Safety Study of I-35W Improvements Done Under Minnesota's Urban Partnership Agreements (UPA) Project

A Thesis
SUBMITTED TO THE FACULTY OF
UNIVERSITY OF MINNESOTA
BY

Jingru Gao

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
MASTER OF SCIENCE

Gary A. Davis

August, 2017

Acknowledgements

I would like to thank all the people who contributed in some way to the work described in this thesis. This accomplishment would not have been possible without them.

First and foremost, I would like to express my sincerest gratitude to my adviser, Professor Gary A. Davis, for his strong support and constant input. Professor Davis' erudition and rigorous attitude as a scientist greatly inspired me and he is always my role model.

In addition, I would like to thank Dr. John Hourdos for his continuous help and guidance on my Master's studies and Professor Chap Le for his excellent courses and interest in my work.

I would also like to thank Brian Kary, Derek Leuer, Jesse Larson, Paul Jung, Lesa Monroe, and Robert Chaucierre from Minnesota Department of Transportation for their assistance and helpful comments throughout the course of the project.

Last but not the least, I must express my deepest gratitude to my parents and my friends for providing me with unfailing support and continuous encouragement throughout my years of Master's studies and the process of researching and writing this thesis.

Abstract

Minnesota's Urban Partnership Agreements (UPA), of which the majority were completed in November, 2010, consisted of a series of improvements addressing the congestion on Interstate highway I-35W corridor and in Downtown Minneapolis. MnDOT Problem Statement NS-329 noted that there was interest in extending some or all of these interventions to other corridors and called for an estimation of their safety effects to assist in making these decisions. Following the UPA, the frequency of rear-end crashes increased substantially in certain regions on I-35W. The objective of this study was to determine if the increase in crash frequency was due to changes in traffic conditions or was a direct effect of the UPA interventions.

A preliminary analysis was done to determine priority crash type and study regions. I-35W, from its start to its junction with I-94, was divided into 17 one-mile sections, and bi-directional (northbound and southbound) crash frequencies in Before-UPA (2006-2008) and After-UPA periods (2011-2013) were compiled for each one-mile section. Rear-end crash turned out to be the most prevalent crash type, but the changing trend of bi-direction rear-end crash frequencies from the Before to After period varied among those one-mile sections. Our interest lay in those regions where there was an outstanding increase in the rear-end crash frequency in the After period, which were approximately the I-35W HOT region (from TH-13 to I-494) and the I-35W PDSL region (from 37th Street to 26th Street).

Both the I-35W HOT and PDSL regions were divided into analysis sections based on constant flow and geometry criteria as well as the loop detector availability. Crash, loop detector, weather condition, and PDSL activation (only for PDSL sections) were compiled for Before and After periods for each analysis section. Rear-end crash records were extracted using MNCMAT, and hard copies of the original crash reports were then reviewed to verify the crash type, location, direction and time for each crash. Traffic conditions came from loop detector data retrieved using MnDOT's DataExtract tool. The source of weather conditions during non-crash hours was MnDOT's RWIS, while that of weather conditions during crash hours were taken from original crash reports. The PDSL historical operation data came from MnDOT's log for the Intelligent Lane Control Signal (ILCS) located at 37th Street.

Logistic regression models were established to estimate the change in rear-end crash risk in a given hour before and after the UPA project controlling for changes in traffic conditions and weather conditions.

The analysis results showed that:

(1) Most analyzed sections in the I-35W HOT region showed no significant change in rear-end crash risk associated with the UPA project except for Section S9 (southbound, just north of Minnesota River). Section N17 actually experienced fewer crashes after the UPA project, but the reduction was not as great as the change in lane occupancy would predict.

(2) The PDSL region experienced substantial increase in traffic congestion following the completion of UPA interventions. This was due to the removal of the old TH 62 & I-35W bottleneck, causing the bottleneck move northward to the I-35W & I-94 junction. The observed increase in rear-end crash risk was not associated with the operation of PDSL when controlling for the changes in traffic conditions.

(3) An “inverted U” relationship between rear-end crash risk and a proxy for traffic density, lane occupancy, when controlling for other factors, were seen in most of the analysis sections. Rear-end crashes were most likely when lane occupancies were approximately 20%-30%.

This study demonstrated a methodology that could be used to evaluate the safety effects of freeway-related projects. To be more specific, this study worked out a way to estimate changes in hourly crash risk while controlling for variations in traffic conditions.

Table of Contents

List of Tables

List of Figures

Chapter 1: Introduction

The Federal Highway Administration's (FHWA) Urban Partnership Agreements (UPA) was aimed to implement a series of congestion reduction strategies that included what is referred to as the "4 Ts": (1) a tolling (congestion pricing) demonstration, (2) enhanced transit services, (3) increased emphasis on telecommuting and flex scheduling, and (4) the deployment of advanced technology (FHWA UPA). In 2007, four metropolitan areas, Miami, Minneapolis/St. Paul, San Francisco, and Seattle were announced to be the UPA Partners.

The local partners of Minnesota's UPA project include the Minnesota Department of Transportation (MNDOT), the Twin Cities Metropolitan Council, Metro Transit, the City of Minneapolis, Minnesota Valley Transit Authority, Anoka, Dakota, Hennepin and Ramsey Counties, Center for Transportation Studies and Hubert H. Humphrey Institute of Public Affairs at the University of Minnesota, and four Transportation Management Organizations (TMOs) (Turnbull et al., 2013).

Minnesota's UPA project consisted of a package of interventions addressing the congestion on I-35W corridor and in Downtown Minneapolis. Generally, interventions of four categories, tolling, transit, telecommuting, and technology, were implemented (Turnbull et al., 2013).

Figure 1. 1 and **Figure 1. 2** show the interventions implemented in Minnesota's UPA project.

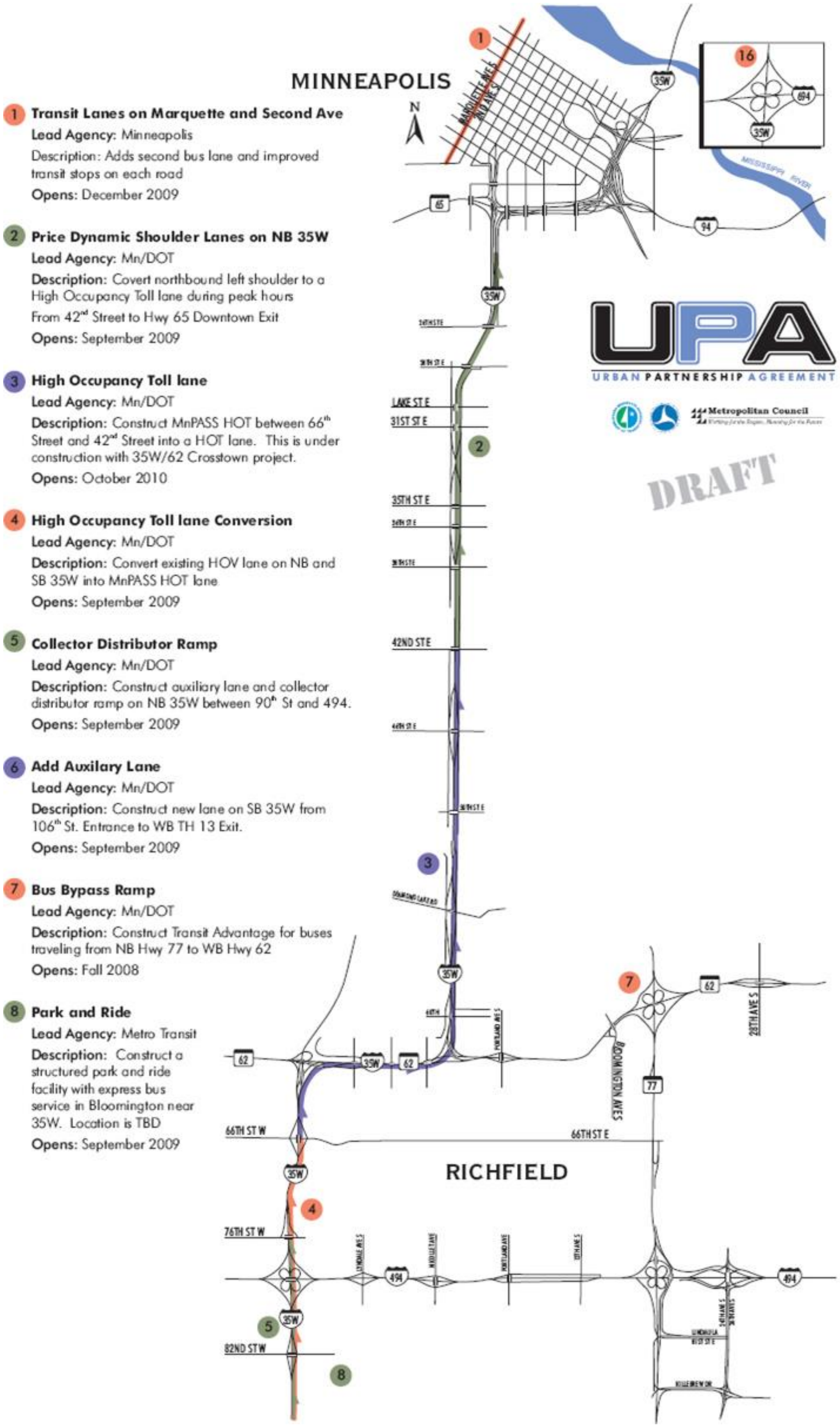

Figure 1. 1 Minnesota's UPA project – Northern Portion.
(source: https://ops.fhwa.dot.gov/congestionpricing/agreements/minneapolismaplg.htm)

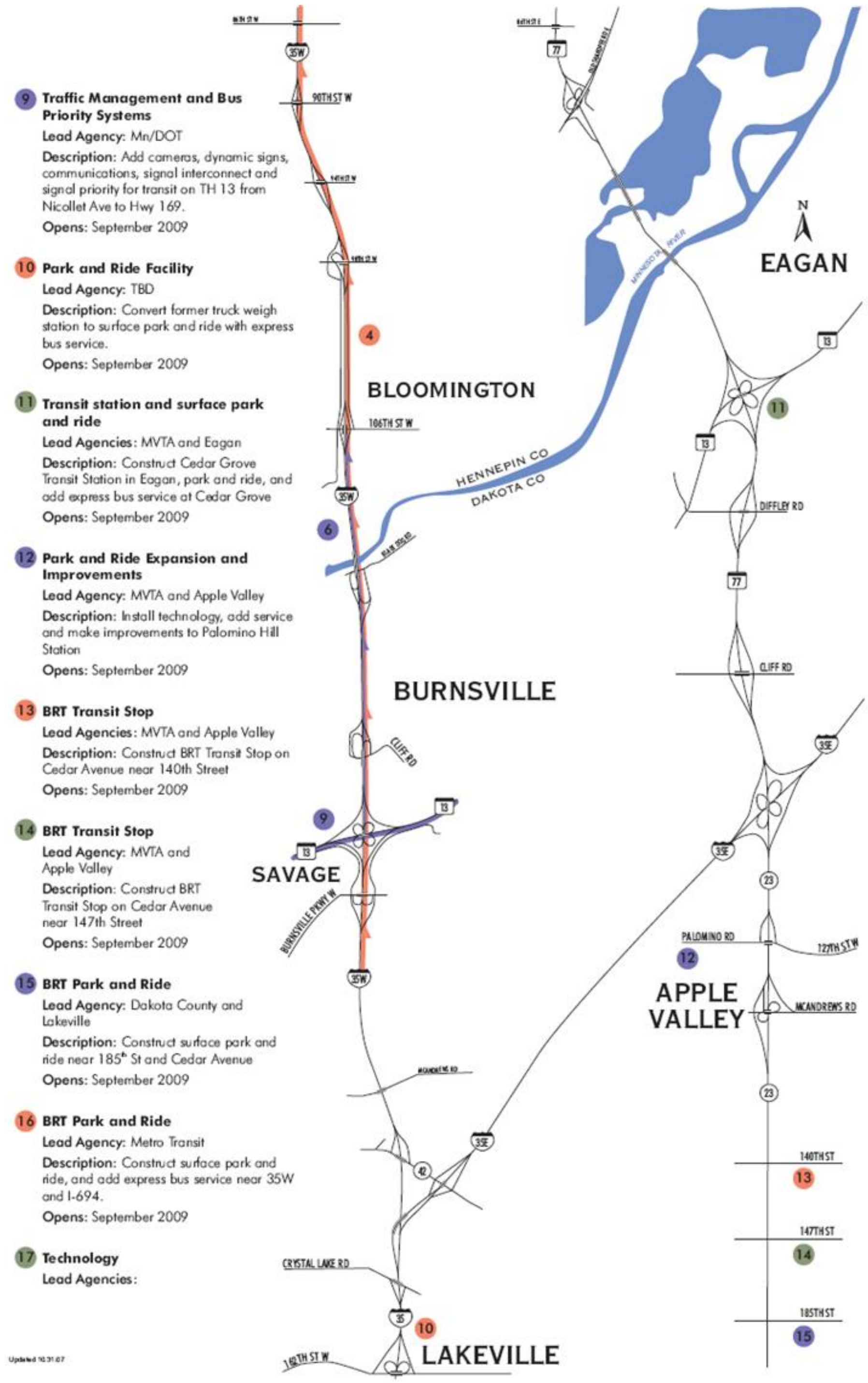

Figure 1. 2 Minnesota's UPA project – Southern Portion.
(source: https://ops.fhwa.dot.gov/congestionpricing/agreements/minneapolismaplg.htm)

The first intervention of Minnesota's UPA project, Transit Advantage Bus Bypass, was implemented in December, 2009, and the last intervention, Real-Time Transit and Traffic

Dynamic Message Signs (DMS), was implemented in May, 2011. The construction of Minnesota's UPA project lasted for approximately 17 months.

Figure 1. 3 is the timeline of Minnesota's UPA project.

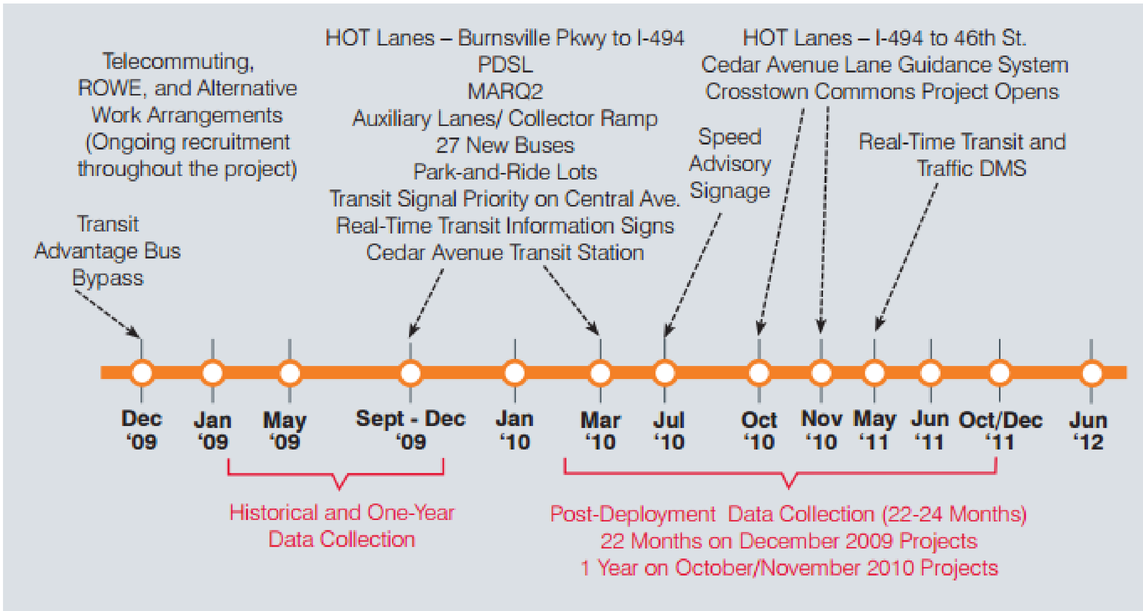

Figure 1. 3 Timeline of Minnesota's UPA project.
(Source: https://ops.fhwa.dot.gov/congestionpricing/docs/fhwajpo11039/)

As is shown, the major improvements of the project, including reconstruction of the Crosstown Commons, provision of high-occupancy toll (HOT) lanes, and priced dynamic shoulder lanes (PDSL) were launched in September, 2009 and completed in November, 2010, and the areas affected by these improvements are shown in **Figure 1. 4**.

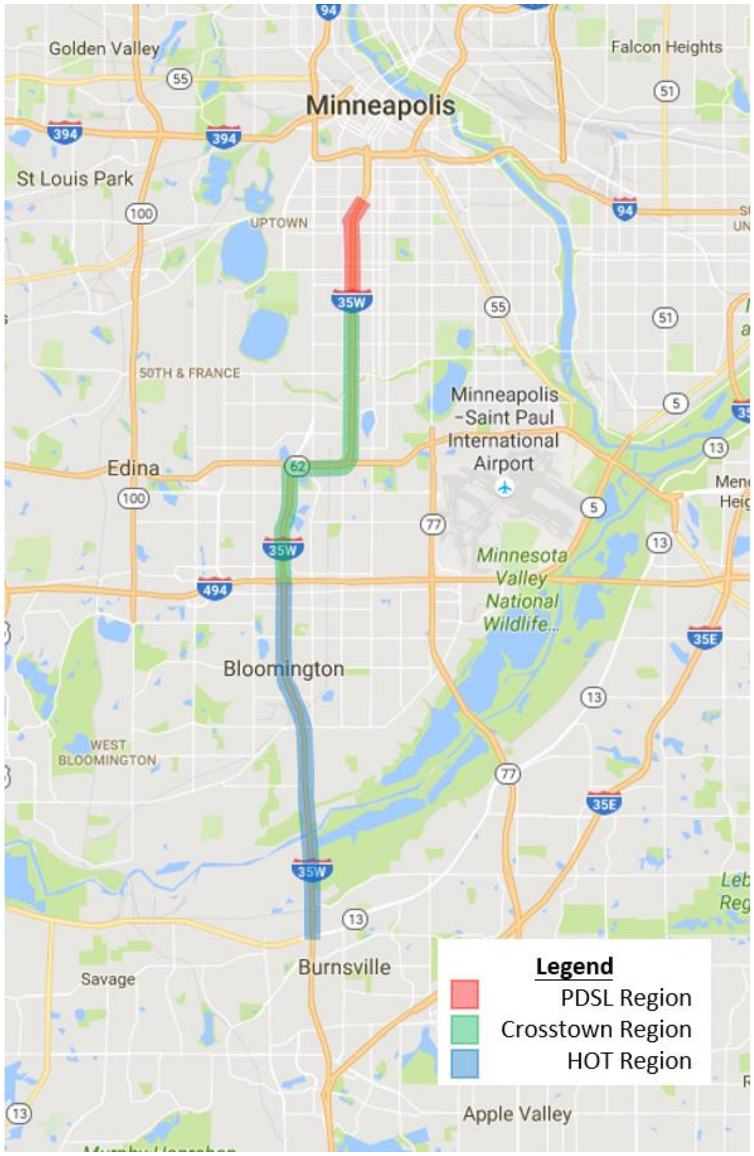

Figure 1. 4 Portion of I-35W receiving UPA improvements. North is up.
(Source: Google Map, modified by Gao)

In 2013, MnDOT released Problem Statement NS-329 noting that there was interest in extending some or all of these interventions implemented in Minnesota's UPA project to other corridors and called for an estimation of their safety effects to help make these decisions. However, since the UPA project consisted of several different interventions, it was desirable to untangle the effects of the interventions. For instance, crash frequencies appeared to have increased where PDSL had been provided, but it was unclear if this is due to the PDSLs themselves or to changes in traffic congestion resulting from the removal of the Crosstown (Trunk Highway 62 & I-35W) bottleneck.

The study aimed to untangle the indirect safety effects due to changes in traffic conditions from the direct effects, if any, due to the UPA improvements.

A preliminary analysis regarding the UPA interventions' safety effects was done. I-35W from its start to I-94 was divided into 17 one-mile sections, as shown in **Figure 1. 5**.

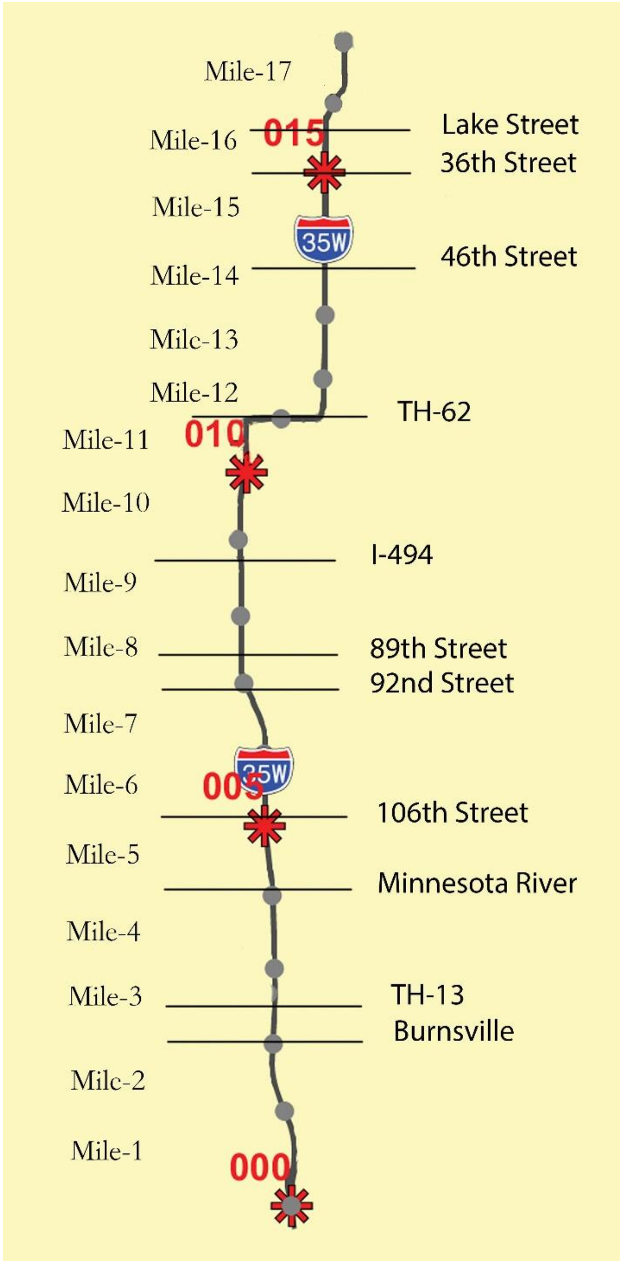

Figure 1. 5 17 1-mile sections of I-35W.

Crash records before (2006-2008) and after (2011-2013) the UPA project were extracted via the Minnesota Crash Mapping Analysis Tool (MNCMAT). Each crash record was allocated to the corresponding one-mile section based on recorded crash location (mileposts), and crash frequencies by crash type were tabulated for each one-mile section. **Table 1. 1** is an example crash frequency summary from one one-mile section, Section Mile-17.

Table 1. 1 Example Crash Summary: Section Mile-17

Crash Code	Crash Type	Crash Frequencies by Crash Type	
		Before	After
0	Unspecified	2	0
1	Rear end	160	220
2	Sideswipe-Same direction	44	53
3	Left turn	0	1
4	Ran off road-Left side	38	79
5	Right angle	7	3
6	Right turn	0	0
7	Ran off road-Right side	15	39
8	Head on	1	1
9	Sideswipe-Opposing	1	0
90	Other	24	25
98	Not applicable	5	4
99	Unknown	0	0

Figure 1. 6 visualizes the crash summary shown in **Table 1. 1**.

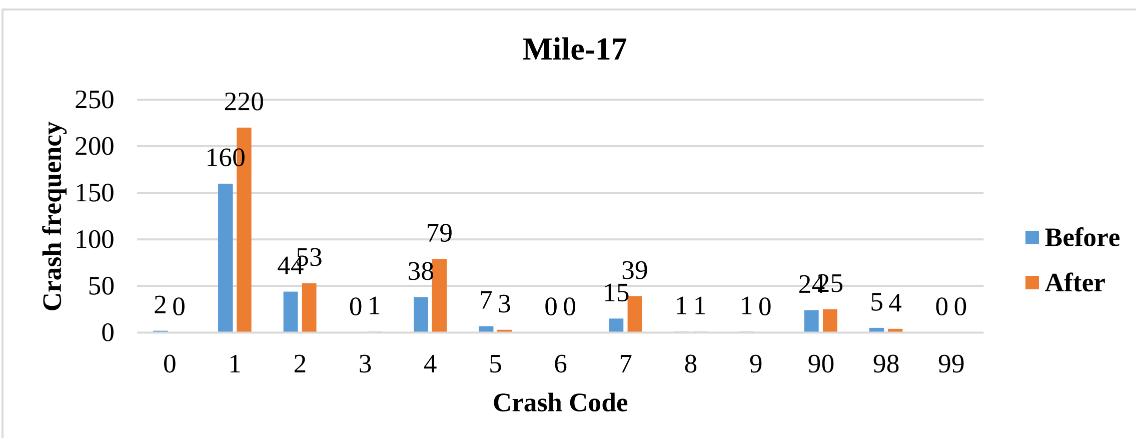

Figure 1. 6 Histograms of crash summary in section Mile-17.

As is shown in **Table 1. 1**, the most frequent crash type in Section Mile-17 was rear-end crashes. Similar analyses were done for the other 16 one-mile sections (crash summaries shown in Appendix) and the rear-end crash turned out to be the most prevalent crash type in those sections as well. Therefore, rear-end crashes became the priority type of crash in this study.

Figure 1. 7 shows Before and After rear-end crash frequencies in each 1-mile section.

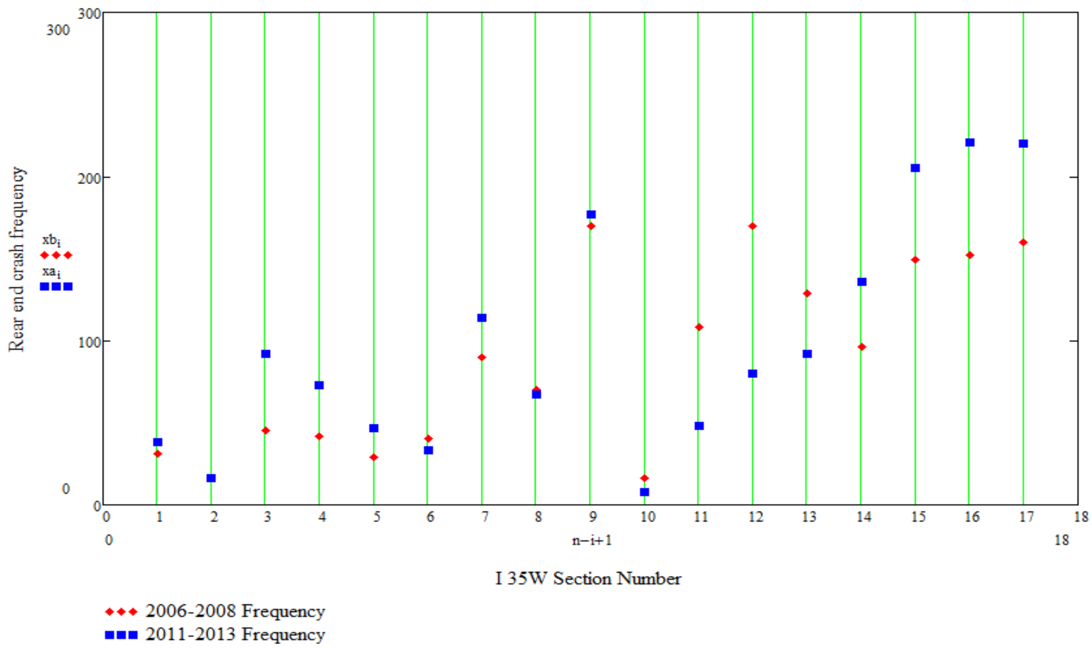

Figure 1. 7 Before and After frequencies of rear-end crashes in each 1-Mile section.

The changing trend of rear-end crash frequencies from the Before to After period varied in each section. Sections Mile-3 to Mile-5, Mile-7, and Mile-14 to Mile-17 experienced an increase in rear-end crash frequency after the UPA project. In Sections Mile-10 to Mile-13, rear-end crash frequency in the After period actually decreased compared to that in the Before period. For the remaining sections, the Before and After rear-end crash frequencies were approximately the same.

The study's interest lay in the sections where there was an obvious increase in the After period's rear-end crash frequency compared to that in the Before period, which is approximately I-35W from TH-13 to I-494 and the I-35W PDSL region.

Therefore, in what follows, the focus will be on two changes that appear to be associated with increases in crash frequencies: the addition of a lane to southbound I-35W as it approaches and crosses the Minnesota River and the conversion of the shoulder on northbound I-35W to a PDSL.

The major ideas from freeway crash risk modeling are reviewed in Chapter Two, including statistical models for crash prediction and contributing factor analyzes. Chapter Three presents the work on data acquisition of crash records, traffic conditions, weather conditions, and PDSL activations, followed by one of the most challenging parts in this study, data preparation, in Chapter Four. Chapter Four described detailed data processing procedure, including verification, screening, and aggregation. Statistical analyses, included model establishing, testing, and, interpretation, are described in Chapter Five. Finally, the research findings and possible extensions are summarized in Chapter Six.

Chapter 2: Literature Review

2.1 Association between Crash Risk and Traffic Conditions

The discussion of the relationship between traffic conditions and traffic accidents can be dated back to 1960s. In the early 1960s, the focus of studies in this area was on the relationship between average daily traffic (ADT) and accident rates (usually defined as the number of accidents per year per million motor vehicle-kilometers). For instance, Lundy (1960) conducted analysis of traffic accidents on 659 miles of four-lane, six-lane, and eight-lane freeways in California from year 1960 to 1962. He proposed a simple linear model depicting the relationship between accident rate as the dependent variable and ADT as the independent variable and found that the accident rate increased with increasing ADT. Staring from mid 1960s, concerned with the effect of ADT on accident measurements, some researchers began to use traffic conditions on hourly basis instead of annual average basis. Gwynn (1967) discovered a U-shaped dependency between accident rate and hourly traffic flow based on studied four-lane divided sections in New Jersey. Ceder and Liven (1982) tried to clarify the interaction between traffic flow and accidents found by Gwynn (1967). The authors studied eight four-lane road sections during an 8-year period and suggested that separation of components of both types of hourly flow and types of accidents could improve the accuracy of accident prediction. Hall and Pendleton (1989) explored the relationship between accident rates and volume to capacity (V/C) ratios on rural highways and reported that V/C ratios may be useful on accident prediction, especially at higher volume levels. Most studies at that time were based on aggregation of crashes and traffic conditions. Mensah and Hauer (1998) looked into two averaging-related problems in those safety performance functions (SPFs), namely argument averaging issue and function averaging issue. They argued that using averaging traffic flows such as annual average daily traffic (AADT) in the estimation of SPFs could bring about large bias depending on the SPFs. Also, discrepancies due to function averaging over time could lower the accuracy of the estimation of expected accident frequencies.

Although it has been recognized that crash risk probably varies as traffic condition varies (Liu 1997), since historical statistical models mainly used the aggregated measurement of traffic condition such as annual average daily traffic and hourly volume, ignoring its time-varying characteristic, the specific association between crash risk and real-time traffic condition could not be pursued. Motivated by such limitation, researchers started to develop real-time crash prediction models based on real-time traffic condition prior to a crash occurrence.

Oh et al. (2001) established a nonparametric Bayesian model, using real-time traffic data, 5-min average and standard deviation of flow, occupancy, and speed, as the measure of accident exposure to estimate real-time accident likelihood. Lee et al. (2002) developed a log-linear model relating crash rates to the selected crash precursors, 5-min variation of speed and traffic density aggregated from loop detector data from 38 sections on a 10-km stretch of the Gardiner expressways in Toronto for a period of 13 months. The results showed the applicability of real-time traffic conditions as precursors on crash potential when controlling for geometry, weather, and time of day. Golob and Recker (2003) conducted both linear and nonlinear multivariate statistical analyses to determine contributing factors, including traffic flow, weather and lighting conditions, for different types of accidents that occurred on six freeway routes in Southern California during 1998. Traffic flow was measured by 30-s data from loop detectors in the vicinity of accident location prior to accident time. The authors concluded that when controlling for weather and lighting conditions, accident severity appeared to be more influenced by traffic volume compared to speed. To identify the freeway traffic flow conditions which were precursors of certain types of crashes, Golob et al. (2004) used 30-s traffic flow data extracted from single loop detectors located on six major freeways in Orange county, California, and found the key traffic flow elements affecting safety were mean volume and median speed, and temporal variations in volume and speed.

Abdel-Aty et al. (2004) introduced the matched case-control logistic regression which has become something of a methodological standard. Police reports were reviewed to identify crashes occurring on a section on Interstate-4 and the real-time traffic data collected from loop detectors were used to characterize the traffic conditions in the vicinity of the crash location. Each crash became a case, while the corresponding non-crashes became controls. Among all selected crash precursors, 5-min upstream average occupancy before a crash and the 5-min downstream coefficient for variation in speed during the crash appeared to have most significant influence on crash occurrence. The authors reported a 69% crash detection rate and a 47% false-positive rate when using those two predictors, 5-min upstream average occupancy before a crash and the 5-min downstream coefficient for variation in speed, in their proposed crash prediction model.

Since 2004, a number of variants on this case-control approach have been reported. Abdel-Aty et al. (2005) further extended former research where matched case-control logistic regression technique was used. The authors developed split models (two separate models) for freeway multivehicle crash prediction under high-speed and low-speed operating conditions by splitting the whole crash data into two data sets based on the 5-min average speeds aggregated from 30-s loop detector data prior to a crash, allowing for different crash mechanisms. Abdel-Aty and Pemmanaboina (2006) used matched case-control logit model to establish a crash likelihood prediction model based on real-time traffic condition data archived in intelligent transportation system and rainfall data from rain stations instrumented on I-4 corridor in Florida. 5-min average occupancy and standard deviation of volume, and 5-min coefficient of variation in speed 5-10 min prior to the crash were found to affect crash occurrence most significantly. Hossain and Muromachi (2012) used high-resolution traffic data from loop detectors on Shibuya 3 and 4 expressways in Tokyo Metropolitan to develop a real-time crash prediction model with 66% crash detection rate and a false-positive rate less than 20%. This study found that the traffic conditions in the upstream and the downstream and the difference in the traffic flow parameters in these locations had great impact on the crash prediction precision. Roshandel et al (2015) reviewed recent studies in this area and provided a meta-analysis. The sensitivity analysis results showed that location of loop detectors can be a confounder when studying the relationship between freeway crash frequencies and traffic characteristics.

The previous studies have an emphasis on “prediction function” of the empirical models for freeway crashes but less on the explanation of the mechanism of freeway crashes. However, there are still several studies which have shed some light on the etiology of freeway crashes.

Hourdos (2005, Hourdos et al. 2006) developed a unique methodology derived from standard case-control study to identify the most relevant real-time traffic metrics for freeway crash risk in the high crash area. Instead of using archived crash records and loop detector data, the he used the top-view video cameras located on a freeway section with high crash rate in Minnesota to visually identify freeway crashes and crash-relevant events and measure traffic variables by machine-vision methods. Besides environmental factors like lighting, several general undesirable traffic flow conditions, including large speed differences among lanes and compression waves leading to abrupt changes in the traffic flow, were identified. Stopping shock waves were found to be necessary precursors for freeway rear-end crashes. With the detection algorithm based crash risk model established in this research, 58% of the crashes were successfully detected with a 6.8% false-positive rate. Davis and Swenson (2006) developed a probabilistic causal model to reconstruct and analyze individual freeway rear-end crashes. Using the trajectory information for individual vehicle extracted from video recordings of crashes, the authors illustrated how Brill's model (Brill, 1972) could be used to explain the occurrence of rear-end crashes. Zheng et. al (2010) explored the relationship between traffic oscillations and traffic safety during congested periods using high-resolution traffic and crash data from a freeway section in the northbound of Interstate-5 in Portland. The authors conducted a matched case-control with conditional logistic regression model. Cases and controls were matched by similar times, presence of congestion, geometry, and weather. The authors reported an 8% increase in rear-end crash risk with a one unit increase in the standard deviation of speed during congested periods. Xu et al. (2012) collected crash and traffic flow data from 47 loop detector stations from a northbound section of the Interstate-880 freeway in California to evaluate the association between freeway traffic states and crash risk. It was a case-control study with conditional logistic regression model and the results showed that the impact of traffic flow parameters on crash risk were different in different traffic states. Xu et al. (2014) reported that the freeway crash risk reached the highest at the level of service (LOS) E, declined in LOS D and F, and was lowest at LOS A. The results indicated that, the relationship between freeway traffic flow states and crash risk should be nonlinear. To be more specific, freeway crash risk should have a roughly concave relationship to traffic density and the maximum crash risk locates at the densities corresponding to the capacity flow. Chatterjee (2016) combined Brill's (Brill, 1972) car-following model with Newell's (Newell, 1993) kinematic wave model, and determined conditions where a stopping wave could result in a rear-end crash. The established theoretical condition was verified with 41 video-recorded shock waves on I-94 in Minneapolis and successfully distinguished between successful brake-to-stop events and rear-end crashes. This also study provided a structural-empirical method for freeway rear-end crash hazard evaluation.

2.2 Association between Crash Risk and Adding Additional Lane(s)

As noted in Chapter One, our study focused on two UPA interventions that may be associated with the increase in rear-end crash frequency, the implementation of HOT lane (number of lanes increased from three to four) and the PDSL (number of lanes increased from four to five when the PDSL was operating). Although our literature search turned up no reports explicitly addressing those two interventions, some researches could provide some clues on the safety effect of adding an additional lane.

McCasland (1978) studied the safety performance of two sections of U.S. 59 Southwest Freeway in Houston where an additional lane was added by narrowing lane widths. The number of lanes changed from three to four in one section, while in the other section, the number of lanes was four before the project and became five after the project. The Poisson comparison of means test results suggested a significant decrease in accident rates (defined as accidents per 100 million vehicle kilometers) at 0.05 significance level. An NCHRP study (Curren, 1995) evaluated the safety performance of the application of shoulder use and narrow lanes on five studied corridors. Although higher accident rates were seen in three of the corridors while two showed a decrease, it was difficult to determine the association between the change in accident rates and conversion. Both studies chose accident rates as the measurement of traffic safety rather than accident frequency.

Bauer et al. (2004) conducted an observational before-and-after evaluation of the safety effects of lane addition projects involving narrow lane and shoulder-use lane conversions on urban freeways in California and reported that projects converting four lanes to five lanes resulted in increases of 10% to 11% in accident frequency. Cao et al. (2012) studied safety benefits associated with conversions of High Occupancy Vehicle (HOV) lanes to HOT lanes on I-394 in Minnesota and found a 5.3% decrease in crash frequency following the conversion. In both of these studies, the dependent variable was annual crash frequency, and neither sought to estimate within-day changes or relate changes in crash frequency to traffic conditions.

2.3 Association between Crash Risk and Weather Conditions

Weather is one of the most studied environmental factors that affects crash risk as it may reduce road friction, impair visibility, and/or make vehicle handling more difficult (Andrey, 2010). According to statistics provided by FHWA's Road Weather Management Program, every year, approximately 22% of crashes occurred in adverse weather or on slick pavement (FHWA Road Weather Management).

In previous research, significant effects of adverse weather conditions on traffic accidents were found (Brijs et al., 2008). Audrey and Yagar (1993) studied the association between rainy weather conditions and crash risk using data from cities of Calgary and Edmonton, Canada, for time period 1979-1983, using a matched sample approach. The overall crash risk was found to be 70% higher in rainy conditions than normal. Khattak et al. (1998) investigated the effects of adverse weather on crash type and injury severity using data for North Carolina for time period 1990 to 1995. The analysis showed that overall crash risk on wet surfaces increased and adverse weather differentially influenced different types of crashes. Knapp et al. (2000) investigated the impact of winter storm on traffic safety at a number of Iowa interstate locations for time period 1995 to 1998. The Poisson modeling results suggested a significant positive association between crash frequency and storm event duration and snowfall intensity. Eisenberg and Warner (2005) estimated the effects of snowfall on US traffic crash rates for time period 1975 and 2000, and the effects were found to be substantial. Although snow days had fewer fatal crashes than dry days, there were more nonfatal-injury crashes and property-damage-only crashes in snow days than dry days. Brijs et al. (2008) used a discrete time-series model to analysis the impact of weather conditions on daily crash frequency in Netherlands. The study showed that apart from exposure, weather condition, namely precipitation, did have significant effect on daily crash number.

2.4 Statistical Models for Crash Likelihood Prediction

At the beginning, researchers assumed a linear relationship between crash frequency or crash rate and factors such as traffic conditions, geometric conditions, and environmental factors. Lundy (1960) conducted analysis of traffic accidents on 659 miles of four-lane, six-lane, and eight-lane freeways in California for time period 1960 to 1962. He proposed a simple linear model depicting the relationship between accident rate (dependent variable) and ADT (independent variable) and found that the positive association between accident rate and ADT. Liu and Popoff (1997) established simple linear regression models to study the relationship between travel speed and collisions. Casualties on provincial highways was the dependent variable while the average speed, the 85th speed, or the speed differential were the independent variables. Average travel speed was found to be closely correlated with both the number of traffic casualties and casualty rates. In addition, speed differentials were found to be associated with casualty rates on provincial highways, and higher speed differentials resulted in higher casualty rates. Ivey et al. (1981) used multiple linear regression models and suggested that traffic, geometric, pavement surface, and rainfall conditions could be used for wet weather accident rate prediction. In addition, they did sensitivity analyses and found that accident rate on highways with speed limits less than 55 mph were sensitive to ADT, highway access, skid number and the time of exposure to rainfall, while that on highways with speed limit equal to 55 mph were sensitive to ADT, highway access, skid number and the standard deviation of traffic speeds about the mean traffic speed.

However, the assumption of a linear relationship requires that crash frequency or crash rate is normally distributed, which is unrealistic, and the non-negative nature of crash frequency or crash rate could not be accounted for (Lee, 2008). To overcome the disadvantages of linear model, count data models and their variants were introduced to crash prediction. Jovanis and Chang (1986) proposed a Poisson regression model relating daily vehicle miles traveled (VMT), hours of snow and rain, and time of a week (weekdays or weekend) to the daily crash frequency on the mainline of the 157-mile Indiana Toll Road and found that VMT, hours of snow and rain were positively associated with crash occurrence. Miaou and Lum (1993) used both linear and Poisson regression models to model the relationship between highway vehicle accidents and geometric design. They suggested that, compared to linear model, Poisson regression models had better statistical properties in developing the relationship. To account for the over-dispersion in crash frequency data where Poisson regression models may not fit, Miaou (1994) applied negative binomial (NB) and zero-inflated Poisson (ZIP) models to the analysis of the relationship between truck accidents and geometric design, and evaluated the performance of both models. Several more variants of the count data models came out recently, such as zero-inflated negative binomial (ZINB) (Shankar et al. 1997), negative binomial Lindley (Geedipally et al., 2012), Poisson-inverse-Gaussion (PIG) (Zha et al., 2016), etc., have been used during past decades. By now, Poisson regression is still one of the most well-accepted methods of modeling roadway crashes (AASHTO, 2010). Loglinear Poisson models of crash frequency, when constrained to take only the values 0 or 1, reduce to logistic regression models (Davis et al., 2016) which makes logistic regression a useful tool for investigating crash risk over short time and space intervals.

Chapter 3: Data Acquisition

The first task was to compile a master data file to conduct statistical analysis. Such data file includes explanatory variables such as relevant traffic volume and lane occupancy, PDSL activation, weather conditions, and crash experience on I-35W during both before (years 2006-2008) and after (years 2011-2013) UPA project. This chapter is dedicated to this first task; that is, the data collection of rear-end crashes, traffic condition, PDSL activation, and weather condition.

3.1 Studied Sections

The scope of this study can be divided into three sections, the HOT region, Crosstown region, and the PDSL region, as shown in **Figure 3. 1**.

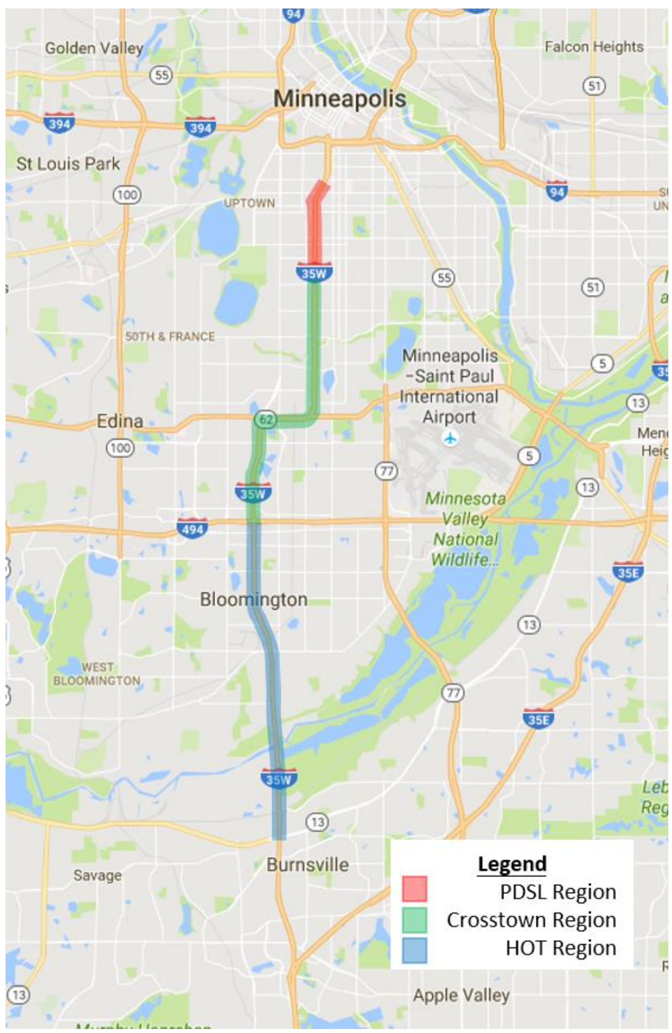

Figure 3. 1 MN UPA project intervention map.
(Source: Google Map, modified by Gao)

All three regions were divided in to sections so that traffic demand and lane geometry were constant within a section. That is, the section boundaries were determined by junctions with on and off ramps or changes in geometric features such as number of lanes. It is worth noting that not all sections had loop detectors installed for all 6 study years (2006-2008 and 2011-2013). The statistical analyses could be done only for those sections with loop detector data available.

The request for the I-35W milepost listing for both before (2006-2008) and after (2011-2013) periods was made to MnDOT, and the lane counts was verified with GoogleEarth.

Our studied regions were the HOT region, I-35W from TH-13 to I-494, and the PDSL region, northbound I-35W from 37th Street to approximately 200 feet before 26th Street.

Table 3. 1, **Table 3. 2**, and **Table 3. 3** are the lists of HOT and PDSL region division results, respectively. The sections with symbol "*" are those sections with loop detectors where statistical analyses could be done.

Table 3. 1 HOT Region Division (Northbound)

Section No.	Start Point		End Point		Number of Lanes	
	Milepost	Location	Milepost	Location	Before	After
N1	002+00.244	NB ENT LOOP FROM BURNSVILLE PKWY MSAS-102	002+00.486	NB EXIT RAMP TO MNTH-13 EB	4	4
N2	002+00.486	NB EXIT RAMP TO MNTH-13 EB	002+00.637	NB ENT LOOP FROM MNTH-13 EB	3	3
N3	002+00.637	NB ENT LOOP FROM MNTH-13 EB	002+00.730	NB EXIT LOOP TO MNTH-13 WB	4	4
N4	002+00.730	NB EXIT LOOP TO MNTH-13 WB	002+00.876	NB ENT RAMP FROM MNTH-13 WB	3	3
N5	002+00.876	NB ENT RAMP FROM MNTH-13 WB	003+00.251	NB EXIT LOOP TO CLIFF RD CSAH-32	4	4
N6*	003+00.251	NB EXIT LOOP TO CLIFF RD CSAH-32	003+00.384	NB ENT RAMP FROM CLIFF RD CSAH-32	3	3
N7	003+00.384	NB ENT RAMP FROM CLIFF RD CSAH-32	003+00.999	NB EXIT RAMP TO BLACKDOG RD M-1	3	3
N8	003+00.999	NB EXIT RAMP TO BLACKDOG RD M-1	004+00.115	NB ENT LOOP FROM BLACKDOG RD M-1	3	3
N9-1*	004+00.115	NB ENT LOOP FROM BLACKDOG RD M-1	004+00.907	LANE CHANGE POINT	3	3
N9-2*	004+00.907	LANE CHANGE POIT	005+00.112	NB EXIT RAMP TO W 106TH ST MSAS-407	4	4

N10	005+00.112	NB EXIT RAMP TO W 106TH ST MSAS-407	005+00.385	NB ENT RAMP FROM W 106TH ST MSAS-407	4	4
N11*	005+00.385	NB ENT RAMP FROM W 106TH ST MSAS-407	006+00.059	NB EXIT RAMP TO W 98TH ST CSAH-1	3	3
N12	006+00.059	NB EXIT RAMP TO W 98TH ST CSAH-1	006+00.360	NB ENT RAMP FROM W 98TH ST CSAH-1	3	3
N13	006+00.360	NB ENT RAMP FROM W 98TH ST CSAH-1	006+00.579	NB EXIT RAMP TO W 94TH ST MSAS-136	4	4
N14	006+00.579	NB EXIT RAMP TO W 94TH ST MSAS-136	006+00.838	NB ENT RAMP FROM W 94TH ST MSAS-136	3	3
N15	006+00.838	NB ENT RAMP FROM W 94TH ST MSAS-136	007+00.164	NB EXIT RAMP TO W 90TH ST MSAS-130	4	4
N16	007+00.164	NB EXIT RAMP TO W 90TH ST MSAS-130	007+00.426	NB ENT RAMP FROM W 90 TH ST MSAS-130	3	3
N17*	007+00.426	NB ENT RAMP FROM W 90 TH ST MSAS-130	008+00.163	NB EXIT RAMP TO W 82ND ST MSAS-354	3	4
N18*	008+00.163	NB EXIT RAMP TO W 82ND ST MSAS-354	008+00.415	NB ENT RAMP FROM W 82ND ST MSAS-354	3	4
N19	008+00.415	NB ENT RAMP FROM W 82ND ST MSAS-354	008+00.602	NB EXIT RAMP TO ISTH-494 EB	4	5

Table 3. 2 HOT Region Division (Southbound)

Section No.	Start Point		End Point		Number of Lanes	
	Milepost	Location	Milepost	Location	Before	After
S1	002+00.253	SB EXIT LOOP TO BURNSVILLE PKWY MSAS-102	002+00.506	SB ENT RAMP FROM MNTH-13 EB	4	4
S2	002+00.506	SB ENT RAMP FROM MNTH-13 EB	002+00.606	SB EXIT LOOP TO MNTH-13 EB	3	3
S3	002+00.606	SB EXIT LOOP TO MNTH-13 EB	002+00.702	SB ENT LOOP FROM MNTH-13 WB	4	4
S4	002+00.702	SB ENT LOOP FROM MNTH-13 WB	002+00.845	SB EXIT RAMP TO MNTH-13 WB; END SB M/O	3	3
S5	002+00.845	SB EXIT RAMP TO MNTH-13 WB; END SB M/O	003+00.270	SB ENT LOOP FROM CLIFF RD CSAH-5	4	5

S6*	003+00.270	SB ENT LOOP FROM CLIFF RD CSAH-5	003+00.414	SB EXIT RAMP TO CLIFF RD CSAH-5	3	4
S7	003+00.414	SB EXIT RAMP TO CLIFF RD CSAH-5	004+00.109	SB ENT RAMP FROM BLACKDOG RD M-1	3	4
S8	004+00.109	SB ENT RAMP FROM BLACKDOG RD M-1	004+00.198	SB EXIT LOOP TO BLACKDOG RD M-1	3	4
S9*	004+00.198	SB EXIT LOOP TO BLACKDOG RD M-1	005+00.113	SB ENT RAMP FROM W 106TH ST MSAS-407	3	4
S10	005+00.113	SB ENT RAMP FROM W 106TH ST MSAS-407	005+00.340	SB EXIT RAMP TO W 106TH ST MSAS-407	3	3
S11*	005+00.340	SB EXIT RAMP TO W 106TH ST MSAS-407	006+00.046	SB ENT RAMP FROM W 98TH ST CSAH-1	3	3
S12	006+00.046	SB ENT RAMP FROM W 98TH ST CSAH-1	006+00.350	SB EXIT RAMP TO W 98TH ST CSAH-1	3	3
S13	006+00.350	SB EXIT RAMP TO W 98TH ST CSAH-1	006+00.581	SB ENT RAMP FROM W 94TH ST MSAS-136	4	4
S14	006+00.581	SB ENT RAMP FROM W 94TH ST MSAS-136	006+00.840	SB EXIT RAMP TO W 94TH ST MSAS-136	3	3
S15	006+00.840	SB EXIT RAMP TO W 94TH ST MSAS-136	007+00.165	SB ENT RAMP FROM W 90TH ST MSAS-130	4	4
S16	007+00.165	SB ENT RAMP FROM W 90TH ST MSAS-130	007+00.400	SB EXIT RAMP TO W 90TH ST MSAS-130	3	3
S17*	007+00.400	SB EXIT RAMP TO W 90TH ST MSAS-130	008+00.182	SB ENT RAMP FROM W 82ND ST MSAS-354	3	3
S18*	008+00.182	SB ENT RAMP FROM W 82ND ST MSAS-354	008+00.387	SB EXIT RAMP TO W 82ND ST MSAS-354	3	3
S19	008+00.387	SB EXIT RAMP TO W 82ND ST MSAS-354	008+00.599	SB ENT RAMP FROM ISTH-494 EB	4	4

Table 3. 3 PDSL Region Division (Northbound)

Section No.	Start Point		End Point		Number of Lanes	
	Location	Milepost	Location	Milepost	Before	After
N37*	013+00.819	NB ENT RAMP FROM 46TH ST CSAH-46	014+00.651	NB EXIT RAMP TO E 36TH ST MSAS-251	4	5
N38*	014+00.651	NB EXIT RAMP TO E 36TH ST MSAS-251	015+00.171	NB ENT RAMP FROM E 35TH ST MSAS-249	4	5
N39	015+00.171	NB ENT RAMP FROM E 35TH ST MSAS-249	015+00.312	NB EXIT RAMP TO E 31ST ST MSAS-366	5	6
N40*	015+00.312	NB EXIT RAMP TO E 31ST ST MSAS-366	016+00.097	200 Feet South of BR#27870 UNDER 26TH ST	4	5

(N37 is the transition area before physical PDSL)

Figure 3. 2 and **Figure 3. 3** visualized the studied sections for HOT region and PDSL region.

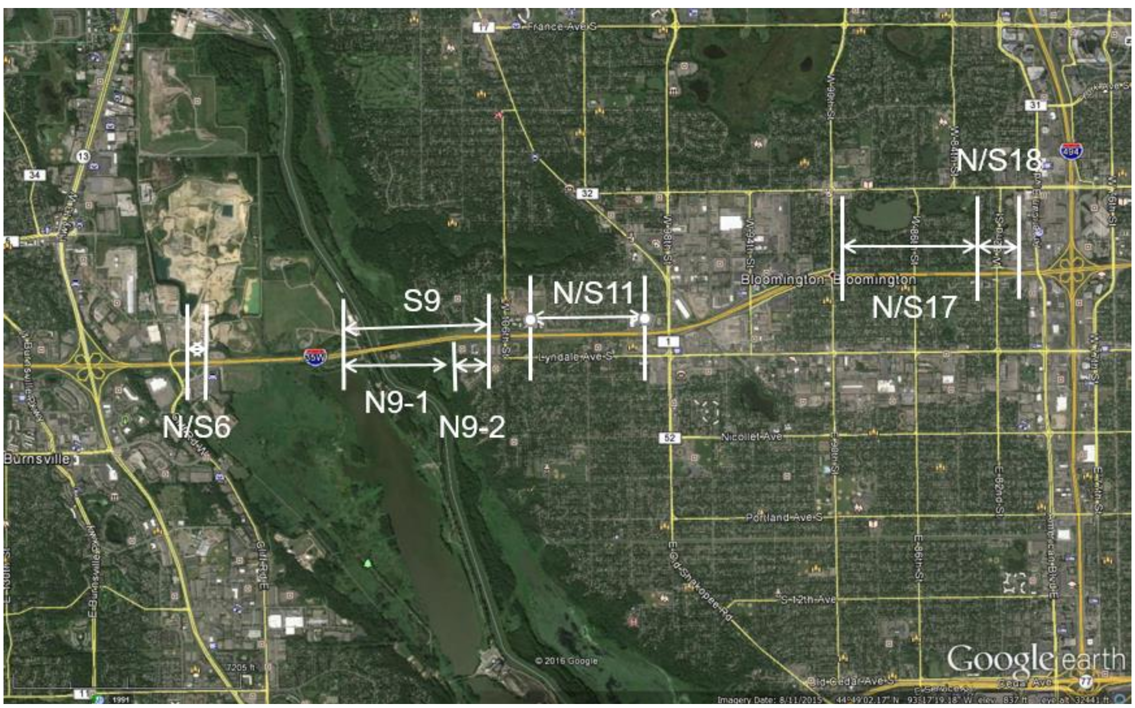

Figure 3. 2 HOT region studied sections.
(Source: Google Earth, modified by Gao)

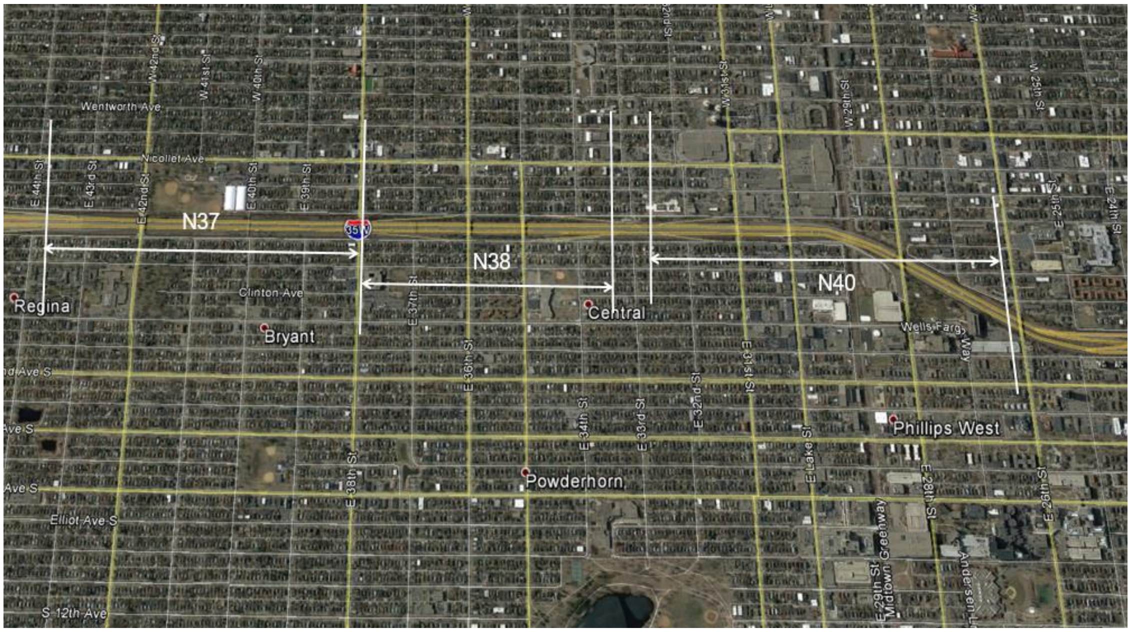

Figure 3. 3 PDSL region studied sections.
(Source: Google Earth, modified by Gao)

3.2 Compiling Crash Data

Crash data were compiled for a before period, year 2006 to 2008, and an after period, year 2011-2013, using MnDOT's Minnesota Crash Mapping Analysis Tool (MNCMAT), and the original crash reports were reviewed to confirm the crash type, crash time, and crash locations.

5545 crash records of all crash types were extracted from the MNCMAT database for the region from the beginning of I-35W to the I-35W & I-94 junction.

For each crash record retrieved from the MNCMAT database, the following details describing the situation when the crash happened were provided:

- Crash location (Sys, Route, and Ref_Point)
- Crash Number consistent with original accident reports (Crash_Num)
- Crash Time (Year, Month, Date, and Time)
- Crash type (Diagram Code)
- Road direction (Rd_Dir)
- Vehicle direction (V1Dir, V2Dir, V3Dir, V4Dir)

- Weather condition (Wthr1, Wthr2)

All crashes were allocated to the corresponding section based on the milepost information provided in MNCMAT database.

Table 3. 4 is a summary table of the number of crashes originally extracted from MNCMAT.

Table 3. 4 Summary Information of Crash Records Extracted Using MNCMAT

Diagram Code	Crash Type	Before				After			
		2006	2007	2008	Total	2011	2012	2013	Total
0	Unspecified	3	6	8	17	4	1	0	5
1	Rear End	567	497	449	1513	508	550	599	1657
2	Sideswipe-Same direction	118	157	150	425	164	141	165	470
3	Left Turn	1	1	1	3	4	1	1	6
4	Ran off road-Left side	66	71	67	204	78	86	107	271
5	Right angle	10	18	24	52	17	18	38	73
6	Right turn	2	0	1	3	0	1	2	3
7	Ran off road-Right side	43	49	68	160	63	86	60	209
8	Head on	13	10	8	31	9	6	12	27
9	Sideswipe-Opposing	2	2	1	5	0	5	1	6
90	Other	67	71	70	208	54	42	60	156
98	Not applicable	14	5	9	28	2	3	5	10
99	Unknown	2	1	0	3	0	0	0	0
Total		908	888	856	2652	903	940	1050	2893

Figure 3. 4 is a bar-chart showing the crash frequencies presented in **Table 3. 4**.

As is shown in both **Table 3. 4** and **Figure 3. 4**, rear-end crashes were the most prevalent crash type on I-35W from its beginning to the I-35W & I-94 junction; thus rear-end crashes were chosen as the priority crash type in this study.

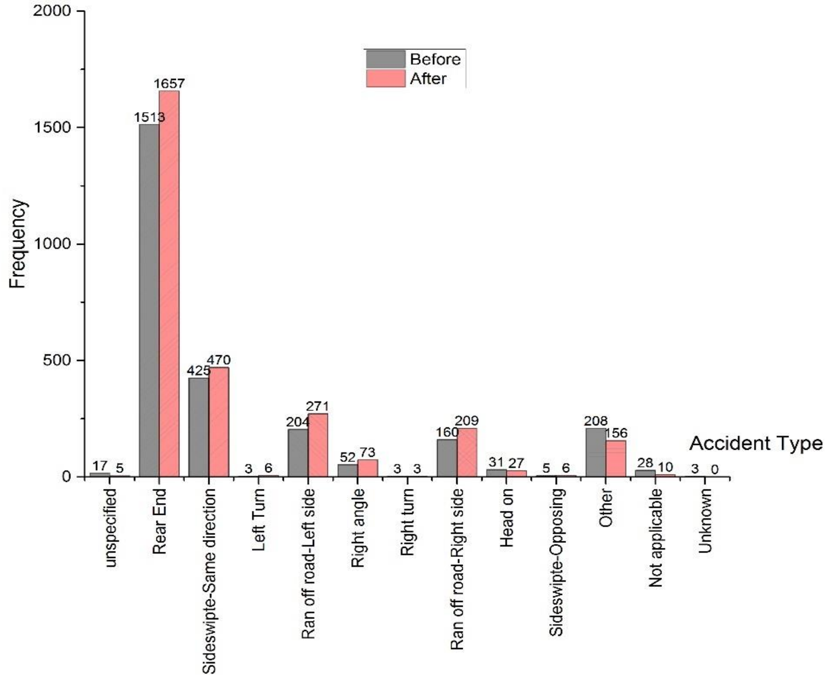

Figure 3. 4 Number of crash records extracted using MNCMAT by crash type.

Table 3. 5 shows the number of rear-end crashes from MNCMAT database for each studied section.

Table 3. 5 Number of Rear-end Crashes from MNCMAT Database for Each Studied Section

Section No.	Number of Rear-end Crashes			
	Before		After	
	With determined* road direction	With unspecified road direction*	With determined road direction	With unspecified road direction
N6	4	4	15	0
N9-1	17	6	17	1
N9-2	1	5	1	0
N11	1	2	1	0
N17	26	6	21	1
N18	65	9	41	6
S6	1	2	0	0
S9	4	11	13	0
S11	0	3	6	0
S17	7	7	23	1
S18	2	9	9	4
N37	8	40	64	6
N38	11	24	69	16
N40	30	56	148	19

(* In crash records extracted from MNCMAT, the "road direction" could be "N" for northbound, "S" for southbound, "E" for eastbound, "W" for westbound, or "Z" for "unknown"; the latter three categories were regarded as "unspecified road direction" and we then tried to determine the actual road direction, northbound or southbound, with the aid of DOT/OTIS accident reports.)

A request was made to Driver & Vehicle Services (DVS) to review the original crash reports of crashes extracted from MNCMAT database. The original crash reports, DOT/OTIS accident reports, provide narratives and sketches in addition to the crash details recorded in MNCMAT database. By reviewing the crash reports, hopefully the following crash information could be verified:

- Crash type
  - If the crash is a rear-end crash
- Crash location
  - If the crash happened on the mainline of I-35W. If not, the crash should be excluded from analysis.
  - The road direction of the crash location.
  - If actual crash location belongs to the section assigned.
- Crash time
  - The actual crash time including year, month, day, and time

It turned out that there were discrepancies between crash reports and crash records in MNCMAT database. In this study, when there was a difference in crash details in MNCMAT database and crash reports, the ones in DOT/OTIS accident reports were adopted.

Although there were specified diagram codes, milepost and road directions in crash records from MNCMAT database, accident reports for all types of crashes and all road directions were still reviewed for all sections, which was about 2000 reports, and crashes with following issues were excluded from the analysis:

- Actual crash type was not “rear-end”.
- Actual crash location was not the mainline of I-35W but on or off ramps.
- Actual crash location or time are unspecified or vague.
- Duplicated crash records where citizen and police reports were merged in DVS motor vehicle crash report database.
- Crash records which were found not to be motor vehicle crashes by MnDOT’s Standards (bike-pedestrian collision, bike-bike collision, suicide, etc.).

Table 3. 6 shows the number of rear-end crashes for studied sections after review of crash reports.

Table 3. 6 Number of Rear-End Crashes for Studied Sections after Crash Reports Review

Section No.	Number of Rear-end Crashes		
	Before	After	Total
N6	7	15	22
N9-1	20	17	37
N9-2	2	2	4
N11	1	1	2
N17	30	21	51
N18	71	44	115
S6	1	0	1
S9	6	13	19
S11	1	5	6
S17	8	24	32
S18	3	10	13
N37	16	62	78
N38	17	73	90
N40	59	148	207

3.3 Compiling Traffic Condition Data

As noted in Chapter Two, it is reasonable to consider traffic conditions as an important factor in rear-end crash risk. The primary data sources for traffic conditions, namely volume and lane occupancy, are:

- All Detector Report (ADR): provided by MnDOT
- DataExtract tool: http://data.dot.state.mn.us/datatools/dataextract.html.

Annual ADRs were used to match the detector stations and lane detector numbers to the corresponding analysis sections. The relevant lane detector numbers were later used to extract traffic condition data using DataExtract tool. Since the locations of loop detectors could have been changed during six studied years, all six annual ADRs were reviewed to guarantee the consistency in loop detector use in traffic condition data collection.

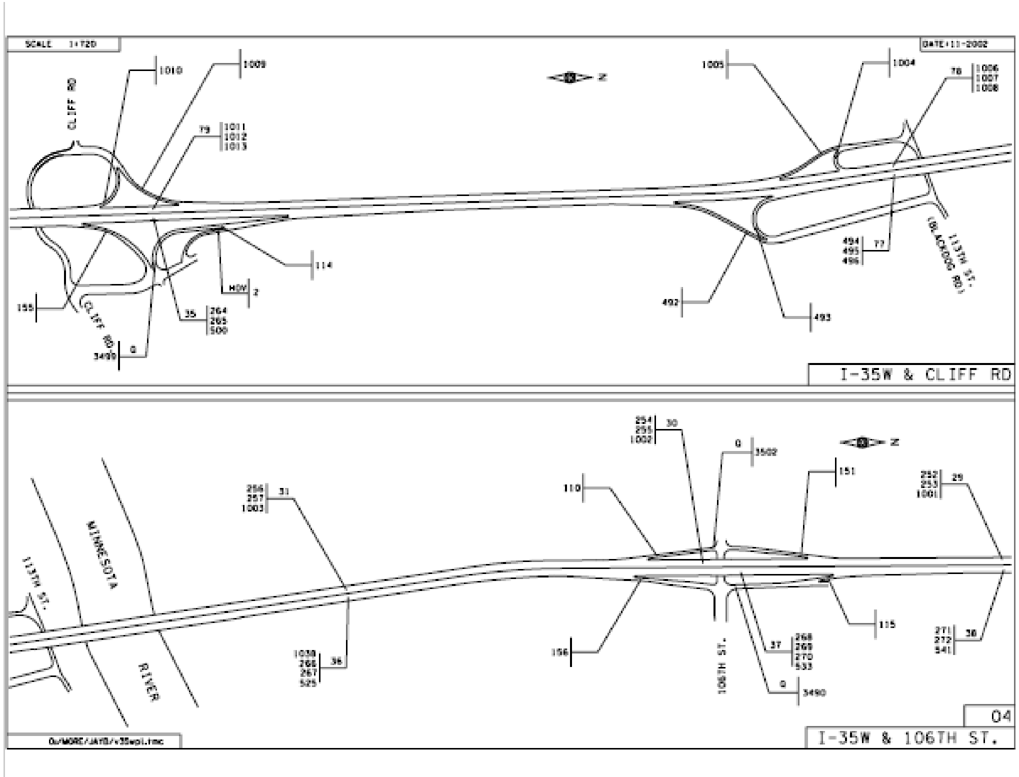

Figure 3. 5 Screenshot of 2006 ADR.

Table 3. 7 is the matching results of detector stations to the studied sections.

Table 3. 7 Matching Results of Detector Stations to Studied Sections

Section No.	Detector Station No.	
	Before	After
N6	35	35
N9-1	77	77&36
N9-2	36	37
N11	38	38&39
N17	42	42&41
N18	43	43
S6	79	79
S9	31	31
S11	29	29&28
S17	25	25&26
S18	24	24
N37	57	57
N38	59	59
N40	60&61	60&61

The original traffic condition traffic data including lane volume and lane occupancy, aggregated over 30-second intervals, the most basic data provided in DataExtract. The total numbers of original lane volume or occupancy records each year are as follows:

- Year 2006, 2007, 2011, and 2013: $\frac{(1\ hour)(3600\frac{second}{hour})}{30\ seconds} \times 24\ hours \times 365\ days =$ 1051200

- Year 2008 and 2012: $\frac{(1\ hour)(3600\frac{second}{hour})}{30\ seconds} \times 24\ hours \times 366\ days = 1054080$

3.4 PDSL Operation Historical Data

The PDSL first opened on Sept 30th, 2009. When initially opened, the standard hours were 6:00-10:00 AM and 2:00-7:00 PM. Hours could be extended for heavy congestion due to weather, events or incidents. Over time, MnDOT noticed high violation rates for the PDSL during the mid-day, thus the PDSL standard hours have been extended to be 6:00 AM-7:00 PM. This change started on April 11th, 2012.

The PDSL activation data came from MnDOT's log of Intelligent Lane Control Signal (ILCS) located at 37th Street, which is right at the beginning of the physical PDSL. In the log, an entry was made when the status of the sign changes and the time was accurate to seconds.

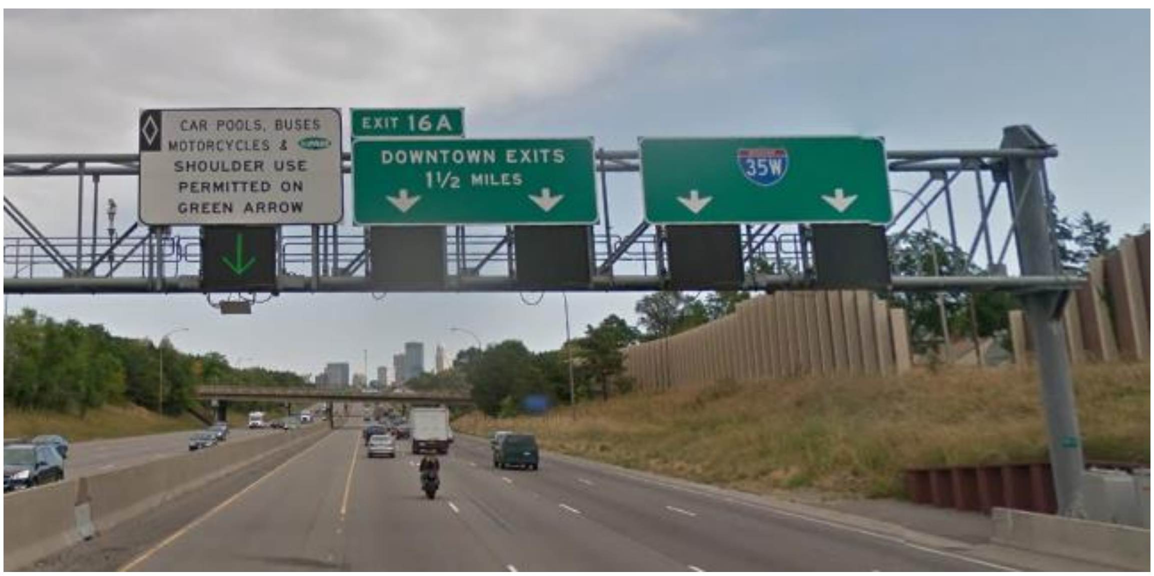

Figure 3. 6 ILCS Located at 37th Street.
(Source: Google Street View)

Below are the standard messages shown on the ILCS:

- LANE_OPEN: Green arrow, lane is open to traffic.

- LANE_CLOSED: Red X ("cross"), lane is closed to traffic.

- CAUTION: Yellow arrow, lane is open but traffic should use caution. It's used when there is an incident in the adjacent lane or shoulder. This could be a crash or stall. It can also be used for debris in the lane that is not closing the lane.

- DARK: ILCS is dark, used for general purpose lanes.

- MERGE_RIGHT: Merge chevron showing lane is closing.

- VSA: Variable speed advisory is displayed.

- UNKNOWN: Lane activation status is unavailable.

As there are 5 ILCS on the structure at 37th Street, the log showed a combination of messages of all five lanes.

	A	B	C	D	E
1	pkey	eventdate	description	deviceid	message
2	75638537	1/3/2011 19:04	LCS DEPLOYED	L35WN48	MERGE_RIGHT DARK DARK DARK DARK
3	75638624	1/3/2011 19:05	LCS DEPLOYED	L35WN48	LANE_CLOSED DARK DARK DARK DARK
4	75677609	1/4/2011 5:40	LCS DEPLOYED	L35WN48	LANE_OPEN DARK DARK DARK DARK
5	75698825	1/4/2011 10:02	LCS DEPLOYED	L35WN48	MERGE_RIGHT DARK DARK DARK DARK
6	75698881	1/4/2011 10:03	LCS DEPLOYED	L35WN48	LANE_CLOSED DARK DARK DARK DARK
7	75702929	1/4/2011 10:57	LCS DEPLOYED	L35WN48	LANE_OPEN DARK DARK DARK DARK
8	75712435	1/4/2011 13:24	LCS DEPLOYED	L35WN48	LANE_OPEN LANE_OPEN USE_CAUTION LANE_CLOSED LANE_CLOSED

Figure 3. 7 Example ILCS log file.

The message should be read from left to right, and as PDSL is the left most lane, the first one listed is the ILCS over the PDSL. In combination, typical messages might appear as follows:

- "LANE_OPEN DARK DARK DARK DARK": PDSL is open. Left most ILCS is displaying the green arrow. Other 4 ILCS are dark.
- "LANE_CLOSED DARK DARK DARK DARK": PDSL is closed. Left most ILCS is displaying the red X. Other 4 ILCS are dark.

After separating PDSL's ILCS from other four, a record of PDSL activation was compiled.

3.5 Compiling Weather Data

In Minnesota, weather conditions can affect the occurrence of crashes, thus it is necessary to control for rainy and snowy weather in analysis. The weather data sources were the following:

- Road Weather Information System (RWIS) managed by MnDOT: http://rwis.dot.state.mn.us/.
- MNCMAT crash records
- Hard copy crash reports

In this study, the weather condition in crash hours were from the weather information shown in MNCMAT crash records and verified with crash reports.

For non-crash hours, RWIS data were used. We first searched for the RWIS sites near our studied sections, and the site "I-35 at Minnesota river" was found to be the closest one. However, as RWIS database could only provide the weather data from site "I-35 at Minnesota river" from year 2011-2013, other sites near "I-35 at Minnesota river" site were needed for the weather data during year 2006-2008. Since the "Minneapolis-St. Paul International Airport" site only had the weather data for year 2010, thus, finally, site "I-35 E Cayuga St. Bridge" was chosen as the source of weather condition for year 2006-2008.

Figure 3. 8 shows the locations of weather information collection sites near studied sections.

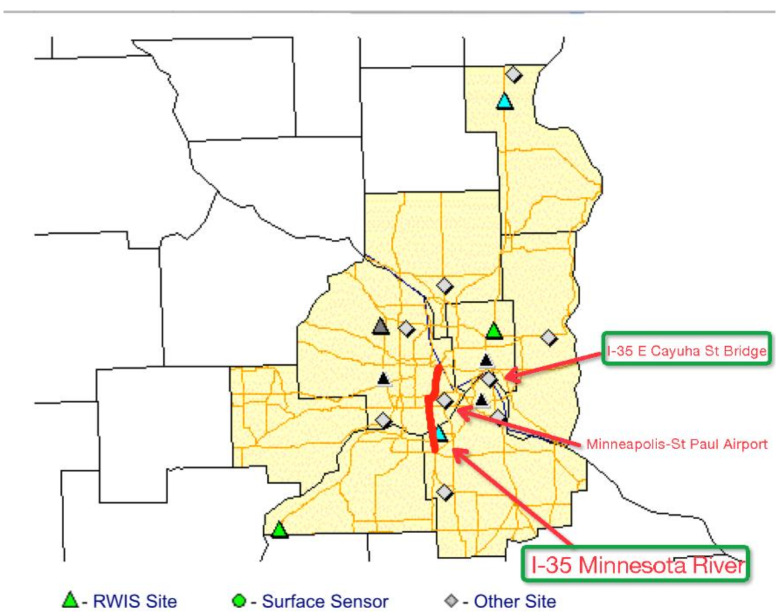

Figure 3. 8 Weather information site locations.

By manually inputting the desired date, historical weather information can be retrieved form the RWIS database by its searching tool. The rainy and snow weather conditions were mainly determined by the precipitation information. When the precipitation information was not available, the surface status served as the supplemental information.

Chapter 4: Data Preparation

In later statistical analyses, for each hour during 2006- 2008 and 2011-2013, the presence or absence of a rear-end crash became the dependent variable while independent variables consisted of traffic volume and lane occupancy, the presence or absence of snowy or rainy conditions, the before or after period indicator, and the presence or absence of PDSL operation. Such analyses required further data processing of the original data, and this chapter describes this process.

Figure 4. 1 and **Figure 4. 2** show the screenshots of example data files for HOT and PDSL regions, respectively.

	A	B	C	D	E	F	G	H	I	J	K	L	M
1	Year	Month	Day	Time	CrashCount01	Rainy	Snowy	BeforeAfter	logvph	occ_dev	occ_dev_sqr	occ_sd	occ_sd_dev
2	2006	1	1	0:00	0	0	0	0	2.997873	-2.9438	8.665936978	1.896227	-1.28196839
3	2006	1	1	1:00	0	0	0	0	3.02006	-2.87893	8.28823179	1.821307	-1.35688876
4	2006	1	1	2:00	0	0	0	0	2.843988	-3.63089	13.18332724	1.353885	-1.82431126
5	2006	1	1	3:00	0	0	0	0	2.696438	-4.08108	16.65525431	1.172871	-2.00532454
6	2006	1	1	4:00	0	0	0	0	2.516913	-4.42374	19.56944602	0.989996	-2.18820012
7	2006	1	1	5:00	0	0	0	0	2.596308	-4.29908	18.48207147	1.012186	-2.16601014
8	2006	1	1	6:00	0	0	0	0	2.810628	-3.735	13.95018847	1.122121	-2.05607483
9	2006	1	1	7:00	0	0	0	0	2.81544	-3.72892	13.90488123	1.146447	-2.03174909

Figure 4. 1 Example data set for statistical analysis for HOT sections.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
1	Year	Month	Day	Time	CrashCount01	Rainy	Snowy	logvph	occ_dev	occ_dev_sqr	occ_sd	occ_sd_dev	Closed	PDSL	DARK	VSA
2	2006	1	1	0:00	0	0	0	7.597621	-3.61259	13.05078169	1.724686	-1.53263162	0	0	0	0
3	2006	1	1	1:00	0	0	0	7.553171	-3.71745	13.81945939	1.659114	-1.59820381	0	0	0	0
4	2006	1	1	2:00	0	0	0	7.246652	-4.30729	18.55276537	1.351492	-1.90582635	0	0	0	0
5	2006	1	1	3:00	0	0	0	6.864771	-4.98073	24.80769995	1.064316	-2.19300201	0	0	0	0
6	2006	1	1	4:00	0	0	0	6.37811	-5.46051	29.81715745	0.798244	-2.4590742	0	0	0	0
7	2006	1	1	5:00	0	0	0	6.512132	-5.34781	28.59903561	0.850528	-2.40679052	0	0	0	0
8	2006	1	1	6:00	0	0	0	6.900098	-4.95179	24.52026915	0.96566	-2.29165807	0	0	0	0
9	2006	1	1	7:00	0	0	0	6.871847	-4.98418	24.84207762	1.007618	-2.24969997	0	0	0	0
10	2006	1	1	8:00	0	0	0	7.163853	-4.65285	21.64898257	1.008766	-2.24855172	0	0	0	0

Figure 4. 2 Example data set for statistical analysis for PDSL sections.

4.1 Crash Data

To avoid possible biases from secondary crashes, in this study, the response variable in the statistical model was the presence absence of a rear-end crash in each hour, which is a binary variable where "1" was assigned to the hours when there was at least one rear-end crash happened and "0" otherwise. There were cases that two consecutive rear-end crashes happened in different hours but the crash times were actually within 1 hour. In this case, the latter ones were regarded as dependent on the previous ones and deleted from the sample.

4.2 Traffic Condition Data

Since hourly traffic condition data were needed for analysis, traffic data were aggregated based on the original 30-second raw data.

To guarantee the effectiveness of the traffic condition data in analysis, a data quality checking process was done before data aggregation. The main task in this data quality checking process was to identify questionable data that should not be used for analysis.

Three types of "bad" data were identified:

- Records with negative volume or occupancy.
- Records with occupancy greater than 100%.
- Records with repeated pattern of data (repeating "0" volume and occupancy records during 0:00-6:00 AM were not regarded as questionable data).

Those 30-second traffic condition records with at least one of the three issues above were excluded from data aggregation.

For each hour, summary measures including lane occupancy mean, lane occupancy variance, and lane volume mean from the remaining 30-second data were computed as follows:

Non-crash hours' traffic data were computed by following formulae:

- Lane occupancy: $\bar{O} = \frac{\sum_{i=1}^{n}\sum_{j=1}^{m} occupancy_{ij}}{n \times m}$
- Lane occupancy variance: $\sigma^2 = \frac{\sum_{i=1}^{n}\sum_{j=1}^{m}(occupancy_{ij} - \bar{O})^2}{n \times m}$
- Lane volume mean: $\bar{V} = \frac{\sum_{i=1}^{n}\sum_{j=1}^{m} volume_{ij}}{n \times m}$

Where

n - number of lanes open in each hour

m - number of records left after abnormal records excluded in each hour

$occupancy_{ij}$ – the j^{th} lane occupancy record for the i^{th} lane

$volume_{ij}$ - the j^{th} lane volume record for the i^{th} lane

It is worth noting that, since crashes often cause substantial changes in traffic conditions, for those hours when crashes occurred the traffic conditions were computed using data from the 30-minute periods preceding the reported times of the crashes.

Finally, traffic flow, centered 1-hour average lane occupancy and centered lane occupancy standard deviation for each section were computed based on the 1-hour lane volume mean, lane occupancy mean, and lane occupancy variance as follows:

- Traffic flow $\mathrm{Q} = 120 \times \mathrm{n} \times \bar{V}$
- Centered average lane occupancy $= \bar{O}_{ij} - \frac{\sum_{k=1}^{52608} \overline{O_k}}{52608}$
- Centered lane occupancy standard deviation $= \sigma_{ij} - \frac{\sum_{k=1}^{52608} \overline{\sigma_k}}{52608}$

4.3 PDSL Activation Data

Proportions of the duration of each PDSL activation status in each hour were calculated after a data quality checking process.

Below are the issues found during the data quality checking process:

- Missing data including blank records and records with message as “UNKNOWN”.
- Vague PDSL activation status indicated by message as “DARK” and “USE_CAUTION”.
- Presence of non-zero occupancy and volume during “PDSL-closed” hours.
- Different messages switched on an ILCS at relative quick pace.

MnDOT provided relevant explanations for the issues above, and data were processed based on this additional information.

First, the missing data in the ILCS log was due to technical problem and there was no other way to regain the PDSL activation data missing in the log. Thus, records with missing data were deleted.

For the second issue, for the log entry “DARK”, it was hard to tell whether the PDSL was running as a general purpose lane or if this was due to a technical issue, such as loss of communication with ILCS, thus we kept this status in our statistical model. For the entry “USE_CAUTION”, mostly it indicated that there was an incident. Obviously, there is correlation between this status and the presence of a crash. Thus, records with “USE_CAUTION” status were deleted from the analysis.

Third, traffic data during closed hours were mainly due to violators, snow or ice events, or to incidents. We differentiated the causes of abnormal traffic data during PDSL-closed hours by looking at the loop detector data trend. If the traffic data in PDSL had "discontinuous" pattern, that is, a non-zero data record presented in the middle of several zero records, this non-zero data record was thought to be caused by violators. The non-zero traffic data due to violators were ignored in this study. If there was a "continuous" trend in the non-zero traffic data during PDSL-closed hours, it is most likely there was snow or ice events, or incidents, and ILCS always had non-closed messages. Such non-zero traffic data have been taken into consideration in this study.

The last issue typically happened because of an evolving incident or human error in commanding the ILCS in response to an evolving incident. We manually picked out the PDSL activation records with such issues and looked at the switching pattern. Those records going back and forth between different messages within 2 minutes were considered due to human error and were smoothed based on the first and last messages in that switch. Otherwise, the original records were kept.

4.4 Weather Data

Rainy and snowy weather condition indicators were introduced to the statistical model established later. Both indicators were binary variables where "1" was assigned to hours when there was rain or snow respectively and "0" otherwise. The weather condition indicator data were processed based on the original weather data collected in section 3.5.

4.5 Other

In addition to crash, traffic condition, PDSL activation, and weather data, another variable, the before or after period indicator was introduced in the statistical model for HOT sections. The before or after indicator was a binary variable, with "0" being assigned for hours in 2011-2013 and "1" for hours in year 2006 to 2008.

Chapter 5: Statistical Analyses

This chapter describes the statistical analysis of how the rear-end crash risk in a given hour varies with respect to other conditions prevailing during that hour. Our goal was to determine the factors associated with rear-end crash risk on the studied sections during the studied years, especially if the change in rear-end crash frequencies was direct effect of UPA interventions. Since different UPA interventions were implemented within I-35W from TH-13 to I-494 and PDSL region, analyses for those two sections were carried out separately. It is worth noting that, for each analyzed section, only those crashes for which we were confident that their actual locations were within this given section were considered.

5.1 Statistical Modeling

5.1.1 Logistic regression model

For a given section of freeway, let Y_i denote "the presence or absence of a rear-end crash during hour i", with "Y_i=1" indicating "the presence of at least one rear-end crash during hour i" and "Y_i=0" indicating "no rear-end crash during hour i".

The variables Y_i are assumed to be Bernoulli outcomes with parameters π_i, the "rear-end crash risk during hour i" in this study's content. That is

$$\Pr\{Y_i = \mathrm{y}\} = {\pi_i}^{y}(1-\pi_i)^{1-y}, \text{ y=0, 1.} \qquad \text{(Eq. 5.1)}$$

Logistic regression was adopted for statistical analysis with binary response.

The logistic regression model assumes that the logarithm of the odds, $\frac{P(Y_i = 1)}{P(Y_i = 0)}$, of an observation Y_i can be expressed as a linear function of K dependent variables:

$$\mathrm{E}\,\{\log\frac{P(Y_i=1)}{P(Y_i=0)}\} = \mathrm{E}\,\{\log\frac{P(Y_i=1)}{1-P(Y_i=1)}\} = E\,\{\log\frac{\pi_i}{1-\pi_i}\} = \beta_0 + \sum_{k=1}^{K}\beta_k x_{ki} \qquad \text{(Eq. 5.2)}$$

Where $\beta_k' s$ are parameters;

X_{ki} 's = observed value of predictor k associated with hour i.

Equation 5.2 can also be written as:

$$\mathrm{E}\,\{\mathrm{P}(Y_i = 1)\} = \pi_i = \frac{\exp(\beta_0+\sum_{k=1}^{K}\beta_k x_{ki})}{1+\exp(\beta_0+\sum_{k=1}^{K}\beta_k x_{ki})} \qquad \text{(Eq. 5.3)}$$

5.1.2 Parameter Estimation

Maximum Likelihood Estimation (MLE) (Dobson and Barnett, 2008) was used to estimate the coefficients, namely the β_k 's, in the logistic regression model presented in Equation 5.3.

The corresponding joint probability function is:

$$\mathrm{L}(\beta_0, \ldots, \beta_K) = \prod_{i=1}^{n} {\pi_i}^{Y_i}(1-\pi_i)^{1-Y_i} \qquad \text{(Eq. 5.4)}$$

Substituting (Eq. 5.3) into (Eq. 5.4) and taking the natural logarithm gives:

$$l(\beta_0, \ldots, \beta_K) = \sum_{i=1}^{n} Y_i\left(\beta_0 + \sum_{k=1}^{K} \beta_k x_{ki}\right) - \sum_{i=1}^{n} \log[1 + \exp(\beta_0 + \sum_{k=1}^{K} \beta_k x_{ki})]$$

(Eq. 5.5)

The estimates of $(\beta_0, \ldots, \beta_K)$ should be the values that maximize the log-likelihood function in (Eq. 5.5).

Interpretation for $\beta_k' s$ are as follows:

i) Intercept β_0

β_0 represents the logarithm of the odds of having the event represented by Y, Y_i=1, when other conditions are fixed at 0.

ii) Slopes $\beta_{k,k\neq 0}$

If X_k is a binary variable representing a presence, β_k represents the logarithm of the odds of having the event represented by Y associated with the exposure adjusted for other predictors.

If X_k is on continuous scale, β_k represents the change in logarithm of the odds associated with one-unit increase in the value of X_k adjusted for other predictors.

5.1.3 Goodness of Fit as Hosmer-Lemeshow Test

The Hosmer-Lemeshow (H-L) (Hosmer and Lemeshow, 2000) test is widely used as a goodness of fit test for logistic regression, especially for risk prediction models.

i) Null and alternative hypotheses

H_0: actual and predicted event rates are similar across G groups (or, the proposed logistic model fits the data)

H_1: actual and predicted event rates are not similar across G groups (or, the proposed logistic model does not fit the data)

ii) Test statistic

To perform the H-L test, after computing maximum likelihood estimates of the coefficients, the data is first grouped into G groups by ordered predicted probabilities of "success" ($Y_i=1$). Then, the H-L test statistic is calculated by the formula below:

$$G^2_{H-L} = \sum_{g=1}^{G} \frac{(O_g - E_g)^2}{E_g(1 - E_g/N_g)} \qquad \text{(Eq. 5.6)}$$

Where

G = the number of subgroups

O_g = number of observed events in the g^{th} group

E_g = number of expected events in the g^{th} group

N_g = the number of observations in the g^{th} group

For large samples, the test statistic, G^2_{H-L}, is approximately Chi-square distribution with (G-2) degrees of freedom ($G^2_{H-L} \sim \chi^2_{G-2}$) under the null hypothesis.

iii) Decision rule:

Given the null hypothesis, the computed value of the test statistic, G^2_{H-L}, is unlikely. Thus:

If Pr ($\chi^2_{G-2} > G^2_{H-L}$) ≤ α, where α is the significance level, for example the conventional 0.05, we reject the null hypothesis and conclude there is not enough evidence to show the model is a good fit;

If Pr ($\chi^2_{G-2} > G^2_{H-L}$) > α, we cannot reject the null hypothesis but conclude that either the fit is good or that there is not enough data to detect a poor fit.

In this study, G=10 was used as the number of subgroups for H-L test and α=0.05 was chosen as the significance level.

5.1.4 Goodness of Fit as Likelihood Ratio Test

The likelihood ratio (L-R) test is commonly used in model comparison for generalized linear models (GLM). L-R test helps compare two nested models where the simpler model is a special case of the more complex model. The test can be described as follows (Agresti, 2014).

i) Null and alternative hypothesis (assuming current model, M_1, holds)

H_0: The reduced (simpler) model, M_0 is equivalent to the more complex model.

H_1: The reduced model is not equivalent to the more complex model.

ii) Test statistic:

L-R test statistic is defined as below:

$$G^2(M_0|M_1) = -2(L_0 - L_1) \quad \text{(Eq. 5.7)}$$

Where

L_0 = the maximum log-likelihood from M_0

L_1 = the maximum log-likelihood from M_1

$G^2(M_0|M_1)$ has an approximately chi-squared null distribution with p (p>0) degrees of freedom, where p is difference in the number of coefficients between M_0 and M_1.

iii) Decision rule:

Given the null hypothesis, the computed value of the test statistic is unlikely. Thus:

If Pr ($\chi_k^2 > G^2(M_0|M_1)) \leq \alpha$ where α is the significance level, we reject the null hypothesis and conclude that the reduced (simpler) model, M_0, is inferior to the more complex model.

If Pr ($\chi_k^2 > G^2(M_0|M_1)) > \alpha$, we cannot reject the null hypothesis but conclude that the simpler model is as good as the more complex) model.

In this study, α=0.05 was chosen as the significance level.

5.2 Logistic Regression Results

5.2.1 I-35W from TH-13 to I-494

To establish the logistic regression model, for each hour during 2006-2008 and 2011-2013, the presence absence of a crash was the dependent variable while independent variables consisted of traffic volume and lane occupancy, the presence or absence of snowy or rainy conditions, and the Before or After period indicator.

Table 5. 1 is a list of variables involved in logistic regression model for sections on I-35W from TH-13 to I-494.

Table 5. 1 Variables Selected for Logistic Regression Analysis for I-35W from TH-13 to I-494

Symbol	Role	Name	Type	Value
Y_i	Response	Rear-end Crash Presence Absence	Binary	The presence absence of a rear-end crash during hour i. 0 = no rear-end crash during hour i; 1 = at least one rear-end crash during hour i.
X_{1i}	Predictor	Rainy	Binary	Rainy weather condition indicator for hour i. 1 – Rainy during hour i; 0 – Otherwise.
X_{2i}	Predictor	Snowy	Binary	Snowy weather condition indicator for hour i. 1 – Snowy during hour i; 0 – Otherwise.
X_{3i}	Predictor	log(vph)	Continuous	Natural logarithm of section traffic flow, in vehicles/hour, during hour i.
X_{4i}^*	Predictor	Lane Occupancy	Continuous	Centered average lane occupancy during hour i.
X_{5i}^*	Predictor	Lane Occupancy^2	Continuous	The square of centered lane occupancy during hour i.
X_{6i}^*	Predictor	Occupancy Standard Deviation	Continuous	Centered lane occupancy standard deviation during hour i.
X_{7i}	Predictor	Before/After	Binary	The time period indicator for hour i. 1 – Hour occurring 2011-2013; 0 – Hour occurring 2006-2008.

For each of the section, N6, N9-1, N9-2, N11, N17, N18, S6, S9, S11, S17, and S18, the generalized linear model routine glm with MLE implemented in statistical analysis package in software R (R, 2015) was used to fit and evaluate logistic regression models containing different combinations of independent variables listed in **Table 5. 1**. However, it turned out that the numbers of rear-end crashes in section N6, N9-2, N11, S6, S11, and S18 were insufficient to obtain reliable parameter estimates, thus the analyses of rear-end crash risk for those 6 sections was not be pursued. Previous research into short-term prediction of freeway crash risk has indicated that risk increases as traffic density increases. Since lane occupancy is, to a first approximation, proportional to density the initial set of analyses included lane occupancy but did not include the square of lane occupancy.

Table 5. 2 is the estimation summary for the initial logistic regression model containing predictors X_1-X_4 and X_6-X_7.

Table 5. 2 Estimation Summary for Initial Model of Rear-End Crash Probability on I-35W from TH-13 to I-494

Section No.	Variable	Estimate	Std. Error	z value	Pr (>|z|)	Signif. codes
N9-1	Constant	-15.76981	4.88482	-3.228	0.00125	**
	Rainy	-0.14798	0.47663	-0.310	0.75621	
	Snowy	-0.16311	0.57710	-0.283	0.77745	
	logvph	0.93626	0.60143	1.557	0.11953	
	Lane Occupancy	0.09135	0.04768	1.916	0.05537	.
	Lane Occupancy^2	--	--	--	--	
	Occupancy Standard Deviation	0.21474	0.07077	3.034	0.00241	**
	Before/After	0.28421	0.40247	0.706	0.48009	
	Null deviance = 491.08			H-L = 12.799, p-value = 0.119		
	Residual deviance = 395.40			AIC: 409.4		
N17	Constant	-23.99466	5.55990	-4.316	1.59E-05	***
	Rainy	0.18507	0.34809	0.532	0.59495	
	Snowy	0.19109	0.60987	0.313	0.75403	
	logvph	2.04723	0.68085	3.007	0.00264	**
	Lane Occupancy	-0.02572	0.05473	-0.470	0.63833	
	Lane Occupancy^2	--	--	--	--	
	Occupancy Standard Deviation	0.29810	0.06377	4.675	2.94E-06	***
	Before/After	-0.28042	0.37570	-0.746	0.45543	
	Null deviance =767.25			H-L =40.814, p-value =2.258E-06		
	Residual deviance = 637.76			AIC: 651.76		
N18	Constant	-16.5466	3.21863	-5.141	2.73E-07	***
	Rainy	0.48311	0.22243	2.172	0.02986	*
	Snowy	0.01569	0.43778	0.036	0.97141	
	logvph	1.20496	0.40353	2.986	0.00283	**
	Lane Occupancy	0.11395	0.02065	5.517	3.44E-08	***
	Lane Occupancy^2	--	--	--	--	
	Occupancy Standard Deviation	0.07927	0.01512	5.242	1.59E-07	***
	Before/After	-0.14611	0.22133	-0.66	0.50916	
	Null deviance =1527.0			H-L = 15.73, p-value = 0.04642		
	Residual deviance = 1270.2			AIC: 1284.2		
S9	Constant	-14.24103	3.09647	-4.599	4.24E-06	***
	Rainy	1.41656	0.61832	2.291	0.021965	*
	Snowy	2.26803	0.59768	3.795	0.000148	***
	logvph	0.59299	0.38376	1.545	0.122295	
	Lane Occupancy	-0.08408	0.06288	-1.337	0.18117	
	Lane Occupancy^2	--	--	--	--	

	Occupancy Standard Deviation	0.37054	0.0791	4.684	2.81E-06	***
	Before/After	1.04037	0.54223	1.919	0.055024	.
	Null deviance = 339.04			H-L =3.7801, p-value = 0.8764		
	Residual deviance =278.91			AIC: 292.91		
S17	Constant	-35.724636	9.351959	-3.820	0.000133	***
	Rainy	-0.007307	0.635975	-0.011	0.990833	
	Snowy	0.797148	0.758764	1.051	0.293448	
	logvph	7.684986	2.628153	2.924	0.003455	**
	Lane Occupancy	0.101498	0.055943	1.814	0.069631	.
	Lane Occupancy^2	--	--	--	--	
	Occupancy Standard Deviation	0.134789	0.084566	1.594	0.110959	
	Before/After	0.477426	0.474381	1.006	0.314214	
	Null deviance = 462.66			H-L = 76.3812, p-value = 2.608e-13		
	Residual deviance = 358.76			AIC: 372.76		

(Signif. codes: 0 '*' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1)**

N9-1 was used as an example to see how to interpret the information provided in **Table 5. 2**:

i) Intercept β_0

The maximum likelihood estimate of the coefficient β_0 was -15.76981 and the standard error associated with this estimate was 4.88482. A test of the null hypothesis $\beta_0 = 0$ yielded a z-statistic equal to -3.228, and the probability of obtaining a z-value at least this large when the null hypothesis was true was 0.0125. Using the rule that we cannot reject a null hypothesis when a P-value is greater than 0.05, which indicates that β_0 is statistically significant different from 0 at 0.05 significance level. The 95% Wald confidence interval of β_0 is (-25.3441, -6.19556), given by (-15.76981-1.96×4.88482, -15.76981+1.96×4.88482). In a situation where all model predictors were equal to zero the estimated probability of a rear-end crash would be exp (-15.76981) / (1+exp (-15.76981)) = 0.000000142 with a 95% Wald confidence interval (9.845×10^{-12}, 0.00203), given by (exp (-25.3441) / (1+exp (-25.3441), exp (-6.19556) / (1+exp (-6.19556)).

ii) Slopes $\beta_{k,k\neq0}$

Table 5. 2 shows that the maximum likelihood estimate of the coefficient β_1 was 0.14798 and the standard error associated with this estimate was 0.47663. A test of the null hypothesis $\beta_1 = 0$ yielded a z-statistic equal to -0.310, and the probability of obtaining a z-value at least this large when the null hypothesis was true was 0.75621. Using the rule

that we cannot reject a null hypothesis when a P-value is greater than 0.05, which indicates that rainy condition does not have statistically significant association with rear-end crash probability at 0.05 significance level. Similar conclusions can be applied to weather variable Snowy, the traffic variables Log Flow, and the Before/After predictor. Those four predictors showed no clear association with rear-end crash risk at 0.05 significance level.

On the other hand, the maximum likelihood estimate of the coefficient β_4 was 0.09135 and the standard error associated with this estimate was 0.04768. A test of the null hypothesis $\beta_4 = 0$ yielded a z-statistic equal to 1.916, and the probability of obtaining a z-value at least this large when the null hypothesis was true was 0.05537. The association that lane occupancy has with rear-end crash is marginally non-significant at 0.05 significance level. The positive value of the estimate of β_4 indicates that rear-end crash probability could increase as lane occupancy increases. The maximum likelihood estimate of the coefficient β_6 was 0.21474 and the standard error associated with this estimate was 0.07077. A test of the null hypothesis $\beta_6 = 0$ yielded a z-statistic equal to 3.034, and the probability of obtaining a z-value at least this large when the null hypothesis was true was 0.00241. Using the rule that we reject a null hypothesis when a P-value is less than 0.05, we see that one should clearly reject the null hypothesis that lane occupancy standard deviation has no association with rear-end crash probability. The positive value of the estimate of β_6 means that rear-end crash probability increases as the standard deviation of lane occupancy increases. The odds of having a rear-end crash in a given hour is $e^{0.21474} = 1.23954$ times the odds of having a rear-end crash in a given hour with one-unit decrease in standard deviation of lane occupancy with a 95% Wald confidence interval (1.0790, 1.4240), holding other conditions constant.

iii) Goodness of fit

The null deviance and residual deviance are 491.08 and 395.40, respectively. The Hosmer-Lemeshow goodness of fit statistic was 12.799 with a p-value of 0.119, which indicates the actual and predicted rear-end crash risk are similar across 10 subgroups at 0.05 significance level.

Since lane occupancy is, to a first approximation, proportional to traffic density, the findings of Xu et al (2014) suggest that the relationship between lane occupancy and crash probability should show an "inverted-U" shape with a maximal point where crash probability is greatest and falling off for lane occupancies both less than and greater than this maximal point. One way to allow for this is to include the square of lane occupancy, our variable X_5, as a predictor and **Table 5. 3** shows the estimation summary for a model using all the predictors listed in **Table 5. 1**.

Table 5. 3 Estimation Summary for I-35W from TH-13 to I-494, with Quadratic Occupancy Effect

Section No.	Variable	Estimate	Std. Error	z value	Pr (>|z|)	Signif. codes
N9-1	Constant	-6.904828	6.686002	-1.033	0.302	
	Rainy	-0.246231	0.478694	-0.514	0.607	
	Snowy	-0.372604	0.597194	-0.624	0.533	
	logvph	-0.218047	0.860995	-0.253	0.800	
	Lane Occupancy	0.298789	0.146963	2.033	0.042	*
	Lane Occupancy^2	-0.005544	0.003969	-1.397	0.162	
	Occupancy Standard Deviation	0.137832	0.084855	1.624	0.104	
	Before/After	0.229706	0.401705	0.572	0.567	
	Null deviance = 491.08			H-L = 11.358, p-value = 0.1822		
	Residual deviance = 393.26			AIC: 409.26		
N17	Constant	-10.524371	7.018016	-1.500	0.1337	
	Rainy	0.183404	0.345895	0.530	0.5960	
	Snowy	0.042289	0.615347	0.069	0.9452	
	logvph	0.254392	0.915177	0.278	0.7810	
	Lane Occupancy	0.382850	0.202204	1.893	0.0583	.
	Lane Occupancy^2	-0.009998	0.005047	-1.981	0.0476	*
	Occupancy Standard Deviation	0.094698	0.118116	0.802	0.4227	
	Before/After	0.642862	0.572757	1.122	0.2617	
	Null deviance = 767.25			H-L = 19.309, p-value = 0.01329		
	Residual deviance = 633.57			AIC: 649.57		
N18	Constant	2.135337	3.36471	0.635	0.52567	
	Rainy	0.343696	0.223785	1.536	0.12458	
	Snowy	-0.142324	0.43293	-0.329	0.74235	
	logvph	-1.286697	0.449032	-2.865	0.00416	**
	Lane Occupancy	0.659326	0.113015	5.834	5.41E-09	***
	Lane Occupancy^2	-0.021735	0.004496	-4.834	1.34E-06	***
	Occupancy Standard Deviation	-0.033361	0.040785	-0.818	0.41337	
	Before/After	0.349962	0.266592	1.313	0.18928	
	Null deviance = 1527			H-L =5.4458, p-value =0.709		
	Residual deviance = 1238			AIC: 1254		
S9	Constant	-3.80281	4.3257	-0.879	0.379336	
	Rainy	1.29209	0.608289	2.124	0.033659	*
	Snowy	2.096371	0.601594	3.485	0.000493	***
	logvph	-0.861893	0.587604	-1.467	0.142433	
	Lane Occupancy	0.340928	0.166932	2.042	0.04112	*
	Lane Occupancy^2	-0.010986	0.005308	-2.07	0.038466	*

	Occupancy Standard Deviation	0.216057	0.088168	2.451	0.014265	*
	Before/After	1.779921	0.612522	2.906	0.003662	**
	Null deviance = 339.04			H-L = 3.3145, p-value = 0.9131		
	Residual deviance =270.99			AIC: 286.99		
S17	Constant	1.091499	6.47764	0.169	0.8662	
	Rainy	-0.089735	0.630595	-0.142	0.8868	
	Snowy	0.344132	0.762821	0.451	0.6519	
	logvph	-1.378298	0.829014	-1.663	0.0964	.
	Lane Occupancy	0.870246	0.194617	4.472	7.76E-06	***
	Lane Occupancy^2	-0.020182	0.00501	-4.028	5.62E-05	***
	Occupancy Standard Deviation	-0.218001	0.139777	-1.56	0.1188	
	Before/After	-0.006204	0.5084	-0.012	0.9903	
	Null deviance = 462.66			H-L =24.339, p-value = 0.00201		
	Residual deviance = 344.13			AIC: 360.13		

(Signif. codes: 0 '*' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1)**

L-R tests were conducted to determine the better model between a reduced model M_0, which does not have a quadratic term for lane occupancy, and the more complex model M_1, which includes a quadratic term of lane occupancy. **Table 5. 4** shows the L-R test results for each analyzed section between TH-13 and I-494.

Table 5. 4 Likelihood Ratio Test Results for Analyzed Sections in TH-13 to I-494 Region

Section No.	Model	Number of coefficients	Log-likelihood	p	$G^2(M_0\|M_1)$	P-value	Significance
N9-1	M_0	7	-197.70	1	2.140	1.44E-01	
	M_1	8	-196.63				
N17	M_0	7	-318.88	1	4.191	4.07E-02	*
	M_1	8	-316.79				
N18	M_0	7	-635.09	1	32.146	1.43E-08	***
	M_1	8	-619.02				
S9	M_0	7	-139.46	1	7.922	4.88E-03	**
	M_1	8	-135.50				
S17	M_0	7	-179.38	1	14.636	1.30E-04	***
	M_1	8	-172.06				

(Signif. codes: 0 '*' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1)**

According to the L-R test results, we reject the null hypothesis that the reduced model (without quadratic term) holds and conclude that adding quadratic term improved model fit for four of the five analyzed sections at 0.05 significance level.

Table 5. 5 shows a summary for the model using only the statistically-significant predictors from **Table 5. 3**.

Table 5. 5 Estimation Summary for I-35W from TH-13 to I-494 for Model Using Only Statistically Significant Predictors from Table 5.3

Section No.	Variable	Estimate	Std. Error	z value	Pr (>|z|)	Signif. codes

N9-1	Constant	-7.8831	0.2495	-31.60	<2E-16	***
	Lane Occupancy	0.1939	0.0177	10.95	<2E-16	***
	Null deviance = 491.08			H-L = 6.7727, p-value = 0.5613		
	Residual deviance = 411.20			AIC: 415.2		
N17	Constant	-8.187874	0.290762	-28.160	<2E-16	***
	Lane Occupancy	0.464298	0.59993	7.739	1.00E-14	***
	Lane Occupancy^2	-0.011120	0.002601	-4.275	1.91E-05	***
	Null deviance = 767.25			H-L = 12.456, p-value = 0.132		
	Residual deviance = 642.59			AIC: 648.59		
N18	Constant	0.249152	2.828839	0.088	0.92982	
	Log Flow	-1.008219	0.369338	-2.730	0.00634	**
	Lane Occupancy	0.580674	0.068091	8.528	<2E-16	***
	Lane Occupancy^2	-0.019245	0.003544	-5.430	5.64E-08	***
	Null deviance = 1527			H-L =5.338, p-value =0.7209		
	Residual deviance = 1242			AIC: 1250		
S9	Constant	-10.150481	0.678043	-14.970	<2E-16	***
	Rainy	1.353125	0.608121	2.225	0.026075	*
	Snowy	2.239326	0.588545	3.805	0.000142	***
	Lane Occupancy	0.143930	0.085466	1.684	0.092169	.
	Lane Occupancy^2	-0.005994	0.002741	-2.187	0.028740	*
	Occupancy Standard Deviation	0.293711	0.074220	3.957	7.58E-05	***
	Before/After	1.487029	0.577951	2.573	0.010084	*
	Null deviance = 339.04			H-L = 2.855, p-value = 0.9431		
	Residual deviance =272.98			AIC: 286.98		
S17	Constant	-7.334567	9.236361	-0.794	0.427139	
	Log Flow	-0.276452	1.158900	-0.239	0.811457	
	Lane Occupancy	0.596468	0.112820	5.287	1.24E-07	***
	Lane Occupancy^2	-0.015055	0.004288	-3.511	0.000446	***
	Null deviance = 462.66			H-L =22.227, p-value = 0.004512		
	Residual deviance = 347.53			AIC: 355.53		

(Signif. codes: 0 '*' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1)**

Among all studied sections on I-35W from TH-13 to I-494, only S9 showed a significant association between Before/After predictor and rear-end crash risk at 0.05 significance level. In S9, the UPA changes were associated with an increase in rear-end crash risk when other conditions were controlled for. For S9, the odds of having a rear-end crash in a given hour in After period is $e^{1.487029} = 4.4239321$ times the odds of having a rear-end crash in a given hour in Before period with a 95% Wald confidence interval (1.4251, 13.7332), holding other conditions constant.

S9 was the only section showing significant weather effect at 0.05 significance level. The odds of having a rear-end crash in a given hour in rainy condition is $e^{1.353125} = 3.869499$ times the odds of having a rear-end crash in a given hour without rain with a 95% Wald confidence interval (1.1749, 12.7438), holding other conditions constant. The odds of having a rear-end crash in a given hour in snow condition is $e^{2.239326} = 9.387002$ times the odds of having a rear-end crash in a given hour without snow with a 95% Wald confidence interval (2.9617, 29.7513), holding other conditions constant.

As is shown in **Table 5. 5**, the two coefficients associated with lane occupancy, β_4 and β_5, are both significantly or marginal significantly different from zero for all analyzed sections except for N9-1; while their signs $\beta_4 > 0$ with $\beta_5 < 0$, imply that rear-end crash probability does have the inverted-U shape.

Using the estimates in **Table 5. 5**, rear-end crash probability for N17, N18, S9, and S17 were maximized when lane occupancy was approximately equal to

i) N17: $o_{max} \approx \bar{o} - \frac{\widehat{\beta_1}}{2\widehat{\beta_2}} = 5.19 - \frac{0.464298}{2(-0.011120)} = 26.07$

ii) N18: $o_{max} \approx \bar{o} - \frac{\widehat{\beta_1}}{2\widehat{\beta_2}} = 6.09 - \frac{0.580674}{2(-0.019245)} = 21.18$

iii) S9: $o_{max} \approx \bar{o} - \frac{\widehat{\beta_1}}{2\widehat{\beta_2}} = 5.04 - \frac{0.143930}{2(-0.005994)} = 17.05$

iv) S17: $o_{max} \approx \bar{o} - \frac{\widehat{\beta_1}}{2\widehat{\beta_2}} = 5.89 - \frac{0.596468}{2(-0.015055)} = 25.70$

Where $\bar{o}$ denotes the average lane occupancy over all hours in the Before and After period for N17, N18, S9, and S17, respectively.

Figure 5. 1 to **Figure 5. 4** show the time-series plots of average lane occupancy in sections N17, N18, S9, and S17 for each hour during both the Before and After periods, with the vertical lines denoting the Before/After change points. The horizontal lines show the approximate values of average lane occupancy where rear-end crash risk is maximal.

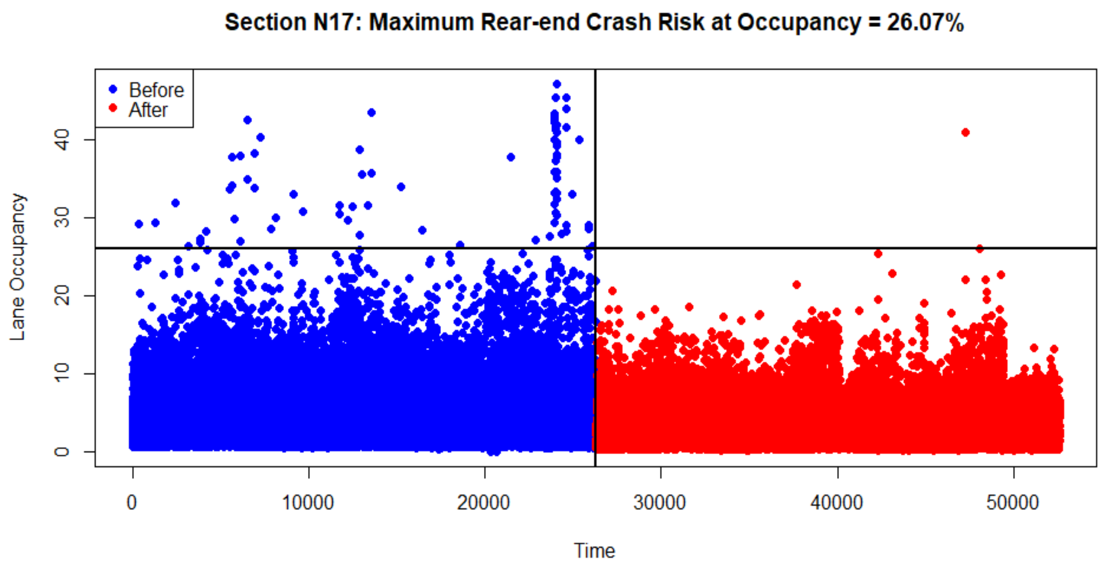

Figure 5. 1 Time-series plot of hourly average lane occupancy for section N17, showing the Before and After UPA periods, and the average lane occupancy with maximal crash probability.

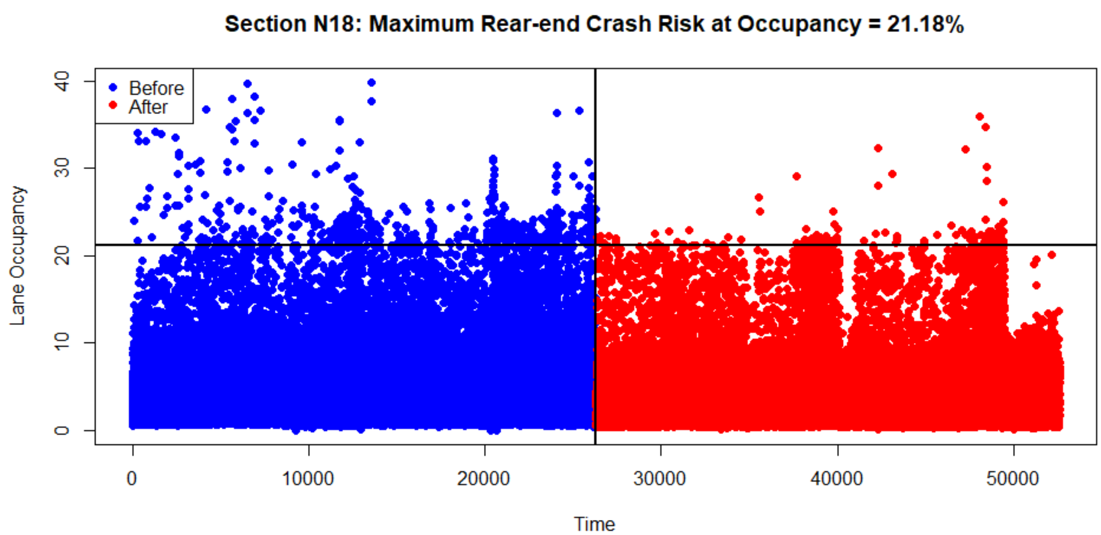

Figure 5. 2 Time-series plot of hourly average lane occupancy for section N18, showing the Before and After UPA periods, and the average lane occupancy with maximal crash probability.

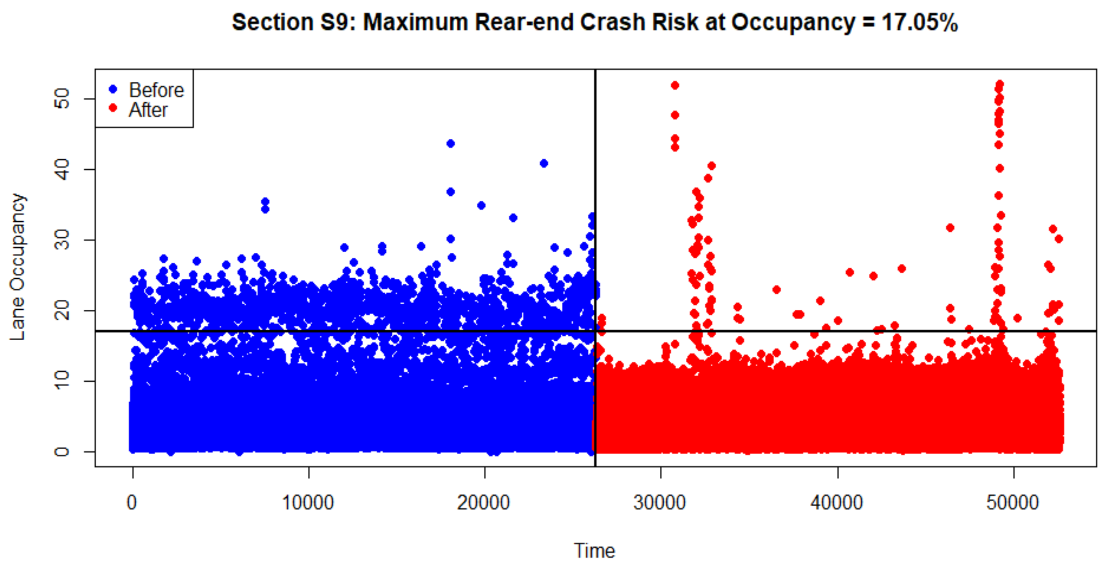

Figure 5. 3 Time-series plot of hourly average lane occupancy for section S9, showing the Before and After UPA periods, and the average lane occupancy with maximal crash probability.

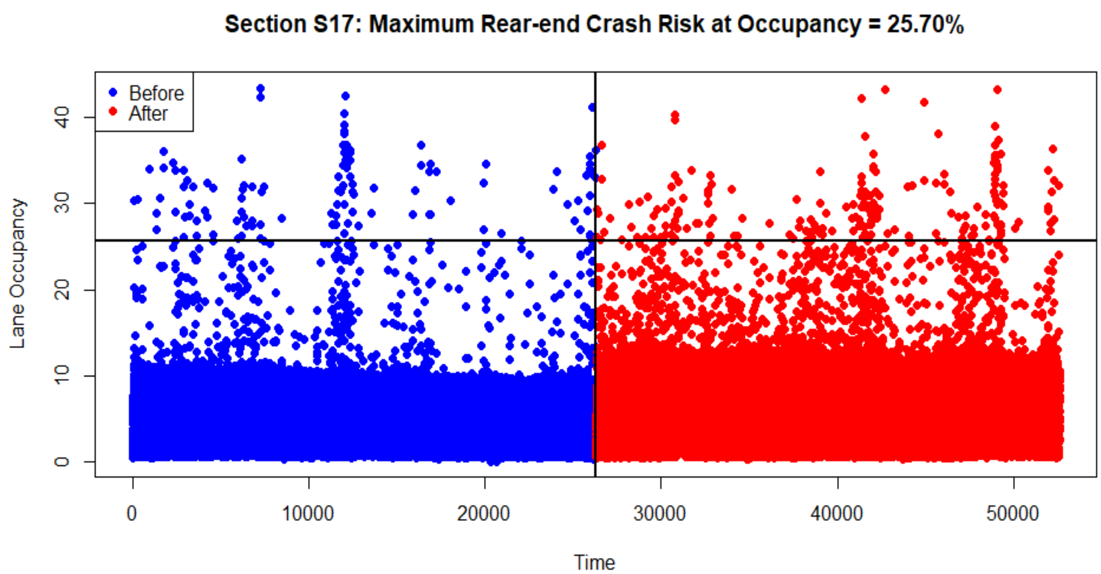

Figure 5. 4 Time-series plot of hourly average lane occupancy for section S17, showing the Before and After UPA Periods, and the average lane occupancy with maximal crash probability.

In summary, most analyzed sections on I-35W from TH-13 to I-494 showed no significant change associated with UPA project except for southbound S9. The rear-end crash risk of S9 (just north of Minnesota River) was a complicated function of weather, traffic and Before vs After the UPA improvements. N17 (just south of I-494) actually experienced fewer crashes after the UPA project, but the reduction was not as great as the change in lane occupancy would predict. In addition, “Inverted U” relationships between lane occupancy and crash risk have been seen in several sections.

5.2.2 I-35W PDSL region

As with the HOT region, for each hour during 2006-2008 and 2011-2013, the presence absence of a crash became the dependent variable while independent variables consisted of traffic volume and lane occupancy, the presence or absence of snowy or rainy conditions, Before or After period indicator, and the presence or absence of PDSL operation.

Table 5. 6 is the list of variables involved in the logistic regression model for I-35W PDSL region.

Table 5. 6 Variables Selected for Logistic Regression Analysis for I-35W PDSL region

Symbol	Role	Name	Type	Value
Y_i	Response	Rear-end Crash Presence Absence	Binary	The presence absence of a rear-end crash during hour i. 0 = no rear-end crash during hour i; 1 = at least one rear-end crash during hour i.
X_{1i}	Predictor	Rainy	Binary	Rainy weather condition indicator for hour i. 1 – Rainy during hour i; 0 – Otherwise.
X_{2i}	Predictor	Snowy	Binary	Snowy weather condition indicator for hour i. 1 – Snowy during hour i; 0 – Otherwise.
X_{3i}	Predictor	log(vph)	Continuous	Natural logarithm of section traffic flow, in vehicles/hour, during hour i.
X_{4i}^*	Predictor	Lane Occupancy	Continuous	Centered average lane occupancy during hour i
X_{5i}^*	Predictor	Lane Occupancy^2	Continuous	The square of lane occupancy during hour i.
X_{6i}^*	Predictor	Occupancy Standard Deviation	Continuous	Centered lane occupancy standard deviation during hour i
X_{7i}	Predictor	Before/After	Binary	The time period indicator for hour i. 1 – Hour occurring 2011-2013; 0 – Hour occurring 2006-2008.
X_{8i}	Predictor	PDSL Closed	Continuous	The proportion of the duration of “PDSL Closed” status during hour i.
X_{9i}	Predictor	PDSL Open	Continuous	The proportion of the duration of “PDSL Open” status during hour i.
X_{10i}	Predictor	Sign Dark	Continuous	The proportion of the duration of “DARK” status during hour i.
X_{11i}	Predictor	VSA	Continuous	The proportion of the duration of “VSA” status during hour i.

For each section, N37, N38, and N40, MLE implemented in a statistical analysis package for the software R (R, 2015) was used to fit and evaluate logistic regression models containing different combinations of independent variables listed in **Table 5. 6**.

Table 5. 7 is the estimation summary for the initial logistic regression model containing predictors X_1-X_4 and X_6-X_7.

Table 5. 7 Estimation Summary for Initial Model of Rear-End Crash Probability on I-35W PDSL Region

Section No.	Variable	Estimate	Std. Error	z value	Pr (>|z|)	Signif. codes
N37	Constant	-7.288	0.2738	-26.616	< 2E-16	***
	Rainy	-0.02299	0.3566	-0.064	0.949	
	Snowy	0.04992	0.4733	0.105	0.916	
	logvph	-0.000050	0.000941	-0.053	0.958	
	Lane Occupancy	0.1451	0.03245	4.471	7.78E-06	***
	Lane Occupancy^2	--	--	--	--	
	Occupancy Standard Deviation	0.05499	0.06699	0.821	0.412	
	Before/After	0.06369	0.3537	0.180	0.857	
	Null deviance = 1075.53			H-L = 13.9621, p-value = 0.08276		
	Residual deviance = 873.95			AIC: 887.95		
N38	Constant	-21.31813	4.48050	-4.758	1.96E-06	***
	Rainy	0.30636	0.32547	0.941	0.34657	
	Snowy	-0.12390	0.52151	-0.238	0.81221	
	logvph	1.64014	0.52621	3.117	0.00183	**
	Lane Occupancy	0.08761	0.03530	2.482	0.01307	*
	Lane Occupancy^2	--	--	--	--	
	Occupancy Standard Deviation	0.08820	0.05542	1.591	0.11151	
	Before/After	0.43536	0.31001	1.404	0.16021	
	Null deviance = 1178.50			H-L = 20.077, p-value = 0.01005		
	Residual deviance = 957.45			AIC: 971.45		
N40	Constant	-6.1828825	0.1505737	-41.062	< 2E-16	***
	Rainy	-0.2567038	0.2407647	-1.066	0.2863	
	Snowy	0.2053577	0.2950196	0.696	0.4864	
	logvph	-0.0000426	0.0006195	-0.069	0.9452	
	Lane Occupancy	0.0370454	0.0207568	1.785	0.0743	.
	Lane Occupancy^2	--	--	--	--	
	Occupancy Standard Deviation	0.1629448	0.0317275	5.136	2.81E-07	***
	Before/After	0.0840515	0.1869440	0.450	0.6530	
	Null deviance = 2437.0			H-L = 52.2, p-value = 1.54E-08		
	Residual deviance = 2069.1			AIC: 2083.1		

(Signif. codes: 0 '*' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1)**

N37 was taken as an example to see how to interpret the information provided in **Table 5. 7**:

i) Intercept β_0

The maximum likelihood estimate of the coefficient β_0 is equal to -7.288 and the standard error associated with this estimate was 0.2378. A test of the null hypothesis $\beta_0 = 0$ yielded a z-statistic equal to -26.616, and the probability of obtaining a z-value at least this large when the null hypothesis was true was less than 0.001. Using the rule that we cannot reject a null hypothesis when a P-value is greater than 0.05, which indicates that β_0 is statistically significant different from 0 at 0.05 significance level. The 95% Wald confidence interval of β_0 is (-7.7541, -6.8219), given by (-7.288-1.96×0.2378, -7.288+1.96×0.2378), indicating β_0 is statistically significantly different from 0 at 0.05 significance level as the interval does not contain 0, which is in consistency with the conclusion drawn from the p-value. In a situation where all model predictors were equal to zero the estimated probability of a rear-end crash would be exp(-7.288)/(1+exp(-7.288))=0.00068 with a 95% Wald confidence interval (0.0004, 0.0011), given by (exp (-7.7541) / (1+exp (-7.7541), exp (-6.8219) / (1+exp (-6.8219).

ii) Slopes $\beta_{k,k\neq0}$

Table 5. 7 shows that, for section N37, the maximum likelihood estimate of the coefficient β_1, associated with the Rainy condition, was -0.02299 and the standard error associated with this estimate was 0.3566. A test of the null hypothesis $\beta_1 = 0$ yielded a z-statistic equal to -0.064, and the probability of obtaining a z-value at least this large when the null hypothesis was true was 0.949. Using the rule that we cannot reject a null hypothesis when a P-value is greater than 0.05, which indicates that rainy condition does not have statistically significant association with rear-end crash probability at 0.05 significance level. Similar conclusions apply to weather variable Snowy, the traffic variables Log Flow, Occupancy Standard Deviation, and the Before/After predictor. Those five predictors showed no clear association with rear-end crash risk at 0.05 significance level.

On the other hand, the maximum likelihood estimate of the coefficient β_4 was 0.1451 and the standard error associated with this estimate was 0.03245. A test of the null hypothesis $\beta_4 = 0$ yielded a z-statistic equal to 4.471, and the probability of obtaining a z-value at least this large when the null hypothesis was true was essentially zero to five decimal places.

The association that lane occupancy has with rear-end crash risk is significant at 0.05 significance level. The positive value of estimate of β_4 indicates that rear-end crash probability increases as lane occupancy increases. The odds of having a rear-end crash in a given hour is $e^{0.1451} = 1.1562$ times the odds of having a rear-end crash in a given hour with one-unit decrease in standard deviation of lane occupancy with a 95% Wald confidence interval (1.0849, 1.2321), holding other conditions constant.

iii) Goodness of fit

The null deviance and residual deviance are 1075.53 and 873.95, respectively. The Hosmer-Lemeshow goodness of fit statistic was 13.9621 with a p-value of 0.08276, which indicates the actual and predicted event rates are similar across 10 deciles at 0.05 significance level.

Similar to what was done for sections in I-35W from TH-13 to I-494, the relationship between lane occupancy and crash probability could show an "inverted-U" shape with a maximal point where crash probability is a highest and falling off for lane occupancies both less than and greater than this maximal point, the square of lane occupancy, variable X_5, was included as a predictor. **Table 5. 8** shows the estimation summary for a model using all the predictors listed in **Table 5. 6**.

Table 5. 8 Estimation Summary for Initial Model of Rear-End Crash Probability on I-35W PDSL region, with Quadratic Occupancy Effect

Section No.	Variable	Estimate	Std. Error	z value	Pr (>|z|)	Signif. codes
N37	Constant	-5.993853	4.975980	-1.205	0.2284	
	Rainy	0.068005	0.354964	0.192	0.8481	
	Snowy	-0.023336	0.479197	-0.049	0.9612	
	logvph	-0.203466	0.598580	-0.340	0.7339	
	Lane Occupancy	0.459360	0.086809	5.292	1.21E-07	***
	Lane Occupancy^2	-0.009862	0.002374	-4.155	3.25E-05	***
	Occupancy Standard Deviation	-0.010092	0.071139	-0.142	0.8872	
	Before/After	-0.669111	0.371041	-1.803	0.0713	.
	Null deviance = 1075.53			H-L = 9.4133, p-value = 0.3086		
	Residual deviance = 833.71			AIC: 849.71		
N38	Constant	-3.457166	4.683427	-0.738	0.460411	
	Rainy	0.239999	0.322963	0.743	0.457411	
	Snowy	-0.287604	0.523075	-0.550	0.582434	
	logvph	-0.562941	0.574510	-0.980	0.327154	
	Lane Occupancy	0.450058	0.099456	4.525	6.03E-06	***
	Lane Occupancy^2	-0.010118	0.002764	-3.661	0.000252	***
	Occupancy Standard Deviation	-0.055956	0.063082	-0.887	0.375062	
	Before/After	0.546829	0.295886	1.848	0.064587	.
	Null deviance = 1178.50			H-L = 3.9974, p-value = 0.8574		
	Residual deviance = 944.62			AIC: 960.62		
N40	Constant	-6.492	0.1806	-35.945	< 2e-16	***
	Rainy	-0.1452	0.2394	-0.606	0.544275	
	Snowy	0.2933	0.2946	0.996	0.319424	
	logvph	-0.000095	0.006008	-0.016	0.987331	
	Lane Occupancy	0.2534	0.03659	6.925	4.35E-12	***
	Lane Occupancy^2	-0.007953	0.001133	-7.022	2.19E-12	***
	Occupancy Standard Deviation	0.1102	0.03066	3.594	0.000326	***
	Before/After	0.2181	0.1809	0.121	0.904052	
	Null deviance = 2437.0			H-L = 13.899, p-value = 0.08444		
	Residual deviance = 2003.6			AIC: 2019.6		

(Signif. codes: 0 '*' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1)**

The changes in rear-end crash risk associated with UPA project were non-significant for all three sections. For N37, rear-end crash risk might have decreased after the UPA project holding other conditions constant, while for N38, rear-end crash risk might have increased after UPA project holding other conditions constant. Weather effect showed significant effect on rear-end crash risk on none of the three analyzed sections in PDSL region at 0.05 significance level.

L-R tests were conducted to determine the better model between the reduced model M_0, the model without quadratic term of lane occupancy, and the more complex model M_1, the model with quadratic term of lane occupancy, for each analyzed section in I-35W PDSL region. **Table 5. 9** shows the L-R test results.

Table 5. 9 Likelihood Ratio Test Results for Analyzed Sections in I-35W PDSL Region

<table>
<tr><th>Section No.</th><th>Model</th><th>Number of coefficients</th><th>Log-likelihood</th><th>p</th><th>$G^2(M_0|M_1)$</th><th>P-value</th><th>Significance</th></tr>
<tr><td rowspan="2">N37</td><td>M_0</td><td>7</td><td>-436.98</td><td rowspan="2">1</td><td rowspan="2">40.246</td><td rowspan="2">2.24E-10</td><td rowspan="2">***</td></tr>
<tr><td>M_1</td><td>8</td><td>-416.85</td></tr>
<tr><td rowspan="2">N38</td><td>M_0</td><td>7</td><td>-478.73</td><td rowspan="2">1</td><td rowspan="2">12.835</td><td rowspan="2">3.40E-04</td><td rowspan="2">***</td></tr>
<tr><td>M_1</td><td>8</td><td>-472.31</td></tr>
<tr><td rowspan="2">N40</td><td>M_0</td><td>7</td><td>-1034.5</td><td rowspan="2">1</td><td rowspan="2">65.49</td><td rowspan="2">5.84E-16</td><td rowspan="2">***</td></tr>
<tr><td>M_1</td><td>8</td><td>-1001.8</td></tr>
</table>

According to the L-R test results, we reject the null hypothesis that the reduced model (without quadratic term) holds and conclude that adding the quadratic term has significantly improved model fit for all three of the PDSL sections at 0.05 significance level.

As is shown in **Table 5. 9**, the two coefficients associated with lane occupancy, β_4 and β_5, are both significantly different from zero while their signs $\beta_4 > 0$ with $\beta_5 < 0$, imply that, for section N37, N38, and N40, rear-end crash probability does have the inverted-U shape. The PDSL does not operate continuously but only when opened by the traffic managers, and the rest of the time the PDSL functions as a shoulder. This means that the After period contained some hours when the PDSL was operating and other hours when it was not. To more clearly isolate the effect of the PDSL operation, the sign logs from MnDOT's Regional Traffic Management Center were reviewed to determine PDSL status, and the After period was divided into four mutually exclusive subsets reflecting PDSL status. An estimation summary for the model with this subdivided After period is shown in **Table 5. 10**. These results are generally similar to those shown in **Table 5. 8** except that in N37 hours when the PDSL was open show a statistically significant reduction in rear-end crash probability and the marginally non-significant UPA project effect in N38 is no longer significant at 0.05 significance level.

Table 5. 10 Estimation Summary for N37, N38, and N40, with PDSL After Period Subdivided According to PDSL Status

Section No.	Variable	Estimate	Std. Error	z value	Pr (>|z|)	Signif. codes
N37	Constant	-7.728	0.3186	-24.259	< 2E-16	***
	Rainy	0.08918	0.3544	0.252	0.8013	
	Snowy	-0.00922	0.477	-0.019	0.9846	
	logvph	-0.000084	0.006789	-0.012	0.9901	
	Lane Occupancy	0.546	0.06441	7.057	1.70E-12	***
	Lane Occupancy^2	-0.009742	0.001833	-5.314	1.07E-07	***
	Occupancy Standard Deviation	-0.003018	0.06474	-0.047	0.9628	
	Before/After	--	--	--	--	
	PDSL Closed	-0.3042	0.5364	-0.567	0.5706	
	PDSL Open	-0.7873	0.3813	-2.065	0.0389	*
	Sign Dark	-0.3658	1.148	-0.319	0.7500	
	VSA	-0.8552	1.568	-0.545	0.5855	
	Null deviance = 1075.53			H-L = 7.8064, p-value = 0.4526		
	Residual deviance = 832.88			AIC: 854.88		
N38	Constant	-4.329209	5.476106	-0.791	0.429199	
	Rainy	0.260577	0.324010	0.804	0.421266	
	Snowy	-0.323182	0.531028	-0.609	0.542792	
	logvph	-0.460368	0.666513	-0.691	0.489747	
	Lane Occupancy	0.442352	0.104363	4.239	2.25E-05	***
	Lane Occupancy^2	-0.009983	0.002875	-3.472	0.000517	***
	Occupancy Standard Deviation	-0.046633	0.066593	-0.700	0.483761	
	Before/After	--	--	--	--	
	PDSL Closed	0.709386	0.482127	1.471	0.141192	
	PDSL Open	0.474689	0.321169	1.478	0.139406	
	Sign Dark	1.288526	0.978760	1.316	0.188010	
	VSA	1.089694	1.051170	1.037	0.299900	
	Null deviance = 1178.50			H-L = 4.3466, p-value = 0.8246		
	Residual deviance = 943.48			AIC: 965.48		
N40	Constant	-6.555	0.1821	-35.987	< 2E-16	***
	Rainy	-0.1391	0.2394	-0.581	0.561289	
	Snowy	0.2837	0.2981	0.952	0.341204	
	logvph	-0.000056	0.002604	-0.021	0.982884	
	Lane Occupancy	0.2669	0.03747	7.124	1.05E-12	***
	Lane Occupancy^2	-0.008277	0.001154	-7.171	7.46E-13	***
	Occupancy Standard Deviation	0.1118	0.03051	3.663	0.000249	***
	Before/After	--	--	--	--	
	PDSL Closed	0.4127	0.2814	1.467	0.142449	
	PDSL Open	-0.08432	0.1932	-0.437	0.662434	

	Sign Dark	-3.585	4.270	-0.840	0.401118	
	VSA	0.1603	0.7982	0.201	0.840889	
	Null deviance = 2437.0			H-L = 7.4029, p-value = 0.4939		
	Residual deviance = 1998.7			AIC: 2020.7		

(Signif. codes: 0 '*' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1)**

Table 5. 11 shows a summary for the model using only the statistically-significant predictors from **Table 5. 10**.

Table 5. 11 Estimation Summary for Analyzed PDSL regions for Model Using Only Statistically Significant Predictors from Table 5.8

Section No.	Variable	Estimate	Std. Error	z value	Pr (>|z|)	Signif. codes
N37	Constant	-7.791773	0.267902	-29.084	< 2E-16	***
	Lane Occupancy	0.444544	0.051367	8.654	< 2E-16	***
	Lane Occupancy^2	-0.009567	0.001802	-5.309	1.1E-07	***
	PDSL Open	-0.635194	0.319932	-1.985	0.0471	*
	Null deviance = 1075.53			H-L = 7.0154, p-value = 0.535		
	Residual deviance = 833.74			AIC: 841.74		
N38	Constant	-7.769015	0.258508	-30.053	< 2E-16	
	Lane Occupancy	0.374028	0.045492	8.222	< 2E-16	***
	Lane Occupancy^2	-0.008324	0.001766	-4.714	2.43E-06	***
	Null deviance = 1178.50			H-L = 10.235, p-value = 0.2489		
	Residual deviance = 949.32			AIC: 955.32		
N40	Constant	-6.489315	0.142749	-45.460	< 2E-16	***
	Lane Occupancy	0.256277	0.036566	7.009	2.41E-12	***
	Lane Occupancy^2	-0.008003	0.001134	-7.054	1.74E-12	***
	Occupancy Standard Deviation	0.108783	0.030181	3.604	0.000313	***
	Null deviance = 2437.0			H-L = 11.553, p-value = 0.1723		
	Residual deviance = 2005.1			AIC: 2013.1		

(Signif. codes: 0 '*' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1)**

Section N37 showed a significant negative association between the operation of PDSL and rear-end crash risk. For N37, the odds of having a rear-end crash in a given hour when PDSL was open in After period is $e^{-0.635194} = 0.530$ times the odds of having a rear-end crash in a given hour in Before period with 95% Wald confidence interval (0.2830, 0.9919), holding other conditions constant. However, unlike section N37, neither N38 nor N40 showed a significant effect of PDSL operation. In section N40, the standard deviation of lane occupancy also showed a definite association with rear-end crash probability, with increases in this standard deviation implying increases in crash probability.

All three sections, N37, N38, and N40 showed the "inverted-U" relationship between rear-end crash probability and lane occupancy. As noted earlier, $\beta_4 > 0$ with $\beta_5 < 0$ imply an "inverted-U" shape to the graph of rear-end crash probability versus lane occupancy. Using the estimates in **Table 5. 11**, rear-end crash probability for N37, N38, and N40 was maximized when lane occupancy is approximately equal to

i) N37: $o_{max} \approx \bar{o} - \frac{\widehat{\beta_1}}{2\widehat{\beta_2}} = 6.22 - \frac{0.444544}{2(-.009567)} = 29.45$

ii) N38: $o_{max} \approx \bar{o} - \frac{\widehat{\beta_1}}{2\widehat{\beta_2}} = 7.26 - \frac{0.374028}{2(-0.008324)} = 29.73$

iii) N40: $o_{max} \approx \bar{o} - \frac{\widehat{\beta_1}}{2\widehat{\beta_2}} = 7.43 - \frac{0.256277}{2(-0.008003)} = 23.44$

Where $\bar{o}$ denotes the average lane occupancy over all hours in the Before and After period for N37, N38, and N40, respectively.

Figure 5. 5 to **Figure 5. 7** show time-series plots of average lane occupancy in sections N37, N38, and N40 for each hour during both the Before and After periods, with the vertical lines denoting the change points. The horizontal lines show the approximate values of average lane occupancy where rear-end crash risk is maximal. As can be seen, during the After period there was a substantial increase in lane occupancy values in the region of maximum rear-end crash probability for all three sections.

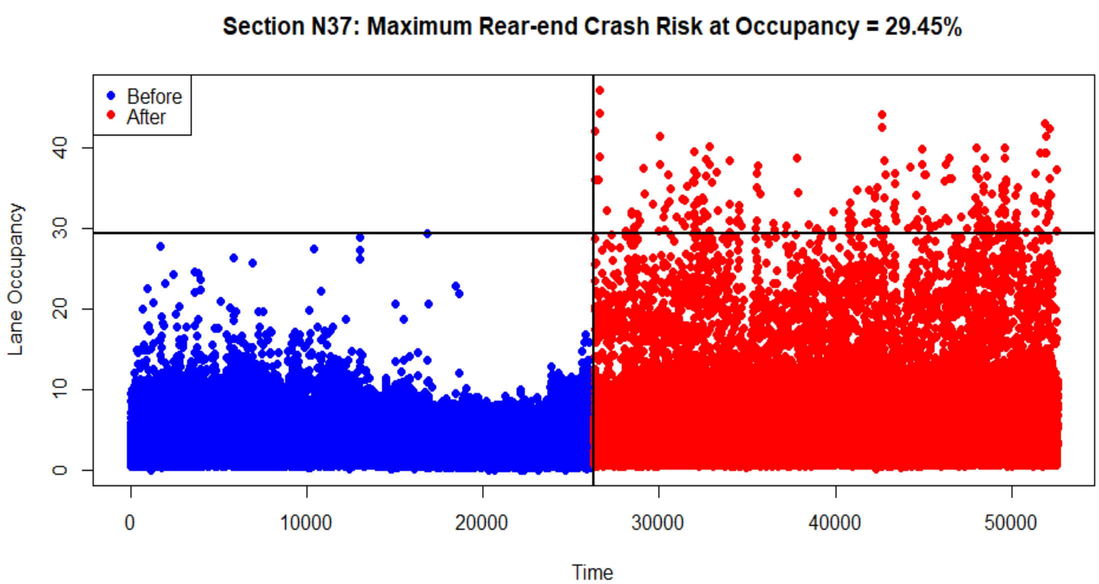

Figure 5. 5 Time-series plot of hourly average lane occupancy for section N37, showing the Before and After PDSL periods, and the average lane occupancy with maximal crash probability.

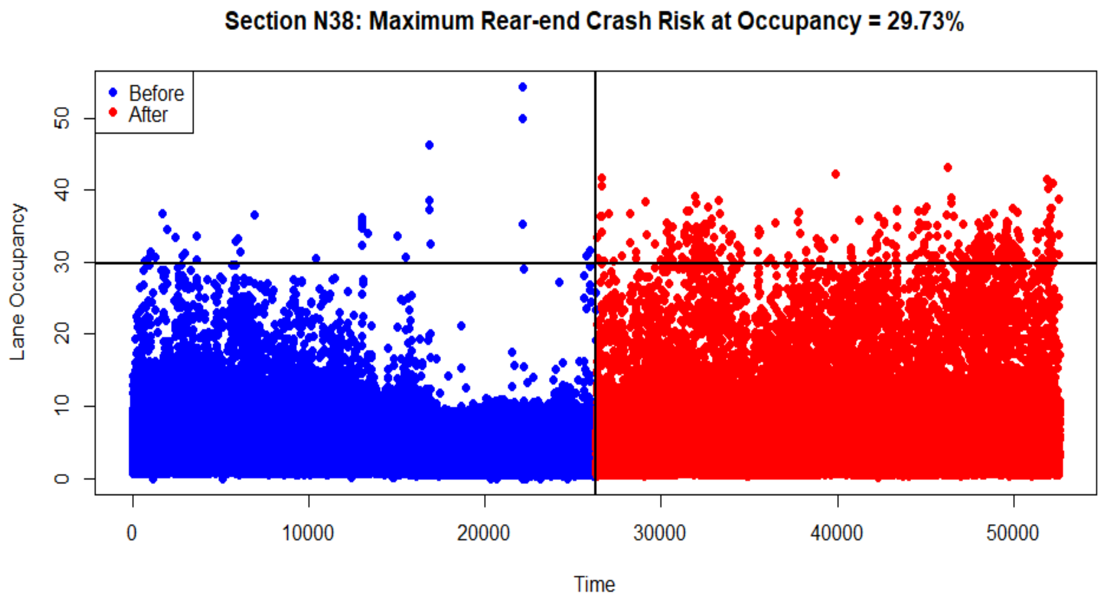

Figure 5. 6 Time-series plot of hourly average lane occupancy for section N38, showing the Before and After PDSL periods, and the average lane occupancy with maximal crash probability.

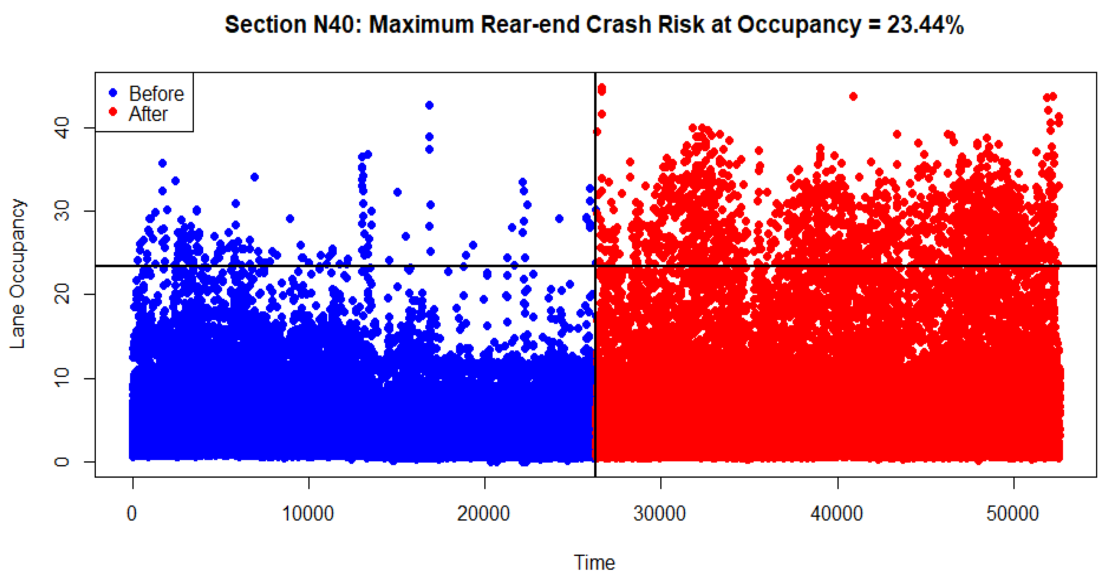

Figure 5. 7 Time-series plot of hourly average lane occupancy for section N40, showing the Before and After PDSL periods, and the average lane occupancy with maximal crash probability.

In summary, all three analyzed sections of the I-35W PDSL region, N37, N38, and N40, showed substantial increases in lane occupancy following UPA project. The observed increases in rear-end crash frequency can be explained by increases in higher-risk traffic conditions. The increase in higher risk traffic conditions were most likely due to removal of I-35W & TH-62 bottleneck.

Chapter 6: Conclusion

The objective of this research was to assess safety-related impacts of Minnesota's UPA project implemented on I-35W. To be more specific, this study is aimed to untangle the indirect safety effects due to changes in traffic conditions from the direct effects, if any, due to the UPA interventions.

A preliminary analysis was done to determine priority crash type and study regions. I-35W, from its start to its junction with I-94, was divided into 17 one-mile sections, and bi-directional (northbound and southbound) crash frequencies in Before-UPA (2006-2008) and After-UPA periods (2011-2013) were compiled for each one-mile section. Rear-end crash turned out to be the most prevalent crash type, and I-35W HOT region (from TH-13 to I-494) and the I-35W PDSL region (from 37th Street to 26th Street) where there was an outstanding increase in the rear-end crash frequency in the After period became our analysis regions.

The I-35W HOT region and the PDSL region were divided into analysis sections based on constant flow and geometry criteria as well as the availability of loop-detector data. Rear-end crash, traffic condition, weather condition, and PDSL historical operation data for Before and After periods were compiled for each analysis section.

The literature review revealed that there was a substantial portion of the research focused on empirical models for both long-term and short-term crash risk prediction, but less on why crashes come about. Based on limited clues in previous studies related to freeway crash mechanism, combined with this project's data availability and quality reality, a logistic regression model was established to estimate the change in rear-end crash risk in a given hour Before versus After period in different analysis sections of I-35W, controlling for the changes in traffic conditions and weather conditions. For each analysis section, for each hour during Before and After periods, the presence or absence of a rear-end crash became the response variable while predictor variables consisted of traffic volume and lane occupancy, the presence or absence of snowy or rainy conditions, before versus after the UPA project, and the presence or absence of PDSL operation.

According to the results of logistic regression analyses, most sections showed no significant change in rear-end crash risk associated with UPA project except for section southbound S9 (just north of Minnesota River). The rear-end crash risk in a given hour within section S9 was a complicated function of UPA intervention, traffic condition, and weather condition. Although section N17 experienced fewer rear-end crashes in After period, the change in lane occupancy itself could not explain such reduction. The PDSL region of I-35W experienced substantial increases in both rear-end crash frequency and traffic congestion (defined as average lane occupancies exceeding 25%) in After period. Rainy condition, snowy condition, and the operation of the PDSL had no direct effect on the likelihood of rear-end crashes when controlling for changes in traffic conditions. In other words, the observed change in crash frequency was explained by the change in traffic conditions. It appeared the operation of PDSL was coincident with removal of the upstream bottleneck in the old I-35W & TH-62 commons and moving the bottleneck northward to the I-35W & I-94 junction.

In addition, this study found evidence for a nonlinear relationship between a proxy for traffic density, lane occupancy, and the probability a rear-end crash occurs during an hour, controlling for traffic volume, weather, and geometry. Rear-end crashes were most likely when lane occupancies were approximately 20%-30%, and crash likelihood tended to decrease for lane occupancies below and above this range.

Limitations did exist in this study, and future research may need to address following issues:

- There were ambiguities in crash locations in crash records and reports reviewed, and crashes with ambiguous crash locations could only be excluded from analyses, which may affect the analysis results.
- Analyses of only a limited number of sections could be conducted due to the availability of loop detectors or the sufficiency of the number of crashes in Before and After periods.
- This study only focused on the dominant crash type, rear-end crashes. For different types of crashes, the relationship between real-time traffic conditions (or weather conditions) and crash risk may be different since they may have different crash mechanisms. However, the methodology demonstrated in this study could be applied and similar analyses could be done for other types of crashes.

In spite of these limitations, this study demonstrated a methodology that could be applied to the evaluation of the safety effects of freeway-related projects. To be more specific, this study worked out a way to estimate changes in hourly crash risk while controlling for variations in traffic conditions.

Bibliography

AASHTO (2010). *Highway Safety Manual, 1st edition*. American Association of State Highways and Transportation Officials Transportation Research Board of the National Academies, Washington, DC.

Abdel-Aty, M., Uddin, N., Pande, A., Abdalla, M.F., Hsia, L. (2004). Predicting freeway crashes from loop detector data by matched case–control logistic regression. *Transportation Research Record: Journal of the Transportation Research Board*, 1897, 88–95.

Abdel-Aty, M., Uddin, N., and Pande, A. (2005). Split models for predicting multivehicle crashes during high-speed and low-speed operating conditions on freeways. *Transportation Research Record: Journal of the Transportation Research Board*, 1908, 51-58.

Abdel-Aty, M. A., and Pemmanaboina, R. (2006). Calibrating a real-time traffic crash-prediction model using archived weather and ITS traffic data. IEEE Transactions on Intelligent Transportation Systems, 7(2), 167-174.

Andrey, J., and Yagar, S. (1993). A temporal analysis of rain-related crash risk. *Accident Analysis and Prevention*, 25(4), 465-472.

Andrey, J. (2010). Long-term trends in weather-related crash risks. *Journal of Transport Geography*, 18(2), 247-258.

Bauer, K., Harwood, D., Hughes, W., and Richard, K. (2004). Safety effects of narrow lanes and shoulder-use lanes to increase capacity of urban freeways. *Transportation Research Record: Journal of the Transportation Research Board*, (1897), 71-80.

Brijs, T., Karlis, D., and Wets, G. (2008). Studying the effect of weather conditions on daily crash counts using a discrete time-series model. *Accident Analysis and Prevention*, 40(3), 1180-1190.

Brill, E (1972). A car-following model relating reaction times and temporal headways to accident frequency. *Transportation Science*, 6, 343-353.

Cao, X. J., Munnich, L., Liu, H., He, X. S., Xu, Z., and Huang, Y. A. (2012). Benefit and Cost Analysis of the I-394 MnPASS Program.

Ceder, A., and Livneh, M. (1982). Relationships between road accidents and hourly traffic flow—I: analyses and interpretation. *Accident Analysis and Prevention*, 14(1), 19-34.

Chatterjee, I. (2016). *Understanding Driver Contributions to Rear-End Crashes on Congested Freeways and Their Implication for Future Safety Measures*. Ph.D. Thesis, University of Minnesota.

Chatterjee, I., and Davis, G. A. (2016). Analysis of Rear-End Events on Congested Freeways by Using Video-Recorded Shock Waves. *Transportation Research Record: Journal of the Transportation Research Board*, 2583, 110-118.

Curren, J. E. (1995). Use of shoulders and narrow lanes to increase freeway capacity (Vol. 369). Transportation Research Board.

Davis, G. A., & Swenson, T. (2006). Collective responsibility for freeway rear-ending accidents?: An application of probabilistic causal models. *Accident Analysis & Prevention*, 38(4), 728-736.

Davis, G. A., Moshtagh, V., & Hourdos, J. (2016). Safety-Related Guidelines for Time-of-Day Changes in Left-Turn Phasing. *Transportation Research Record: Journal of the Transportation Research Board*, (2557), 100-107.

Dobson, A. J., and Barnett, A. (2008). *An introduction to generalized linear models*. CRC press.

FHWA Road Weather Management. Retrieved from https://ops.fhwa.dot.gov/weather/index.asp.

FHWA UPA. Retrieved from https://ops.fhwa.dot.gov/congestionpricing/urb_partner_agree.htm.

Eisenberg, D., and Warner, K. E. (2005). Effects of snowfalls on motor vehicle collisions, injuries, and fatalities. *American journal of public health*, 95(1), 120-124.

Golob, T.F., Recker, W.W. (2003). Relationships among urban freeway accidents, traffic flow, weather, and lighting conditions. *Journal of Transportation Engineering* 129 (4), 342–353.

Golob, T.F., Recker, W.W., Alvarez, V.M. (2004). Freeway safety as a function of traffic flow. *Accident Analysis and Prevention*, 36 (6), 933–946.

Geedipally, S. R., Lord, D., and Dhavala, S. S. (2012). The negative binomial-Lindley generalized linear model: Characteristics and application using crash data. *Accident Analysis and Prevention*, 45, 258-265.

Hall, J. W., and Pendleton, O. J. (1989). *Relationship between V/C ratios and Accident Rates*.

Hosmer, D., and Lemeshow, S. (2000). *Applied Logistic Regression, 2nd edition*, Wiley.

Hossain, M., and Muromachi, Y. (2012). A Bayesian network based framework for real-time crash prediction on the basic freeway sections of urban expressways. *Accident Analysis and Prevention*, 45, 373-381.

Hourdos, J. (2005). *Crash Prone Traffic Flow Dynamics: Identification and Real-Time Detection*. Ph.D. Thesis, University of Minnesota.

Hourdos, J., Garg, V., Michalopoulos, P., and Davis, G. (2006). Real-time detection of crash-prone conditions at freeway high-crash locations. *Transportation Research Record: Journal of the Transportation Research Board*, 1968, 83-91.

Ivey, D. L., Griffin, L. I., Newton, T. M., Lytton, R. L., and Hankins, K. C. (1981). Predicting wet weather accidents. *Accident Analysis and Prevention*, 13(2), 83-99.

Jovanis, P. P., and Chang, H. L. (1986). Modeling the relationship of accidents to miles traveled. *Transportation Research Record: Journal of the Transportation Research Board*, 1068, 42-51.

Khattak, A., Kantor, P., and Council, F. (1998). Role of adverse weather in key crash types on limited-access: roadways implications for advanced weather systems. *Transportation Research Record: Journal of the Transportation Research Board*, (1621), 10-19.

Knapp, K. K., Giese, K., and Kroeger, D. (2000). *Mobility and Safety Impacts of Winter Storm Events in a Freeway Environment. Final Report.*

Lee, C., Saccomanno, F., and Hellinga, B. (2002). Analysis of crash precursors on instrumented freeways, *Transportation Research Record: Journal of the Transportation Research Board*, (1784), 1-8.

Liu, G., and Popoff, A. L. (1997). Provincial-wide travel speed and traffic safety study in Saskatchewan. *Transportation Research Record: Journal of the Transportation Research Board*, (1595), 8-13.

McCasland, W. R. (1984). *The Use of Freeway Shoulders to Increase Capacity: A Review*. Texas Transportation Institute, Texas A&M University System.

Mensah, A., and Hauer, E. (1998). Two problems of averaging arising in the estimation of the relationship between accidents and traffic flow. *Transportation Research Record: Journal of the Transportation Research Board*, (1635), 37-43.

Miaou, S. P. (1994). The relationship between truck accidents and geometric design of road sections: Poisson versus negative binomial regressions. *Accident Analysis and Prevention*, *26*(4), 471-482.

MNDOT (2011). *UPA/CRD Annual Report Minnesota: Innovative Choices for Congestion Relief (FHWA-JPO-11-039).*

Newell, G. F. (1993). A simplified theory of kinematic waves in highway traffic, part I: General theory. *Transportation Research Part B: Methodological*, 27(4), 281-287.

Oh, C., Oh, J-S, Ritchie, S., and Chang, M. (2001). Real-time estimation of freeway 6 accident likelihood. Paper 01-3445 presented at 2001 Transportation Research Board Annual Meeting, Washington, DC.

Roshandel, S., Zheng, Z., and Washington, S. (2015). Impact of real-time traffic characteristics on freeway crash occurrence: Systematic review and meta-analysis. *Accident Analysis and Prevention*, 79, 198-211.

Shankar, V., Mannering, F., and Barfield, W. (1995). Effect of roadway geometrics and environmental factors on rural freeway accident frequencies. *Accident Analysis and Prevention*, 27(3), 371-389.

Team, R. (2015). RStudio: integrated development for R. *RStudio, Inc., Boston, MA. URL http://www.rstudio.com*.

Turnbull, K., Balke, K., Burris, M., Songchitruksa, P., Park, E. Pessaro, B., Samus, J., Saunoi-Sandgren, E., Gopalakrishna, D., Schroeder J., Zimmerman, C., Schreffler, E., Joy, B. (2013). *Urban Partnership Agreement: Minnesota Evaluation Report (FHWA-JPO-13- 048)*.

Xu, C., Liu, P., Wang, W., Li, Z. (2012). Evaluation of the impacts of traffic states on crash risks on freeways. *Accident Analysis and Prevention*, 47 (1), 162–171.

Xu, C., Liu, P., Wang, W., and Li, Z. (2014). Identification of freeway crash-prone traffic conditions for traffic flow at different levels of service. *Transportation research part A: policy and practice*, 69, 58-70.

Zha, L., Lord, D., and Zou, Y. (2016). The Poisson inverse Gaussian (PIG) generalized linear regression model for analyzing motor vehicle crash data. *Journal of Transportation Safety and Security*, 8(1), 18-35.

Zheng, Z., Ahn, S., and Monsere, C. M. (2010). Impact of traffic oscillations on freeway crash occurrences. *Accident Analysis and Prevention*, 42(2), 626-636.

Appendix

Preliminary Analysis Summary

Table A. 1 Crash Summary Table for Section Mile-1

Crash Code	Crash Type	Crash Frequency	
		Before	After
0	Unspecified	0	0
1	Rear end	31	38
2	Sideswipe-Same direction	15	15
3	Left turn	1	0
4	Ran off road-Left side	16	7
5	Right angle	1	3
6	Right turn	0	0
7	Ran off road-Right side	5	13
8	Head on	2	0
9	Sideswipe-Opposing	1	1
90	Other	8	6
98	Not applicable	1	0
99	Unknow	1	0

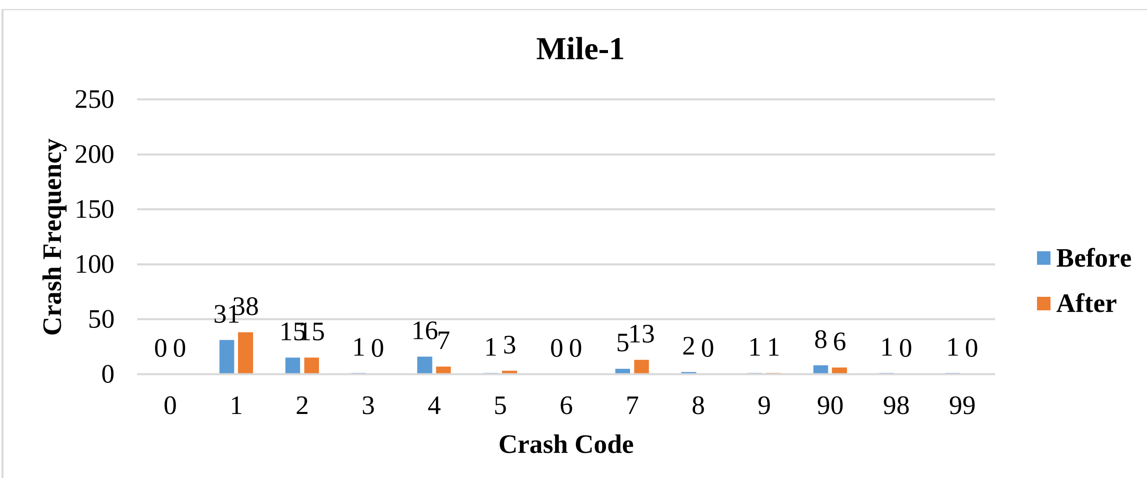

Figure A. 1 Crash Summary Histogram for Section Mile-1

Table A. 2 Crash Summary Table for Section Mile-2

Crash Code	Crash Type	Crash Frequency	
		Before	After
0	Unspecified	0	0
1	Rear end	16	16
2	Sideswipe-Same direction	5	7
3	Left turn	0	0
4	Ran off road-Left side	9	8
5	Right angle	0	1
6	Right turn	0	0
7	Ran off road-Right side	3	12
8	Head on	0	1
9	Sideswipe-Opposing	0	0
90	Other	0	2
98	Not applicable	0	0
99	Unknow	0	0

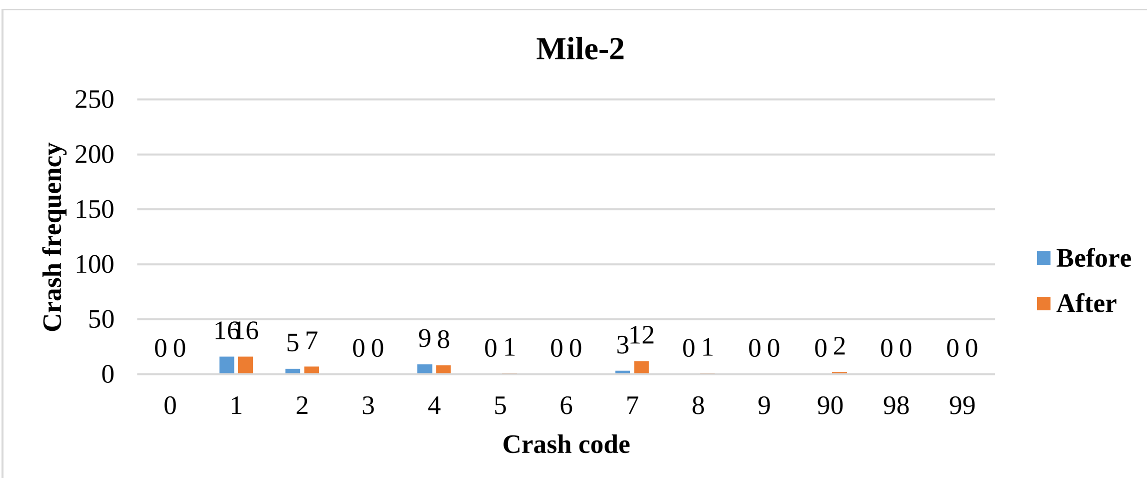

Figure A. 2 Crash Summary Histogram for Section Mile-2

Table A. 3 Crash Summary Table for Section Mile-3

Crash Code	Crash Type	Crash Frequency	
		Before	After
0	Unspecified	1	1
1	Rear end	45	92
2	Sideswipe-Same direction	26	34
3	Left turn	0	0
4	Ran off road-Left side	24	21
5	Right angle	1	3
6	Right turn	0	0
7	Ran off road-Right side	16	10
8	Head on	2	0
9	Sideswipe-Opposing	0	2
90	Other	17	10
98	Not applicable	1	1
99	Unknow	0	0

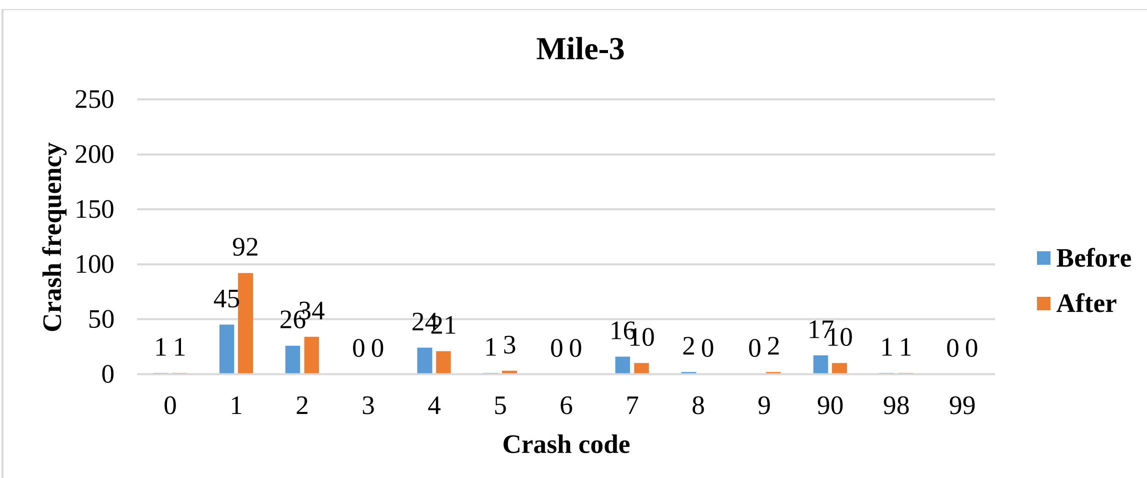

Figure A. 3 Crash Summary Histogram for Section Mile-3

Table A. 4 Crash Summary Table for Section Mile-4

Crash Code	Crash Type	Crash Frequency	
		Before	After
0	Unspecified	1	0
1	Rear end	42	73
2	Sideswipe-Same direction	22	20
3	Left turn	0	0
4	Ran off road-Left side	2	11
5	Right angle	4	4
6	Right turn	0	0
7	Ran off road-Right side	4	10
8	Head on	1	1
9	Sideswipe-Opposing	0	0
90	Other	13	10
98	Not applicable	1	0
99	Unknow	0	0

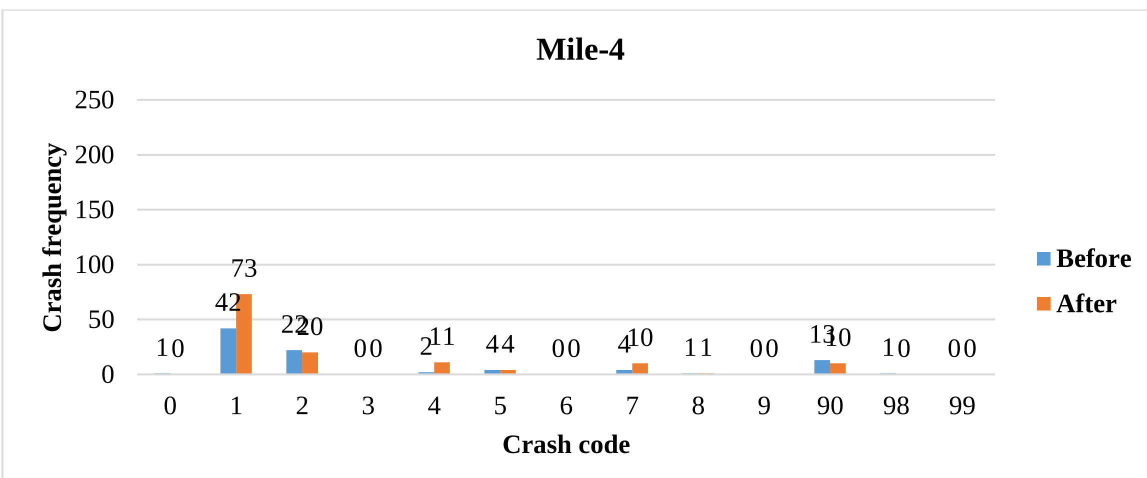

Figure A. 4 Crash Summary Histogram for Section Mile-4

Table A. 5 Crash Summary Table for Section Mile-5

Crash Code	Crash Type	Crash Frequency	
		Before	After
0	Unspecified	0	1
1	Rear end	29	47
2	Sideswipe-Same direction	14	15
3	Left turn	0	0
4	Ran off road-Left side	9	7
5	Right angle	3	5
6	Right turn	0	0
7	Ran off road-Right side	6	7
8	Head on	3	3
9	Sideswipe-Opposing	0	0
90	Other	11	5
98	Not applicable	1	0
99	Unknow	0	0

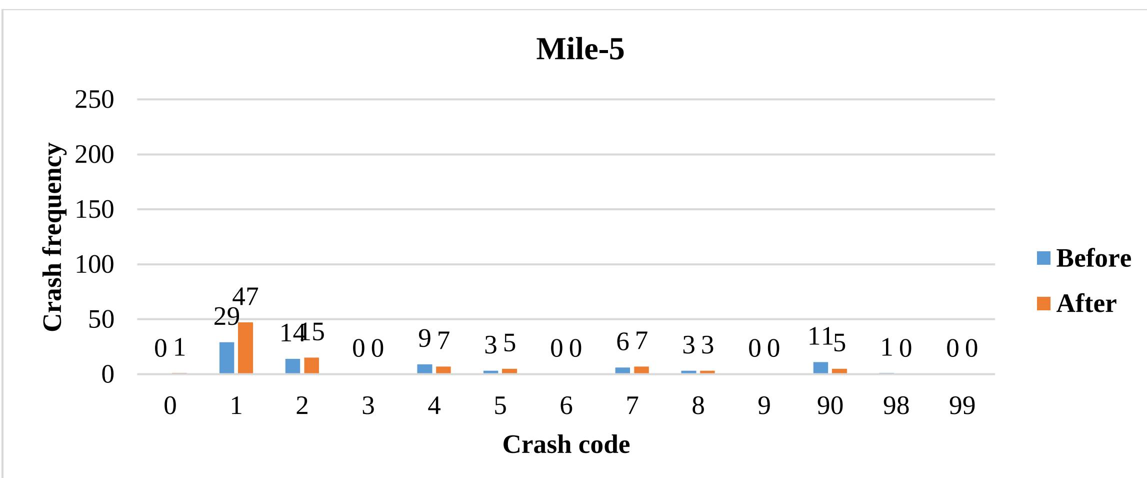

Figure A. 5 Crash Summary Histogram for Section Mile-5

Table A. 6 Crash Summary Table for Section Mile-6

Crash Code	Crash Type	Crash Frequency	
		Before	After
0	Unspecified	0	0
1	Rear end	40	33
2	Sideswipe-Same direction	16	10
3	Left turn	1	0
4	Ran off road-Left side	6	3
5	Right angle	8	2
6	Right turn	0	0
7	Ran off road-Right side	11	9
8	Head on	1	1
9	Sideswipe-Opposing	0	0
90	Other	5	5
98	Not applicable	2	1
99	Unknow	0	0

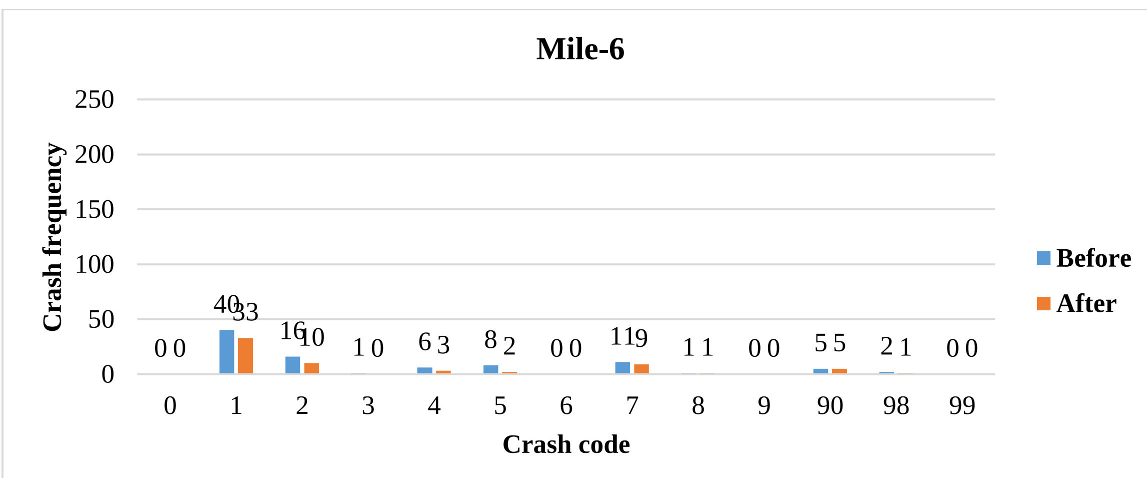

Figure A. 6 Crash Summary Histogram for Section Mile-6

Table A. 7 Crash Summary Table for Section Mile-7

Crash Code	Crash Type	Crash Frequency	
		Before	After
0	Unspecified	2	0
1	Rear end	90	114
2	Sideswipe-Same direction	20	28
3	Left turn	0	0
4	Ran off road-Left side	4	15
5	Right angle	7	5
6	Right turn	1	0
7	Ran off road-Right side	9	6
8	Head on	0	1
9	Sideswipe-Opposing	0	0
90	Other	10	4
98	Not applicable	1	0
99	Unknow	0	0

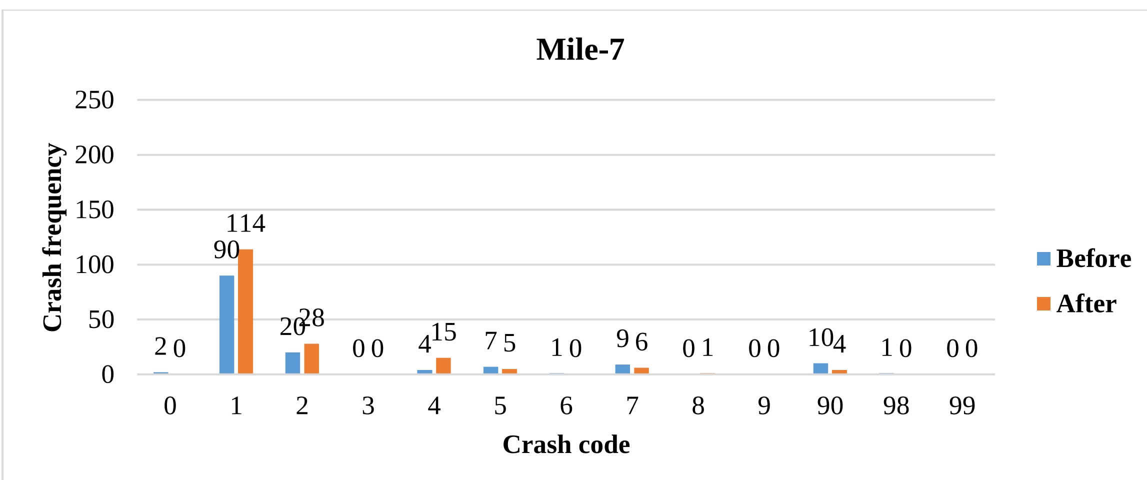

Figure A. 7 Crash Summary Histogram for Section Mile-7

Table A. 8 Crash Summary Table for Section Mile-8

Crash Code	Crash Type	Crash Frequency	
		Before	After
0	**Unspecified**	**1**	**0**
1	Rear end	70	67
2	Sideswipe-Same direction	12	15
3	Left turn	1	1
4	Ran off road-Left side	4	9
5	Right angle	1	5
6	Right turn	0	0
7	Ran off road-Right side	10	6
8	Head on	2	1
9	Sideswipe-Opposing	0	0
90	Other	11	7
98	Not applicable	1	0
99	Unknow	0	0

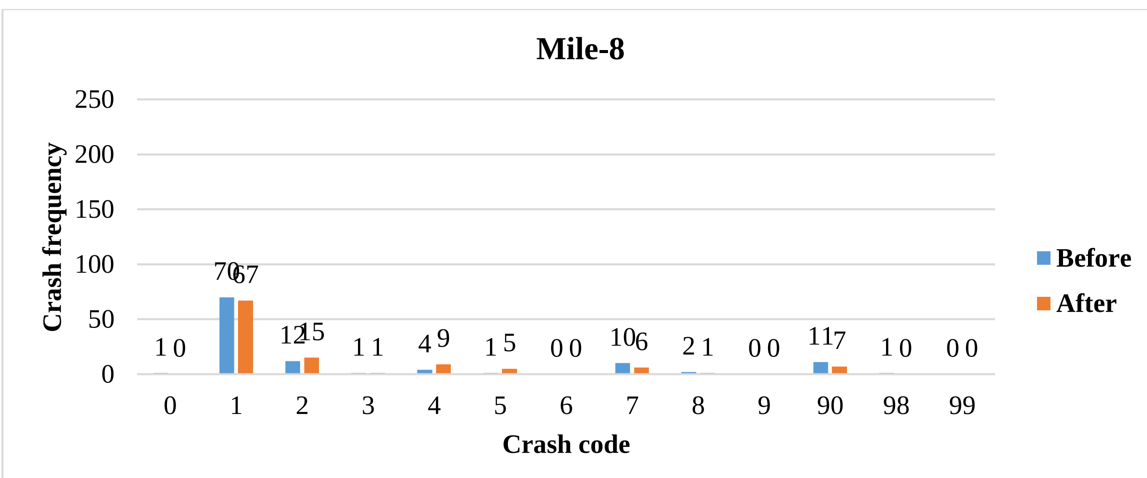

Figure A. 8 Crash Summary Histogram for Section Mile-8

Table A. 9 Crash Summary Table for Section Mile-9

Crash Code	Crash Type	Crash Frequency	
		Before	After
0	Unspecified	0	1
1	Rear end	170	177
2	Sideswipe-Same direction	50	49
3	Left turn	0	1
4	Ran off road-Left side	14	11
5	Right angle	4	6
6	Right turn	1	1
7	Ran off road-Right side	18	18
8	Head on	2	4
9	Sideswipe-Opposing	0	0
90	Other	23	14
98	Not applicable	1	1
99	Unknow	1	0

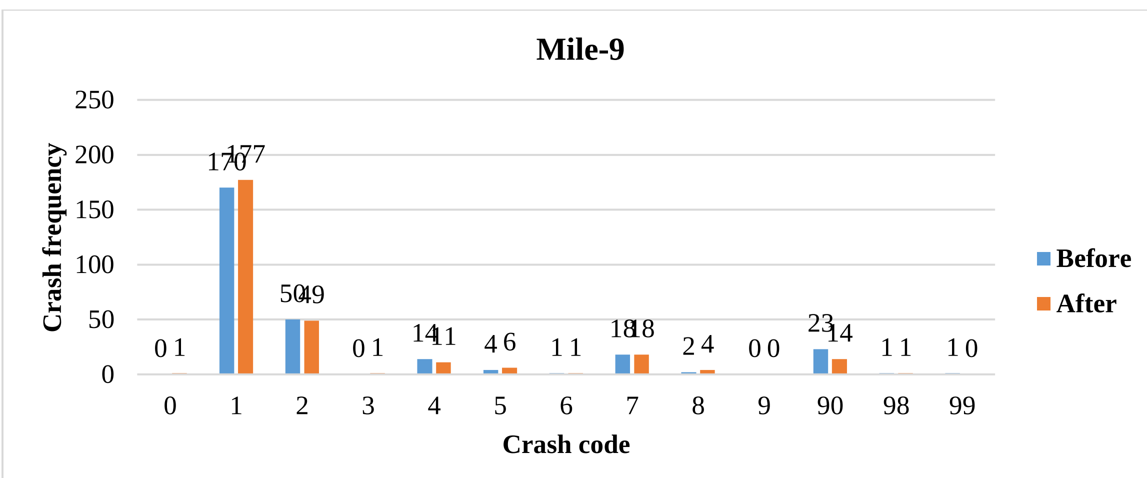

Figure A. 9 Crash Summary Histogram for Section Mile-9

Table A. 10 Crash Summary Table for Section Mile-10

Crash Code	Crash Type	Crash Frequency	
		Before	After
0	Unspecified	1	0
1	Rear end	16	8
2	Sideswipe-Same direction	4	5
3	Left turn	0	0
4	Ran off road-Left side	2	2
5	Right angle	1	0
6	Right turn	0	0
7	Ran off road-Right side	2	3
8	Head on	1	0
9	Sideswipe-Opposing	0	0
90	Other	4	1
98	Not applicable	2	0
99	Unknow	0	0

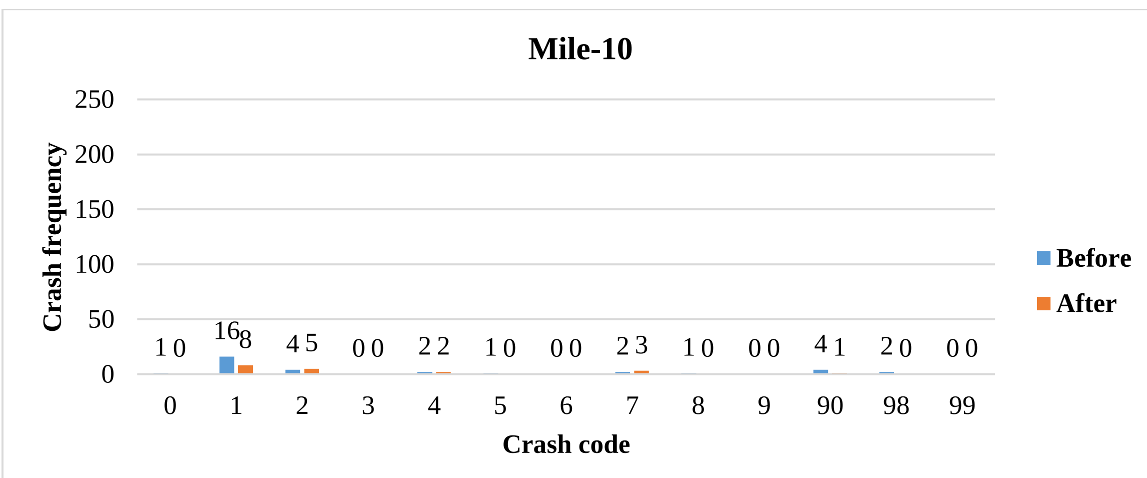

Figure A. 10 Crash Summary Histogram for Section Mile-10

Table A. 11 Crash Summary Table for Section Mile-11

Crash Code	Crash Type	Crash Frequency	
		Before	After
0	Unspecified	2	0
1	Rear end	108	48
2	Sideswipe-Same direction	34	26
3	Left turn	0	2
4	Ran off road-Left side	20	23
5	Right angle	4	3
6	Right turn	0	0
7	Ran off road-Right side	15	19
8	Head on	3	3
9	Sideswipe-Opposing	0	1
90	Other	22	13
98	Not applicable	1	0
99	Unknow	0	0

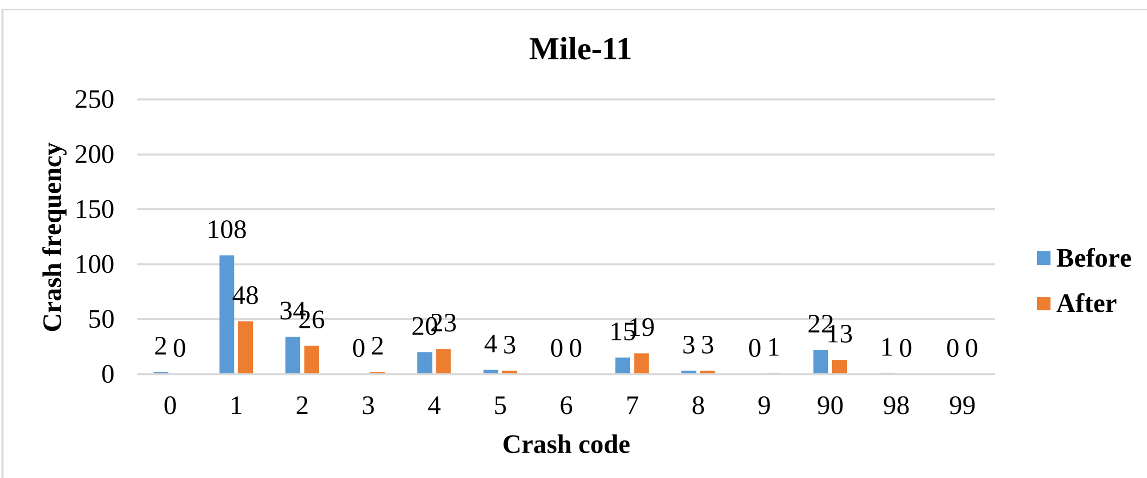

Figure A. 11 Crash Summary Histogram for Section Mile-11

Table A. 12 Crash Summary Table for Section Mile-12

Crash Code	Crash Type	Crash Frequency	
		Before	After
0	Unspecified	2	0
1	Rear end	170	80
2	Sideswipe-Same direction	36	30
3	Left turn	0	0
4	Ran off road-Left side	18	25
5	Right angle	2	6
6	Right turn	0	0
7	Ran off road-Right side	10	18
8	Head on	2	3
9	Sideswipe-Opposing	0	1
90	Other	18	12
98	Not applicable	2	0
99	Unknow	0	0

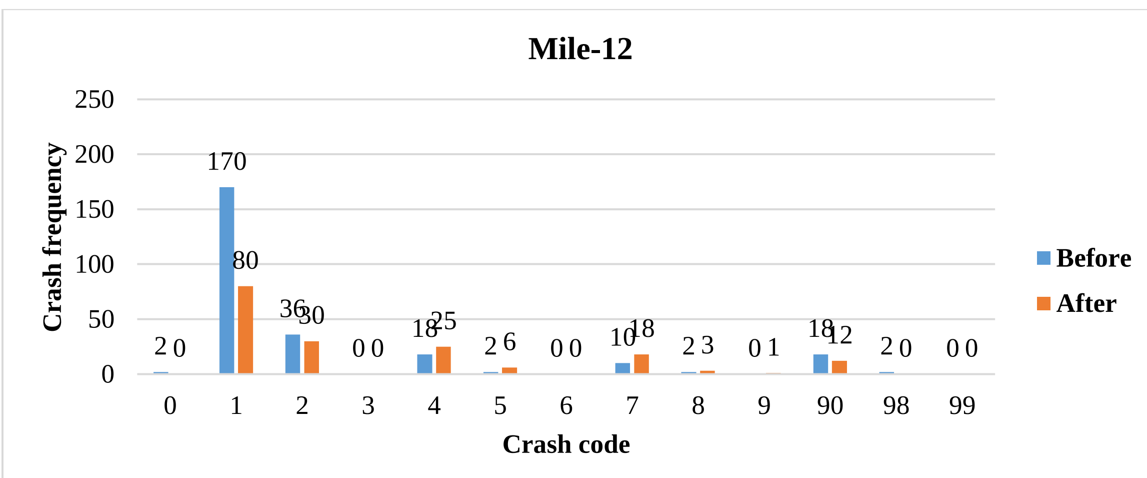

Figure A. 12 Crash Summary Histogram for Section Mile-12

Table A. 13 Crash Summary Table for Section Mile-13

Crash Code	Crash Type	Crash Frequency	
		Before	After
0	Unspecified	1	1
1	Rear end	129	82
2	Sideswipe-Same direction	31	25
3	Left turn	0	0
4	Ran off road-Left side	7	10
5	Right angle	1	11
6	Right turn	0	0
7	Ran off road-Right side	16	10
8	Head on	4	2
9	Sideswipe-Opposing	1	1
90	Other	9	5
98	Not applicable	2	0
99	Unknow	0	0

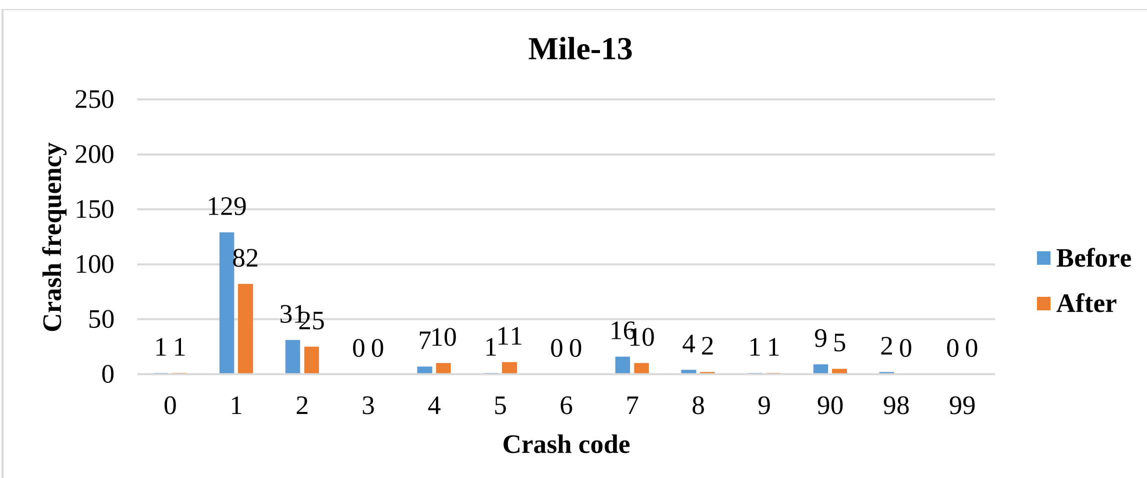

Figure A. 13 Crash Summary Histogram for Section Mile-13

Table A. 14 Crash Summary Table for Section Mile-14

Crash Code	Crash Type	Crash Frequency	
		Before	After
0	Unspecified	1	0
1	Rear end	96	136
2	Sideswipe-Same direction	31	46
3	Left turn	0	0
4	Ran off road-Left side	4	15
5	Right angle	0	10
6	Right turn	0	1
7	Ran off road-Right side	5	7
8	Head on	4	2
9	Sideswipe-Opposing	1	0
90	Other	9	13
98	Not applicable	2	2
99	Unknow	0	0

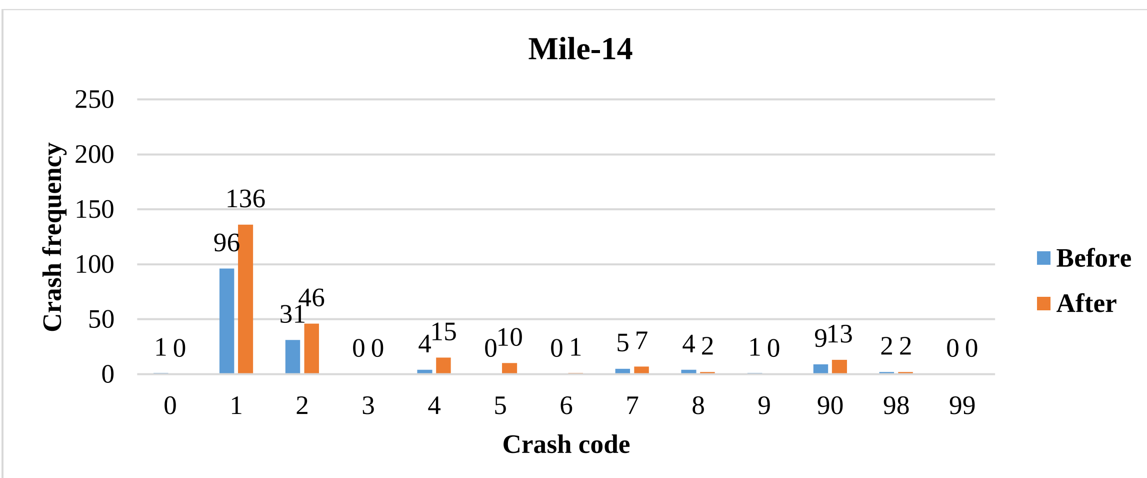

Figure A. 14 Crash Summary Histogram for Section Mile-14

Table A. 15 Crash Summary Table for Section Mile-15

Crash Code	Crash Type	Crash Frequency	
		Before	After
0	Unspecified	0	1
1	Rear end	149	205
2	Sideswipe-Same direction	33	45
3	Left turn	0	0
4	Ran off road-Left side	15	12
5	Right angle	5	3
6	Right turn	0	1
7	Ran off road-Right side	6	12
8	Head on	1	0
9	Sideswipe-Opposing	1	0
90	Other	8	12
98	Not applicable	1	1
99	Unknow	1	0

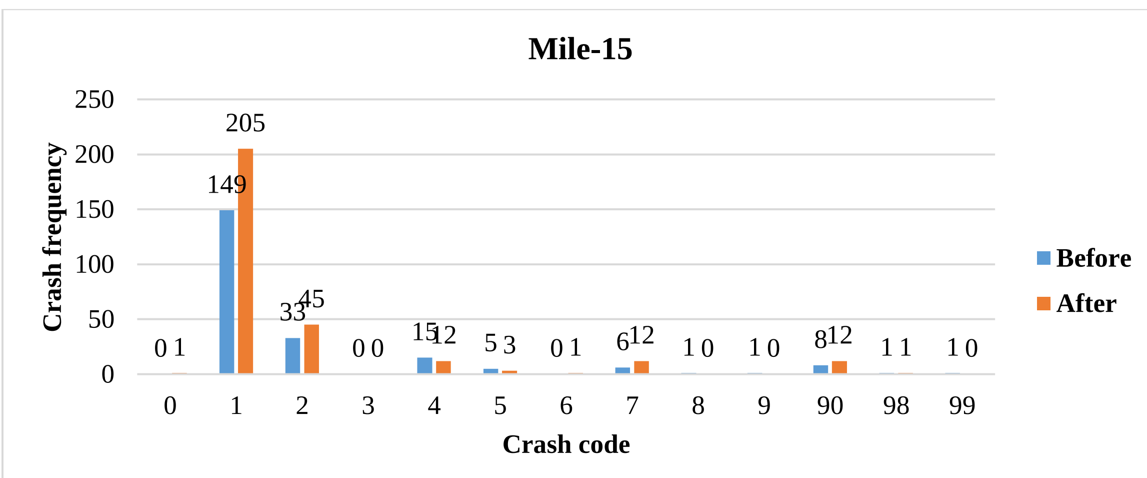

Figure A. 15 Crash Summary Histogram for Section Mile-15

Table A. 16 Crash Summary Table for Section Mile-16

Crash Code	Crash Type	Crash Frequency	
		Before	After
0	Unspecified	3	0
1	Rear end	152	221
2	Sideswipe-Same direction	32	47
3	Left turn	0	1
4	Ran off road-Left side	12	13
5	Right angle	3	3
6	Right turn	1	0
7	Ran off road-Right side	9	10
8	Head on	2	4
9	Sideswipe-Opposing	0	0
90	Other	16	12
98	Not applicable	4	0
99	Unknow	0	0

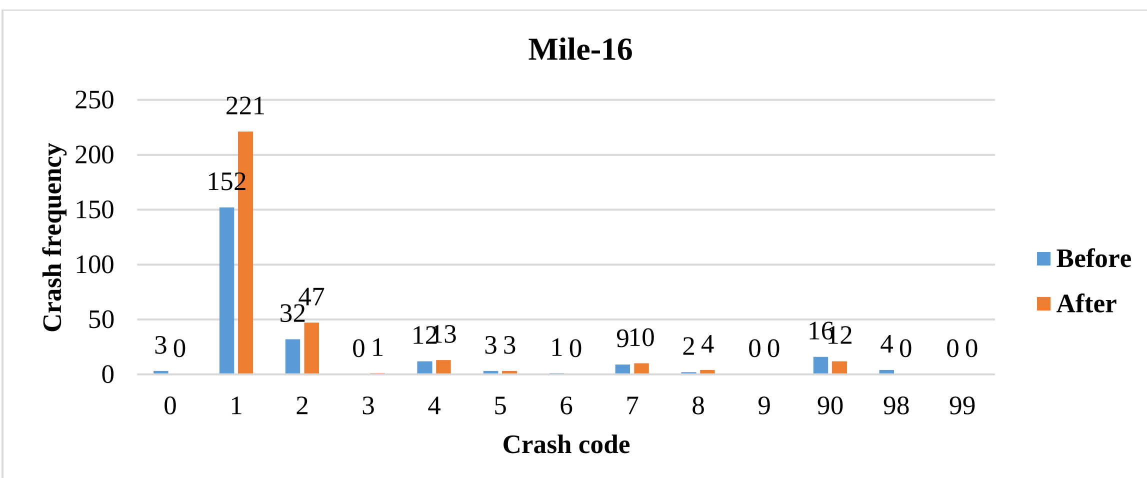

Figure A. 16 Crash Summary Histogram for Section Mile-16

Table A. 17 Crash Summary Table for Section Mile-17

Crash Code	Crash Type	Crash Frequency	
		Before	After
0	Unspecified	2	0
1	Rear end	160	220
2	Sideswipe-Same direction	44	53
3	Left turn	0	1
4	Ran off road-Left side	38	79
5	Right angle	7	3
6	Right turn	0	0
7	Ran off road-Right side	15	39
8	Head on	1	1
9	Sideswipe-Opposing	1	0
90	Other	24	25
98	Not applicable	5	4
99	Unknow	0	0

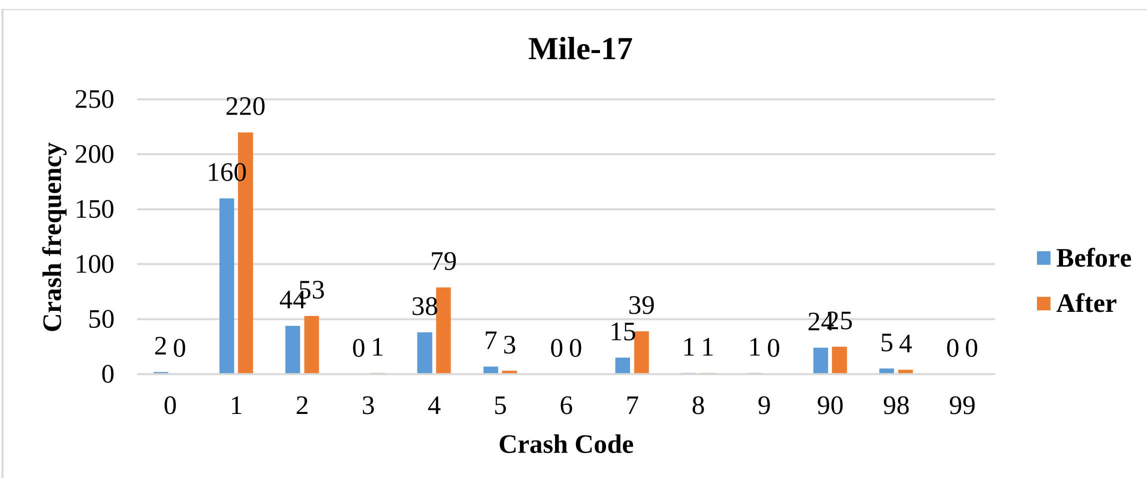

Figure A. 17 Crash Summary Histogram for Section Mile-17