158	453	4.5	640	9.32
4 PM	5	21	80	293	6.5	399	5.81
5 PM	1	12	68	211	4.5	292	4.25
6 PM		6	24	81	5.4	111	1.62
7 PM		4	6	30	10.0	40	0.58
8 PM		1	5	8	7.1	14	0.20
9 PM			3	2	0.0	5	0.07
10 PM				2	0.0	2	0.03
Grand Total	**46**	**267**	**1532**	**5024**		**6869**	**100**

The following two figures, Figures 4-4 and 4-5, respectively, show the percent of severe crashes that occurred in work zones and the overall temporal distribution of work zone crashes based on police judgment. Figure 4-4 shows that 6 a.m. and 8 p.m. are the most likely times to observe severe work zone crashes. Two possible reasons for this result include the presence of peak hour and changes in light conditions. Figure 4-6 depicts the monthly distribution of work zone related crashes between 2009 and 2014.

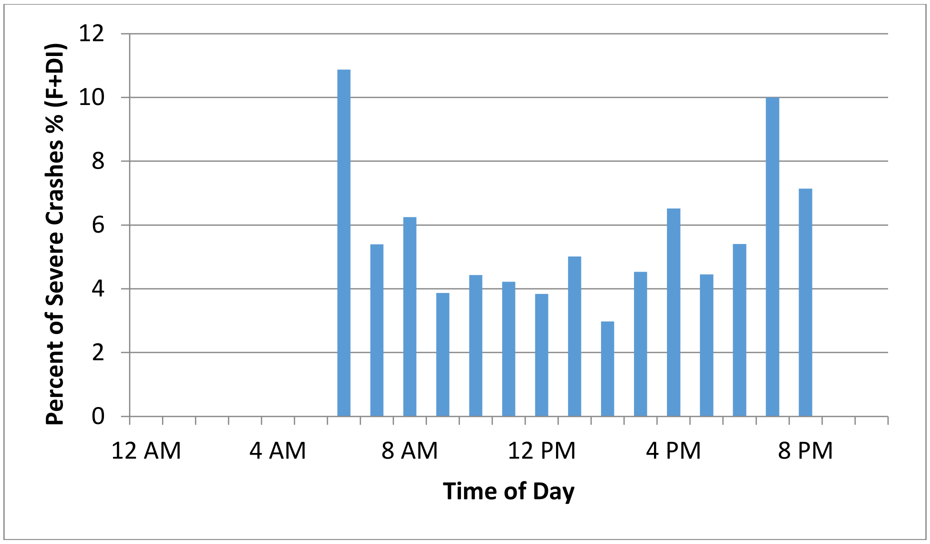


Figure 4-4. Percent of fatal/disabling-injury crashes by time of day

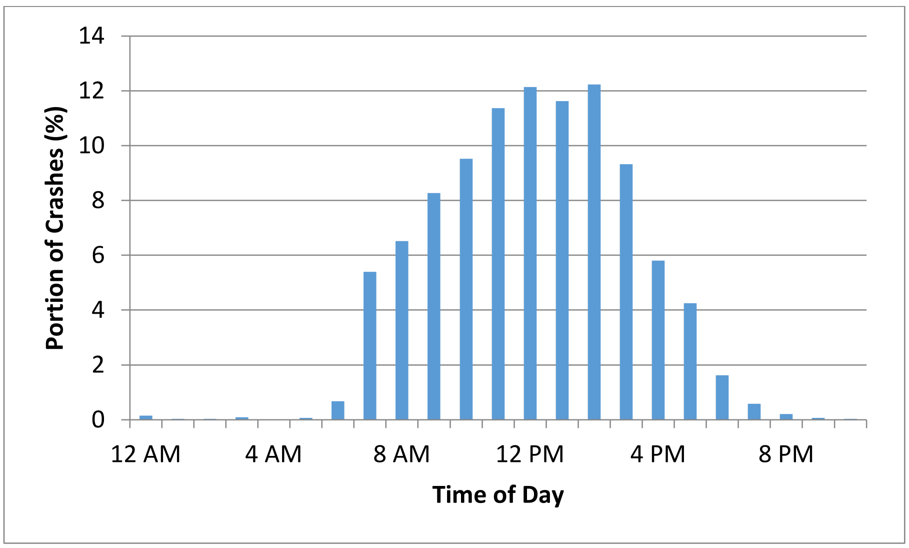


Figure 4-5. Crash distribution by time of day

All the crash statistics above were calculated based on the population of 110,287 work zones between 2009 and 2014 in Missouri. It is noteworthy that the crash statistics were based on police judgment to determine if the crashes were work zone related. In the next sections, the statistics are based on tempo-spatial crash matching using MUTCD mentioned work zone thresholds for freeways, expressways and rural two-lane two-way roads.

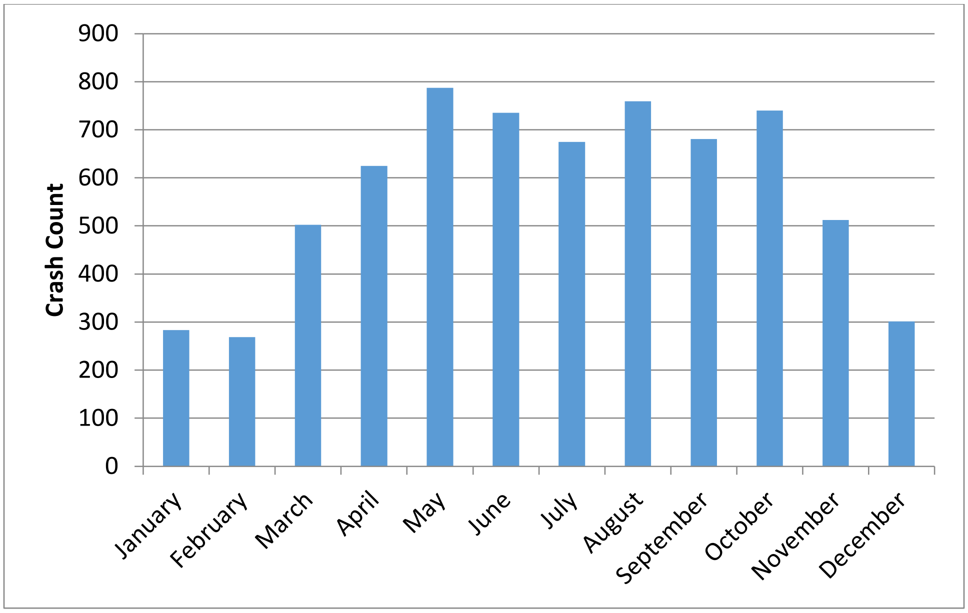


Figure 4-6. Monthly distribution of work zone crashes between 2009 and 2014

4.3.1 Freeway Work Zones

There were 20,873 freeway work zones in Missouri between January of 2009 and December of 2014. As mentioned before, most of these work zones are "small work zones" with short duration and low crash frequencies. Modeling crashes by including all the small work zones is possible, but the high uncertainty of the predictions in the model would limit the usability of the developed models. One solution is to exclude the work zones with short length and duration based on a pre-determined threshold. By dropping these work zones, the sample size would become smaller but would include a greater percentage of work zones with crashes. So, there is a tradeoff between the minimum length and duration of work zones and the resulting sample size.

To find the optimum thresholds for length and duration, different thresholds were tested to find the corresponding sample sizes. Also different models were fitted to the resulting sample, and the average overdispersion was calculated. As will be explained later in the Methodology section of this report, a smaller overdispersion means a more accurate model. The results of different thresholds for work zone duration are in Figures 4-7 and 4-8, and the optimum threshold is at a minimum duration of 10 days.

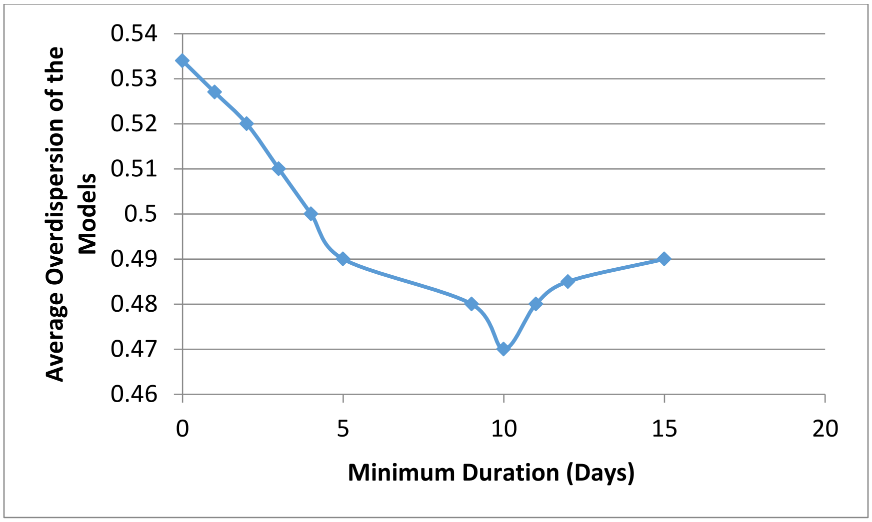


Figure 4-7. Average overdispersion versus minimum duration

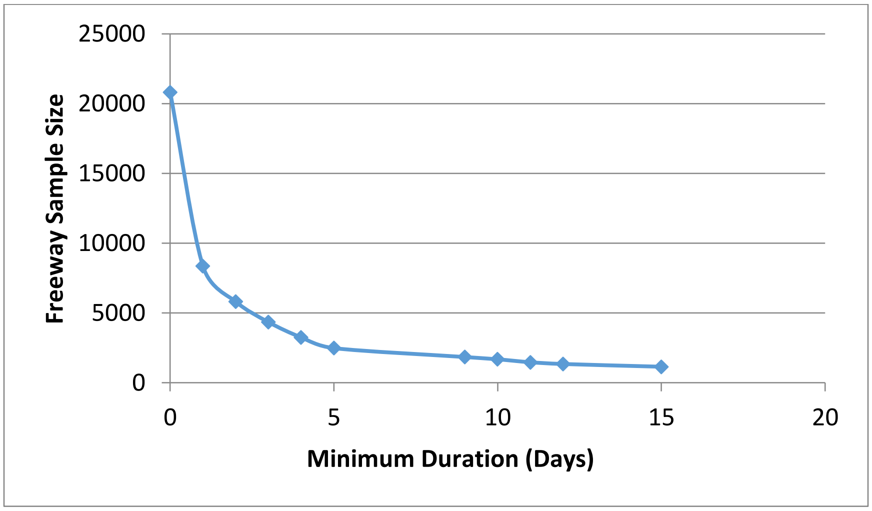


Figure 4-8. Freeway sample size versus minimum duration

A similar process was used for determining the minimum threshold for work zone length. The results in Table 4-13 shows that by using the minimum length of 0.1 mile, the overdispersion decreases from 0.53 to 0.5. By increasing the threshold more than this value, the overdispersion term does not change considerably. So, work zones shorter than 0.1 mile and with duration of fewer than 10 days were omitted. In practical terms, very small work zones have very little traffic and safety impact, thus there is less of a need for using a safety tool for analyzing such work zones.

Table 4-13. Average overdispersion of freeway models and their sample size using different minimum length thresholds

Minimum Length of Work Zones (miles)	Average Overdispersion of Models	Freeway Sample Size
---	0.53	20808
0.1	0.5	19436
0.2	0.5	17460
0.3	0.49	16760
0.4	0.49	15595
0.5	0.483	14000

Table 4-14 shows the descriptive statistics of the sample of 1,546 freeway work zones used in this study.

Table 4-14. Descriptive statistics of the freeway work zone sample

Length, Duration and AADT			
Variables	Average	Min	Max
Length of work zone, mi (km)	5.048 (8.125)	0.101 (0.163)	29.920 (48.151)
AADT (vehicles per day)	30,531	757	128,756
Work Zone Duration (days)	43.4	10	290
Urban/rural percent	69% / 31%		
Number of observations	1,546		

Crashes			
Number of Crashes	All crashes	PDO	Fatal-Injury
Sum	9,199	6,975	2,224
Average	5.950	4.152	1.439
Min/max	0/175	0/136	0/39

The sample contained work zones longer than 0.1 miles and with a duration of more than 10 days, thus decreasing the number of work zones without any crashes. The average length and duration from this reduced sample were 5.048 miles and 43.4 days, respectively. The AADT of the samples ranged from 757 to 128,756 veh/day with an average of 30,531 veh/day.

Table 4-15 depicts the number and percentage of crashes that occurred in the four parts of work zones: advanced warning area, transition area, work and buffer area, and termination area.

Table 4-15. Freeway work zone crash location analysis

Number of Closed Lanes	Severity		Freeway			
			Advanced Warning Area	Transition Area	Work and Buffer Area	Termination Area
0	PDO	Count	266	0	2255	0
		Percent	10.55	0.00	89.45	0.00
	Fatal-Injury	Count	84	0	692	0
		Percent	10.82	0.00	89.18	0.00
1	PDO	Count	795	255	6347	22
		Percent	10.72	3.44	85.55	0.30
	Fatal-Injury	Count	260	84	2055	6
		Percent	10.81	3.49	85.45	0.25
>1	PDO	Count	227	160	1618	15
		Percent	11.24	7.92	80.10	0.74
	Fatal-Injury	Count	86	39	496	1
		Percent	13.83	6.27	79.74	0.16

Between 79.74% and 89.45% of the work zone crashes occurred in work and buffer area, depending on the number of closed lanes. The percent of crashes in advance warning area varied between 10.55% and 13.83%, depending on the number of closed lanes. By increasing the number of closed lanes, the percent of PDO work zone crashes in transition area increased from 0.0% to 7.92%. Termination area crashes were not significantly noteworthy.

4.3.2 Expressway Work Zones

Table 4-16 shows the descriptive statistics of the sample of 1,189 expressway work zones used in this study.

Table 4-16. Descriptive statistics of the expressway work zone sample

Length, Duration and AADT			
Variables	Average	Min	Max
Length of work zone, mi (km)	4.074 (6.557)	0.107 (0.172)	29.606 (47.474)
AADT (vehicles per day)	8,767	713	34,744
Work Zone duration (days)	51.3	10.3	298.3
Urban/rural percent	51% / 49%		
Number of observations	1,189		

Crashes			
Number of Crashes	All crashes	PDO	Fatal-Injury
Sum	3,047	1,624	591
Average	2.563	2.707	0.985
Min/max	0/74	0/42	0/32

The sample contained work zones longer than 0.1 miles with a duration of more than 10 days. The average length and duration were 4.074 miles and 51.3 days, respectively. The AADT of the samples ranged from 713 to 34,744 veh/day with an average of 8,767 veh/day.

Table 4-17 depicts the number and percent of crashes that occurred in the four parts of work zones: advanced warning area, transition area, work and buffer area, and termination area.

Table 4-17. Expressway work zone crash location analysis

Number of Closed Lanes	Severity		Expressway Advanced Warning Area	Expressway Transition Area	Expressway Work and Buffer Area	Expressway Termination Area
0	PDO	Count	96	0	358	0
		Percent	21.15	0.00	78.85	0.00
	Fatal-Injury	Count	24	0	104	0
		Percent	18.75	0.00	81.25	0.00
1	PDO	Count	525	81	2415	10
		Percent	17.32	2.67	79.68	0.33
	Fatal-Injury	Count	188	29	871	8
		Percent	17.15	2.65	79.47	0.73
>1	PDO	Count	30	4	84	7
		Percent	24.00	3.20	67.20	5.60
	Fatal-Injury	Count	12	3	25	2
		Percent	28.57	7.14	59.52	4.76

Between 78.85% and 81.25% of the work zone crashes occurred in work and buffer area. This result may be due to the higher lengths of the work and buffer area compared to the lengths of the other three parts of the work zone. The percentage of crashes in the advance warning area was between 17.15% and 28.57%, depending on the number of closed lanes. By increasing the number of closed lanes, the percentage of crashes in the transition area increased from 0.0% to 7.14%. Termination area crashes were not significantly noteworthy for zero and one closed lane; the percentage of PDO and fatal-injury crashes for work zones with more than one closed lane were 5.60% and 4.76% respectively. The percent of expressway work zone crashes in the work area was less than the freeway, while transition area had more crashes comparing to freeways.

4.3.3 Rural Two-Lane Work Zones

Table 4-18 shows the descriptive statistics of the sample of 6,095 rural two-lane work zones used in this study.

Table 4-18. Descriptive statistics of the rural two-lane work zone sample

Length, Duration and AADT			
Variables	Average	Min	Max
Length of work zone, mi (km)	5.803 (9.339)	0.1 (0.161)	29.897 (48.114)
AADT (vehicles per day)	778.6	50	10,325
Work Zone duration (days)	30.9	10	300
Number of observations	6,095		

Crashes			
Number of Crashes	All crashes	PDO	Fatal-Injury
Sum	1,077	1,077	552
Average	0.267	0.177	0.091
Min/max	0/32	0/23	0/9

The sample contained work zones longer than 0.1 miles with a duration of more than 10 days. The average length and duration were 5.803 miles and 30.9 days, respectively. The AADT of the samples ranged from 50 to 10,325 veh/day with an average of 778.6 veh/day.

Table 4-19 depicts the number and percent of crashes occurred in the four parts of work zones: advanced warning area, transition area, work and buffer area, and termination area.

Table 4-19. Rural two-lane work zone crash location analysis

Number of Closed Lanes	Severity		Rural Two-Lanes Advanced Warning Area	Transition Area	Work and Buffer Area	Termination Area
0	PDO	Count	9	4	394	4
		Percent	2.19	0.97	95.86	0.97
	Fatal-Injury	Count	2	1	146	1
		Percent	1.33	0.67	97.33	0.67
1	PDO	Count	49	10	2414	21
		Percent	1.96	0.40	96.79	0.84
	Fatal-Injury	Count	22	5	1148	3
		Percent	1.87	0.42	97.45	0.25

Between 95.86% and 97.45% of the work zone crashes occurred in the work and buffer area. The percentage of crashes in advance warning area was between 1.33% and 2.19%, depending on the lane closure. Transition and termination area crashes were not significantly noteworthy. Thus for rural work zones, the overwhelming majority of crashes occurred in the work and buffer areas.

4.3.4 Data from Other States

The most challenging part of work zone safety studies is finding good data sources. In addition to Missouri work zone data, the following state DOTs were contacted to determine if they have suitable safety and work zone data: Florida, Idaho, Illinois, Maine, Michigan, New Hampshire, North Carolina, Ohio, Pennsylvania, South Carolina, South Dakota, Tennessee, Virginia, West Virginia, Wyoming, Indiana, Minnesota, Nebraska, Iowa, and Kansas.

The data received from other states were in a format that could not easily be used for work zone safety modeling. New Hampshire data was descriptive statistics of 33 work zone crashes. Iowa DOT sent the list of crashes happened in Iowa. However, the Iowa crash data could not be used for work zone safety modeling due to the lack data regarding work zone characteristics. Most states do not have data in a form to make matching the crashes with work zones possible. Among the states that responded to the request, only Ohio and Kansas data were suitable for the purpose of developing work zone models. However, the number of work zones provided were not enough to develop separate models. However, data from these two states could be used for calibrating the freeway models made by Missouri data.

New Hampshire sent descriptive statistics for 33 freeway work zones. Three figures, Figures 4-9, 4-10 and 4-11, show the hourly, time of day and monthly distribution of the work zone related crashes, respectively.

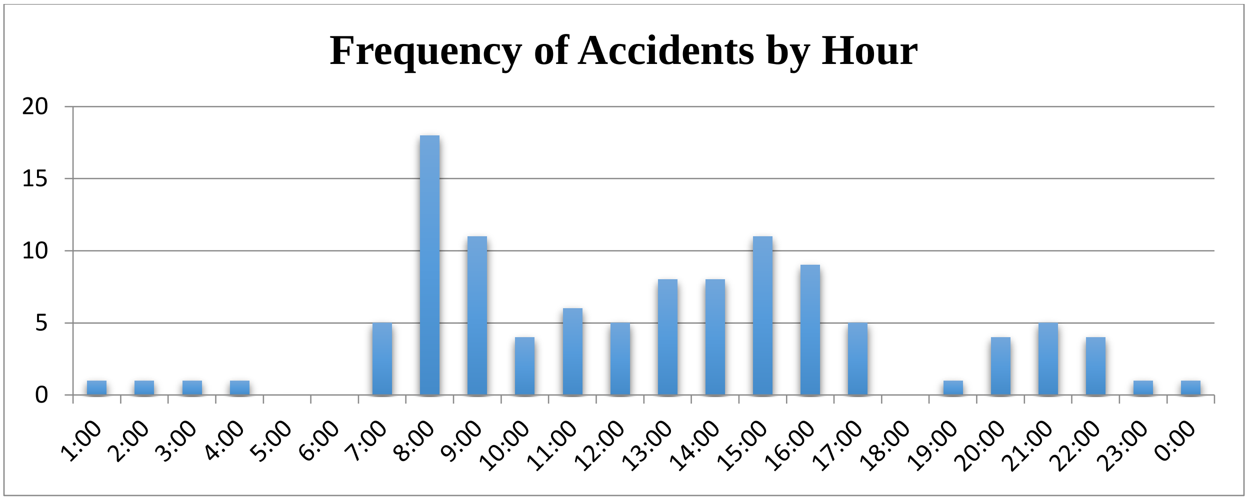


Figure 4-9. Crash distribution by time

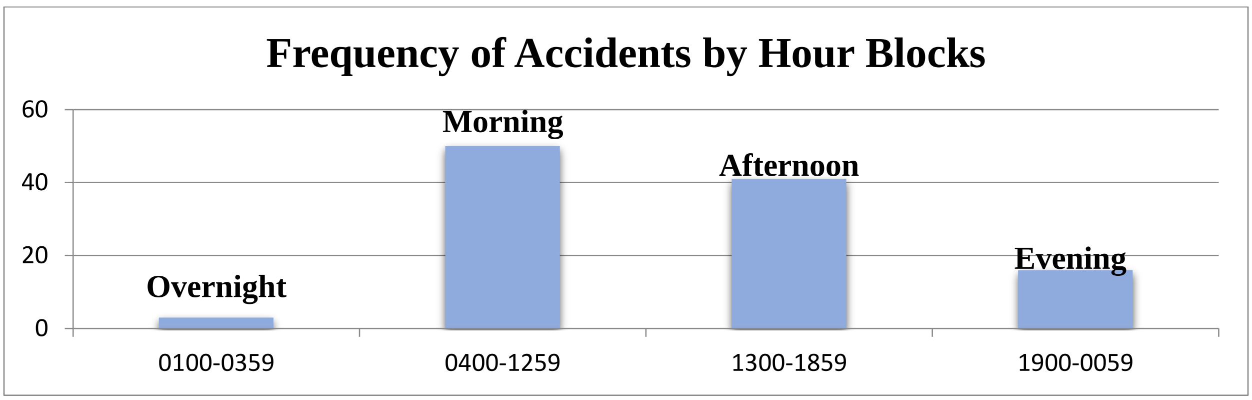


Figure 4-10. Crash distribution by time of day

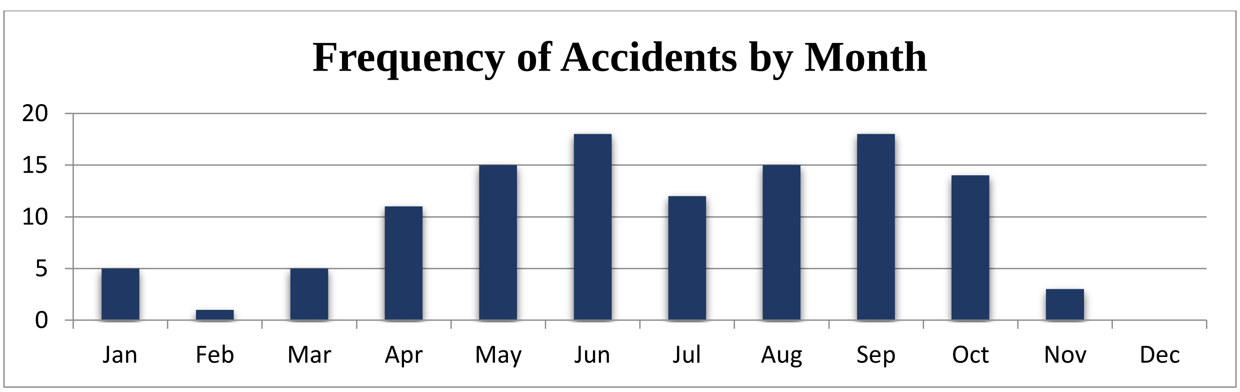


Figure 4-11. Monthly crash distribution

As a comparison, Missouri crashes occurred more frequently on 6 a.m., 7 and 8 p.m., while New Hampshire crashes were more frequently seen around 8 a.m. and 3 p.m. Most of New Hampshire crashes happened during the morning and afternoon. In New Hampshire most of work zone related crashes happened between April and December which is similar to Missouri.

5 MODEL ESTIMATION MEHODOLOGY

Negative Binomial (NB), Zero Inflated Negative Binomial (ZINB), Poisson and Zero Inflated Poisson (ZIP) were investigated as distributions for modeling the dependent variable, crash counts. Akaike Information Criterion (AIC) result was the best for Negative Binomial compared to the other models of ZINB, Poisson and ZIP distributions. The Akaike Information Criterion (AIC) for a given set of data, is a measure of the relative quality of statistical models. The smaller the AIC, the better. So, NB distribution was used for modeling. Most of the existing work zone safety studies used NB models (e.g. Pal and Sinha 1996; Venugopal and Tarko 2000; Tarko and Venugopal 2001; Khattak et al. 2002; Srinivasan et al. 2008; Ozturk et al., 2013; Yang et al. 2013; Sun et al. 2014b).

The NB model, which is the most commonly used model for work zone crash frequency, is explained as follows. Total crashes can be considered the result of a series of Bernoulli trials. Using Bernoulli terminology, the occurrence of a crash is considered a “success” and the alternative a failure. The use of this statistical terminology does not mean that crashes are positive phenomena.

For Y_i independent trials or crashes, there are y_i observed crashes, a negative binomial distribution is appropriate when Y_i is large enough and is given the form of:

$$\mathrm{P}(Y_i = y_i) = \binom{Y_i}{y_i} p^{y_i}(1-p)^{Y_i - y_i}$$

where $\mathrm{p} = \lambda_i / Y_i$, and the negative binomial distribution can be approximated as a Poisson distribution (Rouphail et al. 1988):

$$\mathrm{P}(Y_i = y_i) \cong \frac{\lambda_i}{Y_i\ !} e^{-\lambda_i}$$

If *i* represents a work zone with a specific duration and length, then λ_i is the expected crash frequency of that work zone *i*. Y_i and y_i are all natural numbers.

The explanatory variable x_i is introduced into λ_i (Khattak et al. 2002; Chen and Tarko 2014):

$$\lambda_i = e^{(\beta x_i + \varepsilon_i)}$$

where ε_i is error term, and is used to account for errors such as an omitted explanatory variable. For the negative binomial model, e^{ε_i} is assumed to have a gamma distribution with mean 1 and variance α^2.

In Generalized Linear Models, overdispersion is a situation where the variance of the crash frequency data exceeds the mean (Salkind 2006). If the overdispersion condition exists, then the negative binomial model form should be used instead of the Poisson. With the additional parameter α, the natural form of overdispersion is:

$$\mathrm{Var}[y_i] = \mathrm{E}[y_i]\{1 + \alpha\mathrm{E}[y_i]\}$$

and overdispersion parameter is:

$$\alpha = \frac{\mathrm{Var}[y_i]}{\mathrm{E}[y_i]^2} - \frac{1}{\mathrm{E}[y_i]} = \frac{\mathrm{Var}[y_i] - \lambda_i}{{\lambda_i}^2}$$

and if α is not statistically different from zero, Poisson model is more appropriate than negative binomial. A small α leads to a small variance and a more accurate model. The models in this study were all estimated using the well-known maximum likelihood method.

Many previous studies assumed a constant overdispersion term for NB models; however, a constant overdispersion parameter gives too much weight to short segments (Hauer 2001; Heydecker and Wu 2001). Instead, the overdispersion should be dependent on segment length, i.e. $\frac{\alpha_0}{L}$ where L is length of the segment. Work zones have finite durations, thus the weight of work zones with short duration should also be reduced. The authors recommend that the overdispersion parameter be a function of both work zone length and duration:

$$\alpha = \frac{\alpha_0}{L \times D}$$

where L is the work zone length and D is the duration. The overdispersion parameter, α, is unitless. This study uses three different overdispersion terms as constant $\alpha = \alpha_0$, modified by length $\alpha = \frac{\alpha_0}{L}$ and modified by length and duration as $\alpha = \frac{\alpha_0}{L * D}$. Thus the overdispersion is assumed to be a function of length and duration. The smallest overdispersion means the most accurate model. The safety tool chooses the best model, defined by overdispersion value, based on the input information entered by user.

5.1 Functional Form

Most of the SPFs in safety literature have used length, duration and AADT of the segment as effective exposure variables. Figures 5-1 and 5-2 show the increasing trend of work zone crash frequency by increasing values of these three variables.

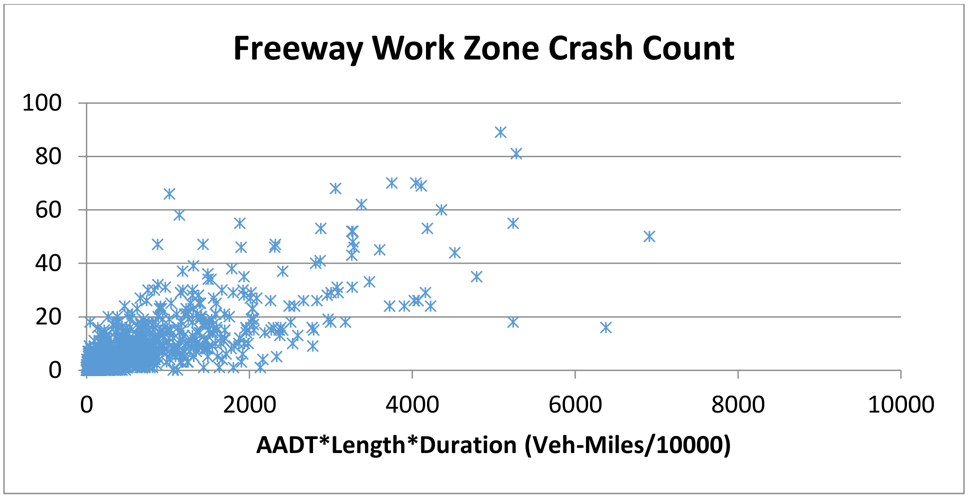


Figure 5-1. Freeway work zone crash trend versus product of AADT, length, and duration

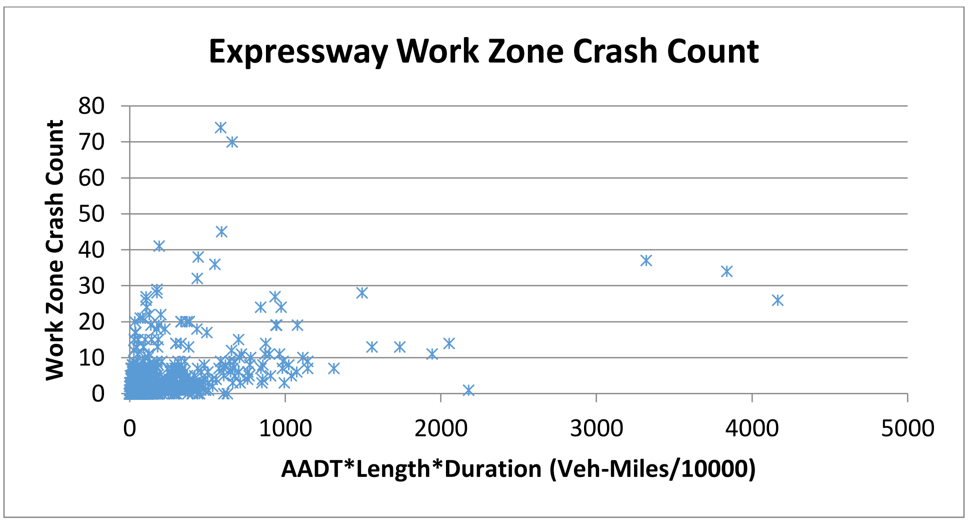


Figure 5-2. Expressway work zone crash trend versus product of AADT, length, and duration

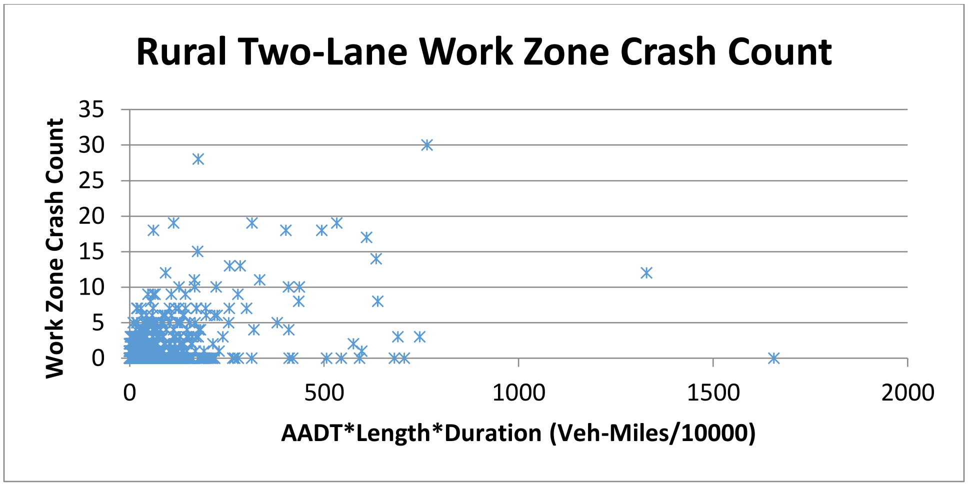


Figure 5-3. Rural two-lane work zone crash trend versus product of AADT, length, and duration

The figures show that the probability of having more crashes increases by Length*Duration*AADT. However, the increasing trend does not have a linear behavior. As these three variables are not the only characteristics of a work zone segment, some of the points do not follow the overall trend; especially for expressway and rural two-lane work zones. So, further analysis and data mining was needed.

Figures 5-4, 5-5, and 5-6 show the trend of work zone crash count based on work zone AADT group, from the aforementioned samples.

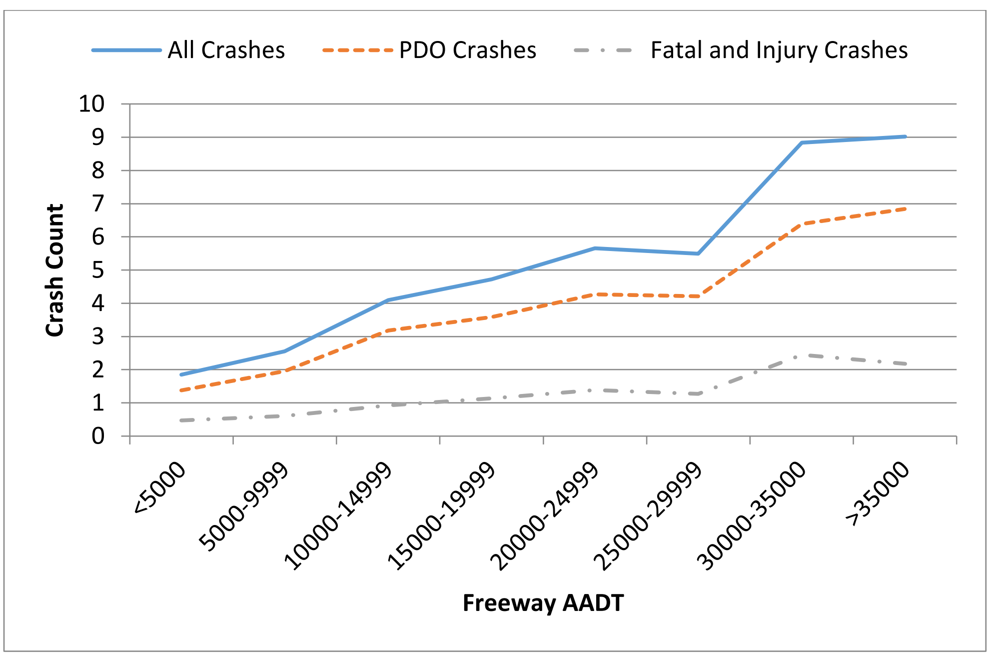


Figure 5-4. Freeway work zone crash trend versus AADT

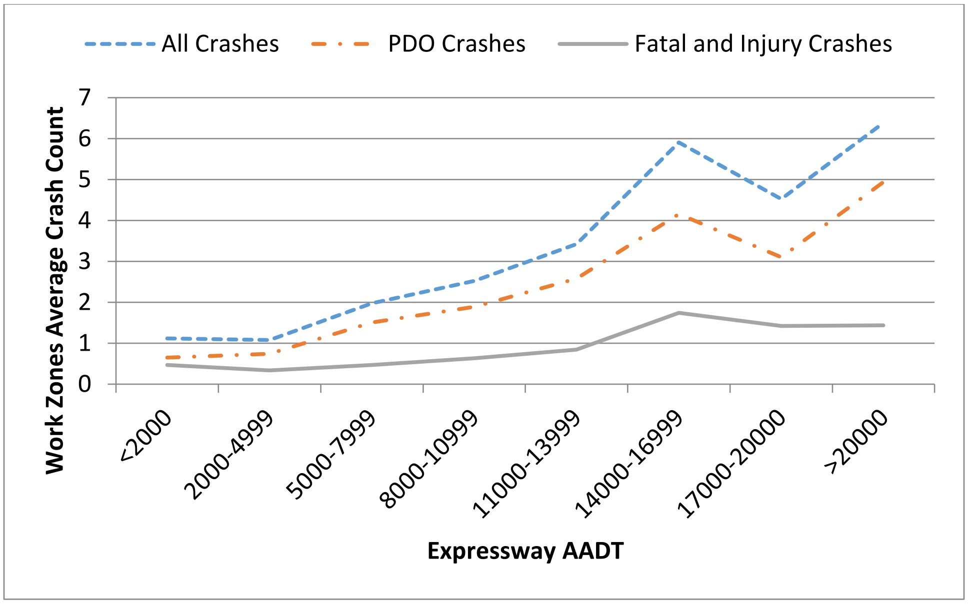


Figure 5-5. Expressway work zone crash trend versus AADT

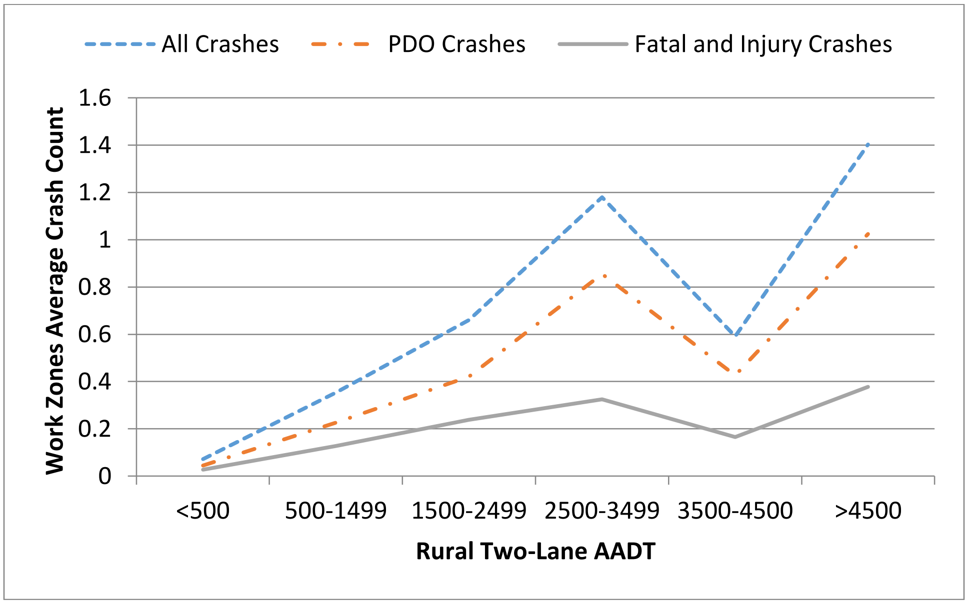


Figure 5-6. Rural two-lane work zone crash trend versus AADT

There is an increasing trend for both PDO and fatal-injury crashes by AADT for all three road functional types. The trend does not seem to be linear. So, the proper functional form for AADT could be a power function: $AADT^{\beta_1}$. For freeway work zones, the number of work zone crashes increases at approximately 30,000 vpd. This increase could be due to a variety of factors such as an increase in the number of available lanes, increased congestion and queuing, or higher interchange density in urban areas.

It is obvious that by increasing the length and duration of the work zones, the probability of a work zone crash occurrence increases. However, the trend could be nonlinear. So the same power functional form is suggested for work zone length and duration.

Many work zone crash prediction models used AADT, length and duration of work zone (Pal and Sinha 1996; Elias and Herbsman 2000; Venugopal and Tarko 2000; Tarko and Venugopal 2001; Khattak et al. 2002; Ozturk et al. 2013; Ozturk et al. 2014; Yang et al. 2013; Sun et al. 2014b) as explanatory variable. Some studies (Venugopal and Tarko 2000; Tarko and Venugopal 2001; Khattak et al. 2002; Srinivasan et al. 2008; Sun et al. 2014b) used urban/rural classification. In utilizing the knowledge of previous studies, the final models functional forms were:

- All variables included freeway combined model for freeway fatal-injury or PDO crashes:

$$N_C = e^{\beta_0} AADT^{\beta_1} L^{\beta_2} D^{\beta_3} e^{\beta_4 * \frac{Closed\ Lanes}{Number\ of\ Lanes}} e^{\beta_5 \frac{On-ramps}{L}} e^{\beta_6 \frac{Off-ramps}{L}} e^{\beta_7 Urban} e^{\beta_8 Injury}$$

- Expressway combined model

$$N_C = e^{\beta_0} AADT^{\beta_1} L^{\beta_2} D^{\beta_3} e^{\beta_4 * \frac{Signal}{L}} e^{\beta_5 Urban} e^{\beta_6 Injury}$$

- Rural two-lane combined model

$$N_C = e^{\beta_0} AADT^{\beta_1} L^{\beta_2} D^{\beta_3} e^{\beta_4 * \frac{Signal}{L}} e^{\beta_5 Injury}$$

- Rural two-lane PDO crash model

$$N_{PDO} = e^{\beta_0} AADT^{\beta_1} L^{\beta_2} D^{\beta_3} e^{\beta_4 * \frac{Signal}{L}}$$

- Rural two-lane fatal-injury crash model

$$N_{Injury} = e^{\beta_0} AADT^{\beta_1} L^{\beta_2} D^{\beta_3} e^{\beta_4 * \frac{Signal}{L}}$$

where the variables are as follows:

N_C	Number of fatal-injury or PDO crashes, based on Injury variable;
N_{PDO}	Number of PDO crashes;
N_{Injury}	Number of fatal-injury crashes;
AADT	Annual Average Daily Traffic (vehicles/day);
D	Duration of observation (days);
L	Segment length (miles);
Closed Lanes	Number of closed lanes in the work zone;

Number of Lanes	Number of lanes in the segment;
On-ramps	Number of on-ramps in the work area of work zone;
Off-ramps	Number of off-ramps in the work area of work zone;
Signal	Number of signalized intersections in the work area of work zone;
Urban	Dummy variable for work zone location, 1= urban, 0 = rural;
Injury	Dummy variable for crash severity, 1 = fatal-injury, 0 = PDO;

Variables were added sequentially, and maximum likelihood was used to estimate parameters.

6 MODEL RESULTS

This chapter summarizes the final results of modeling for three different road functional types: freeway, expressway and rural two-lane work zones. All of the models were developed using a variable-added-in-order method. In this method variables are added to the model one by one; in each stage a variable that improves the model most is added, then significance of variable and the resultong overall model's improvement are tested. If both statistical tests are passed the variable remains in the model, otherwise it is dropped. This process continues for adding other variables. Adding all the variables in all the 15 final models significantly improved the models' performance. The variables that were not significant were dropped from final models.

6.1 Freeway Work Zone Models

This section shows different models that were fitted to the freeway work zone sample with different overdispersion terms and functional forms.

6.1.1 Model 1, Freeway Work Zones with $\boldsymbol{\alpha} = \boldsymbol{\alpha_0}$

This model was made by considering a constant overdispersion. Table 6-1 summarizes the estimated parameters of Model 1 with following functional form:

$$N_C = e^{\beta_0} AADT^{\beta_1} L^{\beta_2} D^{\beta_3} e^{\beta_4 * \frac{Closed\ Lanes}{Number\ of\ Lanes}} e^{\beta_5 Urban} e^{\beta_6 Injury}$$

Table 6-1. Model 1 parameters for fatal-injury or PDO crashes

Explanatory Variable	Parameter Estimates	Standard Error	p-value
Constant	-12.4009	0.3328	<.0001
AADT	0.8826	0.02982	<.0001
L	0.6043	0.01683	<.0001
D	1.0085	0.02077	<.0001
$\frac{\text{Closed Lanes}}{\text{Number of Lanes}}$	0.2322	0.08103	0.0042
Urban	0.3841	0.05399	<.0001
Injury	-1.1394	0.03855	<.0001
Overdispersion, $\boldsymbol{\alpha_0}$	0.3536	0.02346	
Number of Observations		1546	

The model estimates the number of PDO crashes by substituting Injury variable of zero and fatal-injury crashes of 1. Each variable added was statistically beneficial to the model (using the $\aleph^2$ test) and all explanatory variables were statistically significant in the end (i.e. p-value<0.01). From this crash prediction model, CMFs can be derived for all explanatory variables. A 1% increase in AADT, length, and duration leads to the number of crashes increasing by 0.88%, 0.60% and 1.01%, respectively. However, deriving the CMFs from crash models by linearization of a crash prediction model is a controversial method (Hauer 1997; Ozturk et al. 2014). The HSM used this method to extract the work zone CMFs from Khattak's model (Khattak et al. 2002). Since $e^{(0.3841)} = 1.468$, this means urban road segments have 1.468 times the frequency of crashes in comparison to rural roads. The frequency of injury crashes is 32.00% of PDO crashes as calculated by $e^{(-1.1394)} = 0.3200$. The overdispersion is 0.35 which is acceptable. Most of the safety models in the literature have an overdispersion between 0.2 and 0.35 (Ozturk et al. 2013; Ozturk et al. 2014; Srinivasan and Carter 2011). A small α leads to a small standard deviation and better accuracy, but there are some studies with α around 0.8 (Venugopal and Tarko 2000). In road segment safety modeling the duration of the study for each segment is decided by the modeler who can choose even a couple of years. However, work zone studies have smaller sample size and the duration is restricted to work zone plan and phasing. Work zone safety modeling is subjected to this constraint; consequently, sometimes it is subjected to larger overdispersion term, too.

An ideal model is one that predicts the same value as observed, but in practice a model's prediction differs from the observed value. The difference between predicted value and observed value is called residual. One possible way to see how the residuals are distributed with respect to continuous independent variables is to plot cumulative residuals versus that variable. Hauer and Bamfo (1997) and Hauer (2004) suggested the use of Cumulative Residuals (CURE) plot. To make a CURE plot, the sample should be sorted in ascending order with respect to the target exposure variable. Then the cumulative residuals should be computed from the beginning of the sample to each member of the sample. CURE plots should oscillate like a random walk around zero. If CURE plot is decreasing for a range of a variable, it means that the model is overpredicting the results. An increasing CURE plot for the range of a variable indicates underestimation.

As CURE plot is sum of random variables (crash predictions) it is approximately normally distributed (Hauer 2015). In a normal distribution, about 95% of the probability mass should lie between two standard deviations from the mean. So the CURE plot should rarely go beyond the two confidence limits ($\mp 2\sigma^*$). With the same reasoning if significantly more than 40% of CURE plot lie between half of standard deviation limits ($\mp 0.5\sigma^*$), the danger of overfitting problem presents. In an overfitted model, variables coefficients do not show the underlying relationships and a small change in one independent variable could result in an exaggerated change in the dependent variable.

Figures 6-1, 6-2 and 6-3 show the Model 1 CURE plots for AADT, length and duration, respectively.

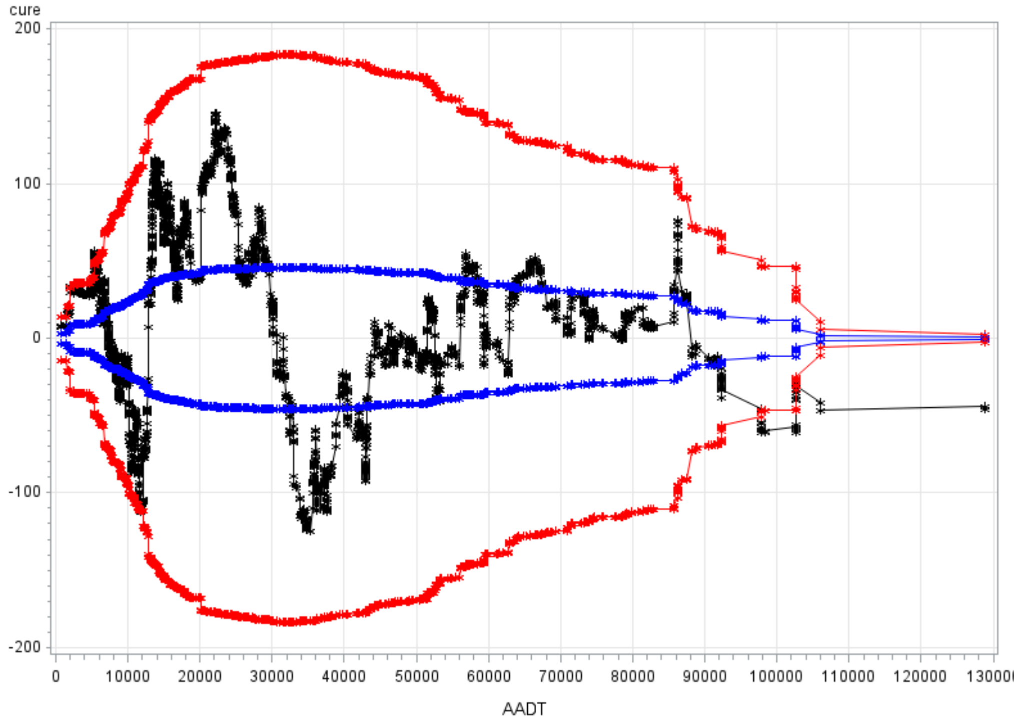


Figure 6-1. Model 1 AADT CURE plot

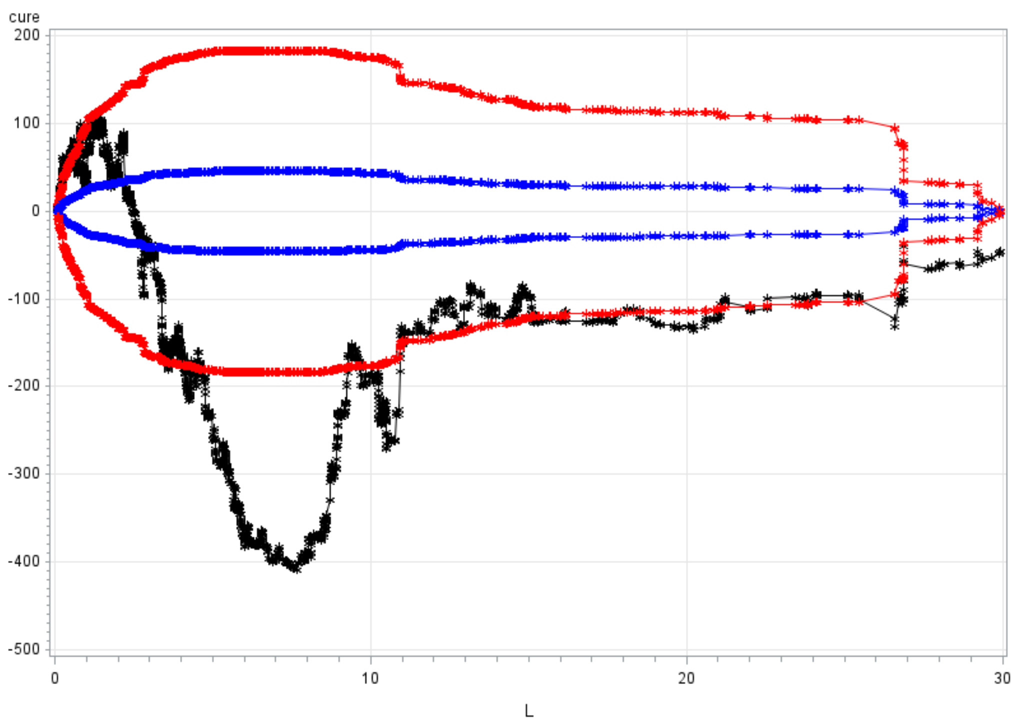


Figure 6-2. Model 1 length CURE plot

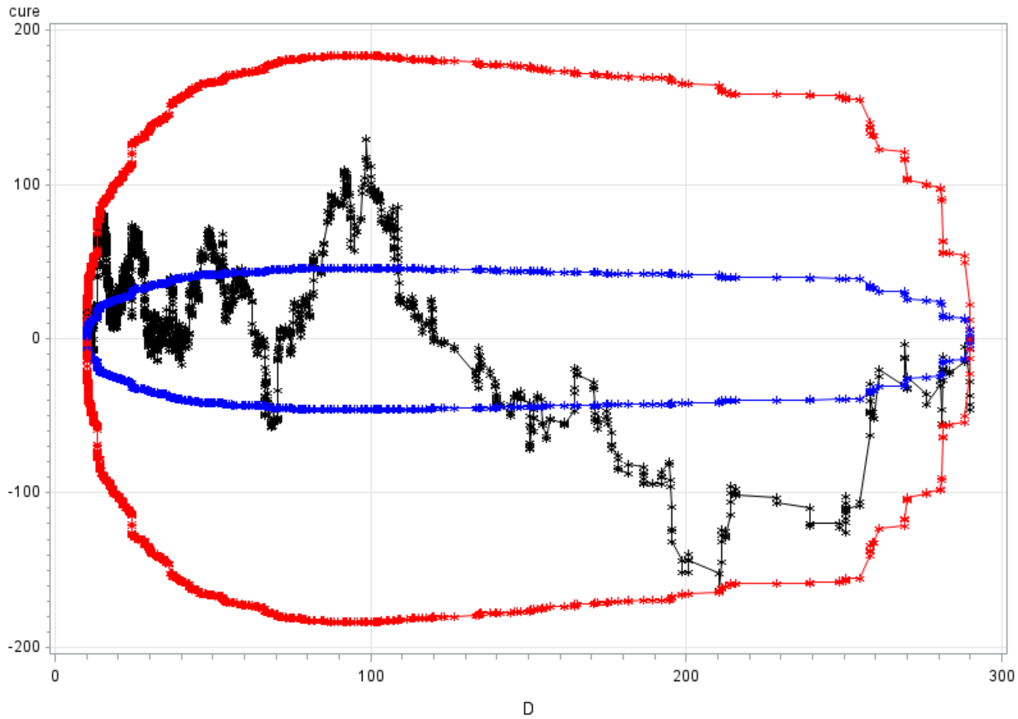


Figure 6-3. Model 1 duration CURE plot

For this model, the AADT and duration CURE plots are acceptable, but the length CURE plot shows the model overpredicts for lengths less than 6 miles. Thus a model was developed for freeway work zones with lengths less than 6 miles and constant overdispersion.

6.1.2 Model 2, Freeway Work Zones Shorter than 6 miles and with $\boldsymbol{\alpha} = \boldsymbol{\alpha_0}$

This model was made by considering a constant overdispersion for work zones shorter than 6 miles. Table 6-2 summarizes the estimated parameters of Model 2 with following functional form:

$$N_C = \mathrm{e}^{\beta_0} AADT^{\beta_1} L^{\beta_2} D^{\beta_3} e^{\beta_4 * \frac{Closed\ Lanes}{Number\ of\ Lanes}} e^{\beta_5 Urban} e^{\beta_6 Injury}$$

Table 6-2. Model 2 parameters for fatal-injury or PDO crashes

Explanatory Variable	Parameter Estimates	Standard Error	p-value
Constant	-13.1689	0.4332	<.0001
AADT	0.9355	0.03961	<.0001
L	0.4457	0.02624	<.0001
D	1.0287	0.02693	<.0001
$\frac{\text{Closed Lanes}}{\text{Number of Lanes}}$	0.3397	0.09840	0.0006
Urban	0.5180	0.08731	<.0001
Injury	-1.1391	0.05058	<.0001
Overdispersion, $\boldsymbol{\alpha_0}$	0.3602	0.03240	
Number of Observations		1092	

Each variable added was statistically beneficial to the model (using the $\aleph^2$ test) and all explanatory variables were statistically significant in the end (i.e. p-value<0.01). From this crash prediction model, CMFs can be derived for all explanatory variables. A 1% increase in AADT, length, and duration leads to the number of crashes increasing by 0.94%, 0.45% and 1.03%, respectively. Since $e^{(0.5180)} = 1.679$, this means urban road segments have 1.679 times the frequency of crashes in comparison to rural road segments. The frequency of injury crashes is 32.01% of PDO crashes as calculated by $e^{(-1.1394)} = 0.3201$. Overdispersion of this model for work zones shorter than 6 miles is 0.36 which is acceptable.

Figures 6-4, 6-5 and 6-6 show the Model 2 CURE plots for AADT, length and duration, respectively. AADT, length and duration CURE plots are all acceptable.

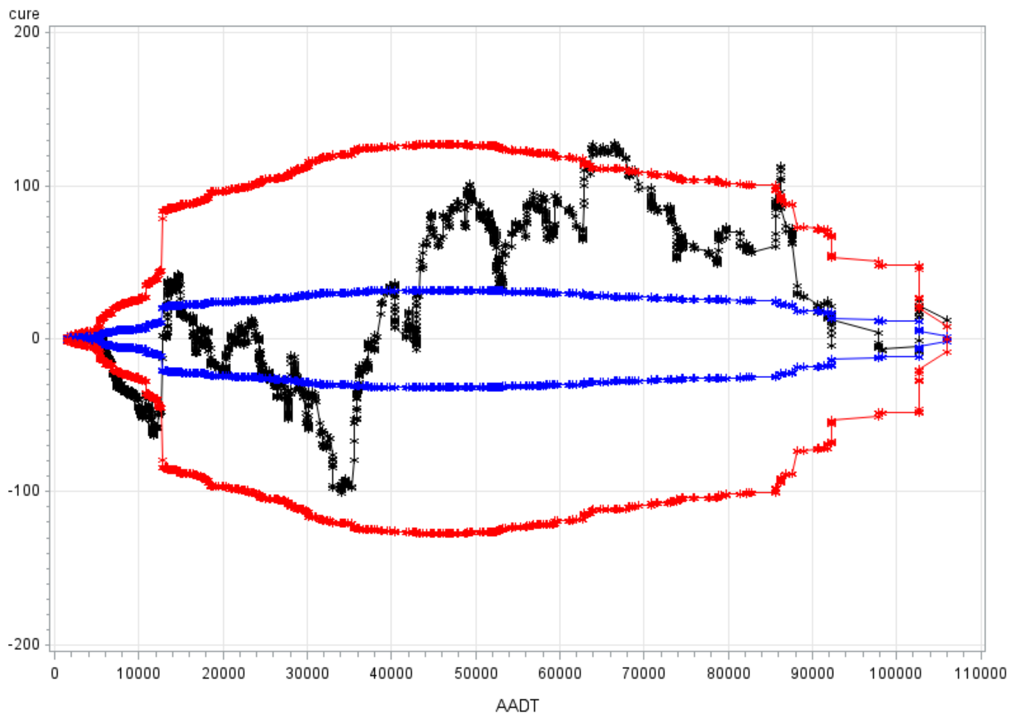


Figure 6-4. Model 2 AADT CURE plot

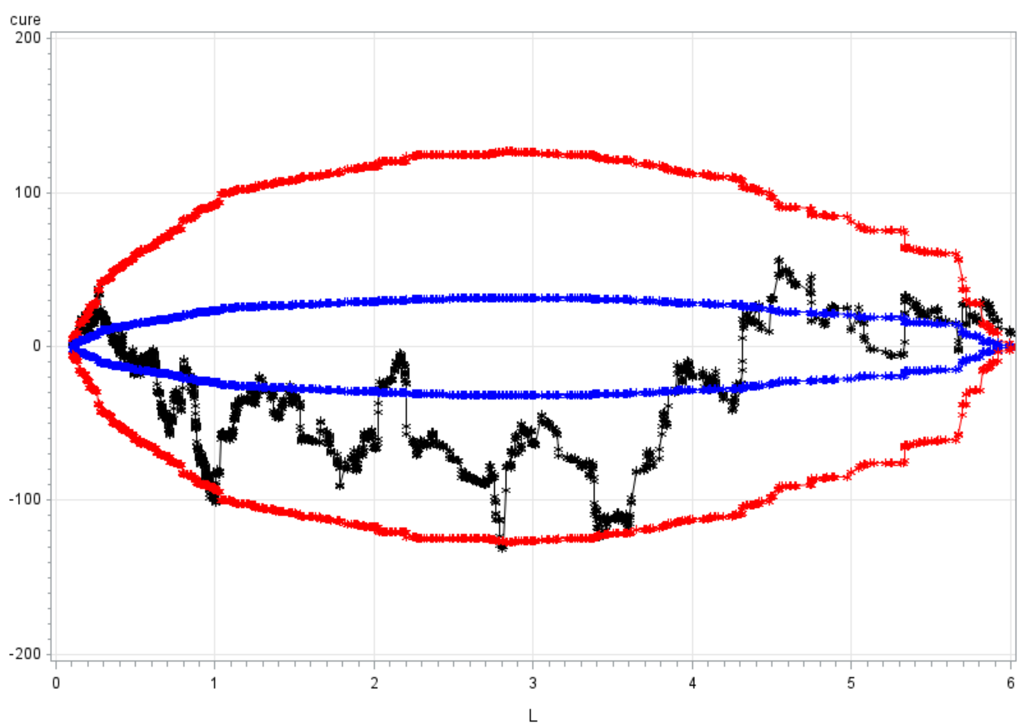


Figure 6-5. Model 2 length CURE plot

Figure 6-6. Model 2 duration CURE plot

6.1.3 Model 3, Freeway Work Zones with $\boldsymbol{\alpha} = \frac{\boldsymbol{\alpha_0}}{\boldsymbol{L}}$

This model was made by considering a length-modified overdispersion for freeway work zones. Table 6-3 summarizes the estimated parameters of Model 3 with following functional form:

$$N_C = \mathrm{e}^{\beta_0} AADT^{\beta_1} L^{\beta_2} D^{\beta_3} e^{\beta_4 * \frac{Closed\ Lanes}{Number\ of\ Lanes}} e^{\beta_5 Urban} e^{\beta_6 Injury}$$

Table 6-3. Model 3 parameters for fatal-injury or PDO crashes

Explanatory Variable	Parameter Estimates	Standard Error	p-value
Constant	-12.5132	0.2991	<.0001
AADT	0.8923	0.02643	<.0001
L	0.6540	0.01699	<.0001
D	0.9986	0.01804	<.0001
$\frac{\text{Closed Lanes}}{\text{Number of Lanes}}$	0.2134	0.07363	0.0038
Urban	0.3506	0.04410	<.0001
Injury	-1.1345	0.03371	<.0001
Overdispersion, $\boldsymbol{\alpha_0}$	0.8928	0.06726	
Number of Observations		1546	

Each variable added was statistically beneficial to the model (using the $\aleph^2$ test) and all explanatory variables were statistically significant in the end (i.e. p-value<0.01). From this crash prediction model, crash modification factors can be derived for all explanatory variables. A 1% increase in AADT, length, and duration leads to the number of crashes increasing by 0.89%, 0.65% and 1.00%, respectively. Since $e^{(0.3506)} = 1.420$, this means urban road segments have 1.420 times the frequency of crashes in comparison to rural roads. The frequency of injury crashes is 32.16% of PDO crashes as calculated by $e^{(-1.1345)} = 0.3216$. Overdispersion of this model is $\alpha = \frac{0.89}{L}$.

Figures 6-7, 6-8 and 6-9 show the Model 3 CURE plots for AADT, length and duration, respectively.

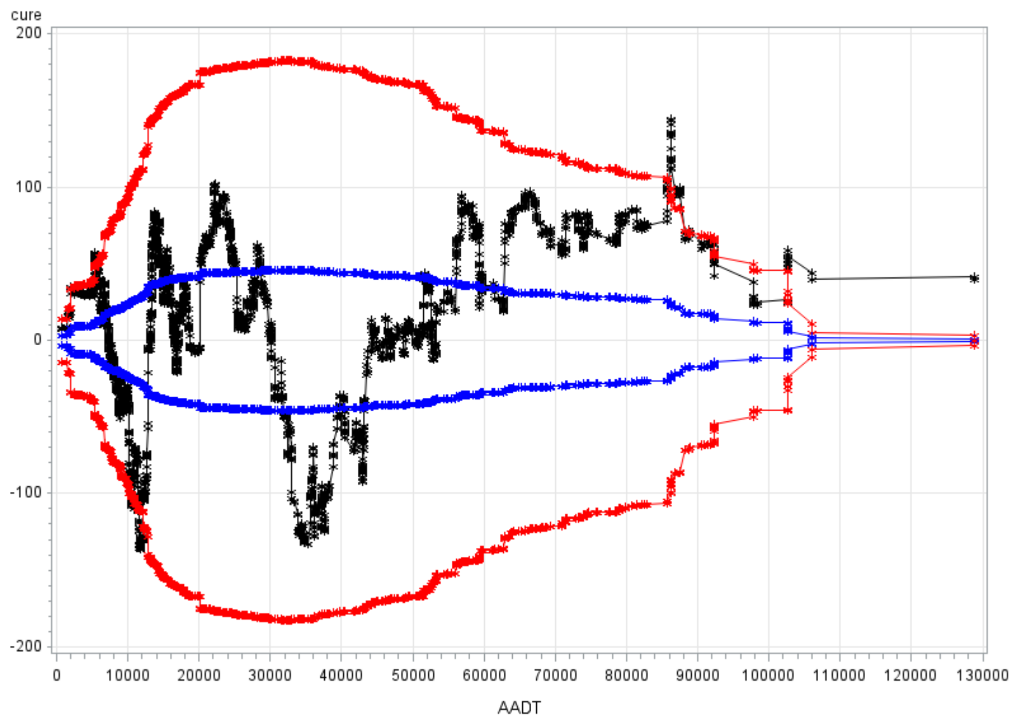


Figure 6-7. Model 3 AADT CURE plot

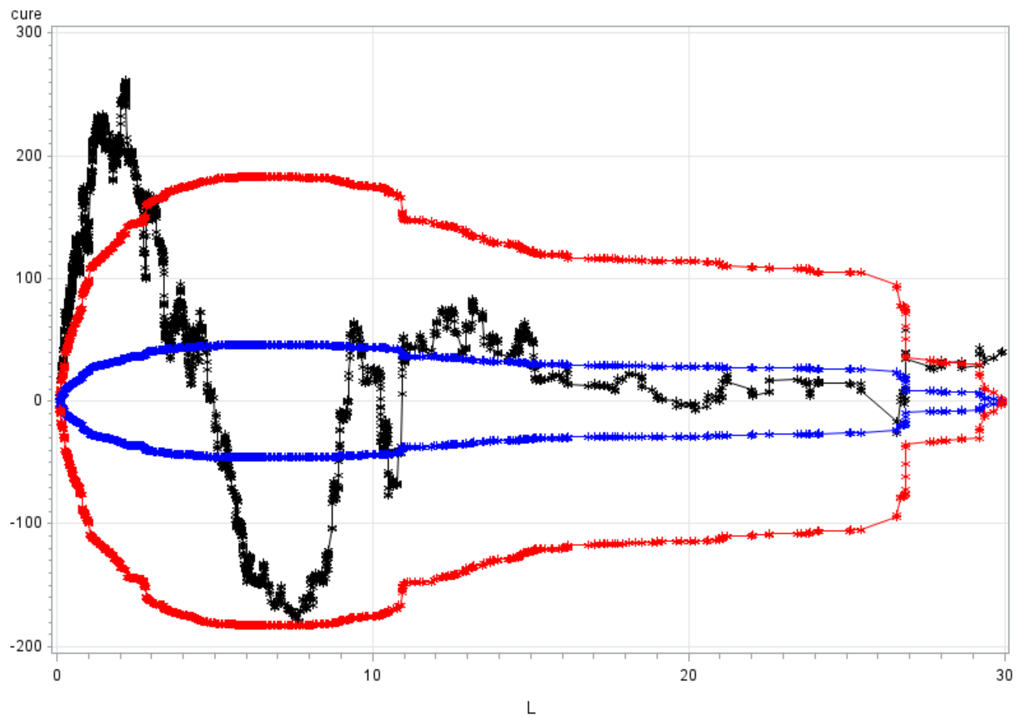


Figure 6-8. Model 3 length CURE plot

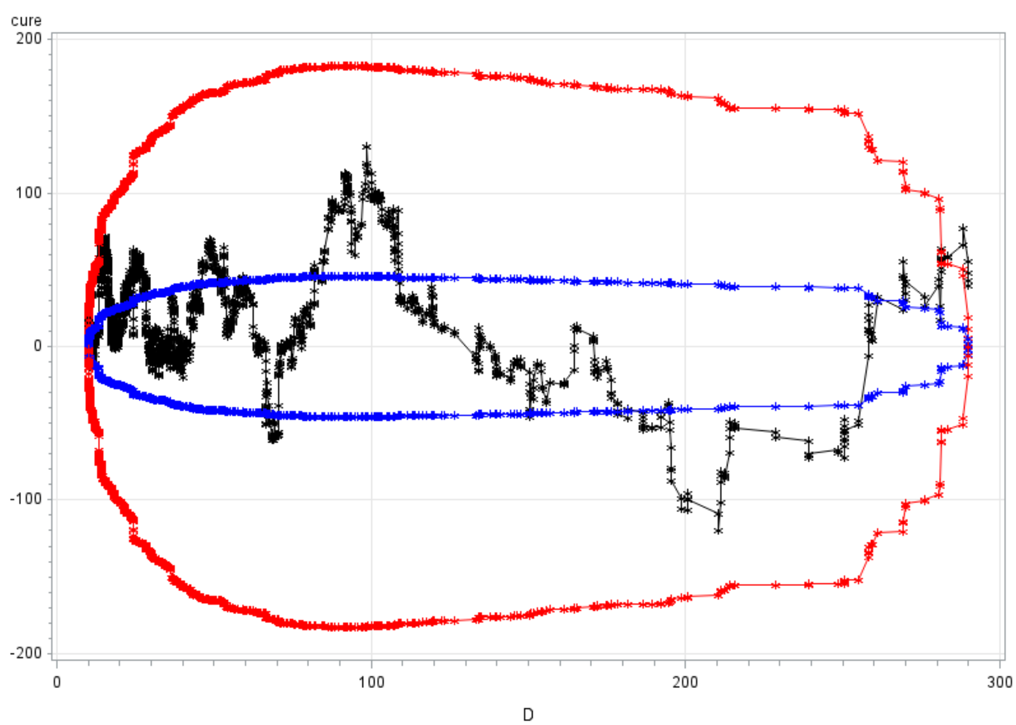


Figure 6-9. Model 3 duration CURE plot

The AADT and duration CURE plots are acceptable, but the length CURE plot shows the model underestimates for work zones shorter than 2 miles and overpredicts for lengths between 2 and 6 miles. Thus, a model was developed for the freeway work zones with lengths less than 6 miles with an overdispersion modified by length.

6.1.4 Model 4, Freeway Work Zones Shorter than 6 miles with $\boldsymbol{\alpha} = \frac{\boldsymbol{\alpha_0}}{\boldsymbol{L}}$

This model was made by considering a length modified overdispersion for freeway work zones shorter than 6 miles. Table 6-4 summarizes the estimated parameters of Model 4 with following functional form:

$$N_C = \mathrm{e}^{\beta_0} AADT^{\beta_1} L^{\beta_2} D^{\beta_3} e^{\beta_4 * \frac{Closed\ Lanes}{Number\ of\ Lanes}} e^{\beta_5 Urban} e^{\beta_6 Injury}$$

Table 6-4. Model 4 parameters for fatal-injury or PDO crashes

Explanatory Variable	Parameter Estimates	Standard Error	p-value
Constant	-13.5250	0.3990	<.0001
AADT	0.9759	0.03672	<.0001
L	0.4595	0.02839	<.0001
D	1.0370	0.02443	<.0001
$\frac{\text{Closed Lanes}}{\text{Number of Lanes}}$	0.3152	0.08704	0.0003
Urban	0.4141	0.08097	<.0001
Injury	-1.1370	0.04678	<.0001
Overdispersion, $\boldsymbol{\alpha_0}$	0.4895	0.04631	
Number of Observations		1092	

Each variable added was statistically beneficial to the model (using the $\aleph^2$ test) and all explanatory variables were statistically significant in the end (i.e. p-value<0.01). From this crash prediction model, crash modification factors can be derived for all explanatory variables. A 1% increase in AADT, length, and duration increases the number of crashes by 0.98%, 0.46% and 1.04%, respectively. Since $e^{(0.4141)}$ = 1.513, this means urban road segments have 1.513 times the frequency of crashes in comparison to rural roads. The frequency of injury crashes is 32.08% of PDO crashes as calculated by $e^{(-1.1370)}$ = 0.3208. Overdispersion of this model is $\alpha = \frac{0.49}{L}$.

Figures 6-10, 6-11 and 6-12 show the Model 4 CURE plots for AADT, length and duration, respectively.

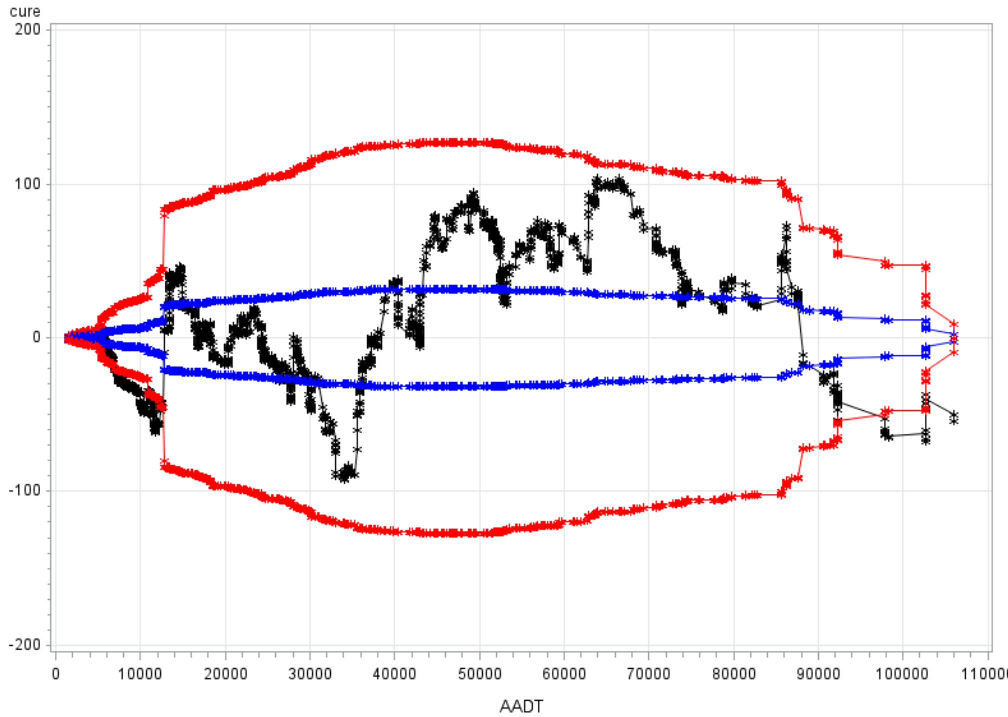


Figure 6-10. Model 4 AADT CURE plot

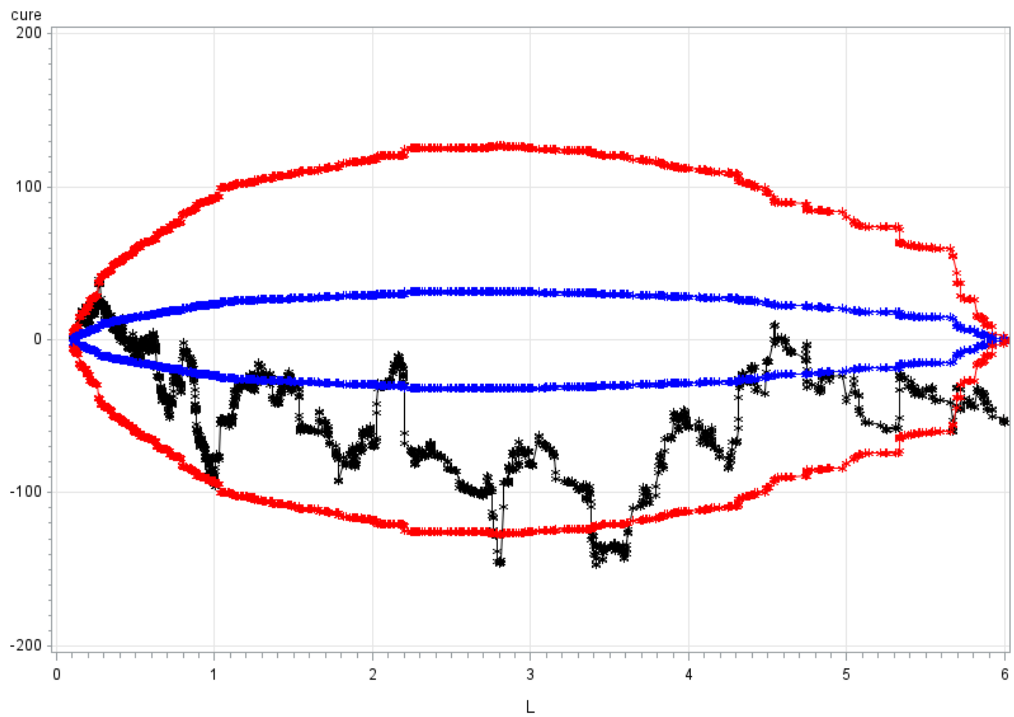


Figure 6-11. Model 4 length CURE plot

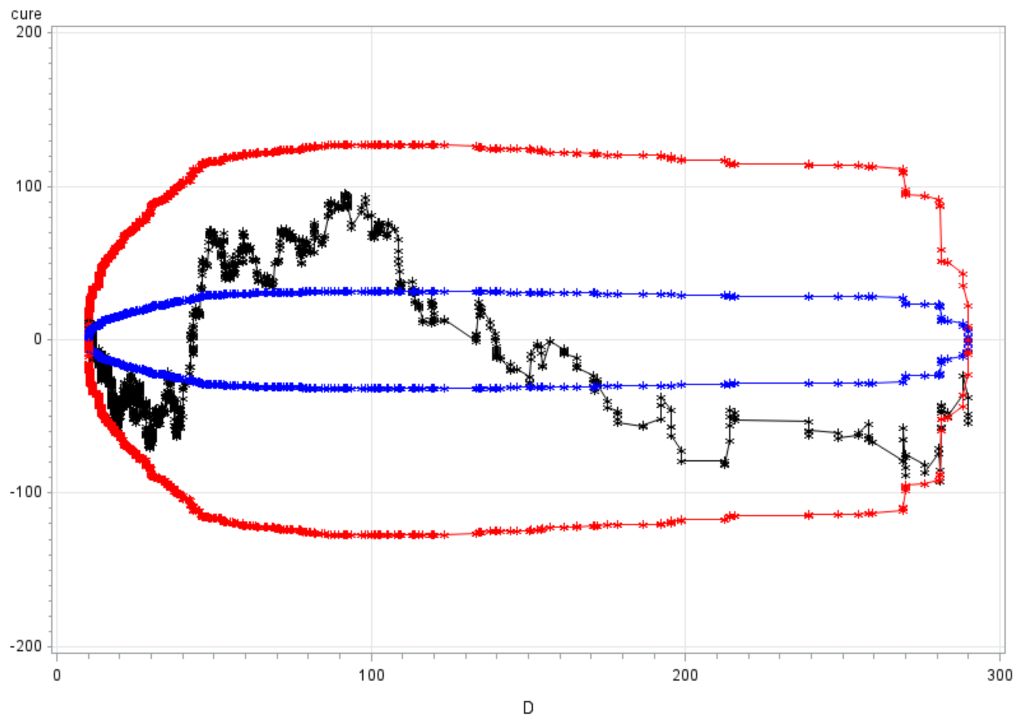


Figure 6-12. Model 4 duration CURE plot

AADT, length and duration CURE plots are acceptable. The next model was fitted for the freeway work zones with an overdispersion modified by length and duration.

6.1.5 Model 5, Freeway Work Zones with $\boldsymbol{\alpha} = \frac{\boldsymbol{\alpha_0}}{\boldsymbol{L*D}}$

This model was made by considering a length and duration modified overdispersion for freeway work zones. Table 6-5 summarizes the estimated parameters of Model 5 with following functional form:

$$N_C = \mathrm{e}^{\beta_0} AADT^{\beta_1} L^{\beta_2} D^{\beta_3} e^{\beta_4 * \frac{Closed\ Lanes}{Number\ of\ Lanes}} e^{\beta_5 Urban} e^{\beta_6 Injury}$$

Table 6-5. Model 5 parameters for fatal-injury or PDO crashes

Explanatory Variable	Parameter Estimates	Standard Error	p-value
Constant	-12.1945	0.2757	<.0001
AADT	0.8638	0.02440	<.0001
L	0.6472	0.01564	<.0001
D	0.9969	0.01577	<.0001
$\frac{\text{Closed Lanes}}{\text{Number of Lanes}}$	0.1419	0.06726	0.0350
Urban	0.3751	0.04053	<.0001
Injury	-1.1423	0.03076	<.0001
Overdispersion, $\boldsymbol{\alpha_0}$	34.3921	2.6134	
Number of Observations		1546	

Each variable added was statistically beneficial to the model (using the $\aleph^2$ test) and all explanatory variables were statistically significant at 1% level but $\frac{\text{Closed Lanes}}{\text{Number of Lanes}}$ coefficient significance level is 5%. From this crash prediction model, crash modification factors can be derived for all explanatory variables. A 1% increase in AADT, length, and duration increases the number of crashes by 0.86%, 0.65% and 1.00%, respectively. Overdispersion of this model is $\alpha = \frac{34.39}{L*D}$.

Figures 6-13, 6-14 and 6-15 show the Model 5 CURE plots for AADT, length and duration, respectively.

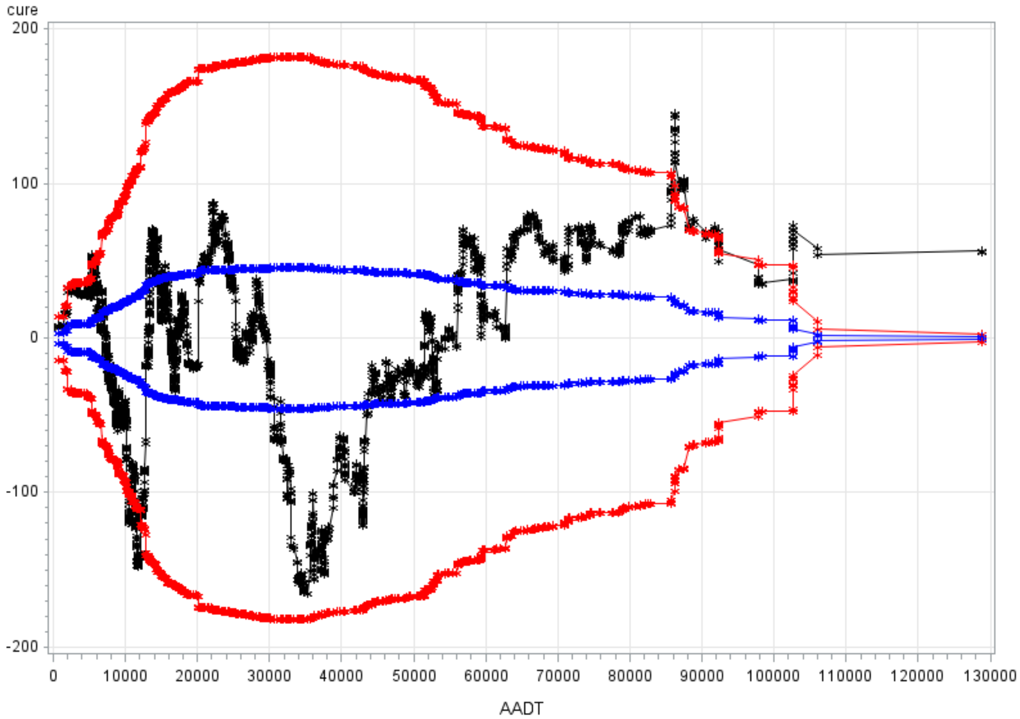


Figure 6-13. Model 5 AADT CURE plot

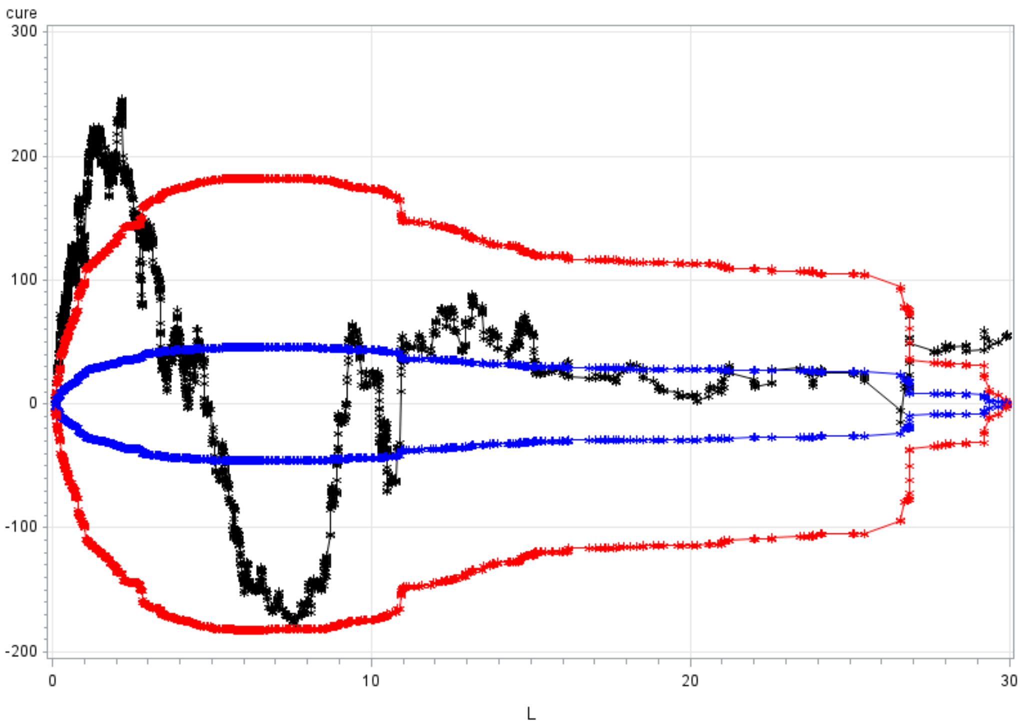


Figure 6-14. Model 5 length CURE plot

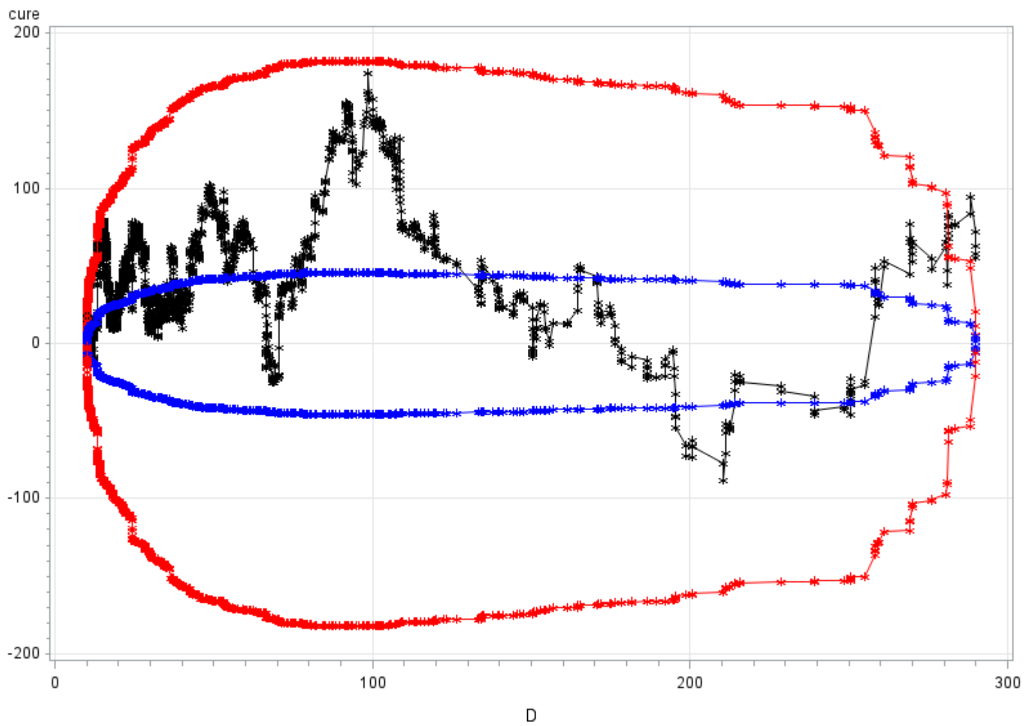


Figure 6-15. Model 5 duration CURE plot

AADT and duration CURE plots are acceptable, but there are still concerns with work zones shorter than 6 miles. Thus the next model was fitted for the freeway work zones with lengths shorter than 6 miles with an overdispersion modified by length and duration.

6.1.6 Model 6, Freeway Work Zones Shorter than 6 miles with $\boldsymbol{\alpha} = \frac{\boldsymbol{\alpha_0}}{\boldsymbol{L*D}}$

This model was made by considering a length and duration modified overdispersion for freeway work zones shorter than 6 miles. Table 6-6 summarizes the estimated parameters of Model 6 with following functional form:

$$N_C = \mathrm{e}^{\beta_0} AADT^{\beta_1} L^{\beta_2} D^{\beta_3} e^{\beta_4 * \frac{Closed\ Lanes}{Number\ of\ Lanes}} e^{\beta_5 Urban} e^{\beta_6 Injury}$$

Table 6-6. Model 6 parameters for fatal-injury or PDO crashes

Explanatory Variable	Parameter Estimates	Standard Error	p-value
Constant	-13.4541	0.3665	<.0001
AADT	0.9730	0.03434	<.0001
L	0.4655	0.02662	<.0001
D	1.0225	0.02111	<.0001
$\frac{\text{Closed Lanes}}{\text{Number of Lanes}}$	0.2924	0.08103	0.0003
Urban	0.4350	0.07651	<.0001
Injury	-1.1322	0.04262	<.0001
Overdispersion, $\boldsymbol{\alpha_0}$	20.5883	1.9007	
Number of Observations		1092	

Each variable added was statistically beneficial to the model (using the $\aleph^2$ test) and all explanatory variables were statistically significant in the end (i.e. p-value<0.01). From this crash prediction model, crash modification factors can be derived for all explanatory variables. A 1% increase in AADT, length, and duration leads to the number of crashes increasing by 0.97%, 0.47% and 1.02%, respectively. Overdispersion of this model is $\alpha = \frac{20.59}{L*D}$.

Figures 6-16, 6-17 and 6-18 show the Model 6 CURE plots for AADT, length and duration, respectively.

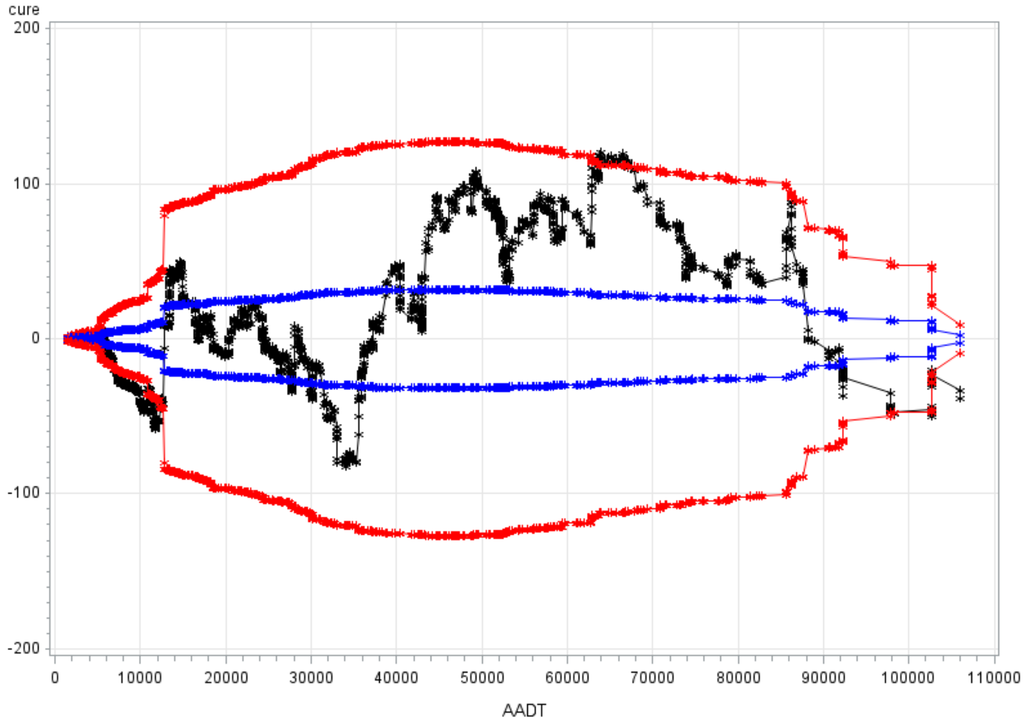


Figure 6-16. Model 6 AADT CURE plot

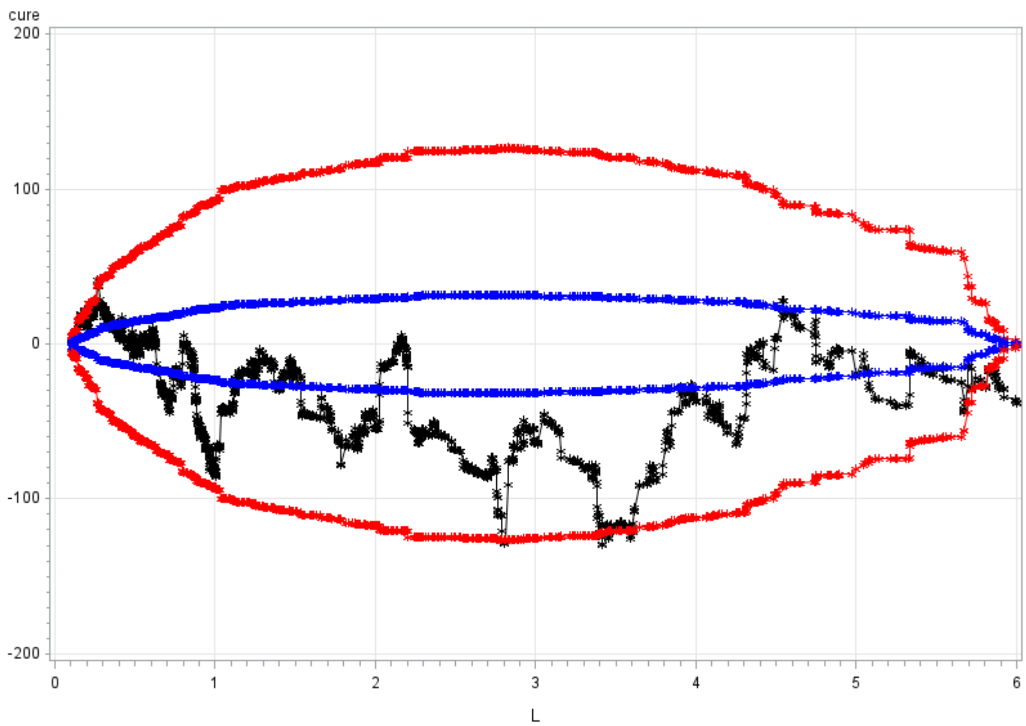


Figure 6-17. Model 6 length CURE plot

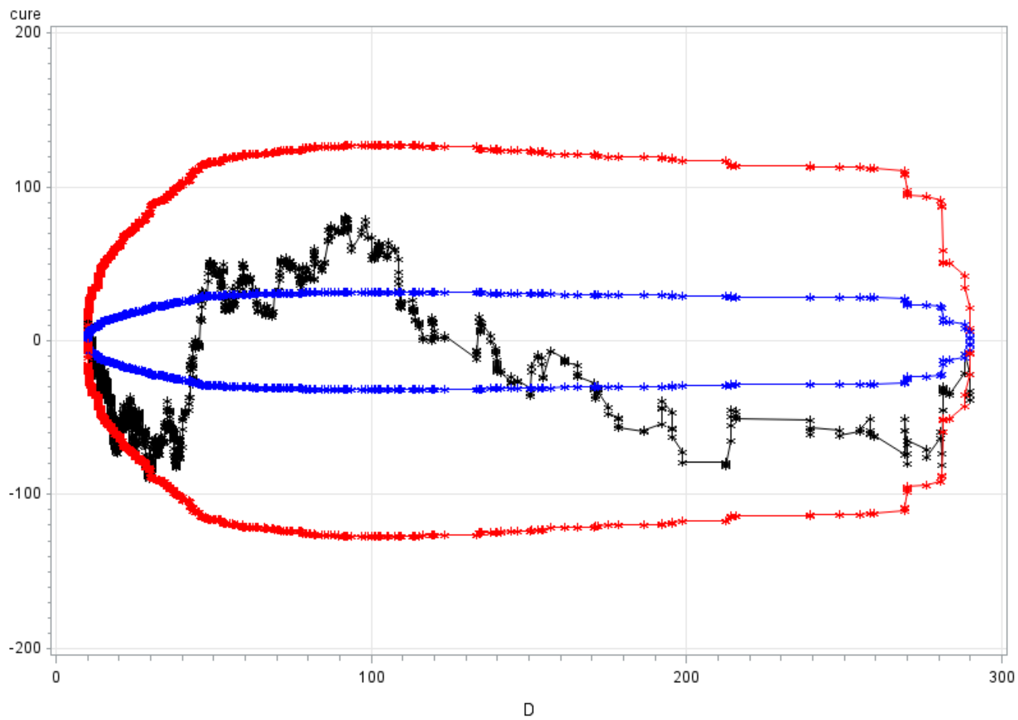


Figure 6-18. Model 6 duration CURE plot

AADT, length and duration CURE plots are acceptable. Next model, 7, uses additional variables but constant overdispersion.

6.1.7 Model 7, Freeway Work Zones with $\boldsymbol{\alpha} = \boldsymbol{\alpha_0}$

This model was made by applying a constant overdispersion term for freeway work zones. Two new variables were used in this model: number of on-ramps and off-ramps in the work area divided by length of the work zone. The number of on-ramps and off-ramps were collected visually by finding the segment on TMS maps and aerial photographs. As the process is time consuming, the sample with these two variables is smaller than the previous sample. Still, the sample size is considerably larger than previous studies in the work zone safety modeling literature. To this end a random sample of 600 freeways were collected and the data was gathered manually. By adding these two variables, the variable $\frac{\text{Closed Lanes}}{\text{Number of Lanes}}$ was not significant anymore. Table 6-7 summarizes the estimated parameters of Model 7 with following functional form:

$$N_C = \mathrm{e}^{\beta_0} AADT^{\beta_1} L^{\beta_2} D^{\beta_3} e^{\beta_4 * \frac{On-ramps}{L}} e^{\beta_5 * \frac{Off-ramps}{L}} e^{\beta_6 Urban} e^{\beta_7 Injury}$$

Table 6-7. Model 7 parameters for fatal-injury or PDO crashes

Explanatory Variable	Parameter Estimates	Standard Error	p-value
Constant	-13.4257	0.5053	<.0001
AADT	0.9577	0.04412	<.0001
L	0.7660	0.03313	<.0001
D	1.0072	0.03587	<.0001
$\frac{On-ramps}{L}$	0.1027	0.05163	0.0470
$\frac{Off-ramps}{L}$	0.1246	0.05269	0.0183
Urban	0.2122	0.07955	0.0078
Injury	-1.1200	0.05509	<.0001
Overdispersion, $\boldsymbol{\alpha_0}$	0.3002	0.03005	
Number of Observations		600	

Each variable added was statistically beneficial to the model (using the $\aleph^2$ test) and all explanatory variables were statistically significant at the 5% level. From this crash prediction model, crash modification factors can be derived for all explanatory variables as shown in previous models. Overdispersion of this model is $\alpha = 0.3$. Figures 6-19, 6-20 and 6-21 show Model 7 CURE plots for AADT, length and duration, respectively; the plots are mostly acceptable but with some small sections exceeding the bounds.

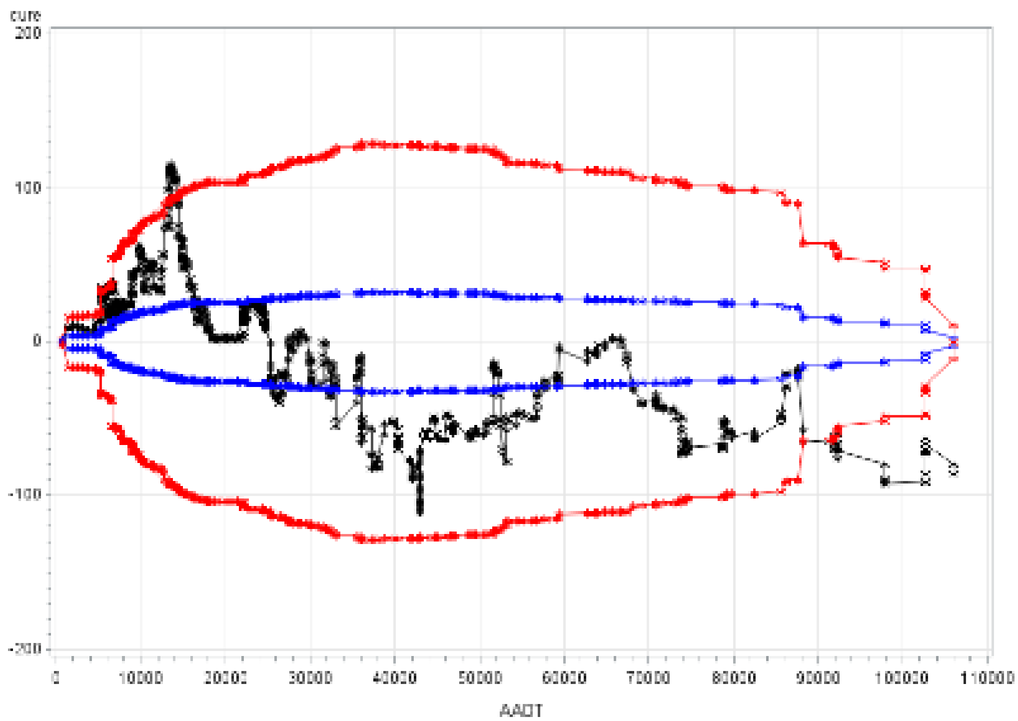


Figure 6-19. Model 7 AADT CURE plot

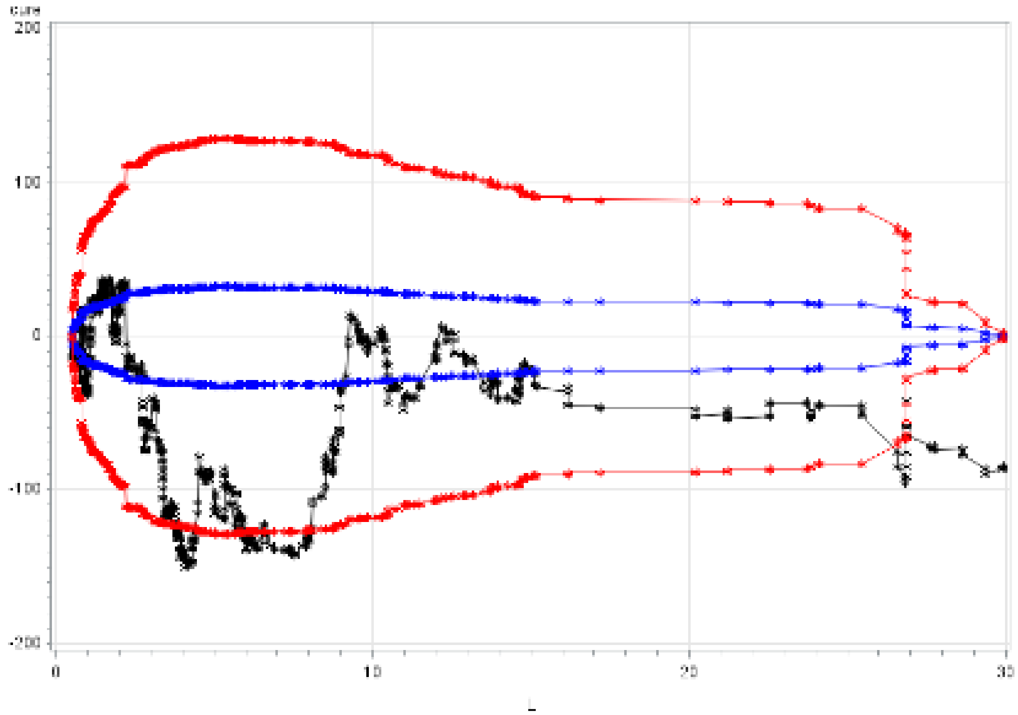


Figure 6-20. Model 7 length CURE plot

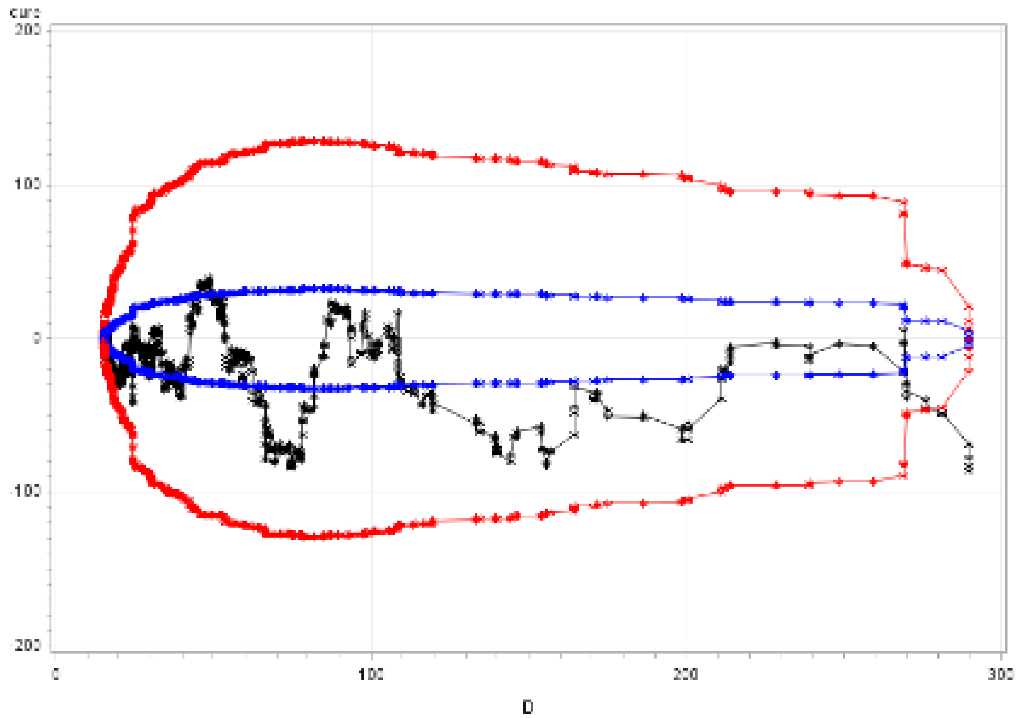


Figure 6-21. Model 7 duration CURE plot

The next attempted model was to use a length modified overdispersion, but the CURE plots were not acceptable. Figures 6-22, 6-23 and 6-24 show the CURE plots and the inadequate fit. However, the model with length and duration modified overdispersion performed well.

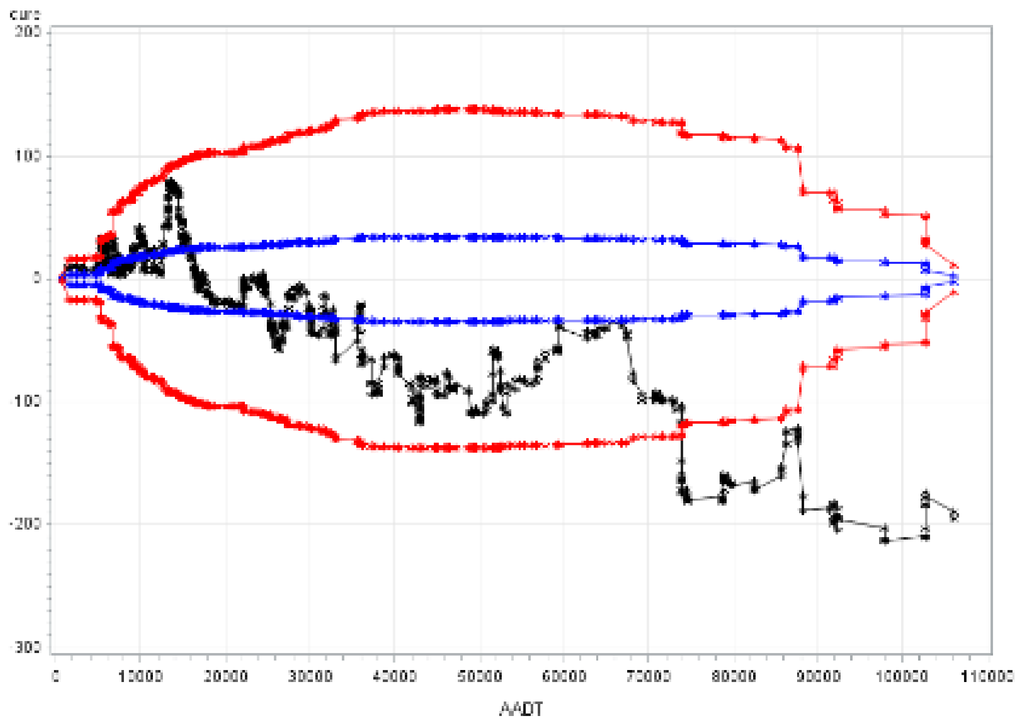


Figure 6-22. Model 7 with length modified overdispersion, AADT CURE plot

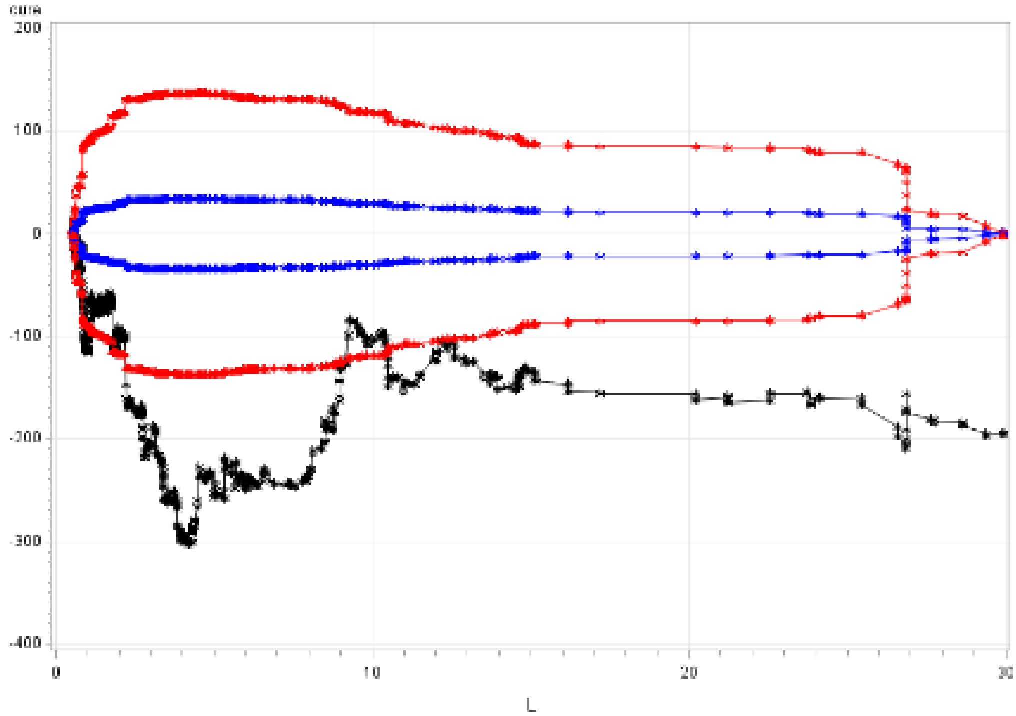


Figure 6-23. Model 7 with length modified overdispersion, length CURE plot

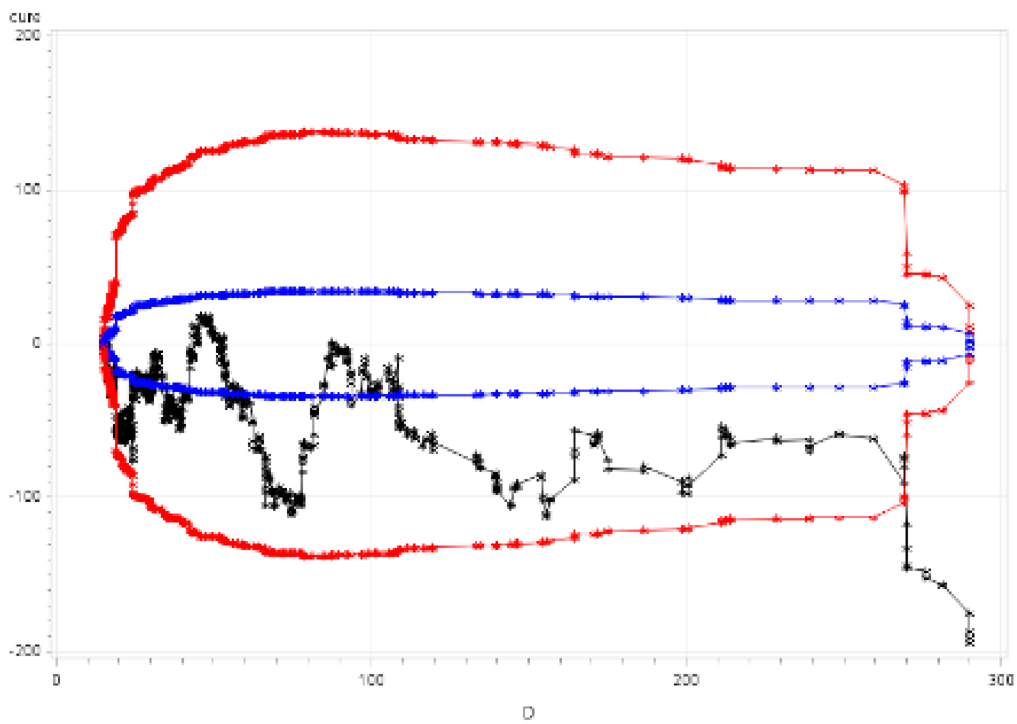


Figure 6-24. Model 7 with length modified overdispersion, duration CURE plot

The next section summarizes the results of this model.

6.1.8 Model 8, Freeway Work Zones with $\boldsymbol{\alpha} = \frac{\boldsymbol{\alpha_0}}{\boldsymbol{L*D}}$

This model was made by considering a length and duration modified overdispersion for freeway work zones. The two new variables, on and off ramps divided by length of work zone, were used to make this model. Table 6-8 summarizes the estimated parameters of Model 8 with following functional form:

$$N_C = \mathrm{e}^{\beta_0} AADT^{\beta_1} L^{\beta_2} D^{\beta_3} e^{\beta_4 * \frac{On-ramps}{L}} e^{\beta_5 * \frac{Off-ramps}{L}} e^{\beta_6 Urban} e^{\beta_7 Injury}$$

Table 6-8. Model 8 parameters for fatal-injury or PDO crashes

Explanatory Variable	Parameter Estimates	Standard Error	p-value
Constant	-12.9446	0.4410	<.0001
AADT	0.8851	0.03886	<.0001
L	0.8264	0.03153	<.0001
D	1.0126	0.02925	<.0001
$\frac{On-ramps}{L}$	0.1805	0.06305	0.0043
$\frac{Off-ramps}{L}$	0.2704	0.06277	<.0001
Urban	0.1488	0.06219	0.0169
Injury	-1.1184	0.04670	<.0001
Overdispersion, $\boldsymbol{\alpha_0}$	45.1352	5.0443	
Number of Observations		600	

Each variable added was statistically beneficial to the model (using the $\aleph^2$ test) and all explanatory variables were statistically significant at the 5% level. From this crash prediction model, crash modification factors can be derived for all explanatory variables. Overdispersion of this model is $\alpha = \frac{45.14}{L*D}$.

Figures 6-25, 6-26 and 6-27 show the Model 8 CURE plots for AADT, length and duration, respectively; the CURE plots were mostly acceptable.

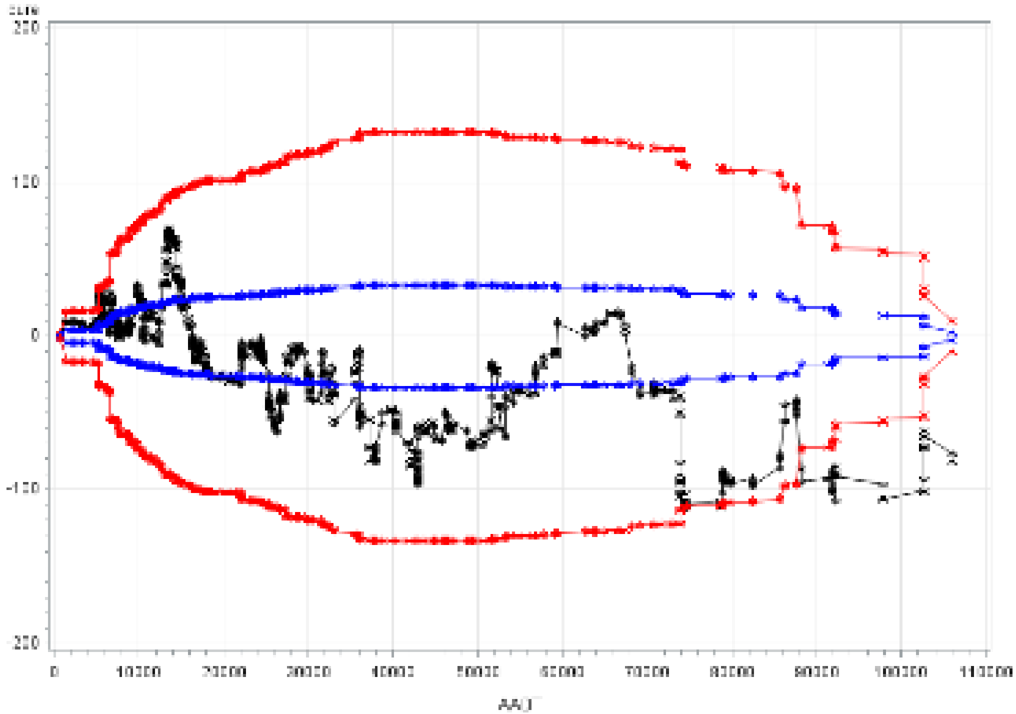


Figure 6-25. Model 8 AADT CURE plot

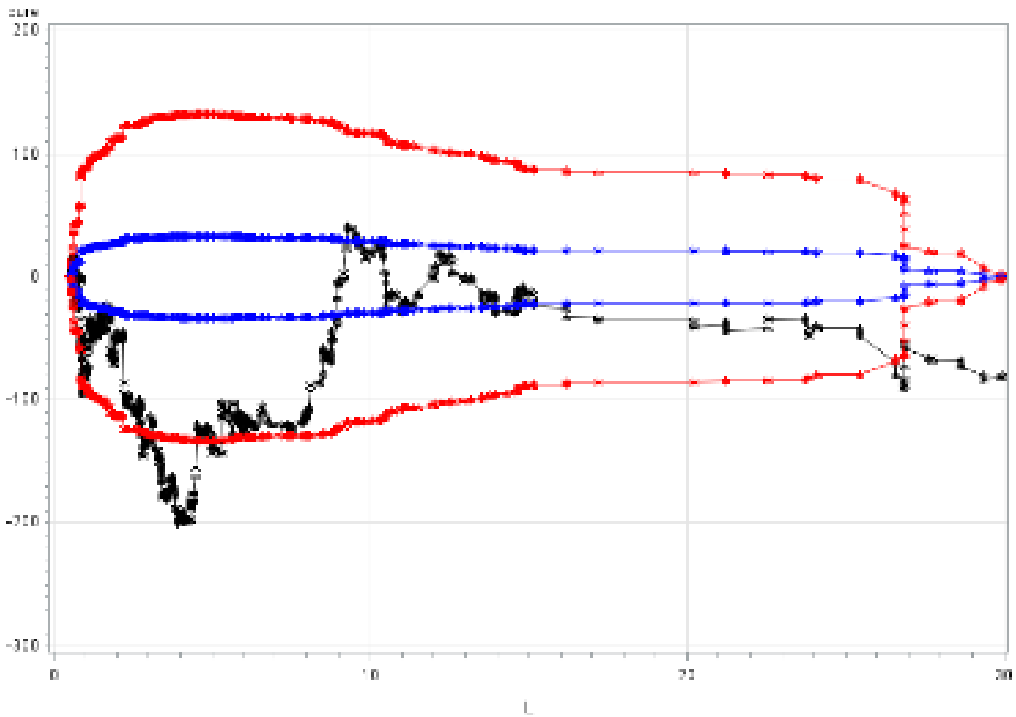


Figure 6-26. Model 8 length CURE plot

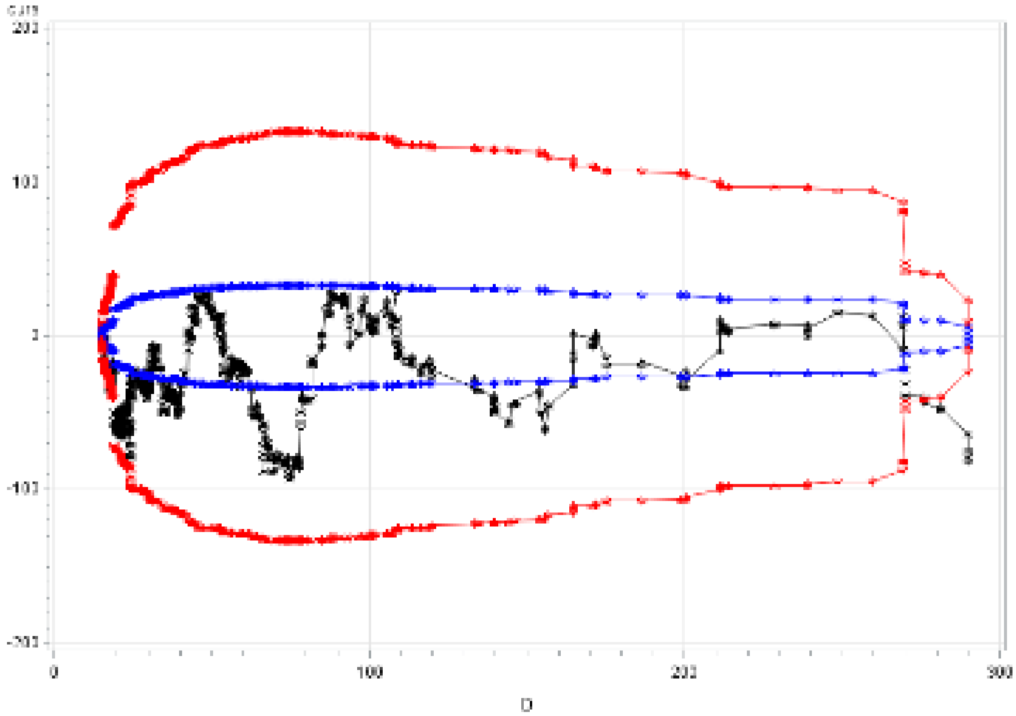


Figure 6-27. Model 8 duration CURE plot

6.2 Expressway Work Zone Models

This section shows different models that were fitted to the expressway work zone sample with different samples and functional forms.

6.2.1 Model 9, Expressway Work Zones with $\boldsymbol{\alpha} = \boldsymbol{\alpha_0}$

This model was made by considering a constant overdispersion. The functional form was similar to previous models. A new variable used was the number of signalized intersections in the work area per mile. The number of signalized intersections in each work zone was available in the road segments database and was collected by through an automated program. Table 6-9 summarizes the estimated parameters of Model 9 with following functional form:

$$N_C = \mathrm{e}^{\beta_0} AADT^{\beta_1} L^{\beta_2} D^{\beta_3} e^{\beta_4 * \frac{Signal}{L}} e^{\beta_5 Urban} e^{\beta_6 Injury}$$

Table 6-9. Model 9 parameters for fatal-injury or PDO crashes

Explanatory Variable	**Parameter Estimates**	**Standard Error**	**p-value**
Constant	-11.9335	0.5399	<.0001
AADT	0.8338	0.05700	<.0001
L	0.6042	0.02803	<.0001
D	0.9990	0.03410	<.0001
$\frac{Signal}{L}$	0.2106	0.01712	<.0001
Urban	0.6584	0.08137	<.0001
Injury	-1.0236	0.06460	<.0001
Overdispersion, $\boldsymbol{\alpha_0}$	0.7154	0.05867	
Number of Observations		1189	

Each variable added was statistically beneficial to the model (using the $\aleph^2$ test) and all explanatory variables were statistically significant at the 1% level. From this crash prediction model, crash modification factors can be derived for all explanatory variables. A 1% increase in AADT, length, and duration leads to the number of crashes increasing by 0.83%, 0.60% and 1.00%, respectively. Since $e^{(0.6584)} = 1.9317$, this means urban road segments have 1.9317 times the frequency of crashes in comparison to rural roads. The frequency of injury crashes is 35.93% of PDO crashes as calculated by $e^{(-1.0236)} = 0.3593$. The overdispersion is 0.7154 which is acceptable.

Figures 6-28, 6-29 and 6-30 show Model 9 CURE plots for AADT, length and duration, respectively. AADT and Length CURE plots were not satisfactory, as they go beyond the limits. Different solutions were devised and tried to improve the fit. The best solution was to separate the data from urban and rural work zones and fit a model to them. Using the length and duration modified overdispersion did not improve the models and led to overfitting.

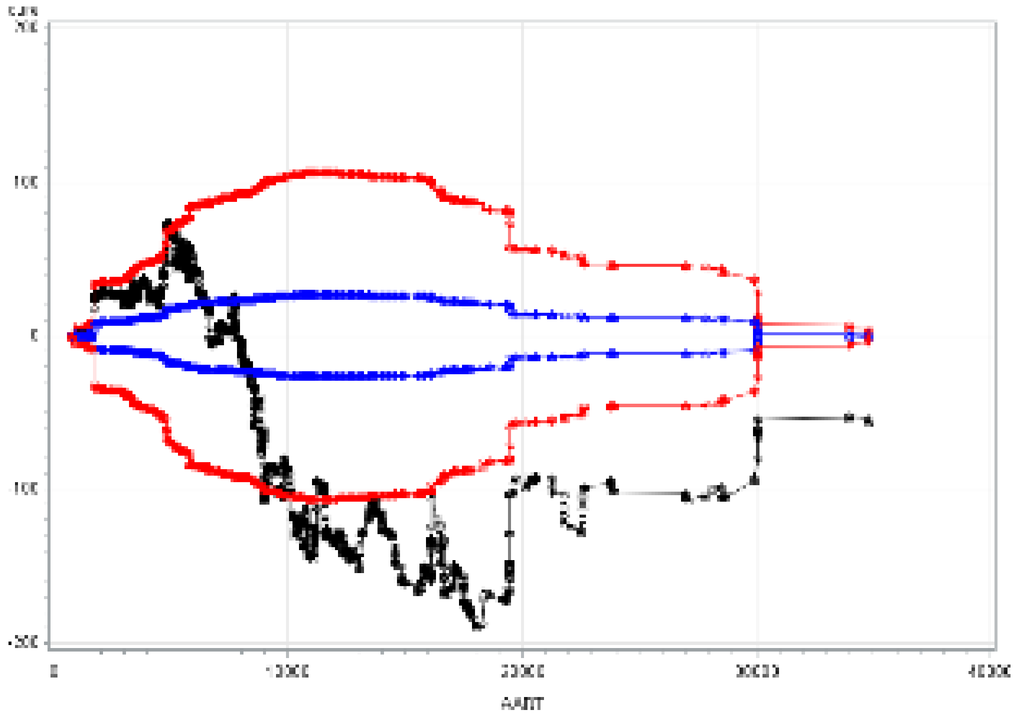


Figure 6-28. Model 9 AADT CURE plot

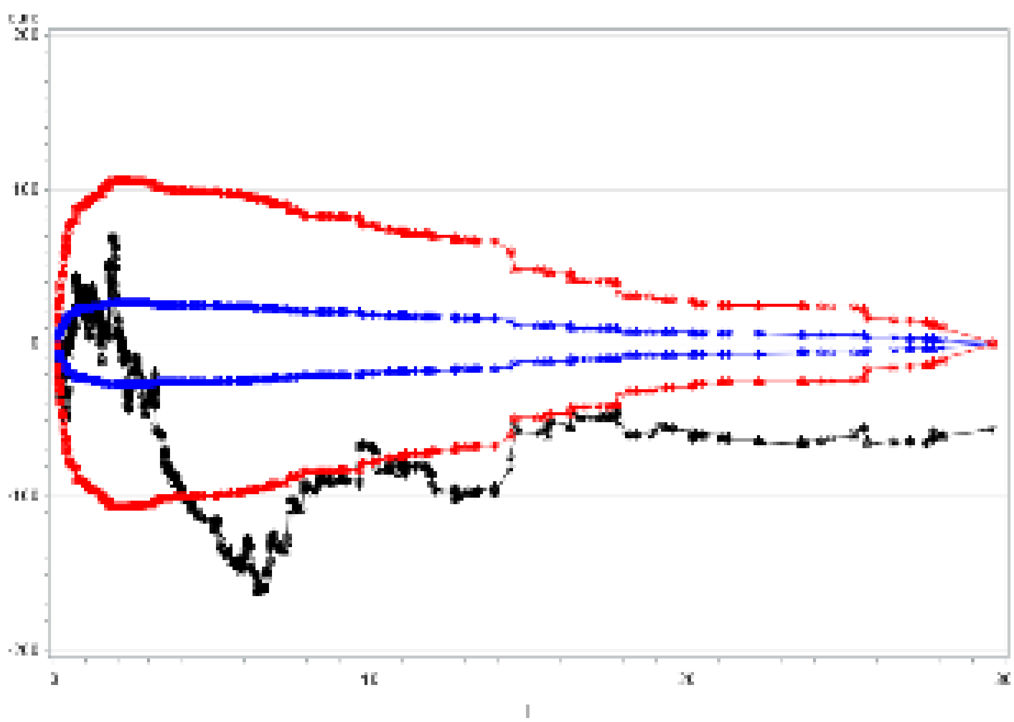


Figure 6-29. Model 9 length CURE plot

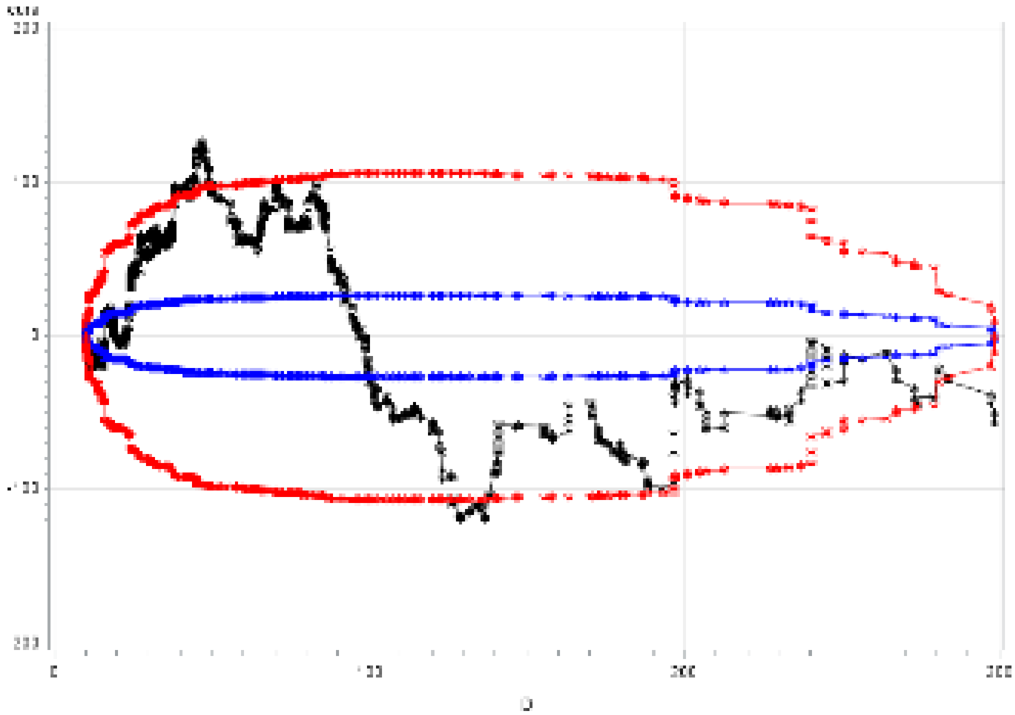


Figure 6-30. Model 9 duration CURE plot

6.2.2 Model 10, Rural Expressway Work Zones with $\boldsymbol{\alpha} = \boldsymbol{\alpha_0}$

This model was made by considering a constant overdispersion. Table 6-10 summarizes the estimated parameters of Model 10 with following functional form:

$$N_C = \mathrm{e}^{\beta_0} AADT^{\beta_1} L^{\beta_2} D^{\beta_3} e^{\beta_4 * \frac{Signal}{L}} e^{\beta_5 Injury}$$

Table 6-10. Model 10 parameters for fatal-injury or PDO crashes

Explanatory Variable	**Parameter Estimates**	**Standard Error**	**p-value**
Constant	-10.9364	0.8186	<.0001
AADT	0.6615	0.09018	<.0001
L	0.6558	0.04296	<.0001
D	1.0952	0.05431	<.0001
$\frac{Signal}{L}$	0.4294	0.09951	<.0001
Injury	-1.0052	0.09893	<.0001
Overdispersion, $\boldsymbol{\alpha_0}$	0.4120	0.08076	
Number of Observations		589	

Each variable added was statistically beneficial to the model (using the $\aleph^2$ test) and all explanatory variables were statistically significant at 1% level. A 1% increase in AADT, length, and duration leads to the number of crashes increasing by 0.66%, 0.66% and 1.10%, respectively. The frequency of injury crashes is 35.93% of PDO crashes as calculated by $e^{(-1.0052)} = 0.3660$. The overdispersion is 0.4120 which is almost half of Model 9's overdispersion (0.7154). Figures 6-31, 6-32 and 6-33 show the Model 10 CURE plots for AADT, length and duration, respectively; all plots are satisfactory.

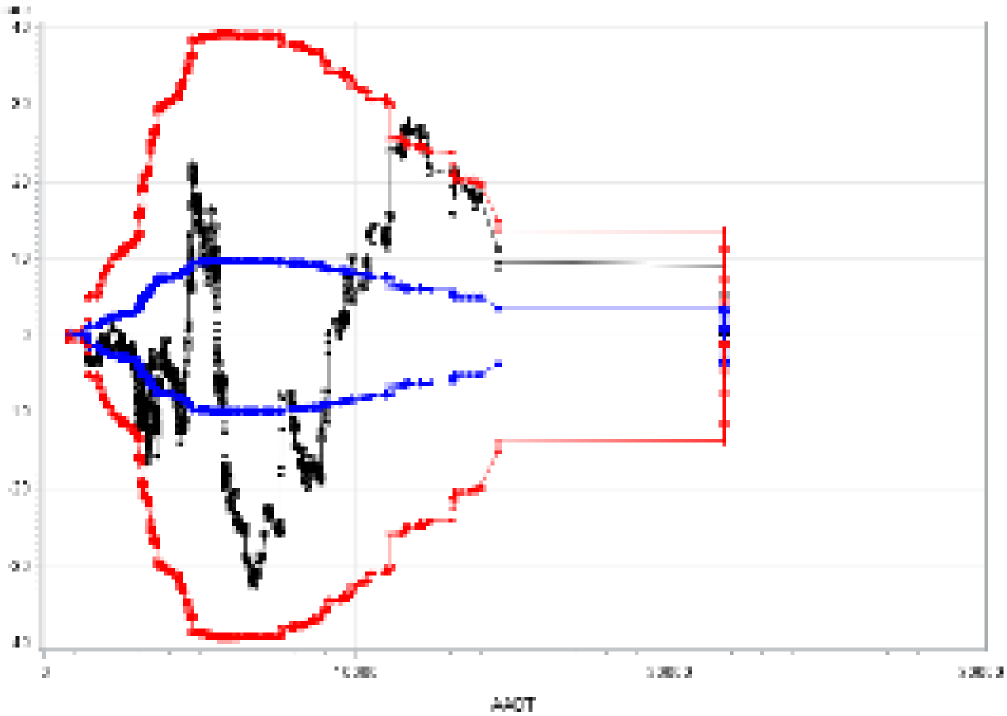


Figure 6-31. Model 10 AADT CURE plot

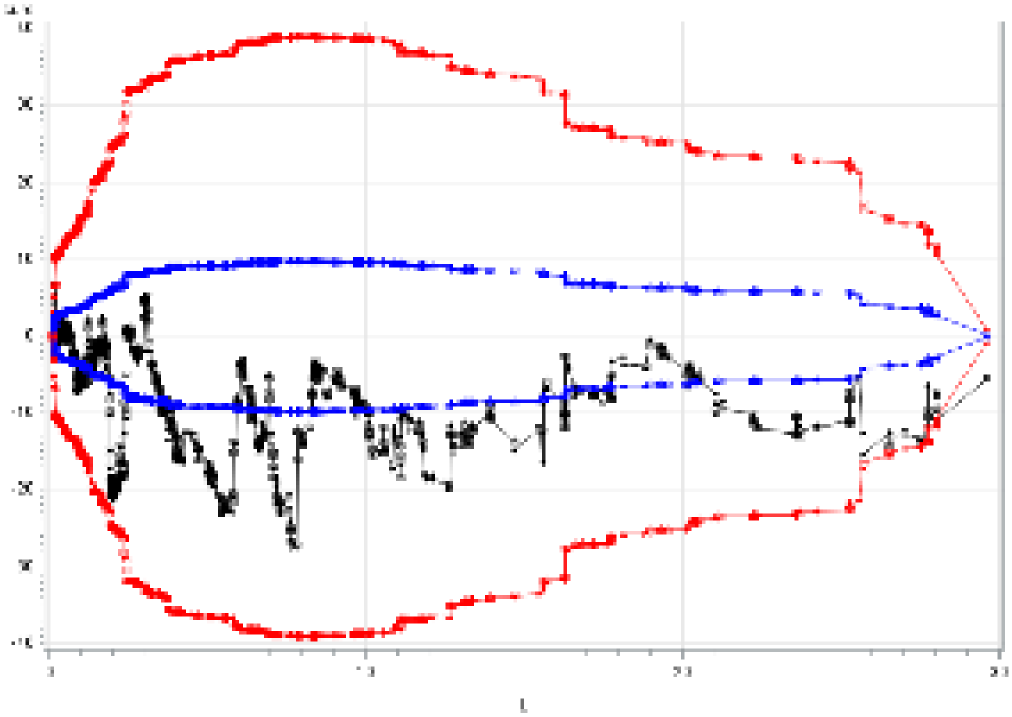


Figure 6-32. Model 10 length CURE plot

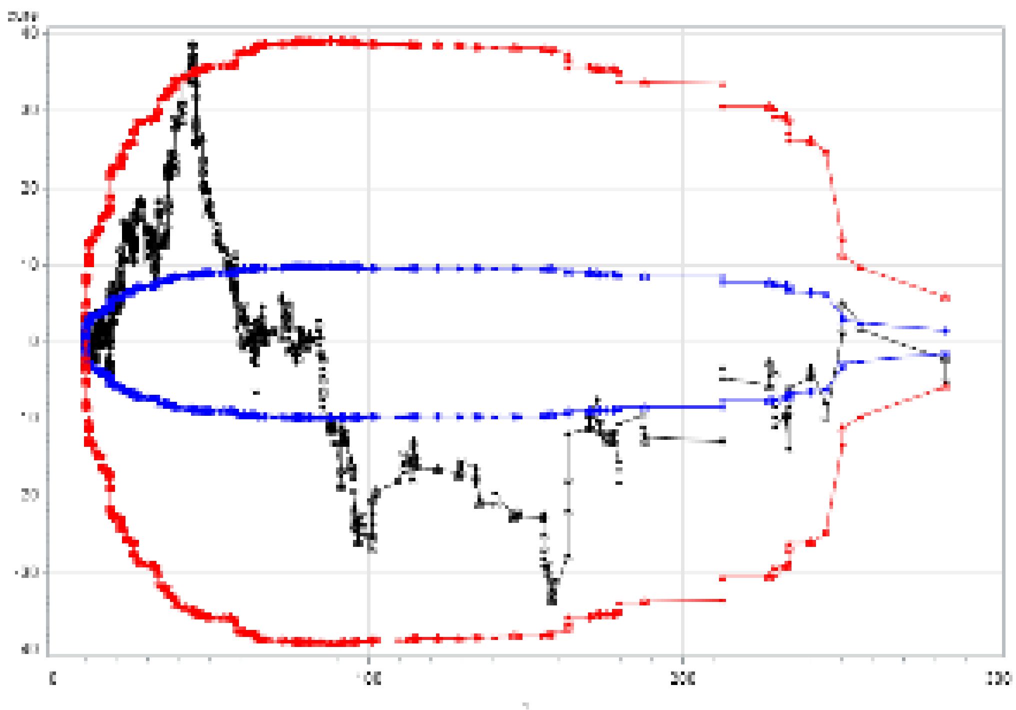


Figure 6-33. Model 10 duration CURE plot

6.2.3 Model 11, Urban Expressway Work Zones with $\boldsymbol{\alpha} = \boldsymbol{\alpha_0}$

This model was made by considering a constant overdispersion. Table 6-11 summarizes the estimated parameters of Model 11 with following functional form:

$$N_C = e^{\beta_0} AADT^{\beta_1} L^{\beta_2} D^{\beta_3} e^{\beta_4 * \frac{Signal}{L}} e^{\beta_5 Injury}$$

Table 6-11. Model 11 parameters for fatal-injury or PDO crashes

Explanatory Variable	**Parameter Estimates**	**Standard Error**	**p-value**
Constant	-11.5982	0.7321	<.0001
AADT	0.8890	0.07261	<.0001
L	0.5858	0.03682	<.0001
D	0.9571	0.04358	<.0001
$\frac{Signal}{L}$	0.1996	0.01852	<.0001
Injury	-1.0330	0.08361	<.0001
Overdispersion, $\boldsymbol{\alpha_0}$	0.8340	0.07864	
Number of Observations		589	

Each variable added was statistically beneficial to the model (using the $\aleph^2$ test) and all explanatory variables were statistically significant at the 1% level. A 1% increase in AADT, length, and duration leads to the number of crashes increasing by 0.88%, 0.59% and 0.96%, respectively. The frequency of injury crashes is 35.60% of PDO crashes as calculated by $e^{(-1.0330)} = 0.3560$. The overdispersion was 0.8340 which was larger than Model 9's overdispersion of 0.7154.

Figures 6-34, 6-35 and 6-36 show the Model 11 CURE plots for AADT, length and duration, respectively. The AADT CURE plot exceeded the band limits, and the length CURE plot was inadequate for work zone smaller than 6 miles. As a solution, a separate model was fitted to the expressway work zones with lengths smaller than 6 miles.

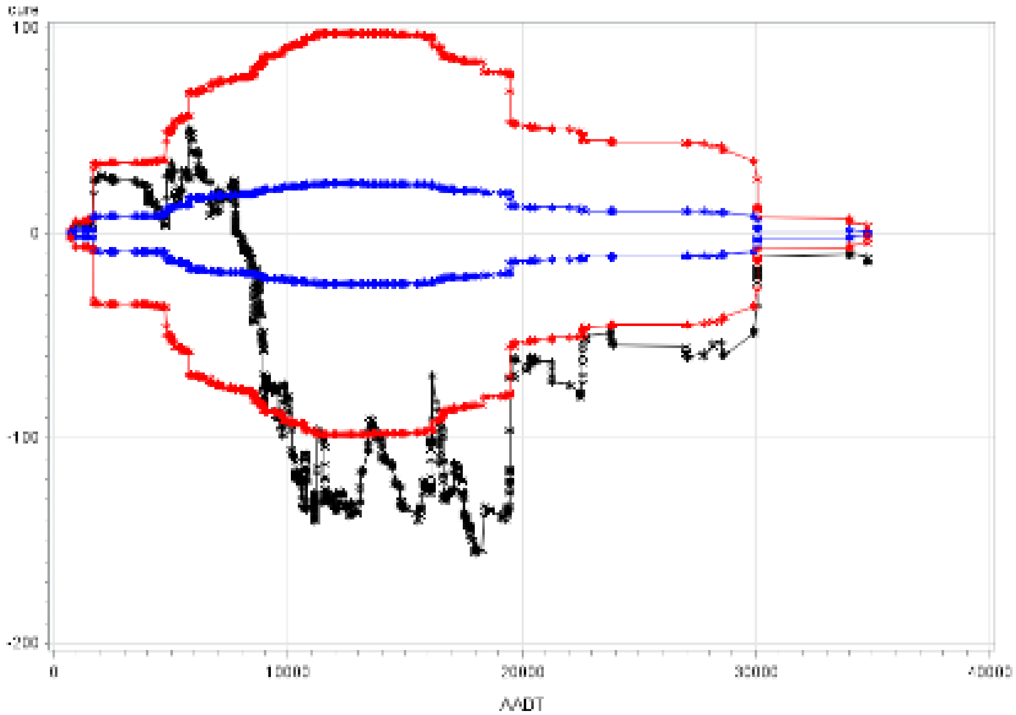


Figure 6-34. Model 11 AADT CURE plot

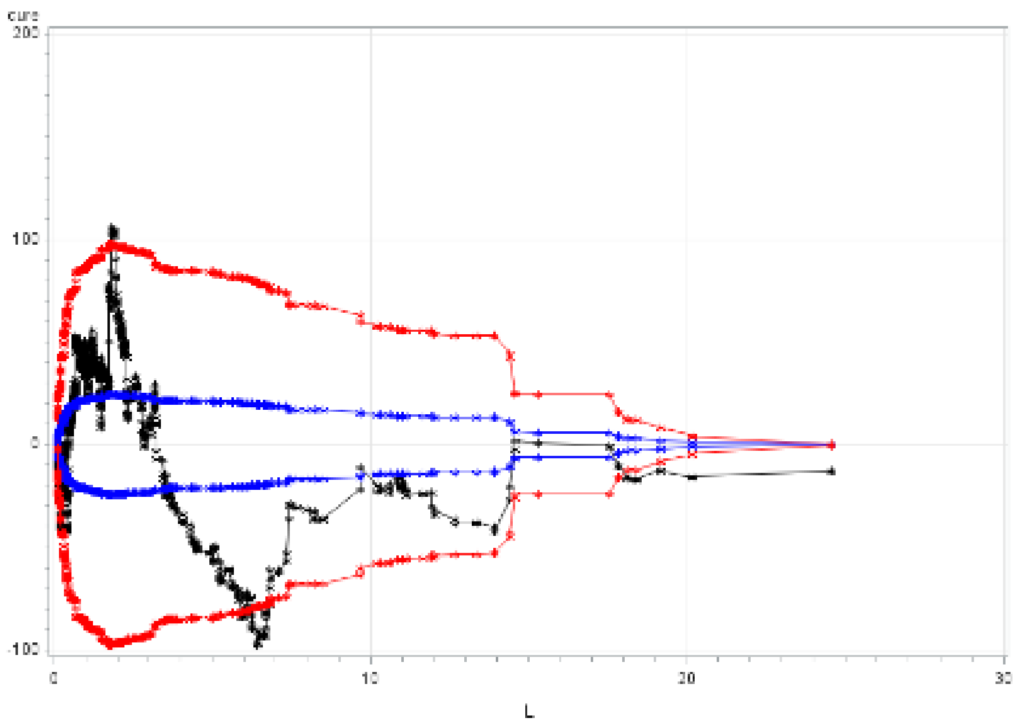


Figure 6-35. Model 11 length CURE plot

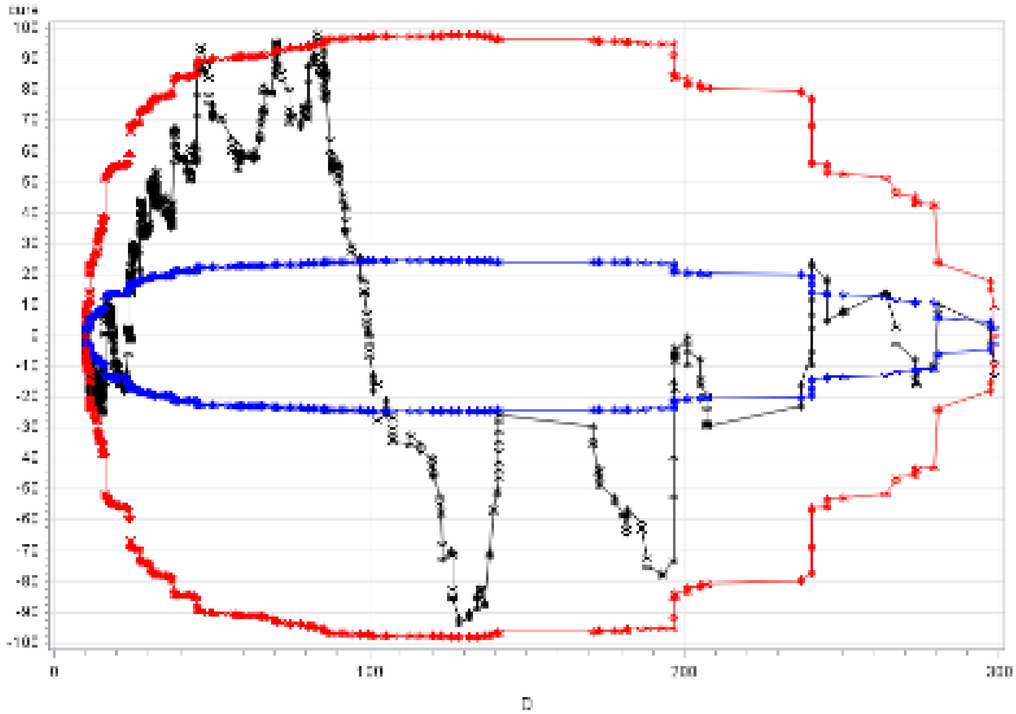


Figure 6-36. Model 11 duration CURE plot

6.2.4 Model 12, Urban Expressway Work Zones Shorter than 6 miles with $\boldsymbol{\alpha} = \boldsymbol{\alpha_0}$

This model was made by considering a constant overdispersion. Table 6-12 summarizes the estimated parameters of Model 12 with following functional form:

$$N_C = \mathrm{e}^{\beta_0} AADT^{\beta_1} L^{\beta_2} D^{\beta_3} e^{\beta_4 * \frac{Signal}{L}} e^{\beta_5 Injury}$$

Table 6-12. Model 12 parameters for fatal-injury or PDO crashes

Explanatory Variable	**Parameter Estimates**	**Standard Error**	**p-value**
Constant	-14.3737	0.8626	<.0001
AADT	1.1486	0.08503	<.0001
L	0.3801	0.04800	<.0001
D	1.0505	0.04513	<.0001
$\frac{Signal}{L}$	0.1613	0.01809	<.0001
Injury	-1.0996	0.08922	<.0001
Overdispersion, $\boldsymbol{\alpha_0}$	0.6954	0.07822	
Number of Observations		549	

Each variable added was statistically beneficial to the model (using the $\aleph^2$ test) and all explanatory variables were statistically significant at 1% level. A 1% increase in AADT, length, and duration leads to the number of crashes increasing by 1.15 %, 0.38% and 1.05%, respectively. The overdispersion is 0.6954. Figures 6-37, 6-38 and 6-39 show the Model 12 CURE plots for AADT, length and duration, respectively; all plots were acceptable.

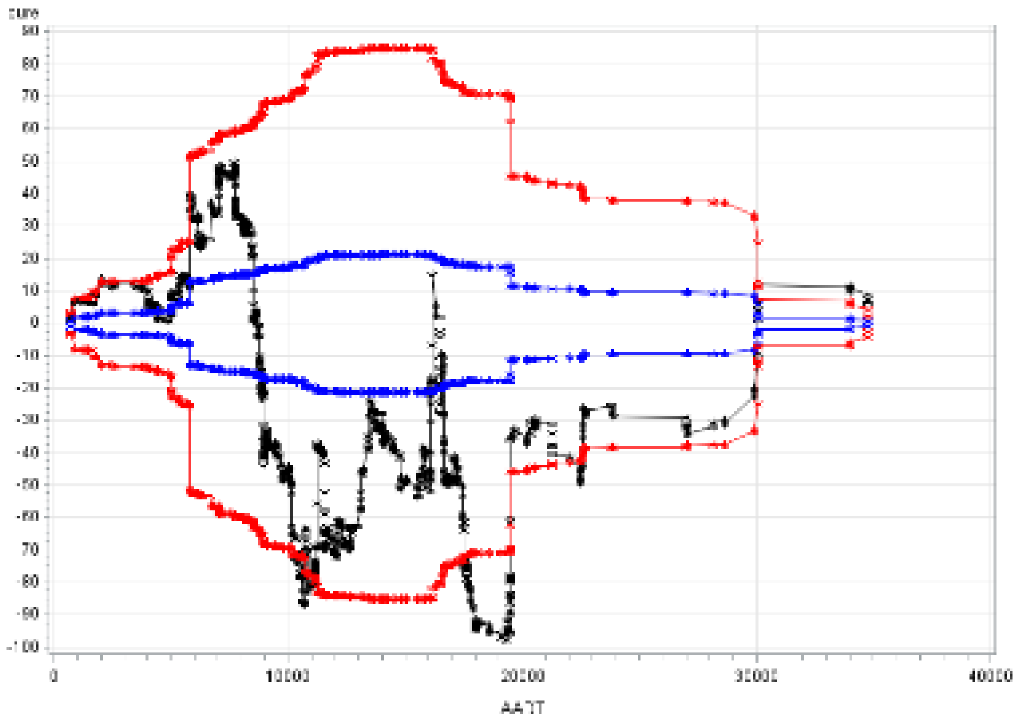


Figure 6-37. Model 12 AADT CURE plot

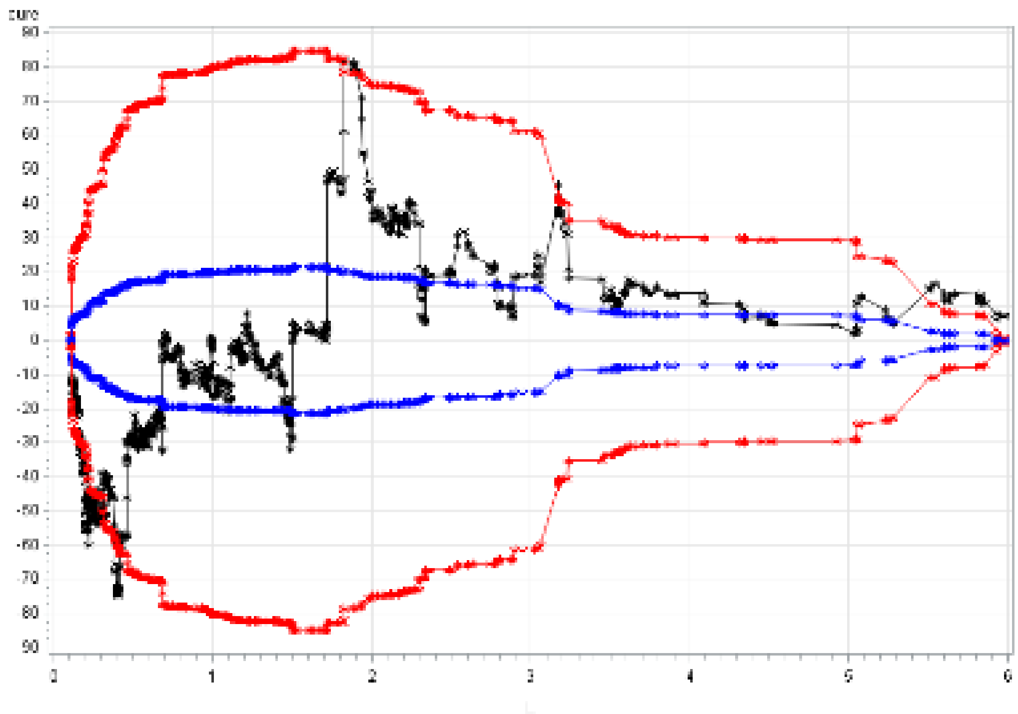


Figure 6-38. Model 12 length CURE plot

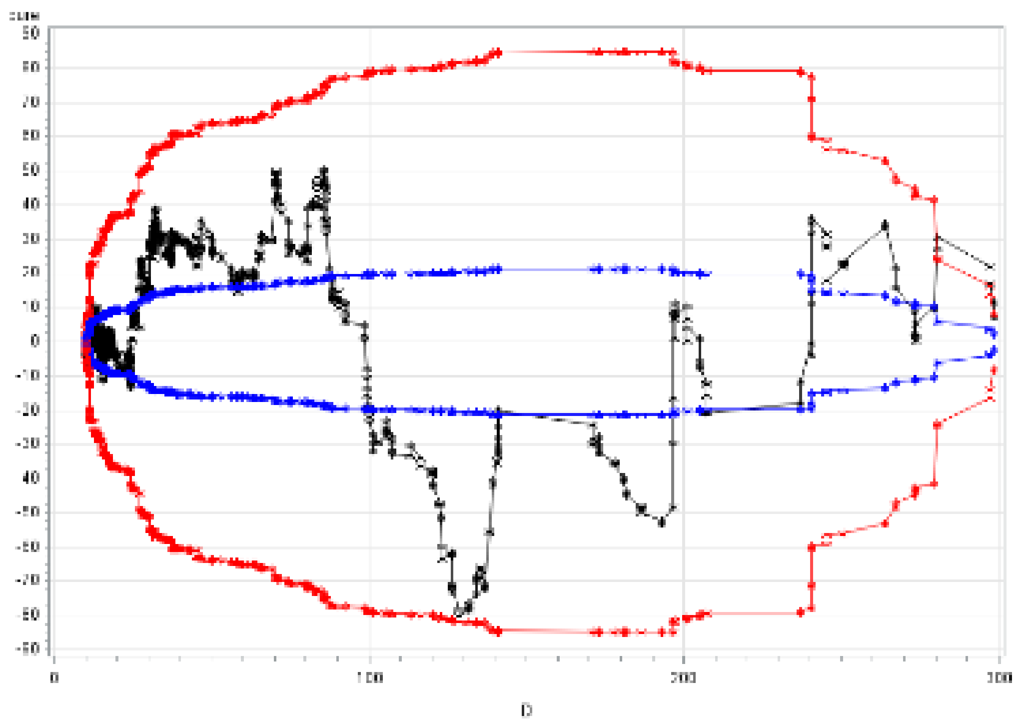


Figure 6-39. Model 12 duration CURE plot

6.3 Rural Two-Lane Highway Work Zone Models

This section shows three different models that were fitted to the rural two-lane work zone sample.

6.3.1 Model 13, Rural Two-Lane Work Zones with $\boldsymbol{\alpha = \alpha_0}$

This model was made by considering a constant overdispersion. The functional form was similar to expressway model 9. Table 6-13 summarizes the estimated parameters of Model 13 with following functional form:

$$N_C = \mathrm{e}^{\beta_0} AADT^{\beta_1} L^{\beta_2} D^{\beta_3} e^{\beta_4 * \frac{Signal}{L}} e^{\beta_5 Injury}$$

Table 6-13. Model 13 parameters for fatal-injury or PDO crashes

Explanatory Variable	**Parameter Estimates**	**Standard Error**	**p-value**
Constant	-12.0750	0.3141	<.0001
AADT	0.8588	0.03807	<.0001
L	0.8426	0.03680	<.0001
D	0.9368	0.04565	<.0001
$\frac{Signal}{L}$	0.5324	0.07441	<.0001
Injury	-0.6445	0.07515	<.0001
Overdispersion, $\boldsymbol{\alpha_0}$	2.5065	0.2006	
Number of Observations		6,095	

Each variable added was statistically beneficial to the model (using the $\aleph^2$ test) and all explanatory variables were statistically significant at 1% level. From this crash prediction model, crash modification factors can be derived for all explanatory variables. A 1% increase in AADT, length, and duration leads to the number of crashes increasing by 0.86%, 0.84% and 0.93%, respectively. The frequency of injury crashes is 52.49% and of PDO crashes as calculated by $e^{(-0.6445)} = 0.5249$. The overdispersion was 2.5065 which was not satisfactory. However, the reason for the poor overdispersion was a function of the nature of data. The few crashes occurring on low volume routes means that there is high uncertainty prediction of such crashes.

Figure 6-40 shows the Model 13 AADT CURE plot and it went beyond the band limits.

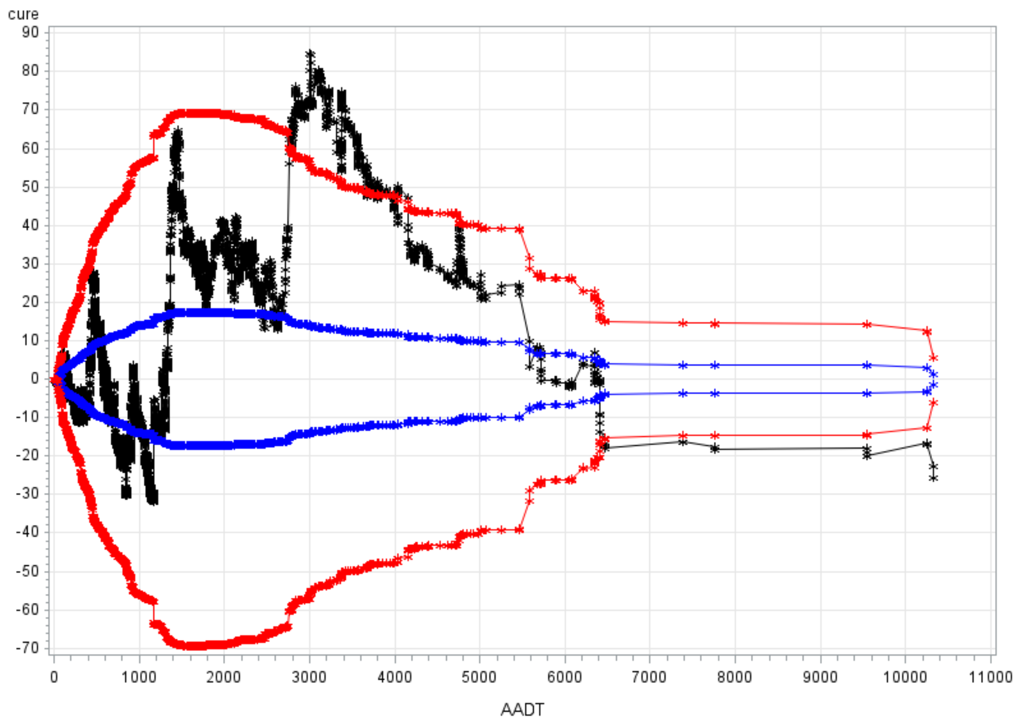


Figure 6-40. Model 13 AADT CURE plot

Different solutions were devised and tested to improve the fit. The best solution was to make different models to predict fatal-injury and PDO crashes. Using length and duration modified overdispersion term did not improve the models and led to overfitting.

6.3.2 Model 14, PDO Crashes of Rural Two-Lane Work Zones with $\boldsymbol{\alpha} = \boldsymbol{\alpha_0}$

This model was made by considering a constant overdispersion term. The functional form was similar to Model 13 but without the "Injury" variable. Table 6-14 summarizes the estimated parameters of Model 14 with following functional form:

$$N_{PDO} = \mathrm{e}^{\beta_0} AADT^{\beta_1} L^{\beta_2} D^{\beta_3} e^{\beta_4 * \frac{Signal}{L}}$$

Table 6-14. Model 14 parameters for PDO crashes

Explanatory Variable	Parameter Estimates	Standard Error	p-value
Constant	-12.4313	0.4188	<.0001
AADT	0.9259	0.05018	<.0001
L	0.7909	0.04635	<.0001
D	0.9322	0.06088	<.0001
$\frac{Signal}{L}$	0.5748	0.1033	<.0001
Overdispersion, $\boldsymbol{\alpha_0}$	2.7476	0.2644	
Number of Observations		6,095	

Each variable added was statistically beneficial to the model (using the $\aleph^2$ test) and all explanatory variables were statistically significant at 1% level. From this crash prediction model, crash modification factors can be derived for all explanatory variables. A 1% increase in AADT, length, and duration leads to the number of PDO crashes increasing by 0.93%, 0.79% and 0.93%, respectively. The overdispersion was 2.7476 which was not satisfactory. As mentioned above, the poor overdispersion results from the nature of data; there were not enough crashes on low volume roads, and the uncertainty of crash occurrence was high. Figure 6-41 shows the Model 14 AADT CURE plot, and it improved in comparison to Model 13.

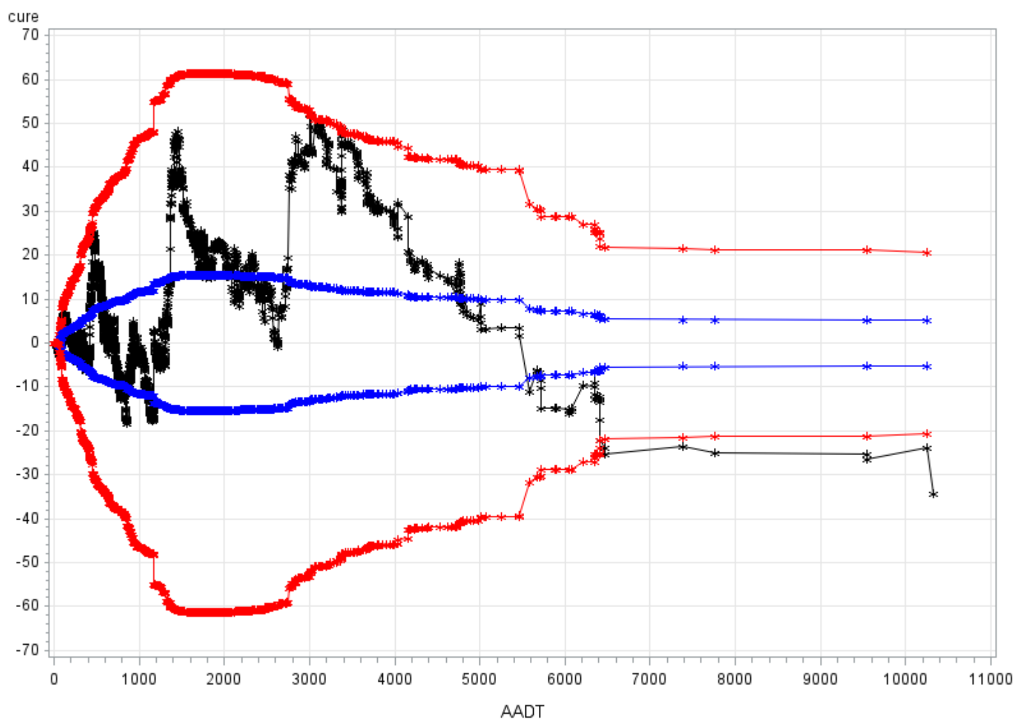


Figure 6-41. Model 14 AADT CURE plot

6.3.3 Model 15, Injury Crashes of Rural Two-Lane Work Zones with $\boldsymbol{\alpha} = \boldsymbol{\alpha_0}$

This model was made by considering a constant overdispersion. The functional form was similar to Model 14. Table 6-15 summarizes the estimated parameters of Model 15 with following functional form:

$$N_{PDO} = \mathrm{e}^{\beta_0} AADT^{\beta_1} L^{\beta_2} D^{\beta_3} e^{\beta_4 * \frac{Signal}{L}}$$

Table 6-15. Model 15 parameters for fatal-injury crashes

Explanatory Variable	Parameter Estimates	Standard Error	p-value
Constant	-12.1802	0.4714	<.0001
AADT	0.7481	0.05830	<.0001
L	0.9382	0.06118	<.0001
D	0.9483	0.06845	<.0001
$\frac{Signal}{L}$	0.4976	0.1139	<.0001
Overdispersion, $\boldsymbol{\alpha_0}$	2.0039	0.2963	
Number of Observations		6,095	

Each variable added was statistically beneficial to the model (using the $\aleph^2$ test) and all explanatory variables were statistically significant at 1% level. From this crash prediction model, crash modification factors can be derived for all explanatory variables. A 1% increase in AADT, length, and duration leads to the number of fatal-injury crashes increasing by 0.75%, 0.94% and 0.95%, respectively. The overdispersion was 2.0039 which was not satisfactory. Again, the high overdispersion was due to the infrequent occurrence of crashes on low volume roads. Figure 6-42 shows the Model 15 AADT CURE plot and it was slightly improved in comparison to Model 13.

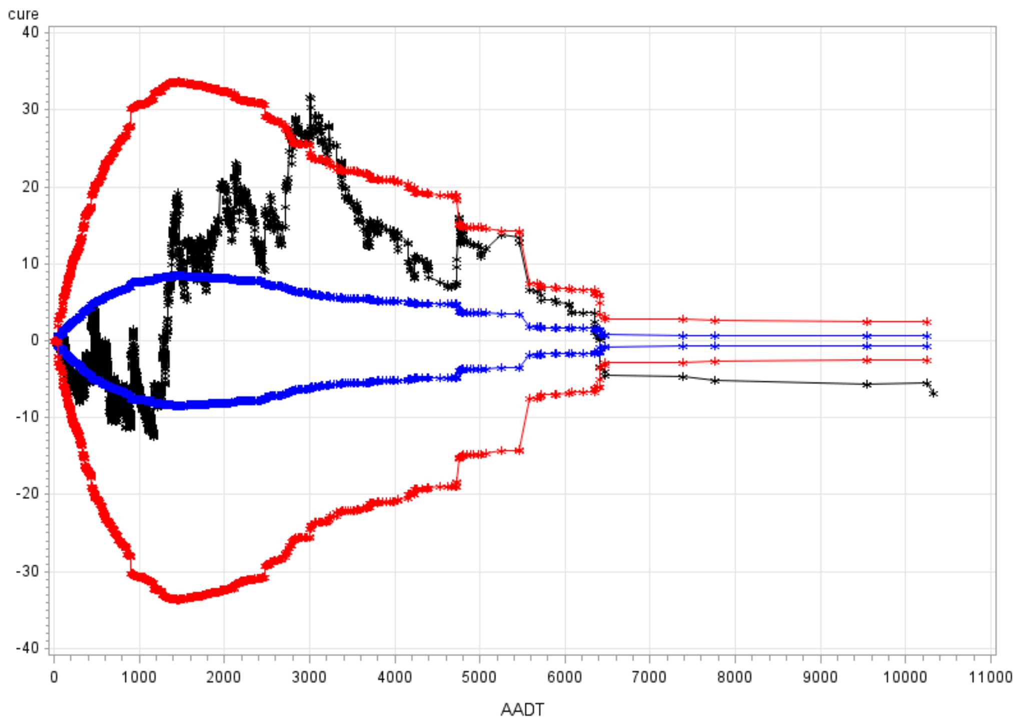


Figure 6-42. Model 15 AADT CURE plot

7 SOFTWARE DEVELOPMENT AND EXAMPLES

This section explains the theoretical basis of the software tool, gives a tutorial on it and shows the applications through few different examples. Section 7.1 focuses on the theoretical basis of the software.

7.1 Assumptions

Practitioners may face some challenges in applying the statistical models described in Chapter 6. One challenge involves the amount of time and computational effort required to generate the crash predictions. Another challenge involves how to compare different alternatives using a quantitative cost approach. A user-friendly spreadsheet tool was developed in this project to address these challenges and facilitate implementation of the developed models.

The software collects the input data from the practitioners in a user-friendly Graphical User Interface (GUI). Based on the facility type (freeway, expressway, and rural two-lane highway) selected by the user, the software chooses the proper and the most accurate model to calculate the results. Freeway models are the first 8 models, expressways include models 9 to 12 and the last three models are for rural two-lane highways.

For freeway work zones with length greater than 6 miles, the software uses models 1, 3, 5 and 7, and for shorter work zones it uses models 2, 4, 6 and 8. For rural expressway work zones it uses model 10, while for urban expressways it chooses among model 11 and 12 based on the length. Model 9 for expressways is not used by the software. Model 14 and 15 are the models that the software uses for rural two-lane highways.

In each category, the software compares the overdispersion term of the models –based on the input data– and selects the smallest; the smaller overdispersion, the more accurate. It estimates the work zone crash count and its standard error based on crash severity. The standard error in a negative binomial model for work zone plan ‘y_i’ is $SE[y_i]$

$$\mathrm{SE}[y_i] = \sqrt{\mathrm{E}[y_i]\{1 + \alpha \mathrm{E}[y_i]\}}$$

where $\mathrm{E}[y_i]$ is the estimated crash count and α is model overdispersion term. The standard errors of rural two-lane highway work zones models are relatively large and the reason is because its models’ overdispersion is larger than first 13 models.’ The main cause of lower accuracy of the rural two-lane highway models is the small crash count of these facilities.

This software tool uses the HSM 2010 crash costs (AASHTO 2010) that are based on a study using data from 2001. To account for the inflation, the discount rate from governmental sources was collected to transform the HSM values to present. HSM suggests \$7,400 and \$158,200 for PDO and fatal/injury crashes respectively. The discount rates used in the software are in Table 7-1.

Table 7-1. Discount rates used in the software

Year	Yearly Discount Rate
Before 1994	3.32%
1995-1999	3.04%
2000-2004	2.43%
2005-2009	3.75%
After 2010	0.75%

Discount rates were considered constant for each five year period. The discount rate after 2010 was constantly 0.75%. This report is published in 2015, so for the years after 2015 the software considers the same discount rate (0.75%) and computes the crash costs to that year automatically. If a user wants to use other crash cost values, he can input his own costs and the year of reference study, and the software converts the crash costs of that study to current year (i.e. 2019).

7.2 Software Design

This section provides a description of the different features of the developed software tool. This software is written in visual basic in Microsoft Excel for Windows environment. By double clicking the file shortcut the main page of the spreadsheet (Figure 7-1) opens. If prompted, the user needs to click the "Enable Editing" button and enable macros in the spreadsheet.

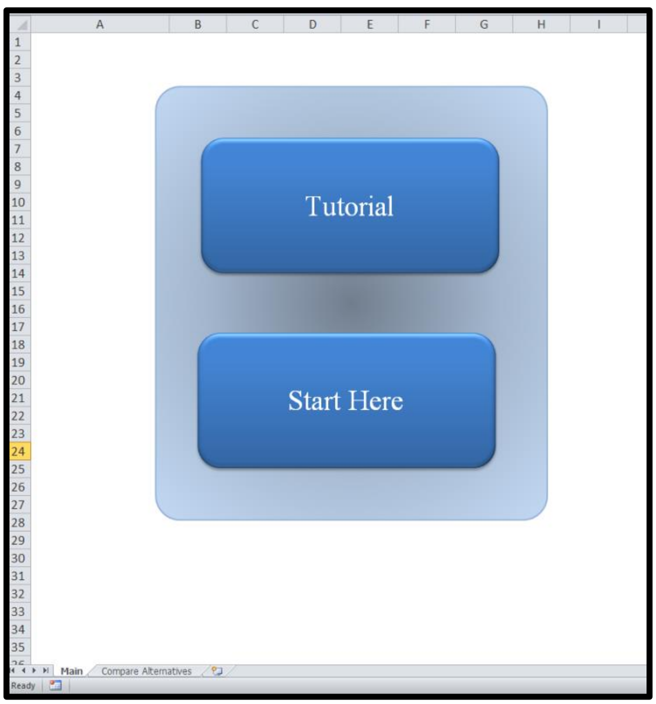


Figure 7-1. Software main page

By clicking on 'Tutorial', the user can see the necessary information about how to use the software. After that, the user can start analyzing the alternatives by clicking on 'Start Here'. This button opens the analyze window as shown in Figure 7-2.

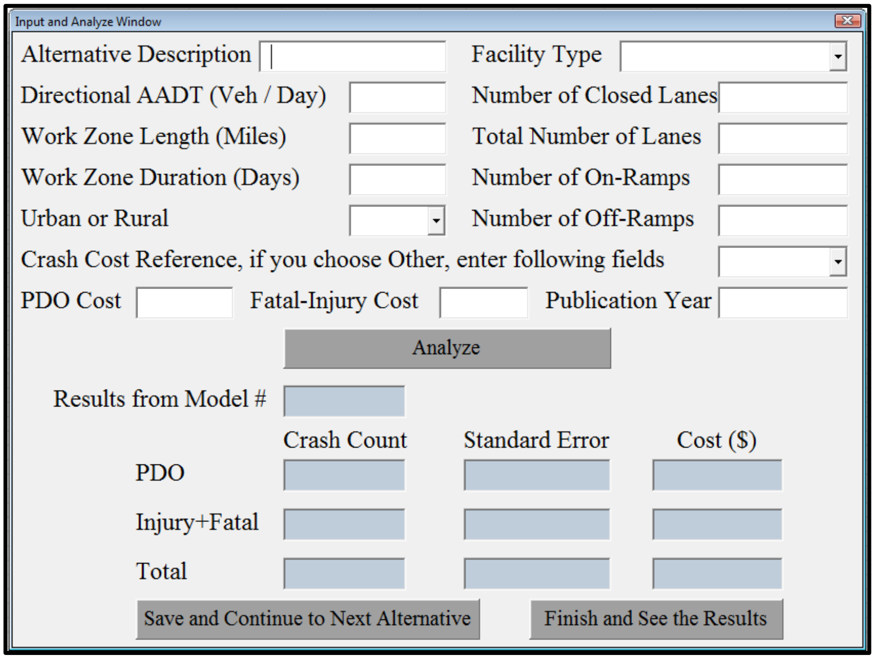


Figure 7-2. Software window for data input and analysis

The user can name the work zone plan alternative in the first box. By choosing the facility type, the variables that are required remain on the window (See Figure 7-3 to 7-5).

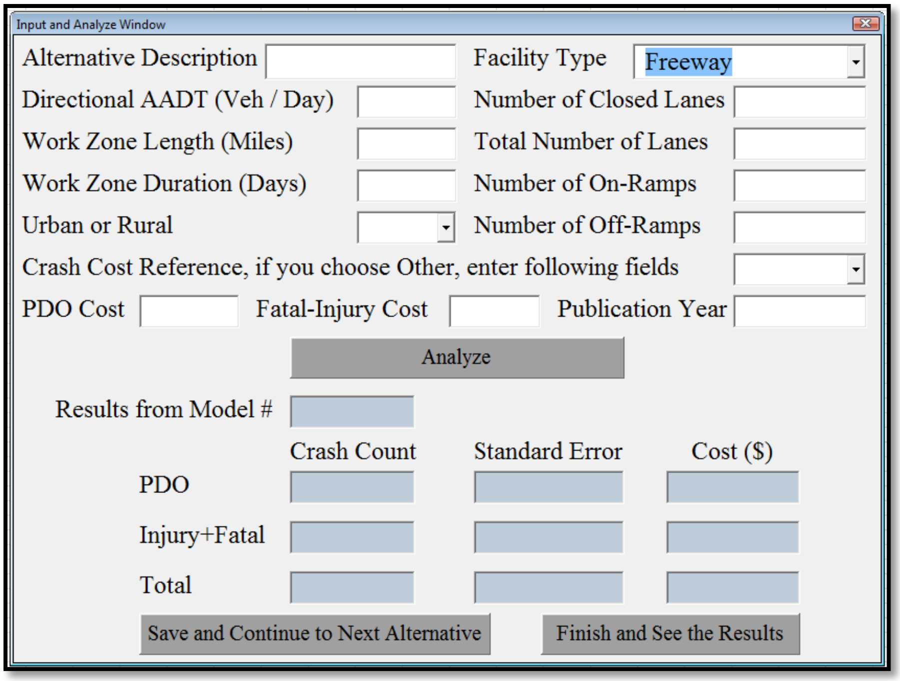


Figure 7-3. Freeway work zone required variables

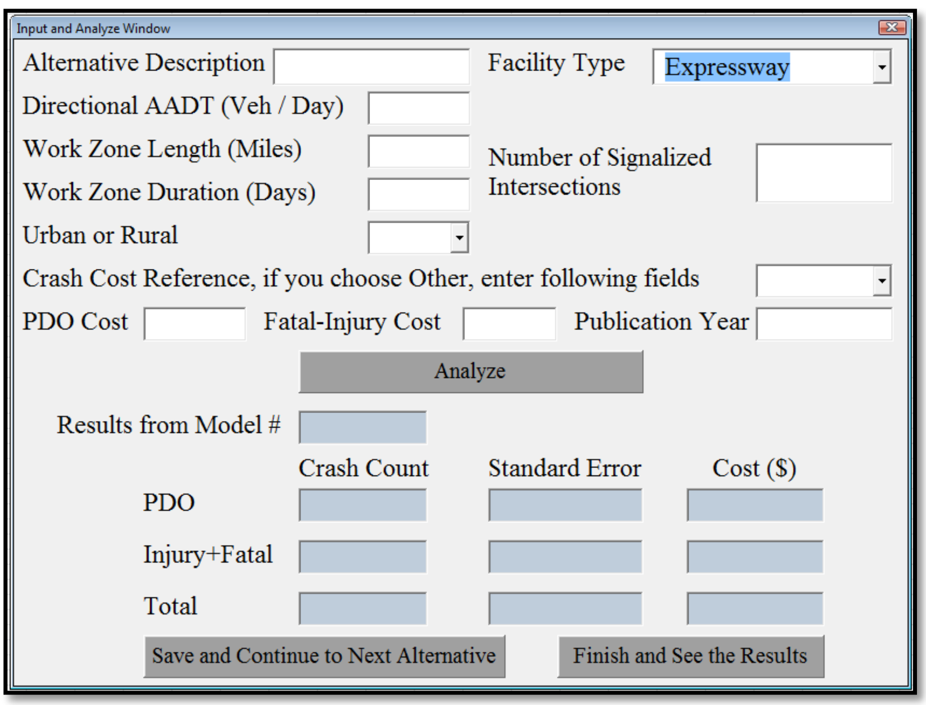


Figure 7-4. Expressway work zone required variables

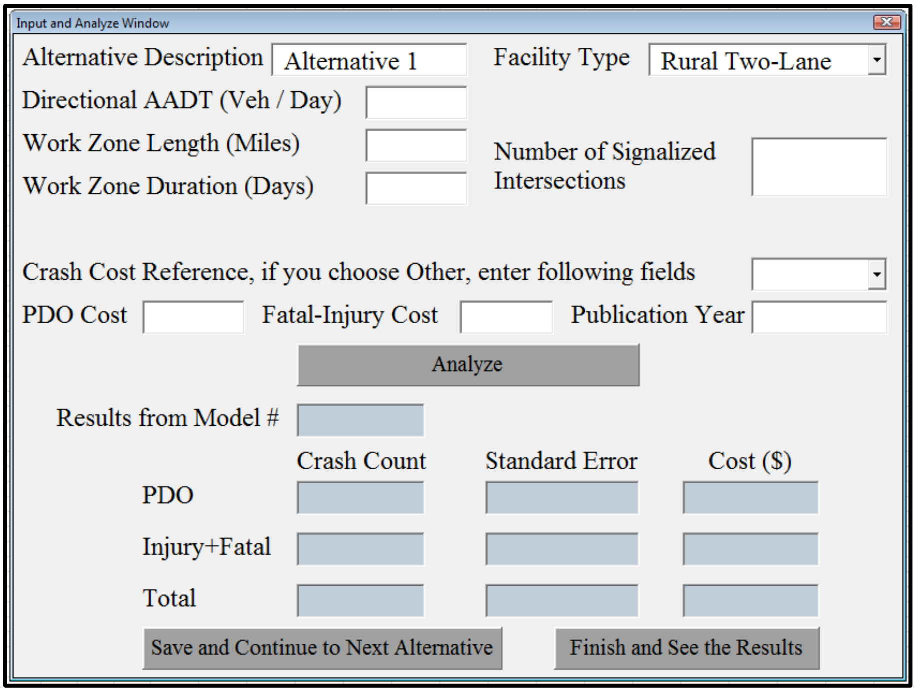


Figure 7-5. Rural two-lane work zone required variables

After entering the required work zone information, the user can select HSM 2010 for estimating work zone alternative crash cost or enter any other reference for computing crash cost (Figure 7-6).

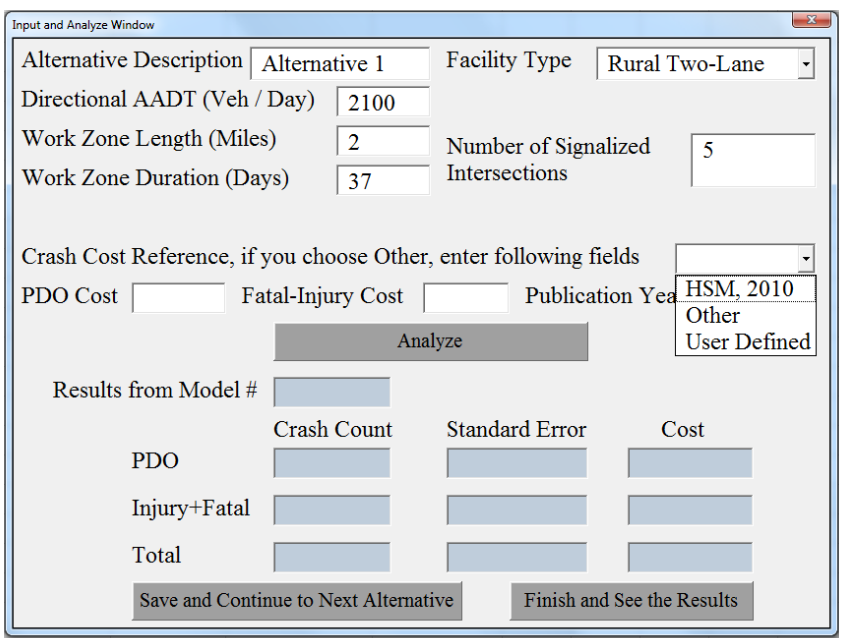


Figure 7-6. Software crash cost

With the “Other” option, the user must enter the values for each alternative. The “User Defined” option is useful when the user wants to consider his or her own crash cost reference multiple times. This option uses the values that the user enters in the “User Defined Crash Cost” worksheet (Figure 7-7).

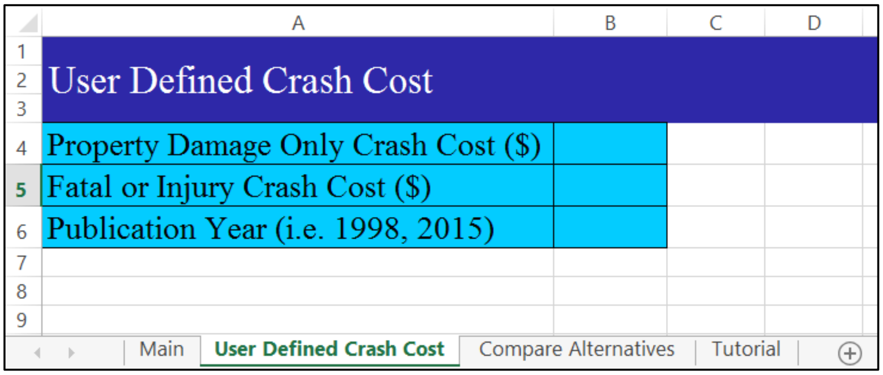


Figure 7-7. User defined crash cost worksheet

Then, the user should click ‘Analyze’ button and the results will be shown like Figure 7-8.

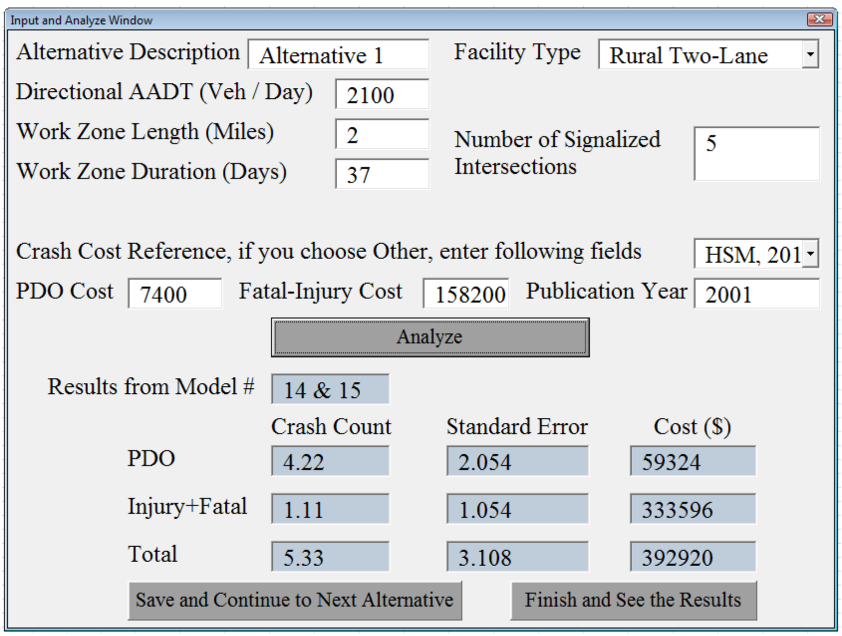


Figure 7-8. Software results window

The software shows the results including the used model number, crash count, standard error and the equivalent cost of crashes for the year used for analysis. By clicking the ‘Save and Continue to Next Alternative’, the software copies the results in the spreadsheet and opens the ‘Input and Analyze’ window for the next alternative plan. This process can be repeated for all alternative plans, and for the last alternative user needs to click on ‘Finish and See the Results’. The results and input data will be shown in ‘Compare Alternatives’ worksheet of the workbook (Figure 7-9). The input variables will also be shown for the user to check if the data entered are correct.

Alternatives Comparison			
	Output		
	FW WZ	Exp WZ	RTL WZ
Expected Number of PDO Crashes	9.8	42.91	10.58
Standard Error of PDO Estimation	4.644	28.311	3.253
Expected Number of Fatal and Injury Crashes	3.16	15.7	2.24
Standard Error of Fatal and Injury Estimation	2.094	10.828	1.497
Total Crash Cost; value in 2016	$1,169,024	$5,720,784	$883,581
Model Used: #	6	10	14 & 15
	Input		
	FW WZ	Exp WZ	RTL WZ
AADT	50000	35000	7000
Duration	56	65	30
Length	3	5	2
Urban/Rural	Urban	Rural	Rural
Number of Closed Lanes	2		
Total Number of Lanes	4		
Number of On-ramps	3		
Number of Off-ramps	2		
Number of Signalized Intersections		5	5
Crash Cost Reference; Publication Year	HSM (2010)	HSM (2010)	HSM (2010)
PDO Crash Cost	$7,400	$7,400	$7,400
Fatal and Injury Crash Cost	$158,200	$158,200	$158,200
Facility Type	Freeway	Expressway	Rural Two-Lane
			Developed by University of Missouri-Columbia; TransZou

Figure 7-9. Sample output of the software

7.3 Sample Applications

This section shows sample applications for using the work zone safety analysis tool described in this report including work zone safety screening, work zone phasing alternative evaluation, and work zone scheduling comparison.

7.3.1 Scheduling Example

A state transportation agency is considering a major shoulder rehabilitation of a 5-mile corridor of a major three lane rural freeway. The freeway AADT is 45,000 vehicles per day; the segment has 2 on-ramps and 3 off-ramps. The agency has short-listed two alternatives based on preliminary analysis of traffic and safety data. The first alternative is to complete the rehabilitation of the entire 5-mile corridor in 100 days with 1 closed lane. The second alternative takes 140 days with no closed lanes. Figure 7-10 shows that the second alternative has 3.50 and 1.13 more PDO and fatal-injury crashes respectively. All in all, alternative 1 estimated crash cost is ($1,909,027 – 1,491,056) = $417,971 less than alternative 2.

Alternatives Comparison		
	Output	
	Alternative 1	Alternative 2
Expected Number of PDO Crashes	12.51	16.01
Standard Error of PDO Estimation	4.354	4.853
Expected Number of Fatal and Injury Crashes	4.03	5.16
Standard Error of Fatal and Injury Estimation	2.168	2.438
Total Crash Cost; value in 2016	$1,491,056	$1,909,027
Model Used: #	6	6
	Input	
	Alternative 1	Alternative 2
AADT	45000	45000
Duration	100	140
Length	5	5
Urban/Rural	Rural	Rural
Number of Closed Lanes	1	0
Total Number of Lanes	3	3
Number of On-ramps	2	2
Number of Off-ramps	3	3
Number of Signalized Intersections		
Crash Cost Reference; Publication Year	HSM (2010)	HSM (2010)
PDO Crash Cost	$7,400	$7,400
Fatal and Injury Crash Cost	$158,200	$158,200
Facility Type	Freeway	Freeway
Developed by University of Missouri-Columbia; TransZou		

Figure 7-10. Work zone scheduling example, software output

7.3.2 Work Zone Screening

An agency wants to participate in a bid for rehabilitating a 4 mile section of an urban expressway road with three signalized intersections. Their schedule is to finish the work in 60 days. The expressway AADT is 35,000 vehicles per day in one direction. What is the number of crashes by severity that this agency should expect?

Results are as shown in Figure 7-11, having 15.45 PDO crashes and 5.14 fatal and injury crashes. Using HSM (AASHTO 2010) values and governmental declared discount rates, the value of these crashes in 2015 is $1,894,102.

Alternatives Comparison	
	Output
	Road Rehabilitating
Expected Number of PDO Crashes	15.45
Standard Error of PDO Estimation	13.47
Expected Number of Fatal and Injury Crashes	5.14
Standard Error of Fatal and Injury Estimation	4.849
Total Crash Cost; value in 2016	$1,894,102
Model Used: #	12
	Input
	Road Rehabilitating
AADT	35000
Duration	60
Length	4
Urban/Rural	Urban
Number of Closed Lanes	
Total Number of Lanes	
Number of On-ramps	
Number of Off-ramps	
Number of Signalized Intersections	3
Crash Cost Reference; Publication Year	HSM (2010)
PDO Crash Cost	$7,400
Fatal and Injury Crash Cost	$158,200
Facility Type	Expressway
	Developed by University of Missouri-Columbia; TransZou

Figure 7-11. Work zone screening example, software output

8 CONCLUSION

The Highway Safety Manual (HSM) (AASHTO 2010) provided many quantitative safety assessment tools for different road facility types and was a great advance in safety. As mentioned in the introduction section of this report, HSM only introduces two work zone Crash Modification Factors (CMFs) for freeway work zone length and duration. So, there is no quantitative method for work zones on other facility types in the HSM.

These two CMFs were extracted from a study by Khattak et al. (2002). Because the data of Khattak et al. (2002) were from California work zones with high traffic volume, there is a need to calibrate using Midwest data. A calibration factor of 3.78 was found by Rahmani et al. (2016) for Missouri data which is significantly larger than 1. When the calibration factor is significantly different from 1, the HSM suggests making new crash prediction models. Therefore, there was a need to develop new freeway work zone models with Missouri data instead of calibrating the HSM model. In this study, models were developed to predict freeway work zone crashes based on data from Missouri. In addition to freeway models, new models were also developed for expressway and rural two-lane work zones. Fifteen different models were made for work zones in the three mentioned facilities. A user-friendly spreadsheet software tool was developed in this project to facilitate implementation of the developed models.

This study also conducted two online surveys of contractors and Department of Transportation (DOT) representatives about the current state of the practice for work zone safety. The respondents included 7 contractors, 27 DOT representatives, and 2 FHWA representatives. The survey results indicate that many agencies look at work zone safety informally using engineering judgment. Respondents indicated that they would like a tool that could help them to quantitatively assess work zone safety.

This report documented the surveys, data, methodology, results and software tutorial of the study. Obtaining useful and appropriate data for work zone safety modeling is a great challenge. This is because the majority of work zones have very short durations with no crashes. In addition, multiple sources of data need to be combined in order to produce the variables needed for modeling. One source is the work zone database that contains information on work zone characteristics such as duration, length, urban/rural, and location. Another source is the crash database that provides information on crashes such as crash location, date/time, and severity. A third source is the traffic data for the vehicles that travel through the work zones.

A large sample of 20,837 freeway, 8,993 expressway and 64,476 rural two-lane work zones in Missouri was analyzed to derive 15 work zone crash prediction models. As mentioned above most of the work zones of short duration and length have few if any crashes. This study developed a way to extract the most appropriate samples. The most appropriate samples of 1,546 freeway, 1,189 expressway and 6,095 rural two-lane work zones longer than 0.1 mile and with a duration of greater than 10 days were used to make eight, four and three models respectively.

In work zone databases, the footprint of a work zone is recorded as the beginning and end of the work area. To account for the work zone signage areas, including the advance warning areas, this study used MUTCD recommended temporary traffic control plans' thresholds for freeway, expressway and rural two-lane work zones. To this end, the road functional type, speed limit, lane width and area designation (urban-rural) were also collected. In contrast, the model used by the HSM (same for most studies in the literature) classified all crashes within 0.5 mile (0.8 km) of the beginning and 0.5 mile (0.8 km) after the end of the work zone as work zone crashes.

By checking different distributions of Poisson, Negative Binomial, Zero Inflated Poisson and Zero Inflated Negative Binomial, the Negative Binomial model was found to perform the best. All 15 of the Negative Binomial models developed in this study included the basic variables of AADT, duration, length, urban/rural, and injury. In addition to these basic variables, the freeway models also had number of closed lanes, total number of lanes, the number of on-ramps, and the number of off-ramps. The expressway and rural two-lane models both only had one additional variable which was the number of signalized intersections. All the models and variables in them were found to be statistically significant.

The research presented in this report can be expanded in several ways. First, Empirical Bayes or even full Bayes can be utilized to address regression-to-the-mean problem. This can be a significant undertaking as each work zone site would need to be calibrated and modeled using HSM Safety Performance Functions. Second, data from other states could be used to account for geographical and driver differences within other states.

REFERENCES

AASHTO. 2010. Highway Safety Manual. Washington, DC.

Chen, E., and Tarko, A.P. 2014. Modeling safety of highway work zones with random parameters and random effects models. *Analytic Methods in Accident Research,* 1, 86-95.

Edara, P. 2009. *Evaluation of Work Zone Enhancement Software*. Missouri DOT Research Report. RI 07-062.

Edara, P., Sun, C., Zhu, Z. 2013. *Calibration of Work Zone Software Programs.* USDOT, MATC and Missouri DOT Final Report. December.

Elias, A.M. and Z.J. Herbsman. 2000. Risk Analysis Techniques for Safety Evaluation of Highway Work Zones. *Transportation Research Record: Journal of the Transportation Research Board.* No. 1715, pp. 10-17.

Federal Highway Administration (FHWA). 1996. *Investigation of Highway Workzone Crashes - Summary Report*. Publication FHWA-RD-96-100. FHWA, U.S. Department of Transportation.

Federal Highway Administration (FHWA). 2009a. *Manual on Uniform Traffic Control Devices for Streets and Highways (MUTCD).* Washington, DC.

Federal Highway Administration (FHWA). 2009b. Work Zone Safety and Mobility Fact Sheet - 10th Annual National Work Zone Awareness Week. FHWA, safety.fhwa.dot.gov/wz/wz_awareness/2009/factsht09.cfm. Accessed May 25, 2014.

Hauer, E. Observational Before/After Studies in Road Safety. 1997. Estimating the Effect of Highway and Traffic Engineering Measures on Road Safety. Emerald Group Publishing Limit.

Hauer, E. 2001. Overdispersion in modelling accidents on road sections and in Empirical Bayes estimation. Accident Analysis & Prevention, 33(6), pp. 799-808.

Hauer, E. 2004. Statistical Road Safety Modeling, Transportation Research Record 1897, pp. 81-87.

Hauer, E. 2015. The Art of Regression Modeling in Road Safety. Springer.

Hauer, E. and Bamfo, J. 1997. Two Tools for Finding What Function Links the Dependent Variable to the Explanatory Variables, Proceedings ICTCT (International Cooperation on Theories and Concepts in Traffic Safety), Lund, Sweden.

Heydecker, B. G., and Wu, J. 2001. Identification of sites for road accident remedial work by Bayesian statistical methods: an example of uncertain inference. Advances in Engineering Software, 32(10), pp. 859-869.

Khattak, A. J., A. J. Khattak, and F. M. Council. 2002. Effects of Work Zone Presence on Injury and Non-Injury Crashes. *Accident Analysis and Prevention*, Vol. 34, pp.19-29.

Lord, D. and Mannering, F. 2010. The Statistical Analysis of Crash-Frequency Data: A Review and Assessment of Methodological Alternatives. *Transportation Research Part A* 44(5), 291–305.

Lord, D., Washington, S. P., and Ivan, J. N. 2005. Poisson, Poisson-Gamma and Zero-Inflated Regression Models of Motor Vehicle Crashes: Balancing Statistical Fit and Theory. *Accident Analysis & Prevention*, 37(1), 35-46.

Ozturk, O., Ozbay, K., Yang, H., and Bartin, B. 2013. Crash Frequency Modeling for Highway Construction Zones. *Transportation Research Board 92nd Annual Meeting,* Washington, DC.

Ozturk, O., K. Ozbay, and H. Yang. 2014. Estimating the Impact of Work Zones on Highway Safety. Transportation Research Board 93rd Annual Meeting, Washington, DC.

Pal, R., and Sinha, K.C. 1996. Analysis of Crash Rates at Interstate Work Zones in Indiana. *Transportation Research Record: Journal of the Transportation Research Board*, 1529, 43-53.

Qi, Y., Srinivasan, R., Teng, H., and Baker, R.F. 2005. Frequency of Work Zone Accidents on Construction Projects. *Report No: 55657-03-15*. University Transportation Research Center City. College of New York, New York.

Rahmani, R., Sun, C., Edara, P., Brown, H., and Zhu, Z. 2016. Freeway Work Zone Crash Prediction Models Using Missouri Data. Accepted for Presentation at TRB 95th Annual Meeting. Washington DC.

Rouphail, N.M., Z.S. Yang, and J. Fazio. 1988. Comparative Study of Short- And Long-Term Urban Freeway Work Zones. Transportation Research Record: Journal of the Transportation Research Board No. 1163, pp. 4-14.

Salkind, N. 2006. *Encyclopedia of Measurement and Statistics*. Sage, Thousand Oaks, California.

Srinivasan, S., Carrick, G., Zhu, X., Heaslip, K., and Washburn, S. 2008. Analysis of Crashes in Freeway Work Zone Queues: A Case Study. 07-UF-R-S3, Final Report. Transportation Research Center, University of Florida, Florida.

Srinivasan, R., and Carter, D. L. 2011. Development of safety performance functions for North Carolina (No. FHWA/NC/2010-09). North Carolina Department of Transportation, Research and Analysis Group.

Srinivasan, R., Ullman, G.L., Finley, M.D., and Council F.M. 2011. Use of Empirical Bayesian Methods to Estimate Temporal-Based Crash Modification Factors for Construction Zones. *Transportation Research Board 90th Annual Meeting*, Washington, DC.

Sun, C., Edara, P., Brown, H. 2014a. *Safety Handbook for Local Agencies.* USDOT MATC and Missouri DOT Final Report.

Sun, C., Edara, P., Brown, H., Zhu, Z. and Rahmani, R. 2014b. Calibration of Highway Safety Manual Work Zone Crash Modification Factors. FHWA InTrans Project 06-277 and SWZDI TPF 5(081). FHWA Smart Work Zone Deployment Initiative Pooled Fund. Final Report. University of Missouri-Columbia, Mo.

Tarko, A. and S. Venugopal. 2001. Safety and Capacity Evaluation of the Indiana Lane Merge System. FHWA/IN/JTRP-2000/19. Purdue University, West Lafayette, Indiana.

Transportation Research Board (TRB). 2010. *Highway Capacity Manual*. Fifth Edition, Transportation Research Board, Washington, DC.

Venugopal, S., and Tarko, A. 2000. Safety Models for Rural Freeway Work Zones. *Transportation Research Record: Journal of the Transportation Research Board* , 1715, 1-9.

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

Yang, H., Ozbay, K., O. Ozturk, and M. Yildirimoglu. 2013. Modeling Work Zone Crash Frequency by Quantifying Measurement Errors in Work Zone Length. Accident Analysis & Prevention 55(June), 2013, pp. 192-201.

APPENDIX A: SOFTWARE TUTORIAL

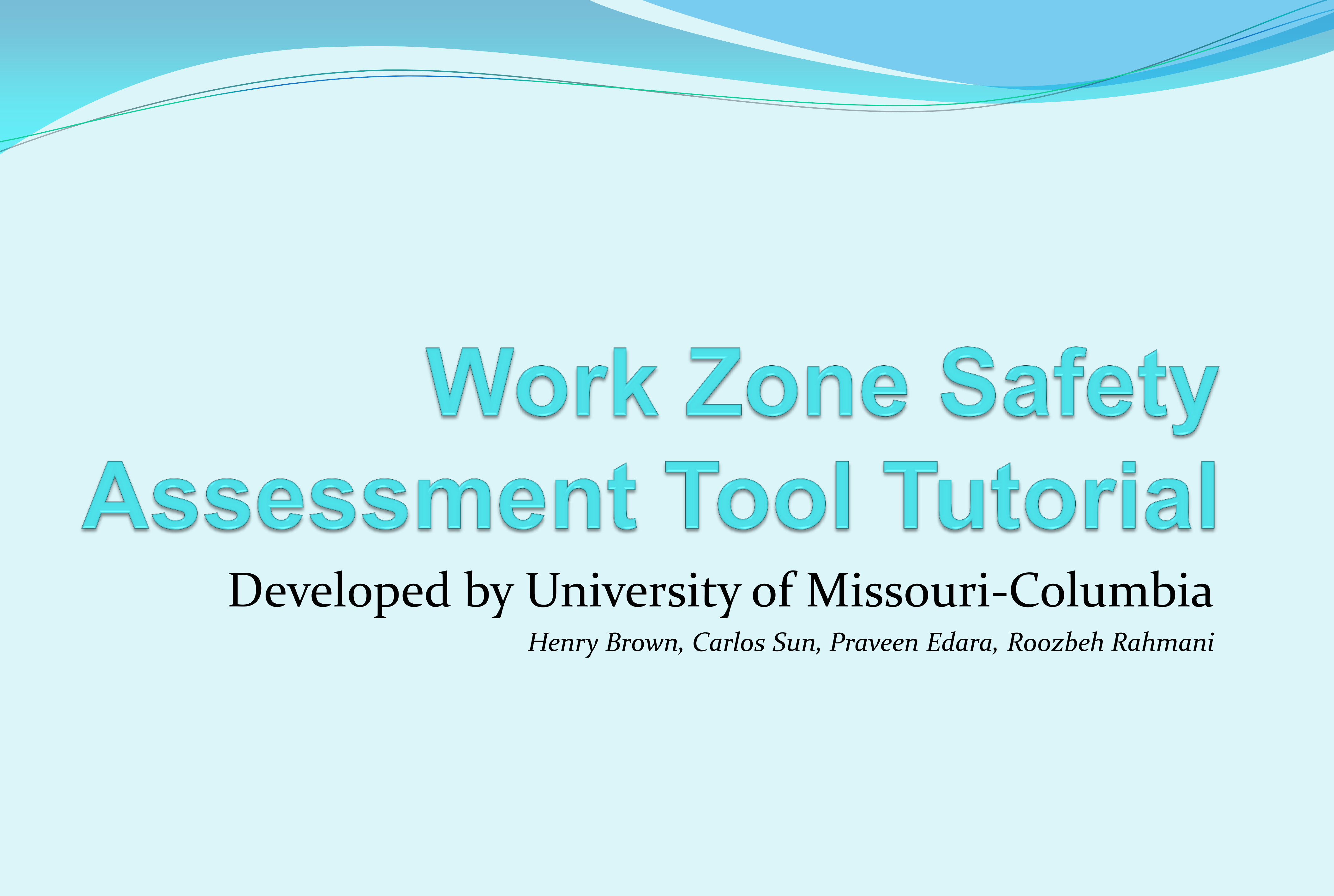


Figure A.1 Tutorial overview

Opening the Software

- Double click the spreadsheet file to open the tool
- If prompted, select the Enable Editing button
- Enable Macros
  - Microsoft Excel 2007: Click Options button and then Enable this content in the Microsoft Office Security Options dialog box
  - Microsoft Excel 2010 (or later): Click Enable Content button

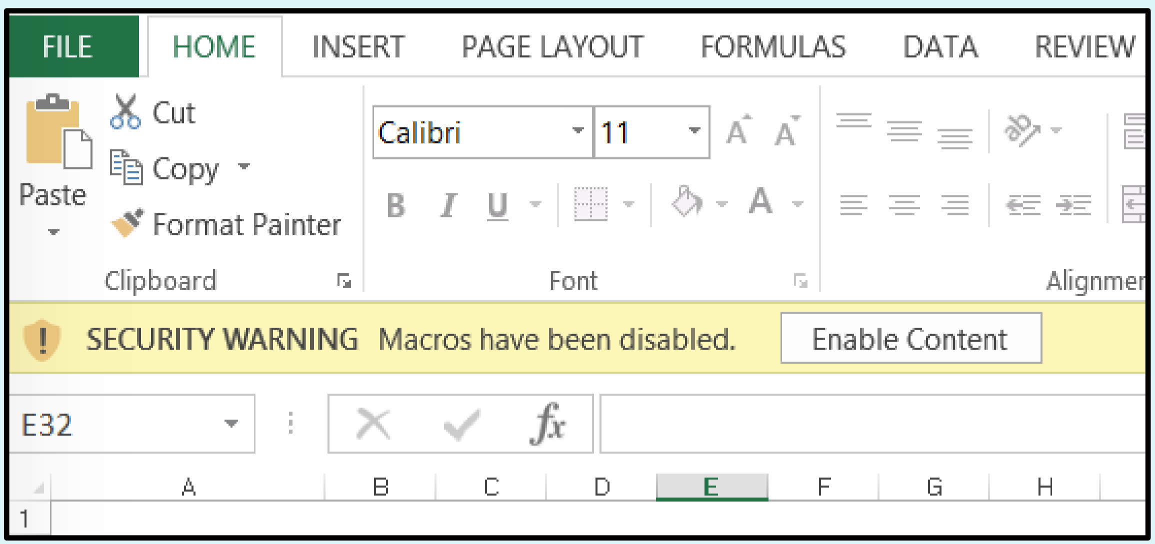


Figure A.2 Opening the software

Software Main Page

- The software is written in Microsoft Excel VBA for Windows

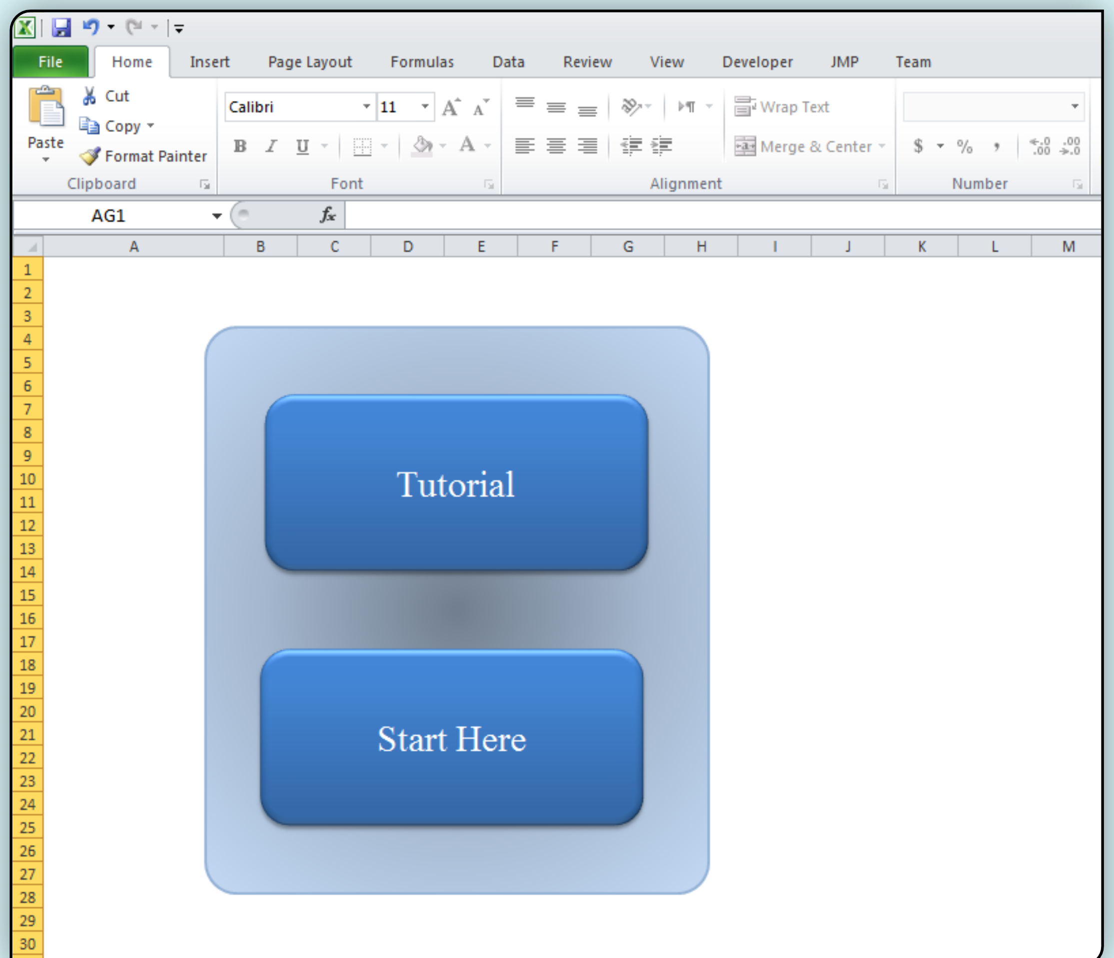


Figure A.3 Software main page

- By clicking on Tutorial button, user can see the software tutorial and by clicking on Start Here, the window for input data and analysis is opened

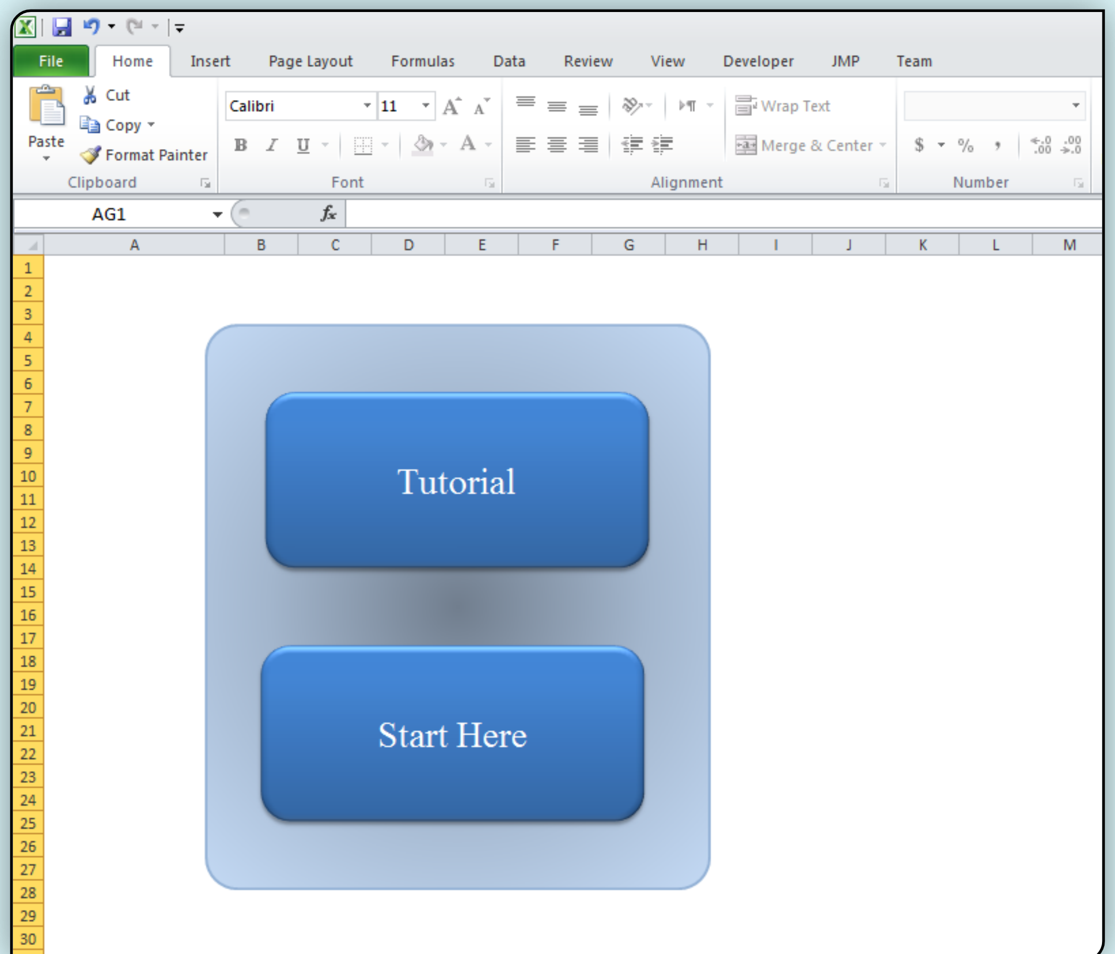


Figure A.4 Screenshot showing how to start analysis or tutorial

Software Input & Analyze Window

- As mentioned previously, by clicking on Start Here button, this window opens
- User can name each work zone plan alternative
- User can choose any of Freeway, Expressway and Rural Two-Lane work zones
- By choosing each facility type the required variables are shown in Input and Analyze Window

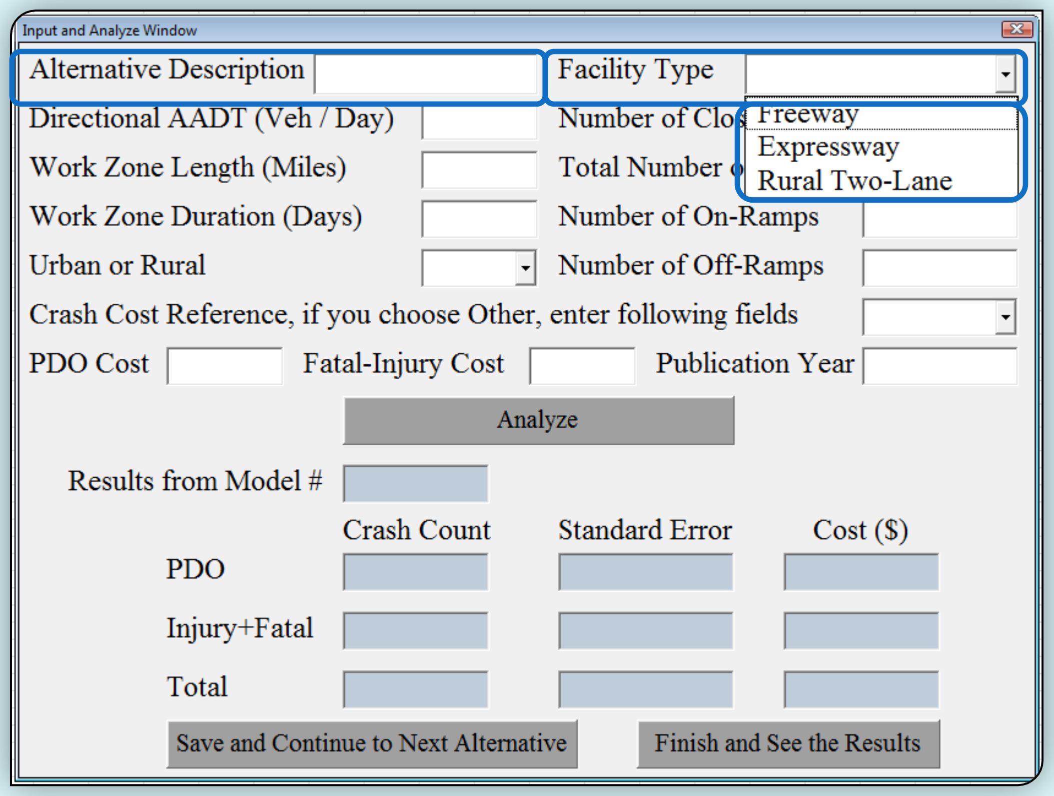


Figure A.5 Software input and analyze window

Freeway Work Zone, Required Input Data

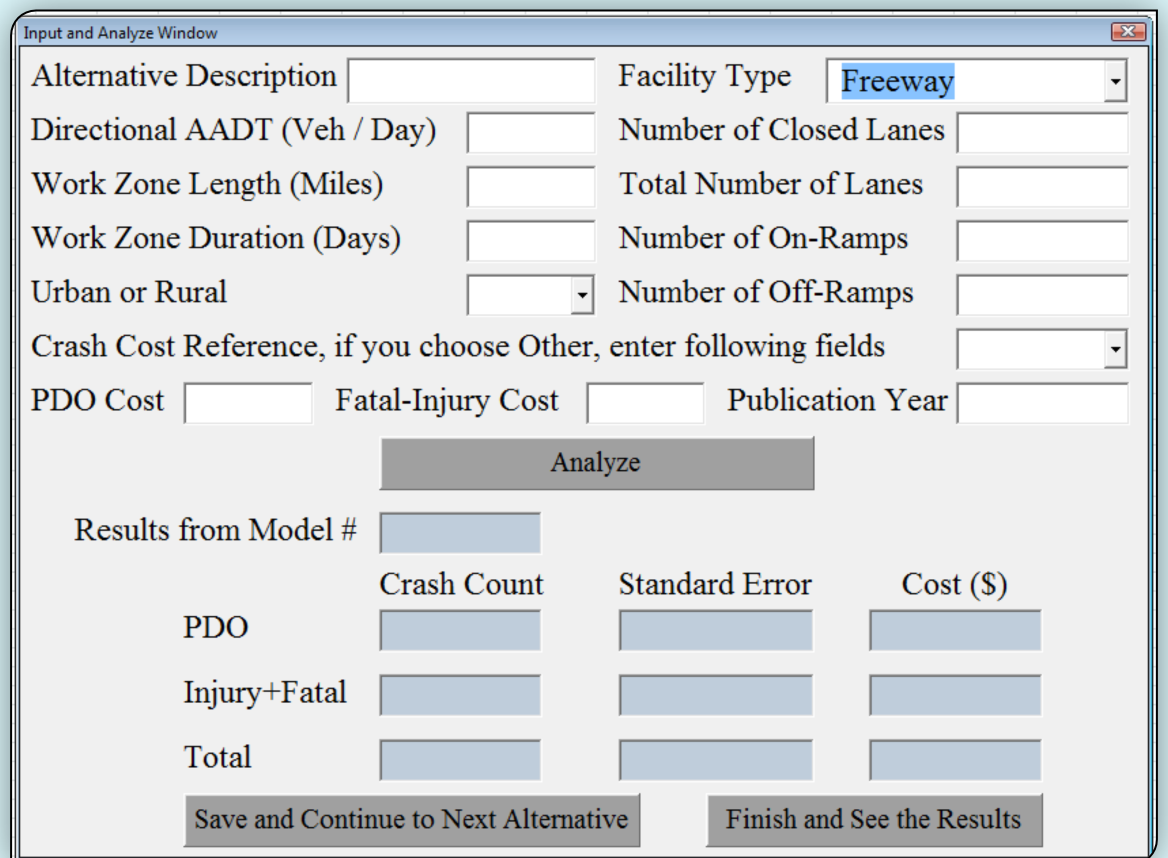


Figure A.6 Input window for freeway work zones

Expressway Work Zone, Required Input Data

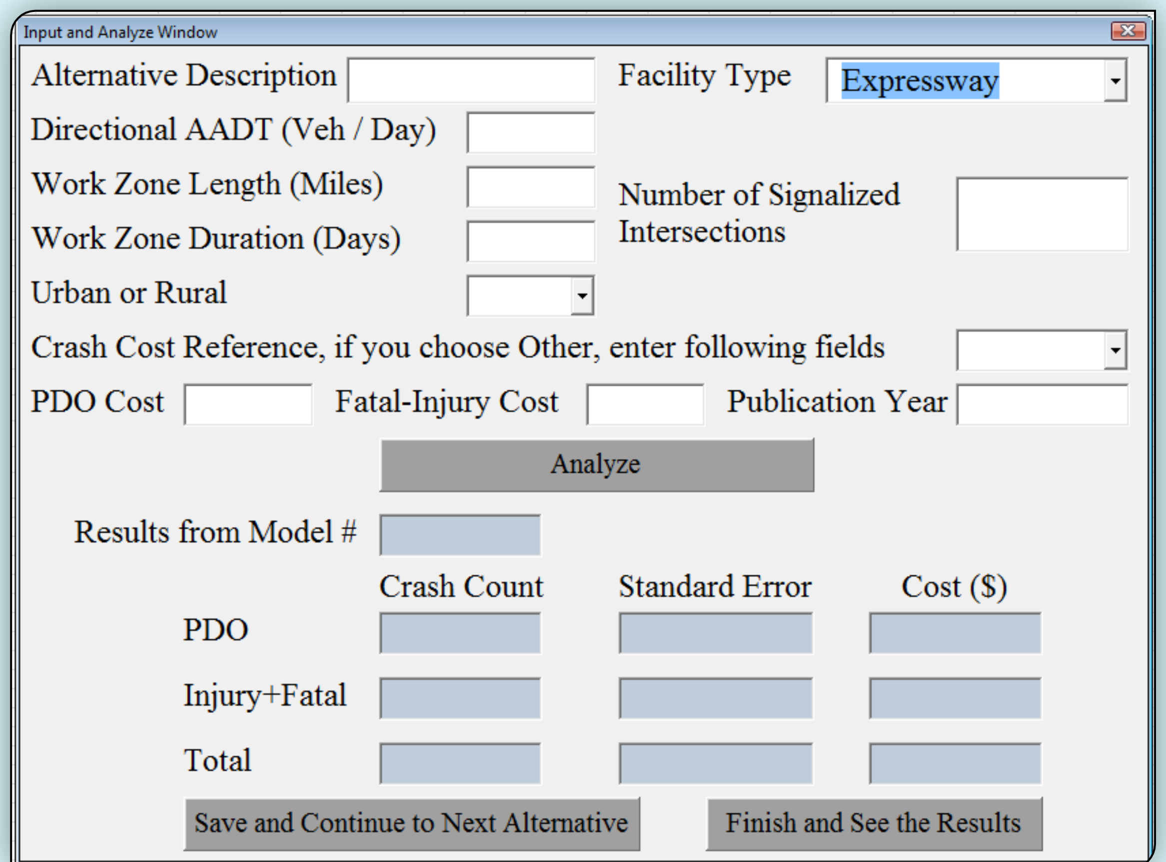


Figure A.7 Input window for expressway work zones

Rural Two-Lane Work Zone, Required Input Data

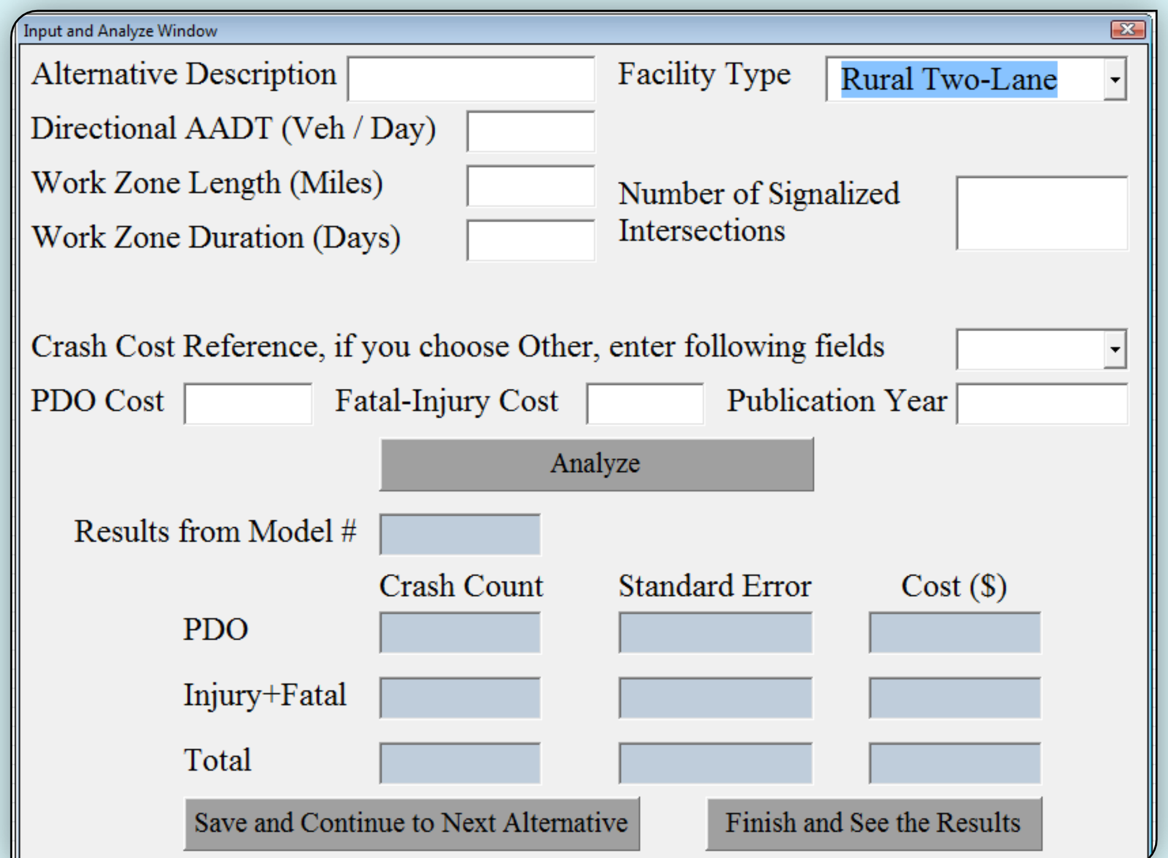
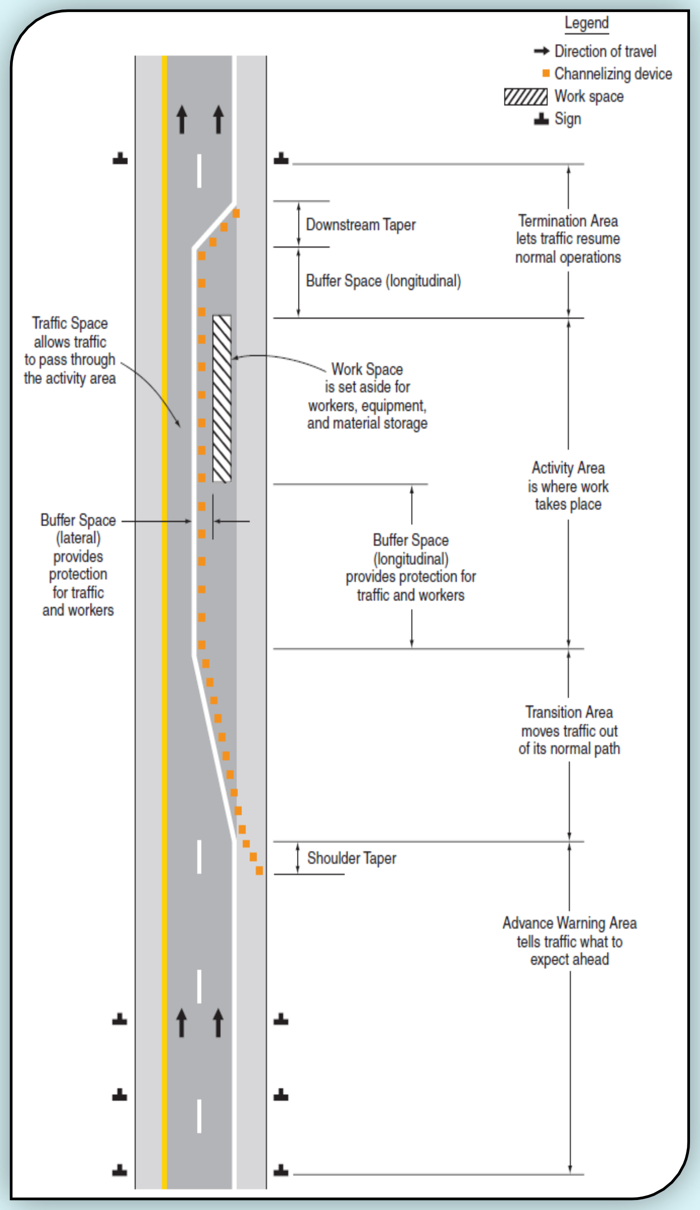


Figure A.8 Input window for rural two-lane work zones

Definitions of Input Variables

- AADT is directional Annual Average Daily Traffic and its unit is vehicles per day.
- Length is the length of Work Area of the work zone in miles
- Duration is the work zone duration in days
- Work zone urban-rural indicator (urban if the city population is more than 5,000 and rural otherwise)
- Number of closed lanes in one direction due to the work zone
- Total number of lanes in one direction where work zone is located
- Number of on-ramps and off-ramps in work area of the work zone (transition and termination areas are not included)
- Number of signalized intersections in work area of the work zone (transition and termination areas are not included)



Work zone components based on MUTCD (2009)

Figure A.9 Definition of input variables

Crash Equivalent Cost

- The software computes the equivalent crash cost of any alternative.
- The software has the equivalent crash costs from the HSM built-in.
- The Other option allows the user to enter his or her own crash cost estimates from any other studies. These values must be entered for each alternative.
- The User Defined option is useful when the user wants to consider a crash cost reference multiple times. This option uses the values that the user enters in the User Defined Crash Cost worksheet.
- The publication year is needed to convert the values to current value.
- Based on the declared US discount rate, ***the software transforms previous values to current year value. For years after 2015 software considers 0.75% discount rate.***

Year	Yearly Discount Rate
Before 1994	3.32%
1995-1999	3.04%
2000-2004	2.43%
2005-2009	3.75%
After 2010	0.75%

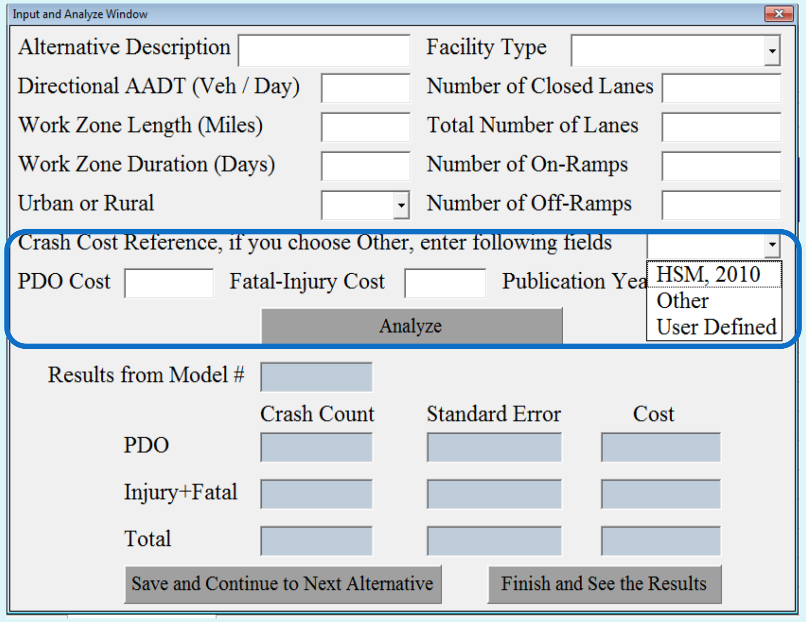
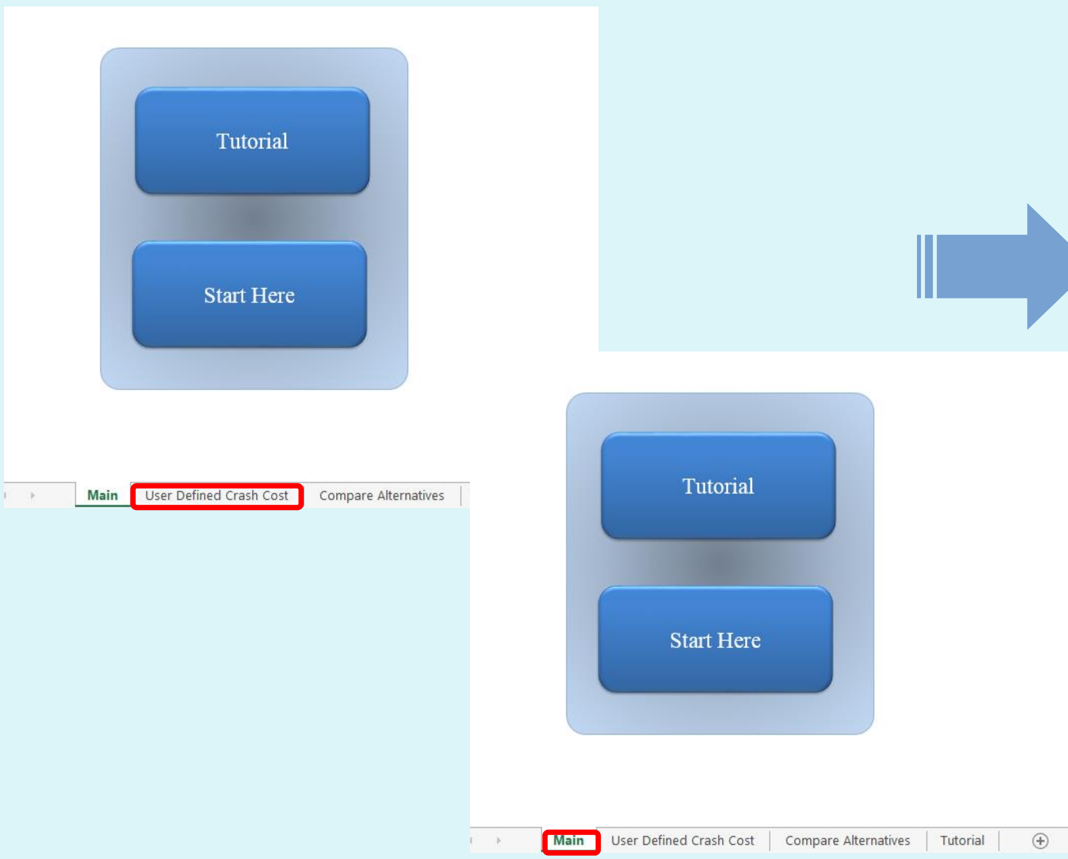


Figure A.10 Crash equivalent cost

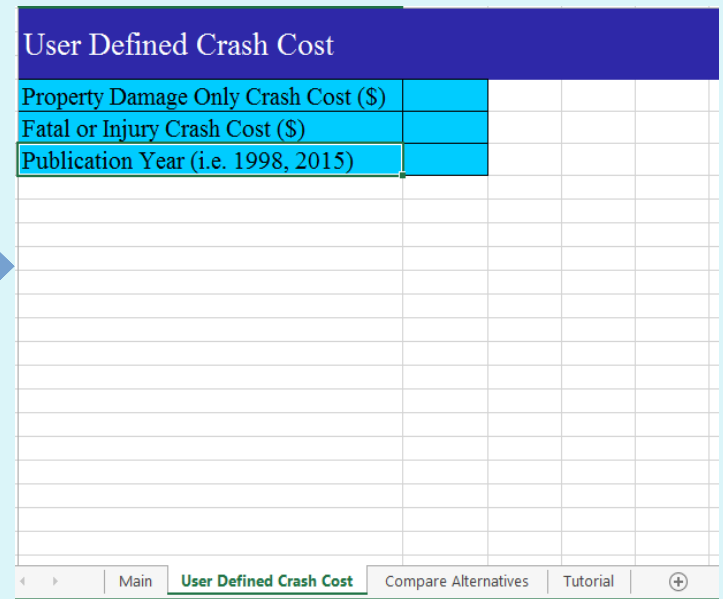
User Defined Crash Cost

- To use a crash cost reference multiple times, the user can enter the values in the User Defined Crash Cost worksheet.

Select User Defined Crash Cost Sheet



Enter values



Go to Main Sheet and continue by choosing User Defined in cost drop-box

Figure A.11 User defined crash cost

Analyze

- After entering all mentioned variables above, the user should click on Analyze button. Results are shown in the bottom half of the window.
- After clicking on Analyze button, the user should click on Save and Continue to Next Alternative button to start entering the next work zone alternative plan.
- For the last alternative instead of Save and Continue to Next Alternative button user should select the Finish and See the Results button to close the window and go to the results page.

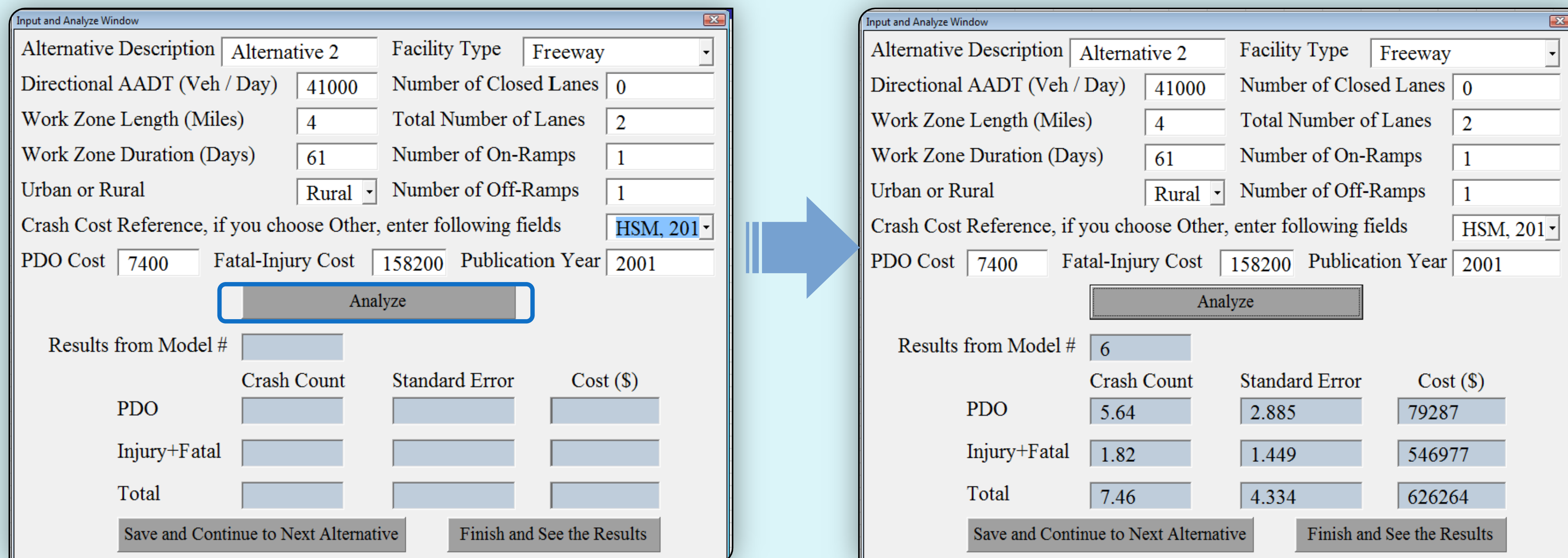
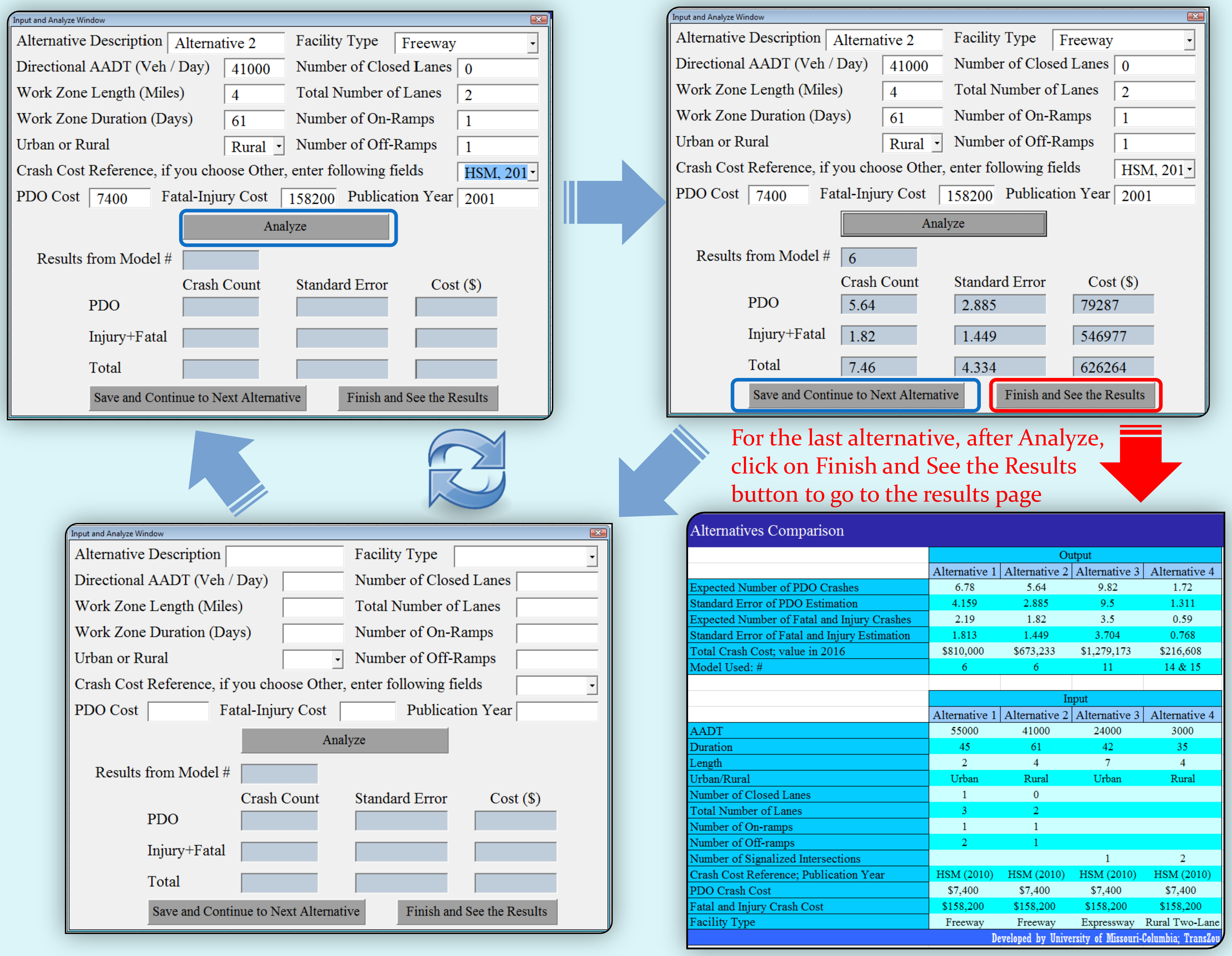


Figure A.12 Analysis window



	Output			
	Alternative 1	Alternative 2	Alternative 3	Alternative 4
Expected Number of PDO Crashes	6.78	5.64	9.82	1.72
Standard Error of PDO Estimation	4.159	2.885	9.5	1.311
Expected Number of Fatal and Injury Crashes	2.19	1.82	3.5	0.59
Standard Error of Fatal and Injury Estimation	1.813	1.449	3.704	0.768
Total Crash Cost; value in 2016	$810,000	$673,233	$1,279,173	$216,608
Model Used: #	6	6	11	14 & 15

	Input			
	Alternative 1	Alternative 2	Alternative 3	Alternative 4
AADT	55000	41000	24000	3000
Duration	45	61	42	35
Length	2	4	7	4
Urban/Rural	Urban	Rural	Urban	Rural
Number of Closed Lanes	1	0		
Total Number of Lanes	3	2		
Number of On-ramps	1	1		
Number of Off-ramps	2	1		
Number of Signalized Intersections			1	2
Crash Cost Reference; Publication Year	HSM (2010)	HSM (2010)	HSM (2010)	HSM (2010)
PDO Crash Cost	$7,400	$7,400	$7,400	$7,400
Fatal and Injury Crash Cost	$158,200	$158,200	$158,200	$158,200
Facility Type	Freeway	Freeway	Expressway	Rural Two-Lane

Developed by University of Missouri-Columbia; TransZou

Figure A-13. Comparison of alternatives