DEPARTMENT OF TRANSPORTATION

Safety Impacts of the I-35W Improvements Done Under Minnesota's Urban Partnership Agreement (UPA) Project

Gary A. Davis, Principal Investigator
Department of Civil, Environmental, and Geo- Engineering
University of Minnesota

June 2017

Research Project
Final Report 2017-22

To request this document in an alternative format, such as braille or large print, call 651-366-4718 or 1-800-657-3774 (Greater Minnesota) or email your request to ADArequest.dot@state.mn.us. Please request at least one week in advance.

Technical Report Documentation Page

1. Report No.	2.	3. Recipients Accession No.
MN/RC 2017-22		

4. Title and Subtitle	5. Report Date
Safety Impacts of the I-35W Improvements Done Under Minnesota's Urban Partnership Agreement (UPA) Project	June 2017
	6.
7. Author(s) Gary A. Davis, Jingru Gao, John Hourdos	8. Performing Organization Report No.
9. Performing Organization Name and Address Dept. of Civil, Mineral, and Environmental Engineering University of Minnesota 500 Pillsbury Drive SE Minneapolis, MN 55455	10. Project/Task/Work Unit No. CTS# 2015005 11. Contract (C) or Grant (G) No. (C) 99008 (wo) 151
12. Sponsoring Organization Name and Address Minnesota Department of Transportation Research Services & Library 395 John Ireland Boulevard, MS 330 St. Paul, Minnesota 55155-1899	13. Type of Report and Period Covered Final Report 14. Sponsoring Agency Code

15. Supplementary Notes

http:// mndot.gov/research/reports/2017/201722.pdf

16. Abstract (Limit: 250 words)

As part of an Urban Partnership Agreement project, the Minnesota Department of Transportation added lanes and began operating a priced dynamic shoulder lane (PDSL) on parts of Interstate 35W. Following the opening of these improvements, the frequency of rear-end crashes increased in certain sections, especially in the PDSL region. The object of this study was to determine if these increases were direct effects of the improvements or were due to changes in traffic conditions. Logistic regression analyses which controlled for changes in traffic conditions indicated no direct effect on the likelihood of rear-end crashes due to operation of the PDSL; the observed change in crash frequency was explained by the change in traffic conditions. This study also found evidence for a nonlinear relationship between a proxy for traffic density, lane occupancy, and the probability of a rear-end crash occurring during an hour. In several sections 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.

17. Document Analysis/Descriptors	18. Availability Statement
Traffic crashes, High occupancy toll lanes, Lane occupancy, Traffic density, Rear end crashes, Logistic regression analysis	No restrictions. Document available from: National Technical Information Services, Alexandria, Virginia 22312

19. Security Class (this report)	20. Security Class (this page)	21. No. of Pages	22. Price
Unclassified	Unclassified	77	

Safety Impacts of the I-35W Improvements Done Under Minnesota's Urban Partnership Agreement (UPA) Project

FINAL REPORT

Prepared by:

Gary A. Davis
Jingru Gao
John Hourdos

Department of Civil, Environmental, and Geo- Engineering
University of Minnesota

June 2017

Published by:

Minnesota Department of Transportation
Research Services & Library
395 John Ireland Boulevard, MS 330
St. Paul, Minnesota 55155-1899

ACKNOWLEDGMENTS

This research was supported by the Minnesota Department of Transportation.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

EXECUTIVE SUMMARY

As part of an Urban Partnership Agreement project, the Minnesota Department of Transportation (MnDOT) added lanes and began operating a priced dynamic shoulder lane (PDSL) on parts of I-35W. Following the opening of these improvements, the frequency of rear-end crashes increased in certain sections, especially in the PDSL region. The object of this study was to determine if this increase was due to changes in traffic conditions or was a direct effect of the improvements.

A preliminary analysis was done to determine the study region. 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. The dominant crash type was rear-end crashes, but the changing trend of bi-direction rear-end crash frequencies from the Before to After period varied among the 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).

The I-35W HOT region and the PDSL region were divided into sections based on constant flow and geometry criteria as well as the availability of loop-detector data. Crash, loop detector, weather condition, and PDSL activation (only for sections in PDSL region) data for the Before and After periods were compiled for each section. Rear-end crash records were extracted using Minnesota Crash Mapping Analysis Tool (MNCMAT), and the original police reports were reviewed to confirm the crash times and locations. Loop-detector data were retrieved using MnDOT's DataExtract tool. The source of weather condition data, namely rain and snow records, was MnDOT's Road Weather Information System (RWIS). The PDSL activation status came from the MnDOT's log of the Intelligent Lane Control Signal (ILCS) located at 37th Street.

Logistic regression analyses estimating the change in rear-end crash risk following the UPA project, controlling for traffic conditions, weather conditions, and PDSL activations, were carried out for each section. 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 (constructed using the loop detector data), the presence or absence of snowy or rainy conditions, before versus after the UPA project, and the presence or absence of PDSL operation.

Results

1) Most analyzed sections in the I-35W HOT region showed no significant change in rear-end crash risk associated with the UPA project. Exceptions were Section N17 (northbound, just south of I-494) and 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. The apparent increase in rear-end crash risk of Section S9 was possibly due to ambiguities in locating crashes during the Before period.
2) An "Inverted U" relationship between a proxy for traffic density, lane occupancy, and the probability of a rear-end crash occurring during an hour were seen in several sections. 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.

3) The I-35W PDSL region experienced a substantial increase in congestion (defined as average lane occupancies exceeding 25%) following completion of the UPA improvements. This was due to removal of the bottleneck in the old I-35W and TH-62 commons, causing the bottleneck to move northward to the I-35W and I-94 junction. When controlling for the change in traffic conditions, there was no significant increase in rear-end crash risk attributable to the PDSL operation.

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. MnDOT is interested in using PDSLs at other freeway locations. The impact of this research is to show that the current implementation of a PDSL did not have an adverse effect on safety.

CHAPTER 1 INTRODUCTION

1.1 BACKGROUND

In 2009, the Minnesota Department of Transportation (MnDOT) began a major set of improvements to Interstate Highway 35W (I-35W) as part of the Federal Highway Administration's Urban Partnership Agreement (UPA). Aimed at reducing congestion on I-35W between the Minnesota River and Interstate 94 (I-94), the UPA intervention consisted of several different improvements, and the areas affected by these are shown in **Figure 1. 1**. These included adding a lane to southbound I-35W as it approaches and crosses the Minnesota River by decreasing the widths of the existing lanes and shoulder; reconstruction to remove a bottleneck in the Crosstown Commons, where I-35W shared right-of-way with Trunk Highway 62 (TH-62); and conversion of the shoulder on a section of northbound I-35W to a price dynamic shoulder lane (PDSL). The PDSL can function either as a normal shoulder or as a high occupancy toll (HOT) lane, depending on demand.

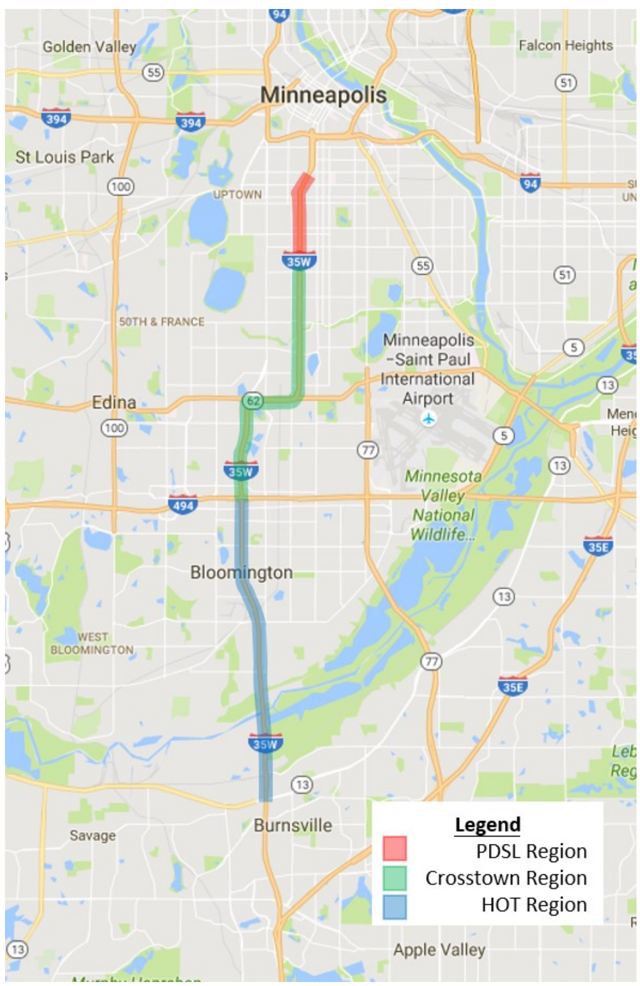

Figure 1. 1 Portion of I-35W receiving UPA improvements. North is up. (Source: Google Map, modified by Gao)

In 2013, MnDOT released Problem Statement NS-329, which noted interest in extending some or all of these interventions to other corridors, but that estimates of their safety effects were needed to assist these decisions. For example, crash frequencies appeared to have increased where the PDSLs were located, but it was unclear if this was due to the PDSLs themselves or to changes in traffic congestion due to removal of an upstream bottleneck at the Crosstown Commons.

The objective of this research was to try and 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. 2.**

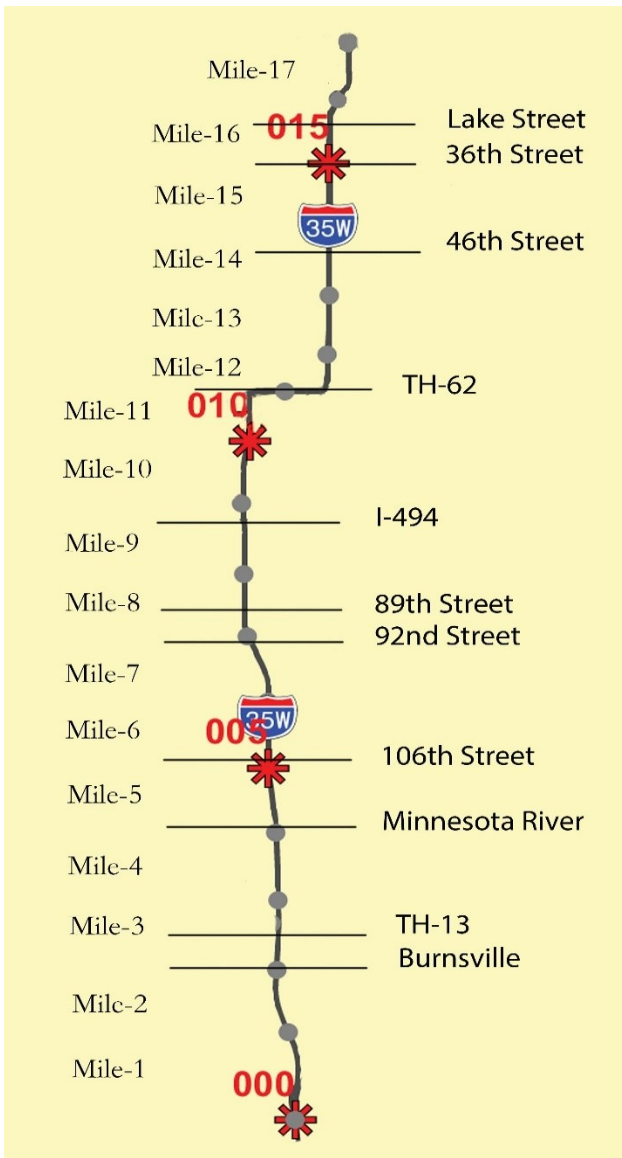

Figure 1. 2 17 1-mile Sections of I-35W

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 (milepost), 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. 3 visualizes the crash summary shown in **Table 1. 1**.

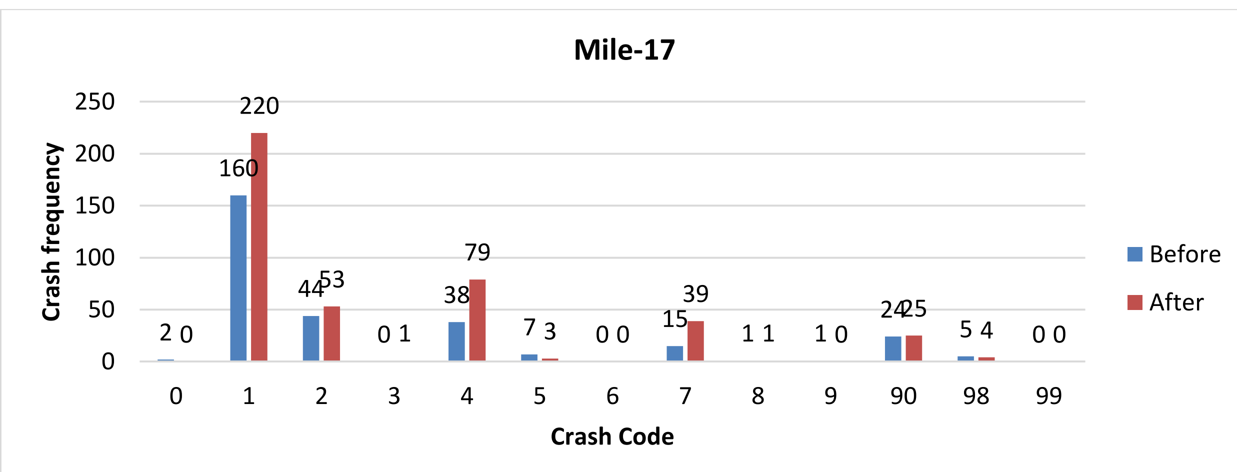

Figure 1. 3 Histograms of Crash Summary in Section Mile-17

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

Figure 1. 4 shows Before and After rear-end crash frequencies in each 1-mile section.

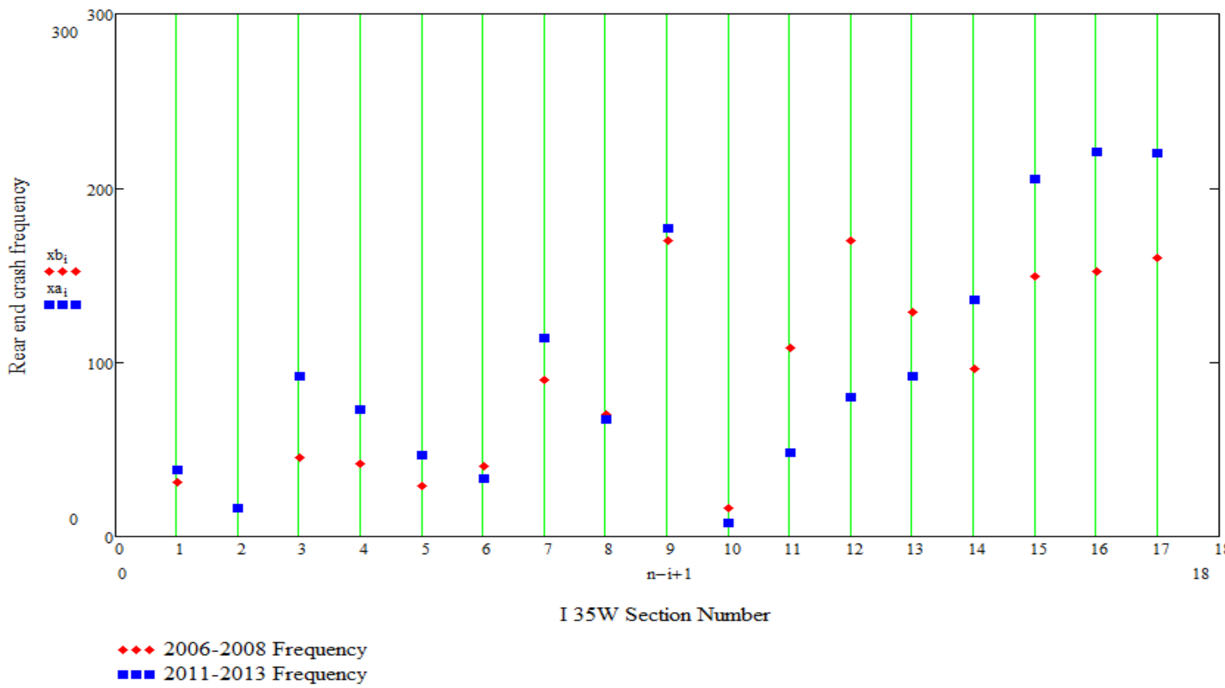

Figure 1. 4 Before And After Frequencies Of Rear-End Crashes In Each 1-Mile Section

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.

1.2 SELECTED LITERATURE REVIEW

As noted above, one of the UPA improvements involved narrowing the lanes and shoulder of a portion of southbound I-35W to increase capacity by adding an additional lane, in this case increasing the number of lanes from three to four. Although our literature search turned up no reports explicitly addressing this intervention, Bauer et al. (2004) reported on the effects of changing from four to five and five to six lanes in one direction. These results have been summarized in the *Highway Safety Manual* (HSM) as an 11% increase in crash frequency for the four-lane to five-lane conversion (AASHTO 2010, p. 13-10). Regarding conversion of HOV lanes to HOT lanes, Cao et al. (2012) looked at changes in crash frequency following conversion of the I-394 HOV lanes to HOT lanes. The authors found weak evidence for a small (~5.3%) decrease in total 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.

The safety performance functions that support much of the *Highway Safety Manual* quantify associations between aggregate measures of traffic, such as AADT and posted speed limit, and crash frequencies accumulated over several years. For some time though, it has been recognized that crash risk probably varies as traffic conditions vary (Liu 1997), leading to a continuing effort at identifying traffic situations where crashes are more likely. Pioneering work in this area was done by Oh et al. (2001) and Lee et al. (2002), with a focus on short-term (c. 1 minute) identification of crash precursors as a possible component of a traveler information system. Subsequently, Abdel-Aty et al. (2004) introduced a case-control approach that has become something of a methodological standard. Police reports were reviewed to identify crashes occurring on a section of Interstate-4 and archived data from inductive loop detectors were then used to characterize traffic conditions in the vicinities of the crashes. These crash-related events provided the cases, while traffic conditions from the same places and similar times, but with no crashes, provided the controls. Logistic regression was then used to identify those traffic measures that discriminated the cases from the controls. The authors reported that a model using the coefficient of variation of traffic speed and the average lane occupancy as crash predictors showed a 69% detection rate and a 47% false-positive rate. Since 2004, a number of variants on this case-control approach have been reported, and Roshandel et al. (2015) have provided a recent review and meta-analysis. Overall, it appears that results can depend on the locations of loop detectors compared to the locations of crashes. For example, crash risk can increase as both traffic density and speed coefficient of variation increase downstream from a crash location, but the effect is much weaker when these are measured upstream from a crash location.

Although the main focus in this area has been on empirical models for short-term prediction of freeway crashes, with less emphasis on explaining why crashes come about, several studies have offered insight into the etiology of freeway crashes. (1) Using loop-detector data, Zhang et al. (2010) reported that traffic oscillations, i.e., the rise and fall of traffic speed and density associated with shock waves, were associated with increased crash likelihood. (2) Xu et al. (2014) reported freeway crash likelihood was highest when traffic is operating at level of service (LOS) E, declined somewhat in LOS D and F, and was lowest in LOS A. Since for freeways the LOS categories A-F are defined by progressively higher traffic densities, this suggests that crash probability could have a roughly concave relationship to traffic density, with a maximum probability occurring near densities characterizing capacity flow. (3) Hourdos (2005, Hourdos et al. 2006) reported an interesting variant on the standard case-control approach. Rather than relying on crash reports and archived loop-detector data, video cameras located on high-rise buildings were used to record traffic on a section of Interstate-94, and the video records were then used to (a) identify crash and near-crash events visually and (b) to measure traffic variables using machine-vision methods. This allowed for a more sensitive testing of location and time-scale effects but, more fundamentally, identified stopping shock waves as necessary precursors to rear-end crashes. Since these waves tend to occur when freeway traffic is transitioning from capacity flow to stop-go conditions, this implies that rear-end crashes should be most likely in these conditions. (4) Using trajectories of individual vehicles involved in stopping shock waves, extracted from video collected by Hourdos, Davis and Swenson (2006) illustrated how Brill's (1972) model of progressively severe braking could explain the occurrence of several actual rear-end crashes. (5) Chatterjee (2016, Chatterjee and Davis 2016) then combined Brill's shock wave model with Newell's (1993) model of the relation between traffic flow and traffic density, producing a structural-empirical method for rating freeway collision hazard.

The theoretical relationship between freeway traffic conditions and probability of a rear-end crash can be illustrated by combining Brill's shockwave model with Greenshield's (1934) linear relationship between traffic density and mean speed. In Brill's model, a sequence of stopping vehicles results in a crash when cumulative differences between driver reaction times and following times exceeds a threshold. That is when

$$S_n = \sum_{i=1}^{n} (r_i - h_i) > \frac{v_0 (a_{\max} - a_0)}{2 a_0 a_{\max}} \tag{1}$$

Where

r_i = braking reaction time of the ith driver in the stopping wave

h_i = following time (headway) of the ith driver

v_0 = common speed of vehicles in wave

a_0 = braking deceleration of driver initiating the wave

a_{max} = maximum available braking deceleration

If driver reaction times are independent random variables, Brill showed that the cumulative difference S_n in equation (1) is a random walk, and the probability that a shockwave results in a collision is related to the distribution of first crossing times for this random walk. For a given traffic density, Greenshield's model can be used to compute the corresponding traffic speeds and following times, and treating the driver reaction times as independent random variables, it is straightforward to evaluate the probability associated with equation (1)'s crossing event. **Figure 1. 5** illustrates the relationship between traffic density and the probability that a rear-end crash occurs on a single lane of freeway 0.5 miles long, when the free-flow speed is 65 mph and capacity flow is 2200 vehicles/lane/mile. For this example, the initial braking deceleration a_0 was 10 feet/sec^2, a_{max} was 20 feet/sec^2 and driver reaction times were normally distributed with a mean of 1.2 seconds and a standard deviation of 0.2 seconds.

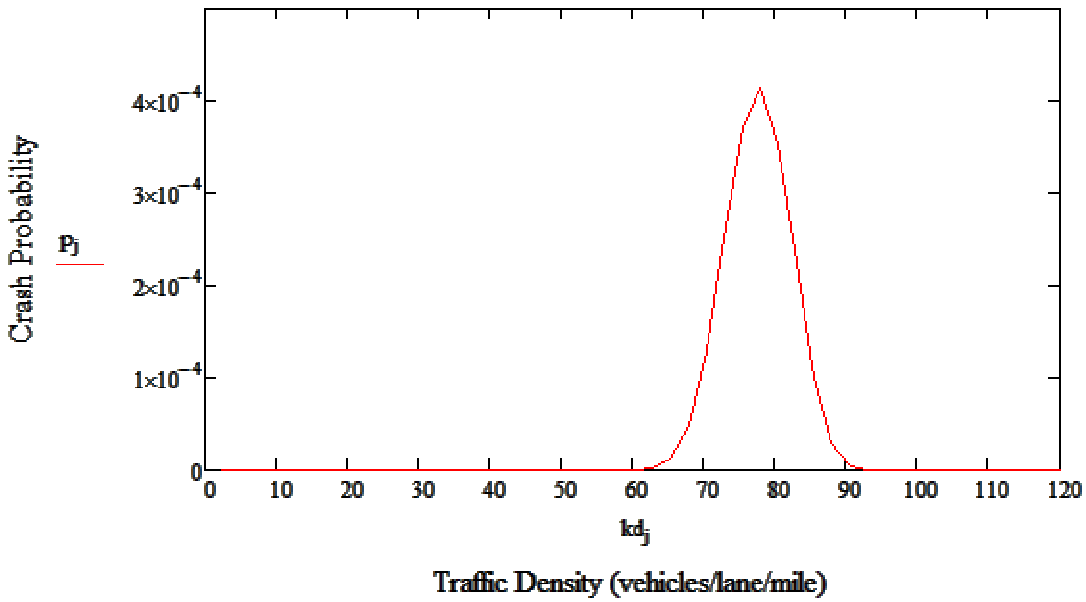

Figure 1. 5 Theoretical relationship between traffic density and the probability a stopping shockwave produces a rear-end crash.

In **Figure 1. 5,** the probability that a shockwave produces a crash is highest when traffic density is about 78 vehicles/lane/mile and decreases for densities above and below this point. The situation depicted in **Figure 1. 5** simplifies what happens in reality; nonetheless, empirical models seeking to relate crash likelihood to traffic density should allow for a possible concave relationship over a range of densities.

CHAPTER 2 DATA ACQUISITION

The first task was to compile the data files needed to conduct statistical analyses. Such data files include explanatory variables such as relevant traffic volume and lane occupancy, weather conditions, and the presence or absence of UPA improvements. They data files also include indicators of crash experience on I-35W during both before (years 2006-2008) and after (years 2011-2013) periods. This chapter describes the data collection of rear-ending crashes, traffic conditions, PDSL activation, and weather condition.

2.1 STUDIED SECTIONS

The scope of this study can be divided into three regions, the HOT region, Crosstown region, and the PDSL region, as shown in **Figure 2. 1**.

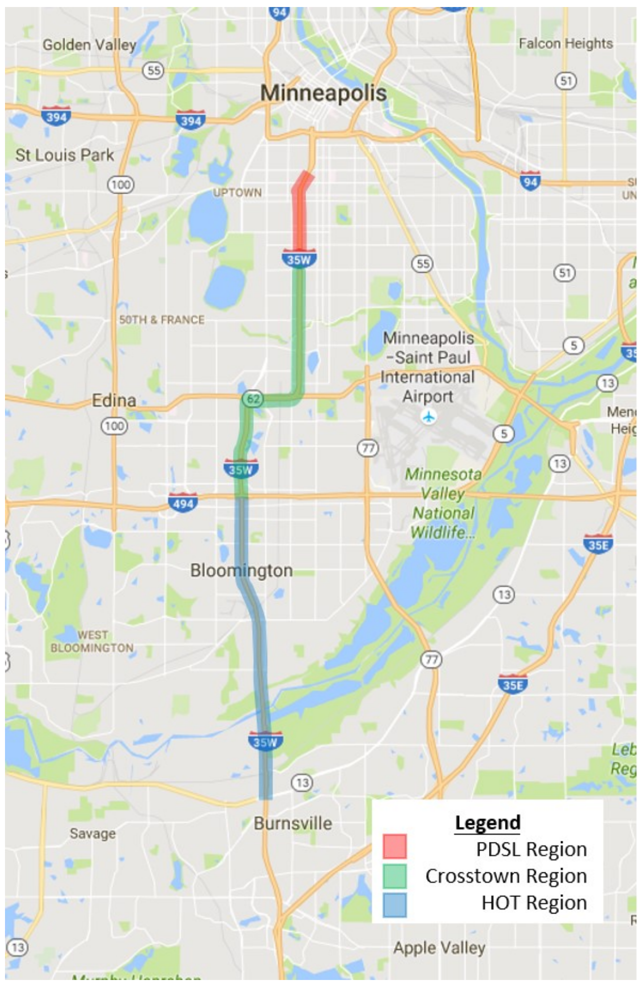

Figure 2. 1 MN UPA Project Intervention Map
(Source: Google Map, modified by Gao)

All three regions were then divided into sections so that traffic demand and lane geometry where constant within a section. That is, lane drops or additions were used to determine section boundaries, as were the junctions with on and off ramps. It is worth noting that not all sections had loop detectors installed for all

6 years (2006-2008 and 2011-2013) of the analysis period. The statistical analyses could be done only for those sections with loop detector data available.

I-35W mileposts for both before (2006-2008) and after (2011-2013) periods were provided by MnDOT, and the lane counts were verified with GoogleEarth.

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

Table 2.1, Table 2.2, and Table 2.3 list the HOT and PDSL Region division results. The sections with symbol "*" are those sections with loop detectors where statistical analyses could be done.

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

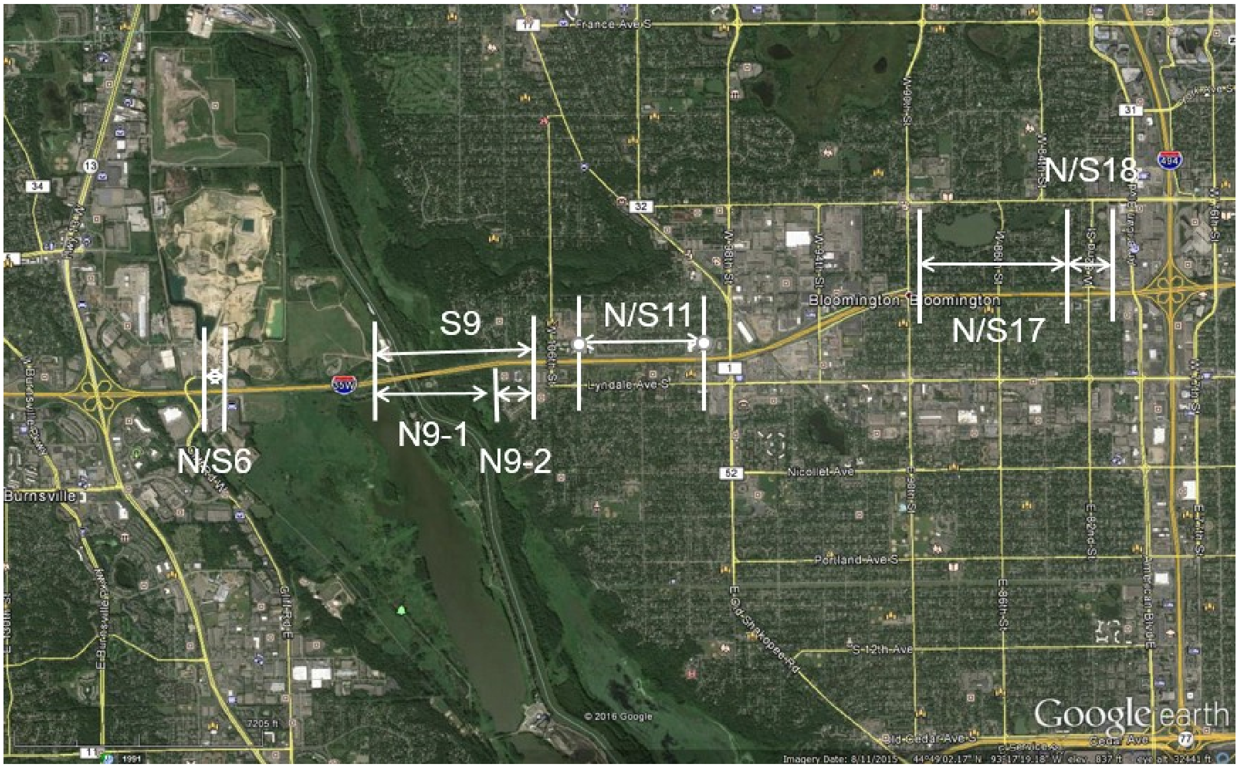

Figure 2. 2 HOT Region Studied Sections
(Source: Google Earth, modified by Gao)

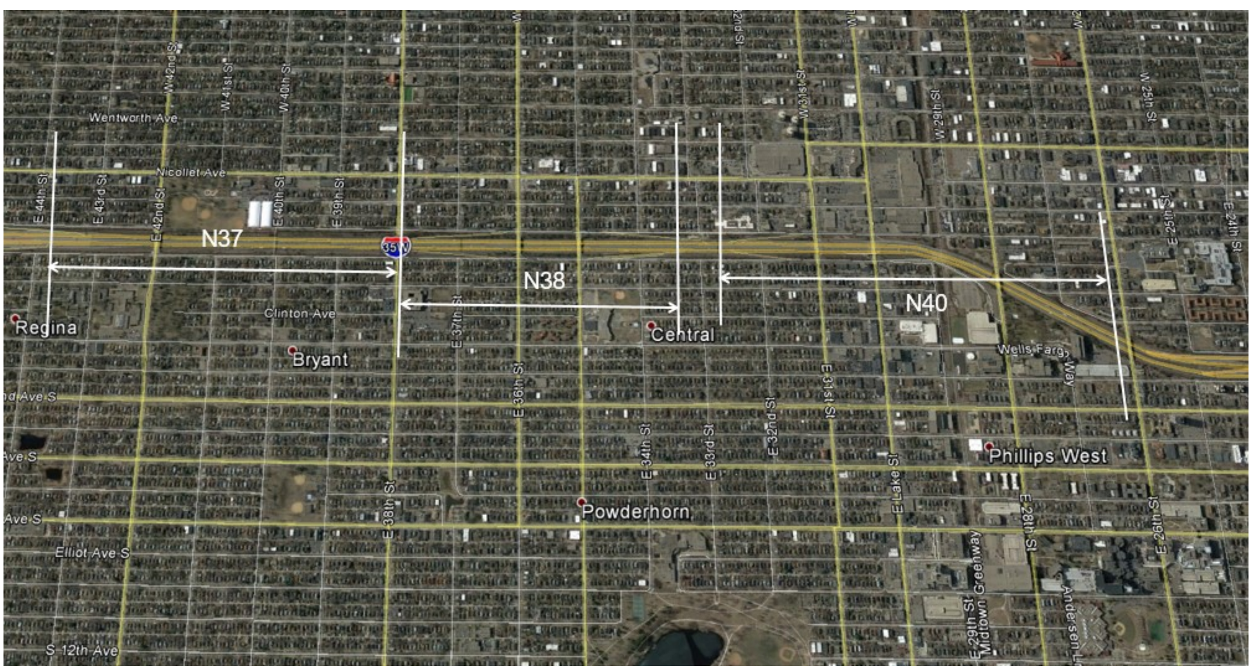

Figure 2. 3 PDSL Region Studied Sections
(Source: Google Earth, modified by Gao)

2.2 COMPILING CRASH DATA

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

5545 crash records of all crash types were extracted from 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(tem), Route, and Ref(erence)_Point)
- Crash Number consistent with original crash 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 2. 4 is a summary table of the number of crashes originally extracted from MNCMAT.

Table 2. 4 Summary Information of Crash Records Extracted from 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 2. 4 is a bar-chart showing the crash frequencies presented in **Table 2. 4**.

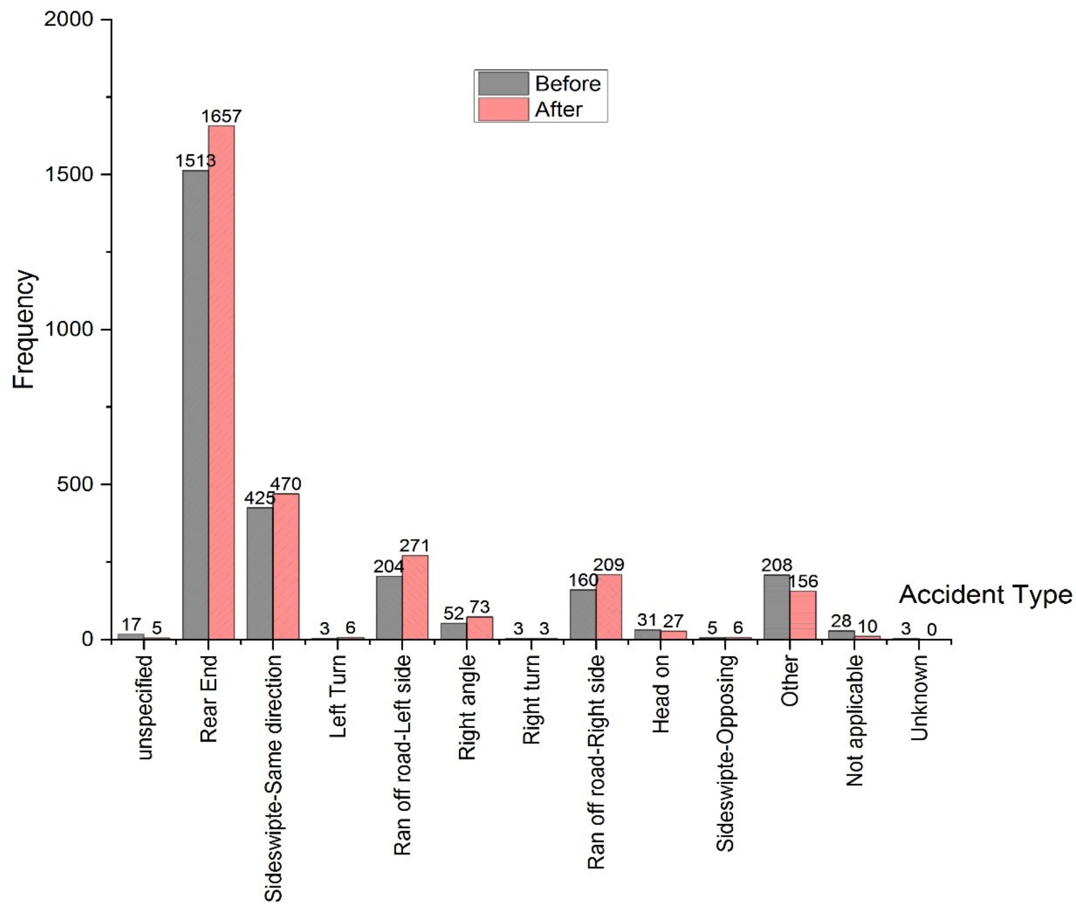

Figure 2. 4 Counts of Crash Records Extracted from MNCMAT by Crash Type

As is shown in both **Table 2. 4** and **Figure 2. 4**, rear-ending crashes were the most prevalent crash type on I-35W from its beginning to the I-35W/I-94 junction; thus rear-ending crashes were chosen as our study priority.

Table 2. ***5*** shows the number of rear-ending crashes from the MNCMAT database for each studied section.

Table 2. 5 Number of Rear-ending Crashes from MNCMAT database for Each Studied Section

Section No.	Number of Rear-ending 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 tried to determine the actual road direction, northbound or southbound, with the aid of DOT/OTIS cCrash reports.

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

- Crash type
    - ➢ If the crash is a rear-ending 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 the crash reports were adopted.

Although there were specified diagram codes, milepost and road directions in crash records from MNCMAT database, we still reviewed crash reports for all types of crashes and all road directions 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 were unspecified or vague.
- Duplicated crash records where citizen and police reports were merged in DVS motor vehicle crash report database.
- Deleted crash records which were found not to be a true crash by MnDOT’s Standards.

Table 2. 6 shows the number of rear-ending crashes for the study sections after review of the crash reports.

Table 2. 6 Number of Rear-ending Crashes for Studied Sections after Review of Crash Reports.

Section No.	Number of Rear-ending 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

2.3 COMPILING TRAFFIC CONDITION DATA

As noted in Chapter 1, there is good reason to consider traffic conditions as affecting rear-ending 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 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 changed during the six study years, all six annual ADRs were reviewed to guarantee the consistency in traffic condition data collection.

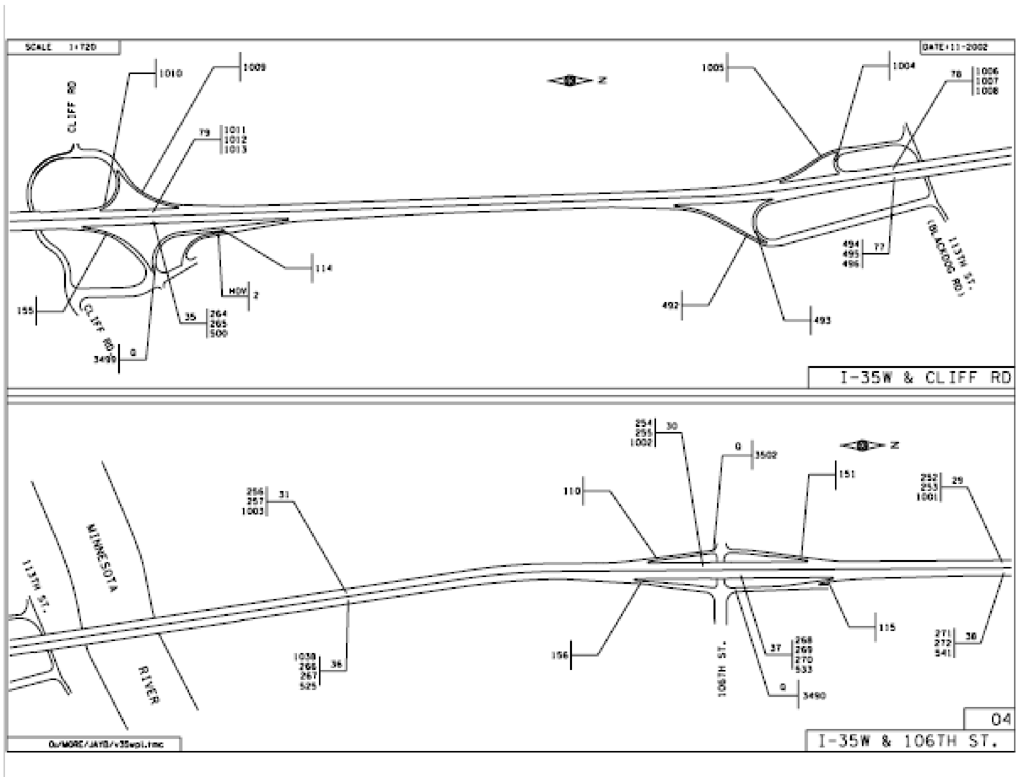

Figure 2. 5 Screenshot from the 2006 ADR.

Table 2. 7 shows the matching of detector stations to study sections.

Table 2. 7 Matching Detector Stations to Study 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 data were for individual loops, aggregated over 30-second intervals, the most basic data provided in DataExtract. The total numbers of original lane volume or occupancy records each year were as follows:

- Years 2006, 2007, 2011, and 2013: $\frac{(1\ hour)(3600\frac{second}{hour})}{30\ seconds} \times 24\ hours \times 365\ days = 1051200$
- Years 2008 and 2012: $\frac{(1\ hour)(3600\frac{second}{hour})}{30\ seconds} \times 24\ hours \times 366\ days = 1054080$

2.4 PDSL OPERATION HISTORY 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, special events or incidents. Over time, MnDOT personnel noticed high violation rates for the PDSL during the mid-day, thus the PDSL standard hours have been extended to 6:00 AM-7:00 PM. This change started on April 11th, 2012.

The PDSL activation data came from MnDOT's log for the 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 changed and the time was accurate to seconds.

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. This is 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 blocking 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 for 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 2. 6 Example Portion of ILCS Log

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

2.5 COMPILING WEATHER DATA

Since it is well-known that adverse weather can lead to crashes on Minnesota freeways, it was also necessary to control for rainy and snowy weather conditions. 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 conditions for crash hours were taken from the weather information shown in MNCMAT crash records and verified with crash reports.

For non-crash hours data from the 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 the RWIS database could only provide the weather data from site "I-35 at Minnesota river" from year 2011-2013, it was necessary to look at other sites near "I-35 at Minnesota river" site for the weather data during year 2006-2008. Since the "Minneapolis-St. Paul International Airport" site only had weather data for year 2010 the site "I-35 E Cayuga St. Bridge" was chosen as the source of weather conditions for the years 2006-2008.

Figure 2. 7 shows the locations of weather information collection sites near studied sections.

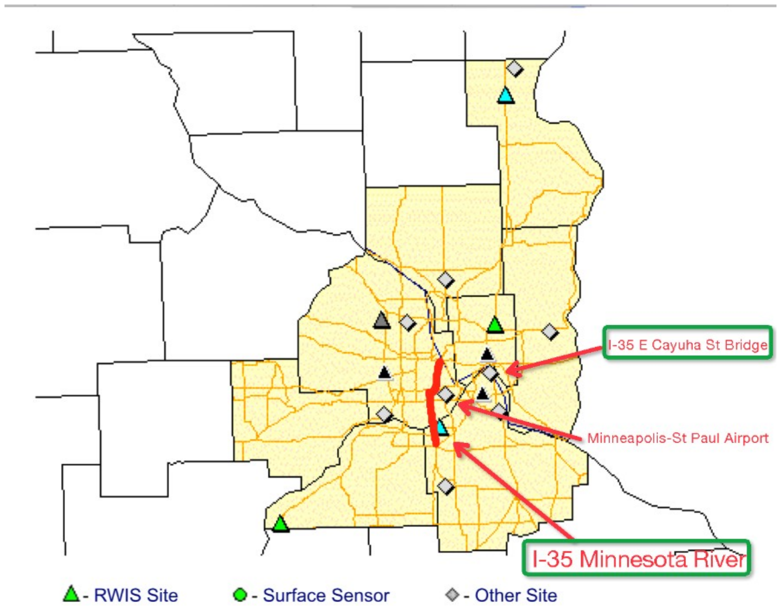

Figure 2. 7   Weather Information Site Locations

By manually inputting the desired date, historical weather information can be retrieved from 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 a supplemental source.

CHAPTER 3 DATA PREPARATION

In the later statistical analyses, for each hour during 2006- 2008 and 2011-2013, the presence or absence of a rear-ending crash became the dependent variable while the independent variables consisted of traffic volume and lane occupancy, the presence or absence of snowy or rainy conditions, and the presence or absence of UPA improvements. This required further processing of the original data and this chapter describes this.

Figure 3. 1 and **Figure 3. 2** show screenshots of example data files for the HOT and PDSL sections, 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 3. 1 Example Data Set for Statistical Analysis of 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 3. 2 Example Data Set for Statistical Analysis of PDSL Sections

3.1 CRASH DATA

To avoid possible biases from secondary crashes, in this study the response variable in the statistical models was the presence absence of a rear-ending crash in each hour, which is a binary variable where "1" was assigned to the hours when at least one rear-ending crash happened and "0" otherwise. There were cases that two consecutive rear-ending 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 deleted from the sample.

3.2 TRAFFIC CONDITION DATA

Since hourly traffic condition data were needed for analysis data aggregation was done using 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 the questionable data that should not be used for analysis.

Three types of questionable data were identified:

- Records with negative volume or occupancy.
- Records with occupancies greater than 100%.
- Records with repeated patterns 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 were then computed from the remaining 30-second data as follows:

- Lane occupancy mean: $\bar{O} = \frac{\sum_{i=1}^{n}\sum_{j=1}^{m} occupancy_{ij}}{n \times m}$
- 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 jth lane occupancy record for the ith lane

$volume_{ij}$ - the jth traffic count record for the ith lane

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 in vehicles/hour, centered lane occupancy and centered lane occupancy standard deviation for each section were computed based from the lane volume mean, lane occupancy mean, and lane occupancy variance as follows:

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

3.3 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 relatively quick pace.

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

First, missing data in the ILCS log was due to technical problems and there was no other way to regain these PDSL activation. Thus missing hours were deleted from the sample.

For the second issue, from the log entry “DARK”, it was difficult to tell whether the PDSL was running as a general purpose lane or it was due to a technical issue, such as loss of communication with ILCS, and so we kept this status in our statistical model. For the entry “USE_CAUTION”, this mainly 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.

3.4 WEATHER DATA

Rainy and snowy weather condition indicators were introduced to the statistical model. Both indicators are binary variables where “1” was assigned to hours when there was rain or snow respectively and “0” otherwise. The weather condition indicator data were determined from the original weather data described in section 2.5.

3.5 OTHER

In addition to crash, traffic condition, PDSL status, 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 year 2006 to 2008, and “1” for hours in 2011-2013.

CHAPTER 4 STATISTICAL ANALYSES

This chapter describes the statistical analyses of how the probability of a rear-end crash occurring in a given hour varies with respect to other conditions prevailing during that hour. The goal is to identify the factors associated with rear-end crash risk on the studied sections during the studied years, especially whether or not rear-end crash risk changed following the UPA interventions. Since different UPA interventions were implemented within I-35W from TH-13 to I-494 and PDSL regions, analyses for those two regions were conducted 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.

4.1 STATISTICAL MODELING

4.1.1 Logistic regression model

For a given section of freeway let Y_i, denote "the presence or absence of a rear-end crash in that section during hour i," with Y_i=1 indicating that at least one rear-end crash occurred during hour i, while Y_i=0 indicates no rear-end crash during hour i. The variables Y_i are assumed to be Bernoulli outcomes with parameters π_i. That is

$$\Pr\{Y_i = \mathrm{y}\} = {\pi_i}^{Y_i}(1-\pi_i)^{1-Y_i} \text{ , y=0, 1.} \tag{2}$$

In order to capture the possible dependence of crash risk on other features, such as traffic or weather conditions, logistic regression was adopted.

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

$$\{\log\frac{P(Y_i=1)}{P(Y_i=0)}\} = \{\log\frac{P(Y_i=1)}{1-P(Y_i=1)}\} = \{\log\frac{\pi_i}{1-\pi_i}\} = \beta_0 + \textstyle\sum_{k=1}^{K}\beta_k x_{ki} \tag{3}$$

Where x_{ki} denotes the observed value of predictor k associated with hour i while the $\beta_k's$ are parameters to be estimated from data. Equation 3 can also be written as:

$$\{\mathrm{P}(Y_i = 1)\} = \pi_i = \frac{\exp(\beta_0+\sum_{k=1}^{K}\beta_k x_{ki})}{1+\exp(\beta_0+\sum_{k=1}^{K}\beta_k x_{ki})} \tag{4}$$

4.1.2 Parameter Estimation

Maximum Likelihood Estimation (MLE) was used to estimate the coefficients, namely the β_k's, in the logistic regression model presented in equation (4). The likelihood function is the probability of obtaining the given sample data as a function of the parameters:

$$\mathrm{L}(\beta_0,\dots,\beta_K) = \textstyle\prod_{i=1}^{n}{\pi_i}^{Y_i}(1-\pi_i)^{1-Y_i} \tag{5}$$

Substituting (4) into (5) and taking the natural logarithm gives

$$l(\beta_0, \dots, \beta_K) = \sum_{i=1}^{n} Y_i(\beta_0 + \sum_{k=1}^{K} \beta_k x_{ki}) - \sum_{i=1}^{n} \log[1 + \exp(\beta_0 + \sum_{k=1}^{K} \beta_k x_{ki})] \qquad (6)$$

The estimates for $(\beta_0, \dots, \beta_K)$ are then the values that maximize the log-likelihood function in (6).

The interpretation of the coefficients $\beta_k' s$ are as follows:

i) The intercept β_0 represents the logarithm of the odds $Y_i = 1$ (i.e. of a crash occurring during hour i) when all other conditions are fixed at zero.

ii) The slope β_k, represents the change in the logarithm of the odds $Y_i = 1$ given a change in predictor X_k. If X_k is a binary variable, β_k represents the change in the logarithm of the odds when X_k changes from 0 to 1. If X_k is continuous, β_k represents the change associated with one-unit increase in the value of X_k.

4.1.3 Goodness of fit as Hosmer-Lemeshow Test

The Hosmer-Lemeshow (H-L) test is widely used as a goodness of fit test for logistic regression models (Hosmer and Lemeshow 2000). The test can be described as follows:

i) Null and alternative hypotheses:

H_0: The proposed logistic regression model fits the data

H_1: The proposed logistic regression 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 are first grouped into G groups in order of the model's predicted probabilities that Y_i=1. Then, the H-L test statistic is calculated by the formula below:

$$G_{H-L}^2 = \sum_{g=1}^{G} \frac{(O_g - E_g)^2}{E_g(1 - E_g/N_g)} \qquad (7)$$

Where

G = the number of subgroups

O_g = observed number of events in the gth group

E_g = expected number of events in the gth group

N_g = the number of observations in the gth group

Given the null hypothesis, for large samples the test statistic, G_{H-L}^2, can be approximated by a Chi-square distribution with (G-2) degrees of freedom.

iii) Decision rule:

If, given the null hypothesis, the computed value of the test statistic is unlikely, that is Pr ($\chi^2_{G-2} > G^2_{H-L}$) ≤ α where α where α is the significance level, for example the conventional 0.05, the null hypothesis is rejected and we infer that the model does not fit the data.

If Pr ($\chi^2_{G-2} > G^2_{H-L}$) > α, we cannot reject the null hypothesis, although this does not prove that the null hypothesis is true.

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.

4.1.4 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:

$$G^2(M_0|M_1) = -2(L_0 - L_1) \tag{8}$$

Where

L_0 = the log-likelihood from M_0

L_1 = the 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:

If, given the null hypothesis, the computed value of the test statistic is unlikely, that is Pr ($\chi^2_k > G^2(M_0|M_1)$) ≤ α where α is the significance level, the null hypothesis is rejected and we conclude that the reduced (simpler) model, M0, is inferior to the more complex model.

If Pr ($\chi^2_k > G^2(M_0|M_1)$) > α, we cannot reject the null hypothesis, and 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.

4.2 LOGISTIC REGRESSION RESULTS FOR 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 or absence of a crash was the dependent variable while the 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 4. 1 is a list of variables involved in logistic regression model for sections on I-35W from TH-13 to I-494.

Table 4. 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-ending crash during hour i 1 = at least one rear-ending 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	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	Standard deviation of lane occupancy 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 following sections, N6, N9-1, N9-2, N11, N17, N18, S6, S9, S11, S17, and S18, the generalized linear model routine glm, implemented in the statistical analysis software R (R, 2015) was used to fit and evaluate logistic regression models containing different combinations of the independent variables listed in **Table 4. 1**. It turned out the numbers of rear-end crashes in sections N6, N9-2, N11, S6, S11, and S18 were insufficient to obtain reliable parameter estimates. 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 4. 2** is the estimation summary for those sections with sufficient crash experience to support analysis, using logistic regression models containing predictors X_1-X_4 and X_6-X_7.

Table 4. 2 Estimation Summaries for Initial Models 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	-15.770	4.885	-3.228	1.25E-03	**
	Rainy	-1.48E-01	4.77E-01	-3.10E-01	7.56E-01	
	Snowy	1.63E-01	5.77E-01	2.83E-01	7.78E-01	
	logvph	2.156	1.385	1.557	1.20E-01	
	Lane Occupancy	9.14E-02	4.77E-02	1.916	5.54E-02	.
	Lane Occupancy^2	--	--	--	--	
	Occupancy Standard Deviation	2.15E-01	7.08E-02	3.034	2.41E-02	**
	Before/After	2.84E-01	4.03E-01	7.06E-01	4.80E-01	
	Null deviance = 491.080			H-L = 12.800, p-value = 1.19 E-01		
	Residual deviance = 395.400			AIC: 409.400		
N17	Constant	-20.962	4.869	-4.305	1.67E-05	***
	Rainy	1.45E-01	3.47E-01	4.17E-01	6.76E-01	
	Snowy	1.28E-01	6.10E-01	2.10E-01	8.34E-01	
	logvph	3.852	1.377	2.797	5.16E-03	**
	Lane Occupancy	-2.25E-02	5.16E-02	4.35E-01	6.64E-01	
	Lane Occupancy^2	--	--	--	--	
	Occupancy Standard Deviation	3.02E-01	6.08E-02	4.978	6.43E-07	***
	Before/After	-1.81E-01	3.65E-01	-4.95E-01	6.21E-01	
	Null deviance =781.210			H-L =24.427, p-value =1.94E-03		
	Residual deviance = 648.900			AIC: 662.900		
N18	Constant	-16.547	3.219	-5.141	2.73E-07	***
	Rainy	4.83E-01	2.22E-01	2.172	2.99E-02	*
	Snowy	1.57E-02	4.38E-01	3.60E-02	9.71E-01	
	logvph	2.775	9.29E-01	2.986	2.83E-03	**
	Lane Occupancy	1.14E-01	2.07E-01	5.517	3.44E-08	***
	Lane Occupancy^2	--	--	--	--	
	Occupancy Standard Deviation	7.93E-02	1.51E-02	5.242	1.59E-07	***
	Before/After	1.46E-01	2.21E-01	-6.60E-01	5.09E-01	
	Null deviance =1527.000			H-L = 15.730, p-value = 4.64E-02		
	Residual deviance = 1270.200			AIC: 1284.200		
S9	Constant	-14.241	3.097	-4.599	4.24E-06	***
	Rainy	1.417	6.18E-01	2.291	2.20E-02	*
	Snowy	2.268	6.00E-01	3.795	1.48E-04	***

	logvph	1.365	8.84E-01	1.545	1.22E-01	
	Lane Occupancy	-8.41E-02	6.29E-02	-1.337	1.81E-01	
	Lane Occupancy^2	--	--	--	--	
	Occupancy Standard Deviation	3.71E-01	7.91E-02	4.684	2.81E-06	***
	Before/After	1.040	5.42E-01	1.919	5.50E-02	.
	Null deviance = 339.040			H-L =3.780, p-value = 8.76 E-02		
	Residual deviance =278.910			AIC: 292.910		
S17	Constant	-35.725	9.352	-3.820	1.33E-04	***
	Rainy	-7.31E-03	6.36E-01	-1.10E-02	9.91E-01	
	Snowy	7.97E-01	7.59E-01	1.051	2.94E-01	
	logvph	7.685	2.628	2.924	3.46E-03	**
	Lane Occupancy	1.02E-01	5.59E-02	1.814	6.96E-02	.
	Lane Occupancy^2	--	--	--	--	
	Occupancy Standard Deviation	1.35E-01	8.46E-02	1.594	1.11E-01	
	Before/After	4.77E-01	4.74E-01	1.006	3.14E-01	
	Null deviance = 462.660			H-L = 76.381, p-value = 2.61E-13		
	Residual deviance = 358.760			AIC: 372.760		

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

Here we use N9-1 as an example to see how to interpret the information provided in **Table 4. 2**:

i) Intercept β_0

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.

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

Table 4. 2 shows that the maximum likelihood estimate of the coefficient β_1 was -0.148 and the standard error associated with this estimate was 0.477. A test of the null hypothesis $\beta_1 = 0$ yielded a z-statistic equal to -0.148/0.477 = -0.31, and the probability of obtaining a z-value at least this large when the null hypothesis was true was 0.463. Using the rule that we cannot reject a null hypothesis when a P-value is greater than 0.05, the rainy condition does not show a statistically significant association with rear-end crash probability. Similar conclusions apply 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 , associated with lane occupancy, was 0.091 and the standard error associated with this estimate was 0.048. 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 0.055. This is slightly higher than the cutoff of 0.05 which we interpret as marginal evidence for an association between lane occupancy and rear-end crash risk. The positive value of estimate of β_4 indicates that rear-end crash probability increases as lane occupancy increases. The maximum likelihood estimate of the coefficient β_6 , associated with the standard deviation of lane occupancy, was 0.215 and the standard error associated with this estimate was 0.071. 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 0.002. 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, and the positive value of the estimate of β_6 means that rear-end crash probability increases as the standard deviation of lane occupancy increases.

iii) Goodness of fit

The null and residual deviance is 491.08 and 395.40, respectively. The H-L 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 deciles at 0.05 significance level.

As indicated in Chapter 1, there are both empirical and theoretical reasons to consider an "inverted U" shaped relationship between lane occupancy and crash probability, with a maximal point where crash probability is greatest, and falling off for lane occupancies both less than and greater than this maximal point. On way to allow for this is to include the square of lane occupancy, the variable X_5, as a predictor. **Table 4. 3** shows estimation summaries for models which include this predictor.

Table 4. 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	-6.905	6.686	-1.033	3.02E-01	
	Rainy	-2.46 E-01	4.79E-01	-5.14E-01	6.07E-01	
	Snowy	-3.73E-01	5.97E-01	-6.24E-01	5.33E-01	
	logvph	-5.02E-01	1.983	-2.53E-01	8.00E-01	
	Lane Occupancy	3.00E-01	1.47E-01	2.033	4.20E-02	*
	Lane Occupancy^2	-5.54E-03	3.97E-03	-1.397	1.62E-01	
	Occupancy Standard Deviation	1.38E-01	8.49E-02	1.624	1.04E-01	
	Before/After	2.30E-01	4.02E-01	5.72E-01	5.67E-01	
	Null deviance = 491.080			H-L = 11.359, p-value = 1.82 E-01		
	Residual deviance = 393.260			AIC: 409.260		
N17	Constant	-5.541	5.431	-1.020	3.08E-01	
	Rainy	1.54E-01	3.44E-01	4.46E-01	6.56E-01	
	Snowy	-5.91E-02	6.14E-01	-9.60E-02	9.23E-01	
	logvph	-9.35E-01	1.654	-5.65E-01	5.72E-01	
	Lane Occupancy	5.15E-01	1.89E-01	2.724	6.49E-03	**
	Lane Occupancy^2	-1.34E-02	4.96E-03	-2.710	6.72E-03	**

	Occupancy Standard Deviation	3.81E-02	1.11E-01	3.43E-01	7.32E-01	
	Before/After	9.83E-01	5.35E-01	1.838	6.61E-02	.
	Null deviance = 781.210			H-L = 12.799, p-value = 1.19E-01		
	Residual deviance = 640.440			AIC: 409.400		
N18	Constant	2.135	3.365	6.35E-01	5.26E-01	
	Rainy	3.44E-01	2.24E-01	1.536	1.25E-01	
	Snowy	-1.42E-01	4.33E-01	-3.29E-01	7.42E-01	
	logvph	-2.963	1.034	-2.865	4.16E-03	**
	Lane Occupancy	6.59E-01	1.13E-01	5.834	5.41E-09	***
	Lane Occupancy^2	-2.17E-02	4.50E-03	-4.834	1.34E-06	***
	Occupancy Standard Deviation	-3.34E-02	4.08E-02	-8.18E-01	4.13E-01	
	Before/After	3.50E-01	2.67E-01	1.313	1.89E-01	
	Null deviance = 1527.000			H-L =24.427, p-value =1.94E-03		
	Residual deviance = 1238.000			AIC: 662.900		
S9	Constant	-3.803	4.326	-8.79E-01	3.79E-01	
	Rainy	1.292	6.08E-01	2.124	3.37E-02	*
	Snowy	2.096	6.02E-01	3.485	4.93E-02	***
	logvph	-1.985	1.353	-1.467	1.42E-01	
	Lane Occupancy	3.41E-01	1.67E-01	2.042	4.11E-02	*
	Lane Occupancy^2	-1.10E-02	5.308E-03	-2.070	3.85E-02	*
	Occupancy Standard Deviation	2.16E-01	8.82E-02	2.451	1.43E-02	*
	Before/After	1.779	6.13E-01	2.906	3.66E-03	**
	Null deviance = 339.040			H-L = 3.316, p-value = 9.10E-01		
	Residual deviance =270.990			AIC: 286.99		
S17	Constant	1.092	6.478	1.69E-01	8.66E-01	
	Rainy	-8.97E-02	6.31E-01	-1.42E-01	8.87E-01	
	Snowy	3.44E-01	7.63E-01	4.51E-01	6.52E-01	
	logvph	-3.174	1.909	-1.663	9.64E-02	.
	Lane Occupancy	8.70 E-01	1.95E-01	4.472	7.76E-06	***
	Lane Occupancy^2	-2.02E-02	5.01E-03	-4.028	5.62E-05	***
	Occupancy Standard Deviation	-2.18E-01	1.40E-01	-1.560	1.19E-01	
	Before/After	-6.20E-03	5.08E-01	-1.20E-02	.9.90E-01	
	Null deviance = 462.660			H-L =24.340, p-value = 2.01E-03		
	Residual deviance = 344.130			AIC: 360.130		

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

Table 4. 4 shows a summary for the model using only the statistically-significant predictors from **Table 4. 3**.

Table 4. 4 Estimation Summary for I-35W from TH-13 to I-494 for Model Using Only Statistically Significant Predictors from Table 4.3

Section No.	Variable	Estimate	Std. Error	z value	Pr (>|z|)	Signif. codes
N9-1	Lane Occupancy	2.99E-01	1.47E-01	2.033	4.20E-02	*
	Null deviance = 491.080			H-L = 11.359, p-value = 1.82E-01		
	Residual deviance = 393.260			AIC: 409.260		
N17	Lane Occupancy	5.15E-01	1.89E-01	2.722	6.49E-03	**
	Lane Occupancy^2	-1.34E-02	4.96E-03	-2.710	6.72E-03	**
	Before/After	9.83E-01	5.35E-01	1.838	6.61E-02	.
	Null deviance = 781.21 0			H-L = 12.799, p-value =1.19E-01		
	Residual deviance = 640.440			AIC: 409.4		
N18	Log Flow	-2.963	1.034	-2.865	4.16 E-03	**
	Lane Occupancy	6.59E-01	1.13E-01	5.834	5.41E-09	***
	Lane Occupancy^2	-2.17E-02	4.50E-03	-4.834	1.34E-06	***
	Null deviance = 1527			H-L =24.427, p-value =1.94E-03		
	Residual deviance = 1238			AIC: 662.900		
S9	Rainy	1.292	6.08E-01	2.124	3.37E-02	*
	Snowy	2.096	6.02E-01	3.485	4.93E-04	***
	Lane Occupancy	3.41E-01	1.67E-01	2.042	4.11E-02	*
	Lane Occupancy^2	-1.10E-02	5.31E-03	-2.070	3.85E-02	*
	Occupancy Standard Deviation	2.16E-01	8.82E-02	2.451	1.43E-02	*
	Before/After	1.780	6.13E-01	2.906	3.66E-03	**
	Null deviance = 339.04			H-L = 3.315, p-value = 9.13E-01		
	Residual deviance =270.990			AIC: 286.990		
S17	Log Flow	-3.174	1.909	-1.663	9.64E-02	.
	Lane Occupancy	8.70E-01	1.95E-01	4.472	7.76E-06	***
	Lane Occupancy^2	-2.02E-02	5.01E-03	-4.028	5.62E-05	***
	Null deviance = 462.660			H-L =24.340, p-value = 2.01E-03		
	Residual deviance = 344.130			AIC: 360.130		

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

Among all studied sections on I-35W between TH-13 to I-494, there is a statistically significant association between Before/After predictor and rear-end crash risk for Section S9, and for Section N17 the coefficient of the Before/After predictor was marginally significant. In Section N17 the frequency of rear-ending crashes actually decreased from 30 in the before period to 21 in the after period (see Table 2.6) but this decrease was smaller than what the change in lane occupancy would predict. In Section S9 the frequency of clearly locatable rear-ending crashes went from 6 in the before period to 13 in the after period and this change was statistically significant. Before identifying this as a causal effect of the UPA improvements, however, one should note that Table 2.5 suggests that this difference could also be due to poor identification of vehicle directions in the Before period.

As is shown in **Table 4. 4**, the two coefficients associated with lane occupancy, β_4 and β_5, are both 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 suggested by **Figure 1. 5**.

L-R tests were conducted to determine the better model between a reduced model M0 which does not have a quadratic term for lane occupancy, and the more complex model M1, which includes a quadratic term of lane occupancy. **Table 4. 5** shows the L-R test results for each analyzed section between TH-13 and I-494

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

Section No.	Model	Number of coefficients	Log-likelihood	p	$G^2(M_0 \mid M_1)$	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	-324.45	1	8.460	3.63E-03	**
	M_1	8	-320.22				
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 can 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

Using the estimates in **Table 4. 3**, rear-end crash probabilities for N17, N18, S9, and S17 are maximized when lane occupancy is approximately equal to

i) N17: $o_{max} \approx \bar{o} - \frac{\widehat{\beta_1}}{2\widehat{\beta_2}} = 5.19 - \frac{0.515}{2(-0.0134)} = 24.41$

ii) N18: $o_{max} \approx \bar{o} - \frac{\widehat{\beta_1}}{2\widehat{\beta_2}} = 6.09 - \frac{0.659}{2(-0.0217)} = 21.27$

iii) S9: $o_{max} \approx \bar{o} - \frac{\widehat{\beta_1}}{2\widehat{\beta_2}} = 5.04 - \frac{0.341}{2(-0.0110)} = 20.54$

iv) S17: $o_{max} \approx \bar{o} - \frac{\widehat{\beta_1}}{2\widehat{\beta_2}} = 5.89 - \frac{0.870}{2(-0.0202)} = 27.42$

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 4. 1 to **Figure 4. 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 line denoting the change point between the Before and After periods. The horizontal line shows the approximate value of average lane occupancy where rear-end crash risk is maximal.

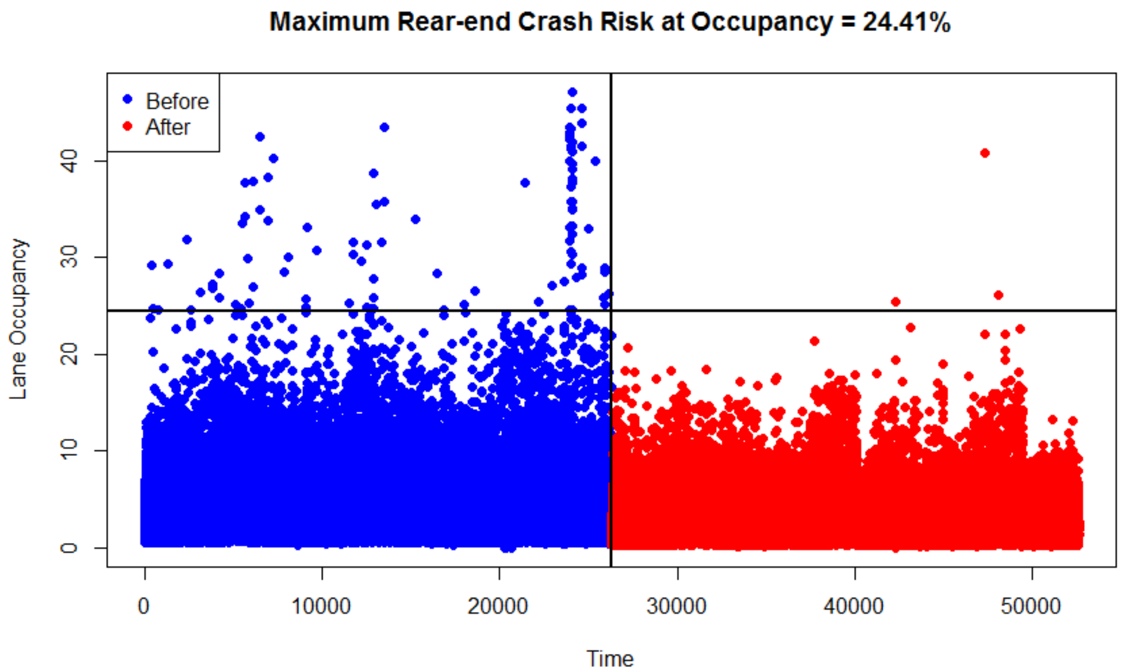

Figure 4. 1 Time-Series Plot of Hourly Average Lane Occupancy for Section N17, Showing the Before and After UPA Periods, and the Average Lane Occupancy with Maximal Crash Probability.

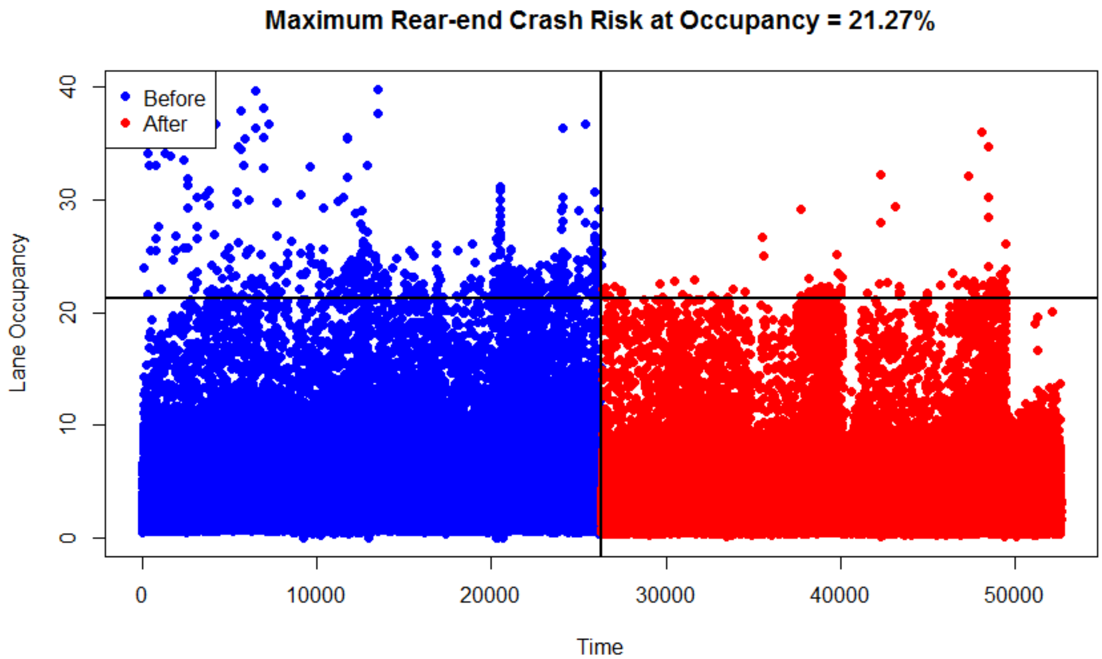

Figure 4. 2 Time-Series Plot of Hourly Average Lane Occupancy for Section N18, Showing the Before and After UPA Periods, and the Average Lane Occupancy with Maximal Crash Probability.

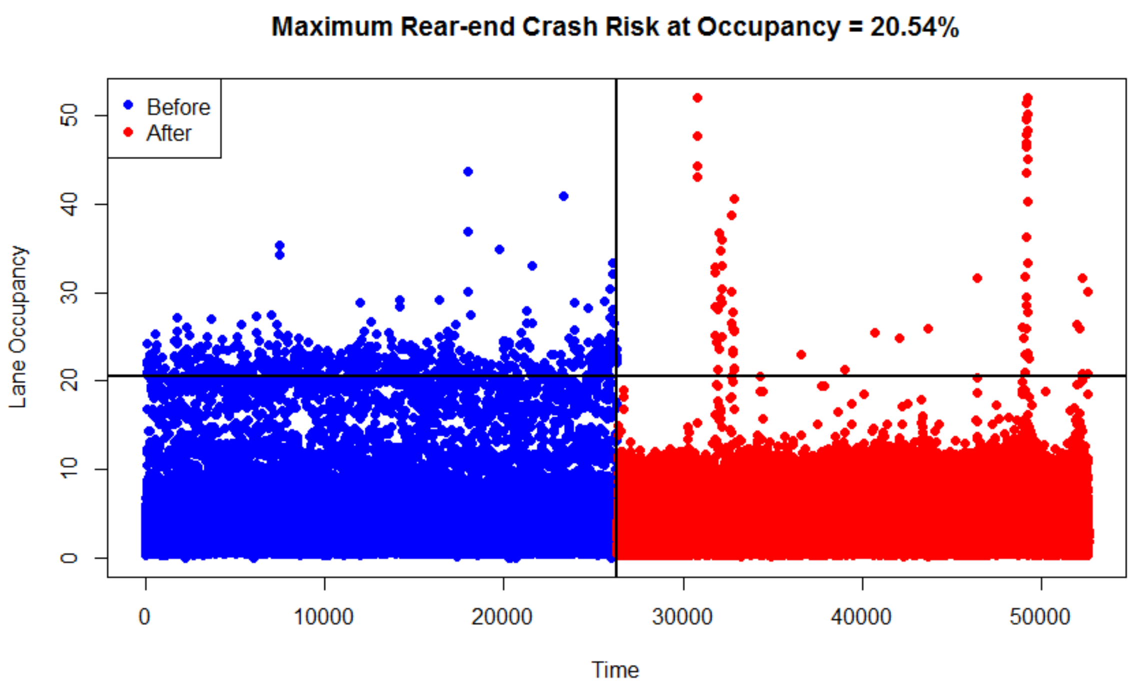

Figure 4. 3 Time-Series Plot of Hourly Average Lane Occupancy for Section S9, Showing the Before and After UPA Periods, and the Average Lane Occupancy with Maximal Crash Probability.

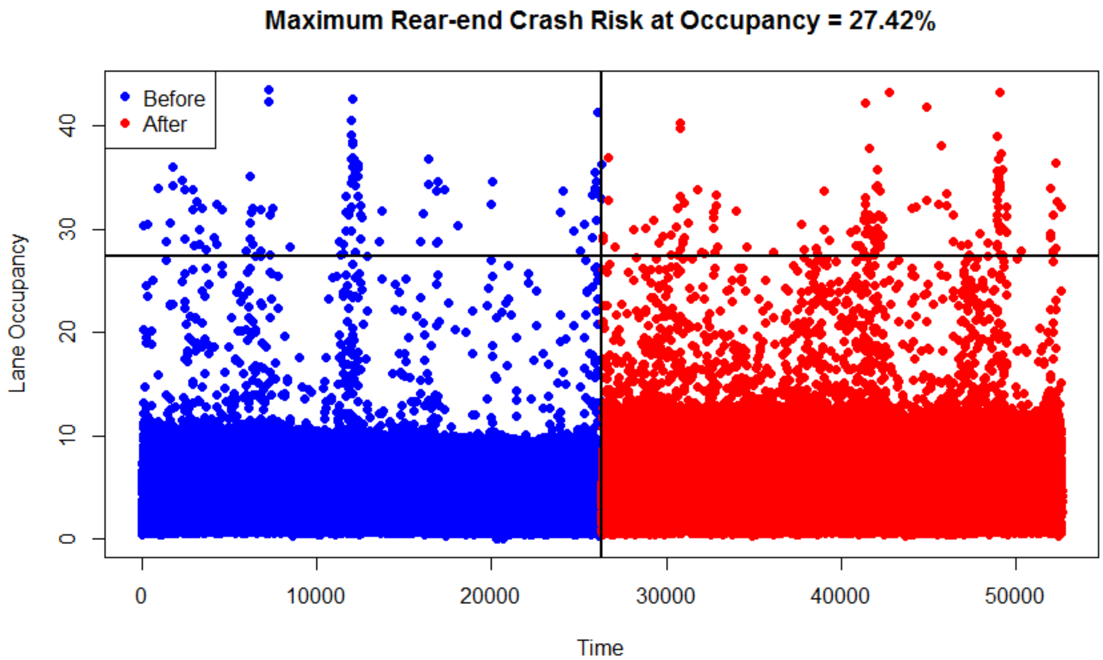

Figure 4. 4 Time-Series Plot of Hourly Average Lane Occupancy for Section S17, Showing the Before and After UPA Periods, and the Average Lane Occupancy with Maximal Crash Probability.

In summary, most analyzed sections on I-35W from TH-13 to I-494 showed no significant change in rear-end crash risk associated with UPA project. The exceptions were northbound N17 and southbound S9. 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. The apparent change in rear-end crash risk on S9 (just north of Minnesota River) was possibly due to ambiguous locations of crashes during the Before period. In addition, an "Inverted U" relationship between lane occupancy and crash risk was seen in several sections

4.3 LOGISTIC REGRESSION RESULTS FOR PDSL REGION

Analyses similar to those done for the TH-13 to I-494 region were also done the region of I-35W affected by PDSL operation. For each hour during 2006-2008 and 2011-2013, the presence or 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, a Before or After period indicator, and the variables characterizing the status of PDSL operation. **Table 4. 6** is the list of variables involved in the logistic regression model for I-35W PDSL sections.

Table 4. 6 Variables Selected for Logistic Regression Analysis for I-35W PDSL Sections

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-ending crash during hour i 1 = at least one rear-ending 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	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	Standard deviation of lane occupancy 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 the statistical analysis package R was used to fit and evaluate logistic regression models containing different combinations of independent variables listed in **Table 4. 6**. To start, **Table 4. 7** presents results using a set of predictors similar to that used in the initial analyses done for the HOT region. That is, with the square of lane occupancy deleted and a simple Before/After effect for the PDSL.

Table 4. 7 Estimation Summary for Initial Model of Rear-End Crash Probability on I-35W PDSL Sections

Section No.	Variable	Estimate	Std. Error	z value	Pr (>|z|)	Signif. codes
N37	Constant	-7.29E+00	2.74E-01	-26.616	< 2.00E-16	***
	Rainy	-2.30E-02	3.57E-01	-6.40 E-02	9.49E-01	
	Snowy	4.99E-02	4.73E-01	1.05E-01	9.160	
	logvph	-5.00E-05	9.41E-04	-5.30E-02	9.58E-01	
	Lane Occupancy	1.45E-01	3.25E-02	4.471	7.78E-06	***
	Lane Occupancy^2	--	--	--	--	
	Occupancy Standard Deviation	5.50E-02	6.70E-02	8.21E-01	4.12	
	Before/After	6.37E-02	3.54E-01	1.80E-01	8.57E-01	
	Null deviance = 1075.530			H-L = 13.962, p-value = 8.28E-02		
	Residual deviance = 873.950			AIC: 887.950		
N38	Constant	-21.318	4.481	-4.758	1.96E-06	***

	Rainy	3.06E-01	3.25E-01	9.41E-01	3.47E-01	
	Snowy	-1.24E-01	5.22E-01	-2.38E-01	8.12	
	logvph	1.640	5.26E-01	3.117	1.83E-03	**
	Lane Occupancy	8.761E-02	3.53E-02	2.482	1.307E-02	*
	Lane Occupancy^2	--	--	--	--	
	Occupancy Standard Deviation	8.82E-02	5.54E-02	1.591	1.12E-01	
	Before/After	4.36E-01	3.10E-01	1.404	1.60E-01	
	Null deviance = 1178.500			H-L = 20.077, p-value = 1.01E-02		
	Residual deviance = 957.450			AIC: 971.450		
N40	Constant	-6.183	1.51E-01	-41.062	< 2.00E-16	***
	Rainy	-2.57E-01	2.41E-01	-1.066	2.86E-01	
	Snowy	2.05E-01	2.95E-01	6.96E-01	4.86E-01	
	logvph	-4.30E-05	6.20E-01	-6.90E-02	9.45E-01	
	Lane Occupancy	3.71E-02	2.08E-02	1.785	7.43E-02	.
	Lane Occupancy^2	--	--	--	--	
	Occupancy Standard Deviation	1.63E-01	3.17E-02	5.136	2.81E-07	***
	Before/After	8.41E-02	1.87E-01	4.50E-01	6.53E-01	
	Null deviance = 2437.000			H-L = 52.200, p-value = 1.54E-08		
	Residual deviance = 2069.100			AIC: 2083.100		

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

As an example of how to interpret the information provided in **Table 4. 7** consider section N37. First, the estimate for the constant term β_0 is equal to -7.290. In a situation where all model predictors were equal to zero the estimated probability of a rear-end crash occurring during that hour would be

$$\exp(-7.290)/(1+\exp(-7.290))=0.00068.$$

Next, **Table 4. 7** shows that, for section N37, the maximum likelihood estimate of the coefficient β_1, associated with the Rainy condition, was -0.023 and the standard error associated with this estimate was 0.357. 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 is true was 0.949. Using the rule that we cannot reject a null hypothesis when a P-value is greater than 0.05, this indicates that the rainy condition did not have a statistically significant association with rear-end crash probability. Similar conclusions apply to weather the variable Snowy, the traffic variables Log Flow and Occupancy Standard Deviation, and the Before/After predictor.

On the other hand, the maximum likelihood estimate of the coefficient β_4, associated with lane occupancy, was 0.145 and the standard error associated with this estimate was 0.033. 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, indicating a statistically significant association between lane occupancy and crash probability. The positive value of estimate of β_4 indicates that rear-end probability increases as lane occupancy increases. The H-L goodness of fit statistic was 13.962 with a p-value of 0.083, which indicates a marginally acceptable fit between the observed and predicted event frequencies.

Similar to what was done for the HOT region, the possibility that the relationship between lane occupancy and crash probability shows an "inverted-U" shape should be considered. As before the square of lane occupancy, variable X5, was included as a predictor and **Table 4. 8** shows estimation summaries for when this predictor is added to those used previously.

Table 4. 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	-5.994	4.976	-1.205	2.28E-01	
	Rainy	6.80E-02	3.55E-01	1.92E-01	8.48E-01	
	Snowy	-2.33E-02	4.79E-01	-4.90E-02	9.61E-01	
	logvph	-2.04E-01	5.99E-01	-3.40E-01	7.34E-01	
	Lane Occupancy	4.59E-01	8.68E-02	5.292	1.21E-07	***
	Lane Occupancy^2	-9.86E-03	2.37E-03	-4.155	3.25E-05	***
	Occupancy Standard Deviation	-1.01E-02	7.11E-02	-1.42E-01	8.87E-01	
	Before/After	-6.69E-01	3.71E-01	-1.803	7.13E-02	.
	Null deviance = 1075.530			H-L = 9.413, p-value = 3.09E-01		
	Residual deviance = 833.710			AIC: 849.710		
N38	Constant	-3.457	4.683	-7.38E-01	4.60E-01	
	Rainy	2.40E-01	3.23E-01	7.43E-01	4.57E-01	
	Snowy	-2.88E-01	5.23E-01	-5.50E-01	5.82E-01	
	logvph	-5.63E-01	5.75E-01	-9.80E-01	3.27E-01	
	Lane Occupancy	4.50E-01	9.95E-02	4.525	6.03E-06	***
	Lane Occupancy^2	-1.01E-02	2.76E-03	-3.661	2.52E-04	***
	Occupancy Standard Deviation	-5.60E-02	6.31E-02	-8.87E-01	3.75E-01	
	Before/After	5.47E-01	2.96E-01	1.848	6.46E-02	.
	Null deviance = 1178.500			H-L = 3.997, p-value = 8.57E-01		
	Residual deviance = 944.620			AIC: 960.620		
N40	Constant	-6.49E+00	1.81E-01	-35.945	< 2.00E-16	***
	Rainy	-1.45E-01	2.39E-01	-6.06E-01	5.44E-01	
	Snowy	2.93E-01	2.95E-01	9.96E-01	3.19E-01	
	logvph	-9.54E-05	6.01E-03	-1.60E-02	9.87E-01	
	Lane Occupancy	2.53 E-01	3.66E-02	6.925	4.35E-12	***

	Lane Occupancy^2	-7.95E-03	1.13E-03	-7.022	2.19E-12	***
	Occupancy Standard Deviation	1.10E-01	3.07E-02	3.594	3.26E-04	***
	Before/After	2.18E-02	1.81E-01	1.21E-01	9.04E-01	
	Null deviance = 2437.000			H-L = 13.899, p-value = 8.44E-02		
	Residual deviance = 2003.600			AIC: 2019.600		

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

The change in rear-end crash risk associated with UPA project were marginal in Section N37 and N38 at 0.05 significance level. For N37, rear-end crash risk decreased after the UPA project, holding other conditions constant, while for N38, rear-end crash risk increased after UPA project, holding other conditions constant. None of three analyzed PDSL sections showed significant weather effects at 0.05 significance level. For all three sections the two coefficients associated with lane occupancy, β_4 and β_5, were 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 suggested by **Figure 1. 5**.

L-R tests were conducted to determine the better model between the reduced model M0, the model without quadratic term of lane occupancy, and the more complex model M1, the model with quadratic term of lane occupancy, for each analyzed section in I-35W PDSL region. **Table 4. 9** shows the L-R test results.

Table 4. 9 Likelihood Ratio Test Results for Analyzed Sections in I-35W PDSL Region

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

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

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

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 contains some hours when the PDSL is operating and other hours when it is 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 models with this subdivided After period is shown in **Table 4. 10**. These results are generally similar to those shown in **Table 4. 8** except that in Section N37, the difference between the coefficient of "PDSL Open" and 0 was marginal at 0.05 significance level. The negative sign of this coefficient indicates that in Section N37, during hours when the PDSL was open, rear-end crash risk was actually reduced, holding other conditions constant. The UPA project's effect on rear-end crash risk in Section N38 was no longer significant at 0.05 significance level.

Table 4. 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	-7.73E+00	3.19E-01	-24.259	< 2.00E-16	***
	Rainy	8.92E-02	3.54E-01	2.52E-01	8.01E-01	
	Snowy	-9.22E-03	4.77E-01	-1.90E-02	9.85E-01	
	logvph	-8.40E-05	6.79E-03	-1.20E-02	9.90E-01	
	Lane Occupancy	4.55E-01	6.44E-02	7.057	1.70E-12	***
	Lane Occupancy^2	-9.74E-03	1.83E-03	-5.314	1.07E-07	***
	Occupancy Standard Deviation	-3.02E-03	6.470	-4.70E-02	9.63E-01	
	Before/After	--	--	--	--	
	PDSL Closed	-3.04E-01	5.36E-01	-5.67E-01	5.71E-01	
	PDSL Open	-7.87E-01	3.81E-01	-2.065	3.89E-01	*
	Sign Dark	-3.66E-01	1.150	-3.19E-01	7.50E-01	
	VSA	-8.55E-01	1.570	-5.45E-01	5.86E-01	
	Null deviance = 1075.530			H-L = 7.806, p-value =4.53E-01		
	Residual deviance = 832.880			AIC: 854.880		
N38	Constant	-4.329	5.476	-7.91E-01	4.29E-01	
	Rainy	2.61E-01	3.24E-01	8.04 E-01	4.21E-01	
	Snowy	-3.23E-01	5.31E-01	-6.09E-01	5.43E-01	
	logvph	-4.60E-01	6.67E-01	-6.91E-01	4.90E-01	
	Lane Occupancy	4.42E-01	1.04E-01	4.239	2.25E-05	***
	Lane Occupancy^2	-9.98E-03	2.88E-03	-3.472	5.17E-04	***
	Occupancy Standard Deviation	-4.66E-02	6.66E-02	-7.00 E-01	4.84 E-01	
	Before/After	--	--	--	--	
	PDSL Closed	7.09E-01	4.82E-01	1.471	1.41E-01	
	PDSL Open	4.75E-01	3.21E-01	1.478	1.39E-01	
	Sign Dark	1.289	9.79E-01	1.316	1.88E-01	
	VSA	1.090	1.051	1.037	3.00E-01	
	Null deviance = 1178.500			H-L = 4.347, p-value = 8.25E-01		
	Residual deviance = 943.480			AIC: 965.480		
N40	Constant	-6.560	1.82E-01	-35.987	< 2.00E-16	***
	Rainy	-1.39E-01	2.39E-01	-5.81E-01	5.61E-01	
	Snowy	2.84E-01	2.98E-01	9.52E-01	3.41E-01	

	logvph	-5.59E-05	2.60E-03	-2.10E-02	9.83E-01	
	Lane Occupancy	2.67E-01	3.75E-02	7.124	1.05E-12	***
	Lane Occupancy^2	-8.28E-03	1.15E-03	-7.171	7.46E-13	***
	Occupancy Standard Deviation	1.12E-01	3.05E-02	3.663	2.49E-04	***
	Before/After	--	--	--	--	
	PDSL Closed	4.13E-01	2.81E-01	1.467	1.43E-01	
	PDSL Open	-8.43E-02	1.93E-01	-4.37E-01	6.62E-01	
	Sign Dark	-3.590	4.270	-8.40E-01	4.01E-01	
	VSA	1.60E-01	7.98E-01	2.01E-01	8.41E-01	
	Null deviance = 2437.000			H-L = 7.4029, p-value = 4.94E-01		
	Residual deviance = 1998.700			AIC: 2020.700		

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

Table 4. 11 shows a summary for the model using only the statistically-significant predictors from **Table 4. 10.**

Table 4. 11 Estimation Summary for Analyzed PDSL Sections for Model Using Only Statistically Significant Predictors from Table 4.10

Section No.	Variable	Estimate	Std. Error	z value	Pr (>|z|)	Signif. codes
N37	Constant	-7.73E+00	3.19E-01	-24.259	< 2.00E-16	***
	Lane Occupancy	4.55E-01	6.44E-02	7.057	1.70E-12	***
	Lane Occupancy^2	-9.74E-03	1.83E-03	-5.314	1.07E-07	***
	PDSL Open	-7.87E-01	3.81E-01	-2.065	0.0389	*
	Null deviance = 1075.530			H-L = 7.806, p-value = 4.53E-01		
	Residual deviance = 832.880			AIC: 854.880		
N38	Lane Occupancy	4.42E-01	1.04E-01	4.239	2.25	***
	Lane Occupancy^2	-9.98E-03	2.88E-03	-3.472	5.17E-04	***
	Null deviance = 1178.500			H-L = 4.347, p-value = 8.25E-01		
	Residual deviance = 943.480			AIC: 965.480		
N40	Constant	-6.560	1.82E-01	-35.987	< 2.00E-16	***
	Lane Occupancy	2.67E-01	3.75E-02	7.124	1.05E-12	***
	Lane Occupancy^2	-8.28E-03	1.15E-03	-7.171	7.46E-13	***
	Occupancy Standard Deviation	1.12E-01	3.05E-02	3.663	2.49E-04	***
	Null deviance = 2437.000			H-L = 7.403, p-value = 4.94 E-01		
	Residual deviance = 1998.700			AIC: 2020.700		

Section N37 showed a significant negative association between the operation of the 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 the After period was $e^{-0.787} = 0.455$ times the odds of having a rear-end crash in a given hour in Before period, 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 PDSL 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 4. 10**, rear-end crash probability for N37, N38, and N40 was maximized when lane occupancy was approximately equal to

i) N37: $o_{max} \approx \bar{o} - \frac{\widehat{\beta_1}}{2\widehat{\beta_2}} = 6.22 - \frac{0.455}{2(-.00974)} = 29.58$

ii) N38: $o_{max} \approx \bar{o} - \frac{\widehat{\beta_1}}{2\widehat{\beta_2}} = 7.26 - \frac{0.442}{2(-.00998)} = 29.40$

iii) N40: $o_{max} \approx \bar{o} - \frac{\widehat{\beta_1}}{2\widehat{\beta_2}} = 7.43 - \frac{0.267}{2(-.00828)} = 23.55$

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

Figure 4. 5 to **Figure 4. 7** show the 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 line denoting the change point. The horizontal lines show the approximate values of average lane occupancy where rear-end crash risk was 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 risk for all three sections.

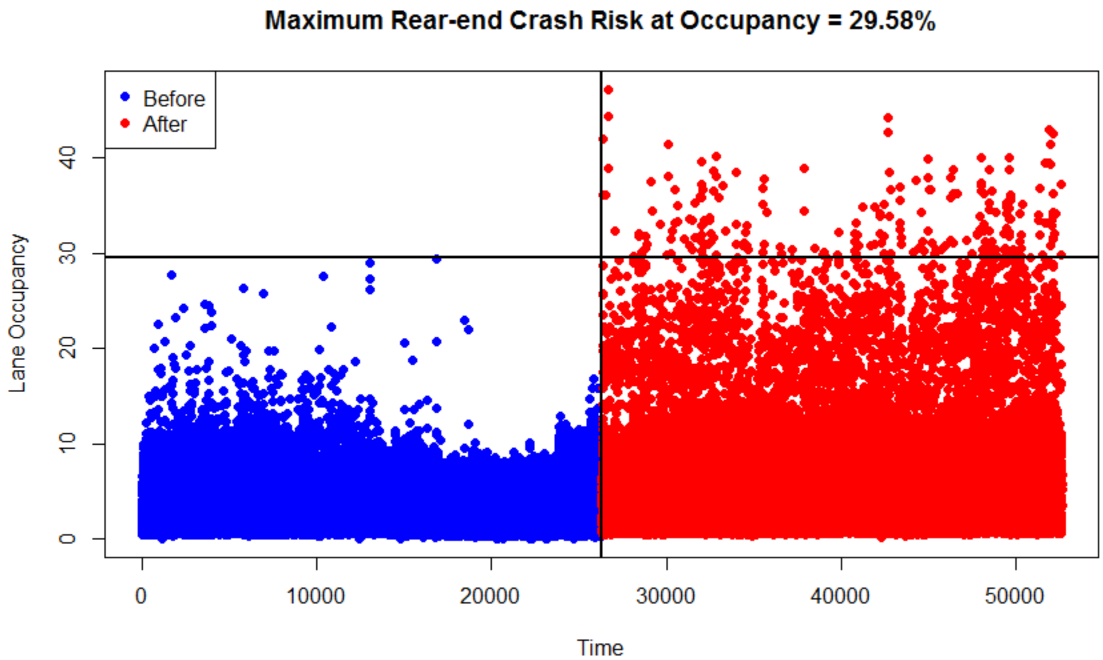

Figure 4. 5 Time-Series Plot of Hourly Average Lane Occupancy for Section N37, Showing the Before PDSL and After PDSL Periods, and the Average Lane Occupancy with Maximal Crash Probability

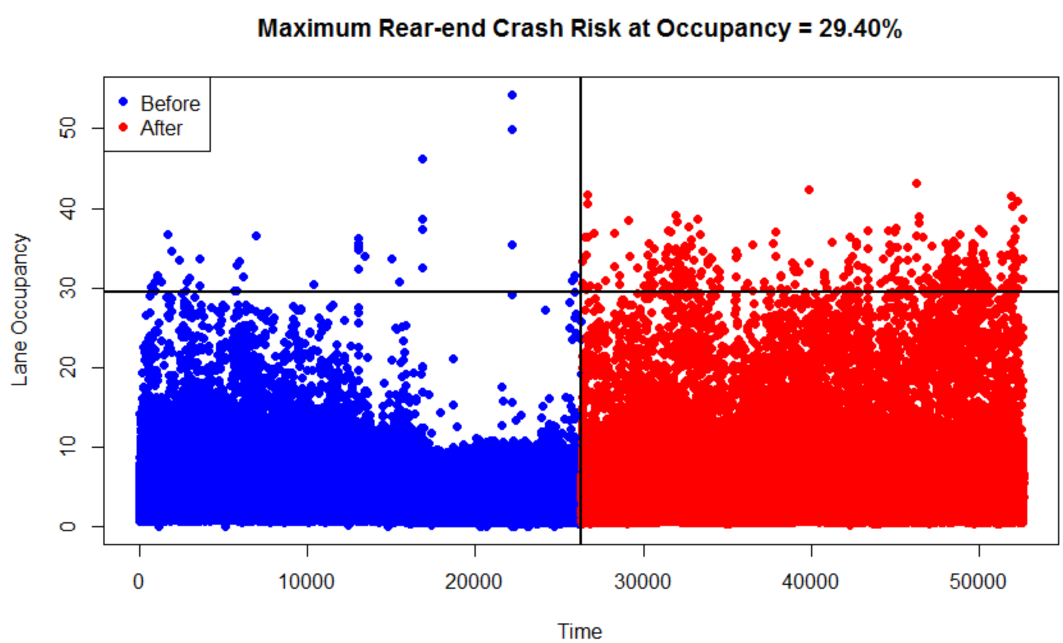

Figure 4. 6 Time-Series Plot of Hourly Average Lane Occupancy for Section N38, Showing the Before PDSL and After PDSL Periods, and the Average Lane Occupancy with Maximal Crash Probability

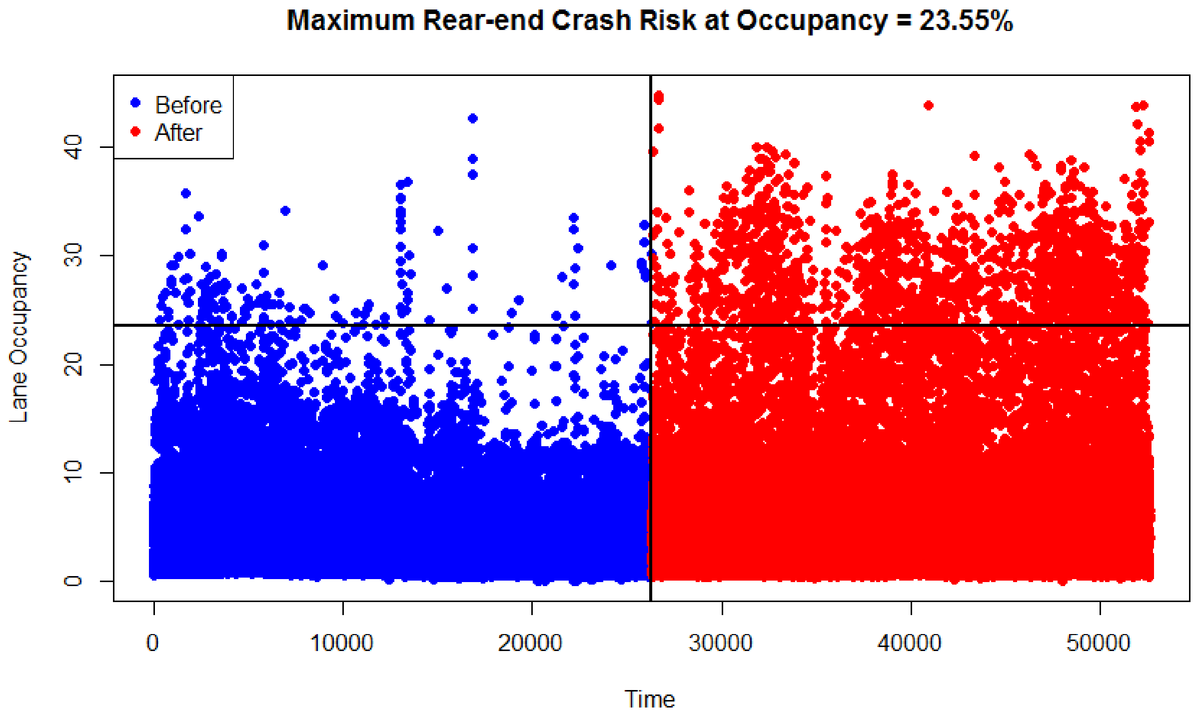

Figure 4. 7 Time-Series Plot of Hourly Average Lane Occupancy for Section N40, Showing the Before PDSL and After PDSL Periods, and the Average Lane Occupancy with Maximal Crash Probability

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 TH-62/ I-35W bottleneck.

REFERENCES

1. AASHTO (2010). *Highway Safety Manual*, 1st edition. American Association of State Highways and Transportation Officials Transportation Research Board of the National Academies, Washington, DC.
2. 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.
3. Agresti, A. (2014). *Categorical Data Analysis*. Hoboken: Wiley.
4. Brill, E (1972). A car-following model relating reaction times and temporal headways to accident frequency*, Transportation Science*, 6, 343-353.
5. Cao, J., Munnich., Liu, H., He., S., Xu, Z., Huang, A., and Munnich, L*. Benefit and Cost Analysis of the I-394 MnPASS Program*, Report CTS 12-03, Intelligent Transportation Systems Institute, University of Minnesota, 2012.
6. 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.
7. Chatterjee, I., & Davis, G. A. (2016). Analysis of Rear-End Events on Congested Freeways by Using Video-Recorded Shock Waves*. Transportation Research Record*, 2583, 110-118.
8. Dobson, A. J., & Barnett, A. (2008). *An Introduction to Generalized Linear Models*. CRC press.
9. Hosmer, D., and Lemeshow, S. (2000). *Applied Logistic Regression*, 2nd edition, Wiley.
10. Hourdos, J. (2005). *Crash Prone Traffic Flow Dynamics: Identification and Real-Time Detection*, Ph.D. Thesis, University of Minnesota.
11. 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*, 1968, 83-91.
12. Newell, G (1993). A simplified theory of kinemaic waves in highway traffic, part I: General theory, *Transportation Research*, Part B, 27, 281-287.
13. Oh, C., Oh, J-S, Ritchie, S., and Chang, M. (2005). Real-time estimation of freeway accident likelhood for Safety Enhancement. *Journal of Transportation Engineering*, 131(5), 358-363.
14. 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.
15. Team, R. (2015). *RStudio: integrated development for R. RStudio*, Inc., Boston, MA URL http://www. rstudio. com.
16. 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).
17. 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.
18. Yeo, H., Jang, K., Skabardonis, A., and Kang., S. Impact of traffic states on freeway crash involvement rates, *Accident Analysis and Prevention*, 50, 713-723, 2013.

APPENDIX A: 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

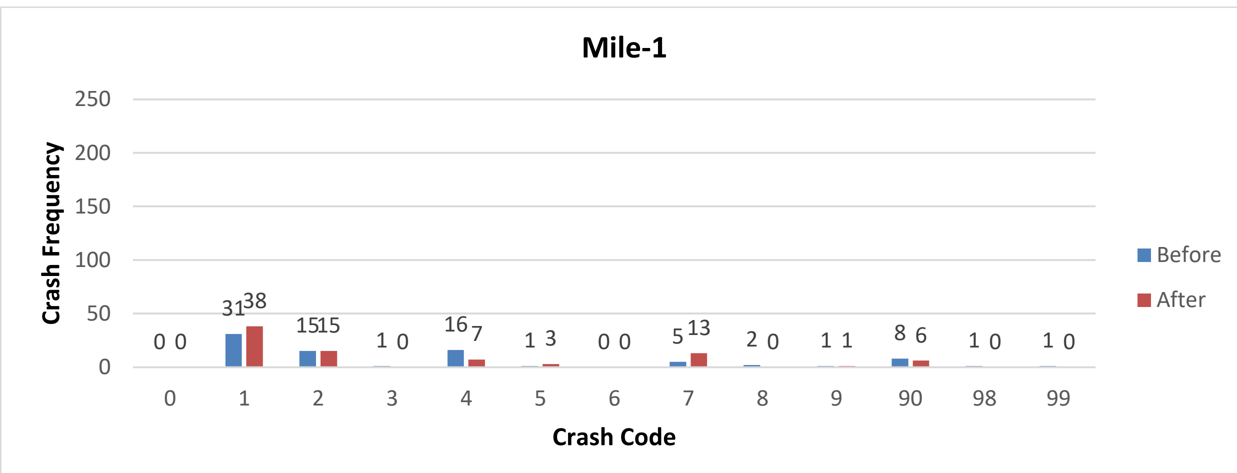

Table A. 2 Crash Summary Table for Section Mile-2

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

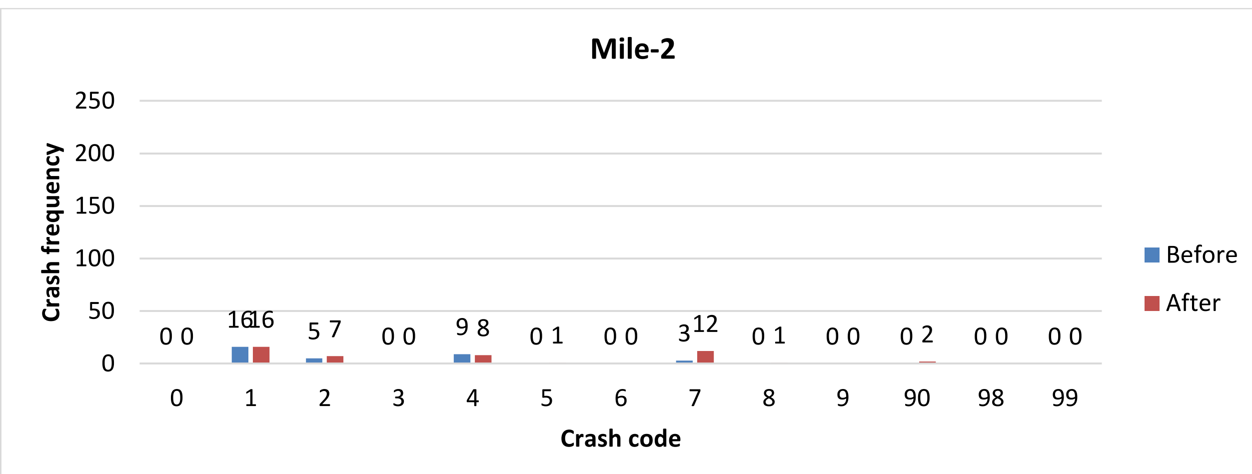

Table A. 3 Crash Summary Table for Section Mile-3

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

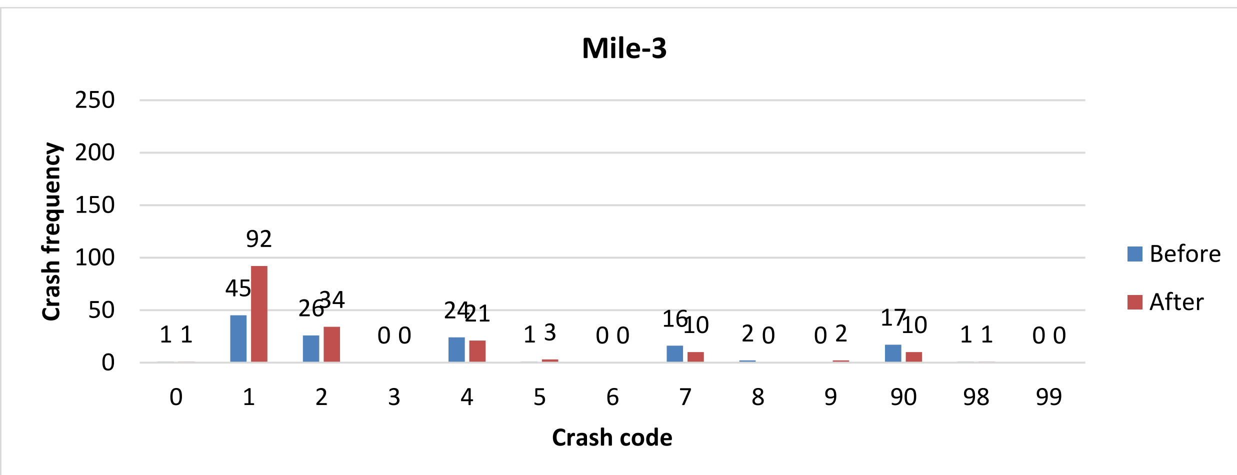

Table A. 4 Crash Summary Table for Section Mile-4

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

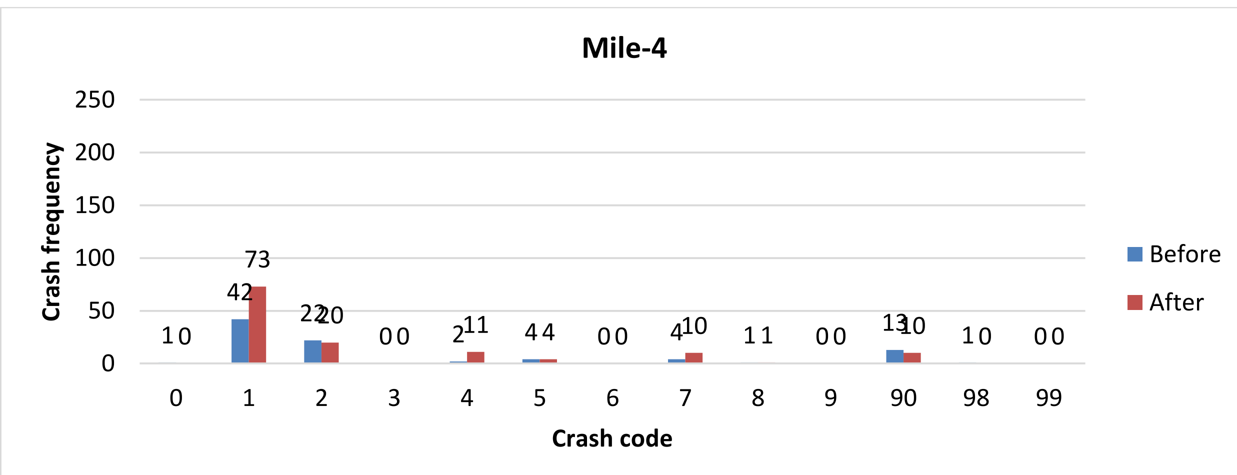

Table A. 5 Crash Summary Table for Section Mile-5

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

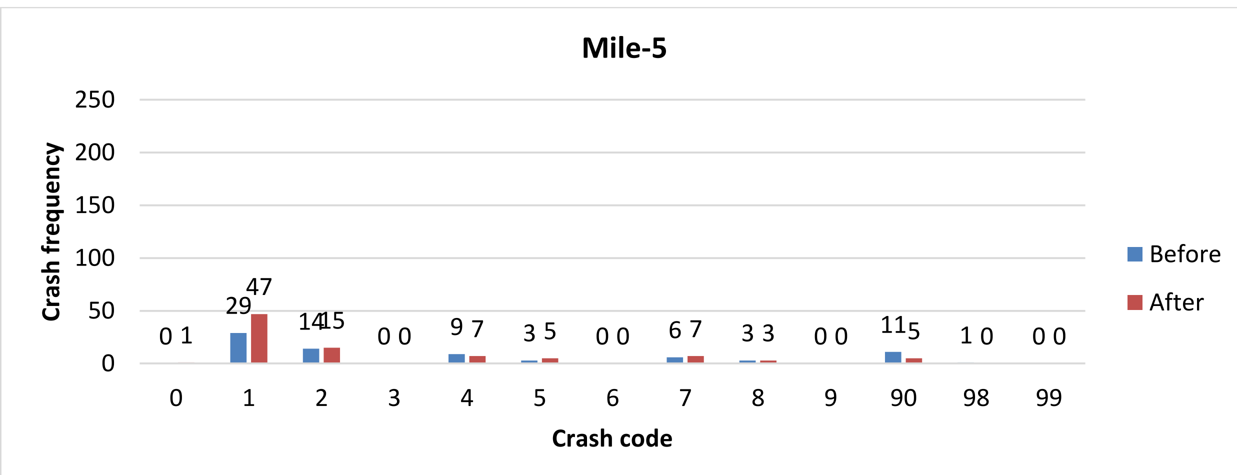

Table A. 6 Crash Summary Table for Section Mile-6

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

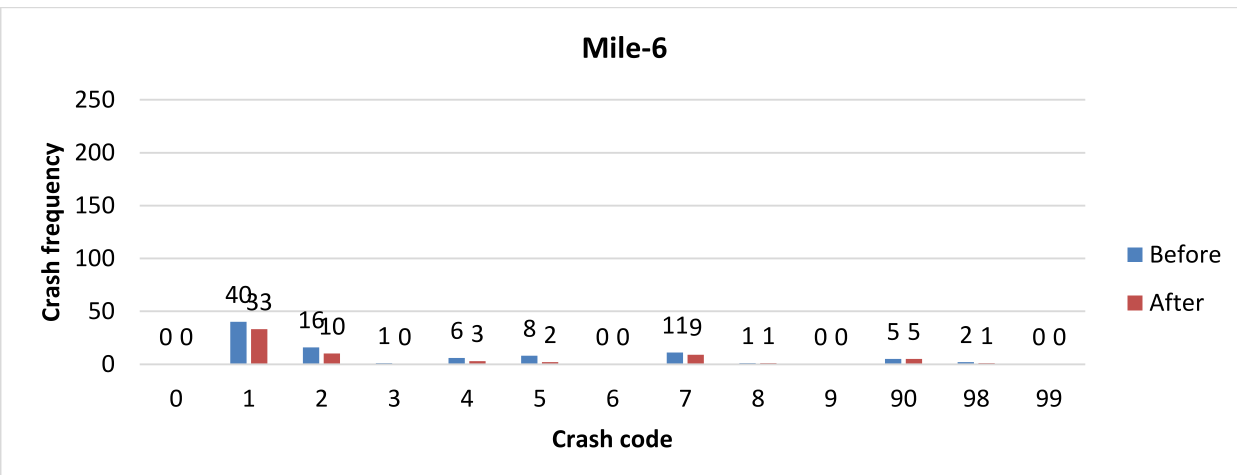

Table A. 7 Crash Summary Table for Section Mile-7

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

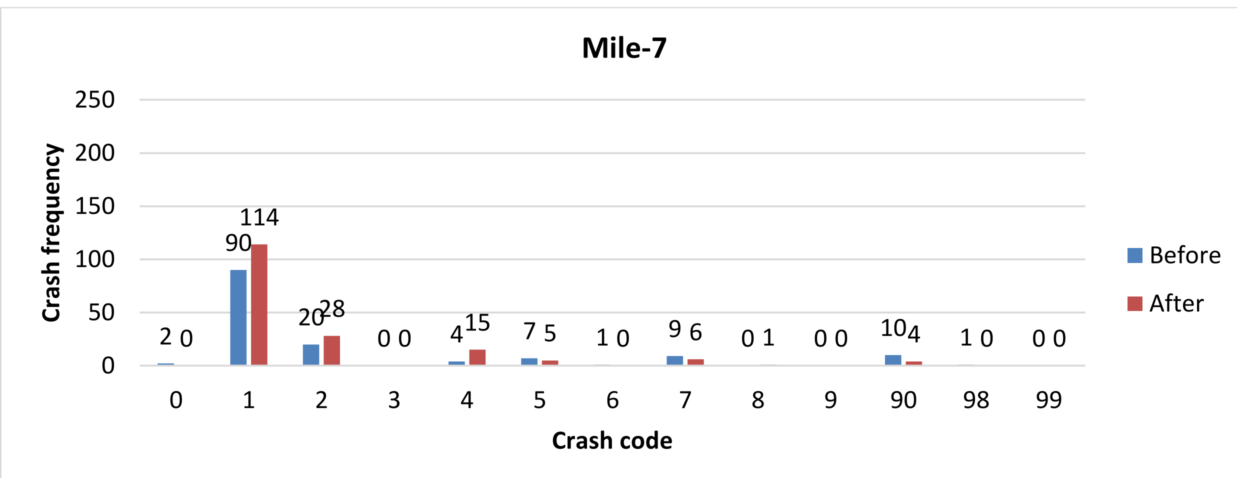

Table A. 8 Crash Summary Table for Section Mile-8

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

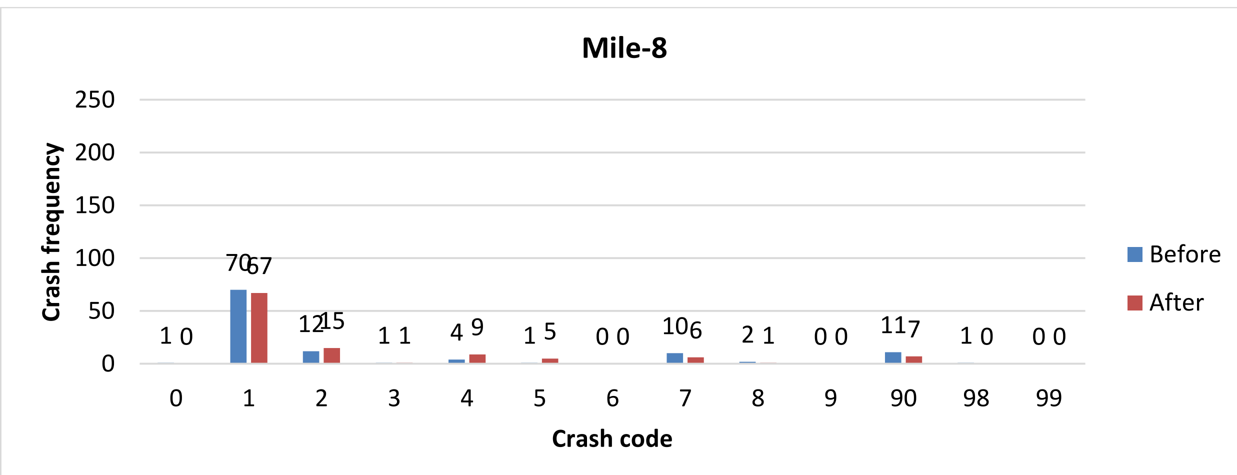

Table A. 9 Crash Summary Table for Section Mile-9

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

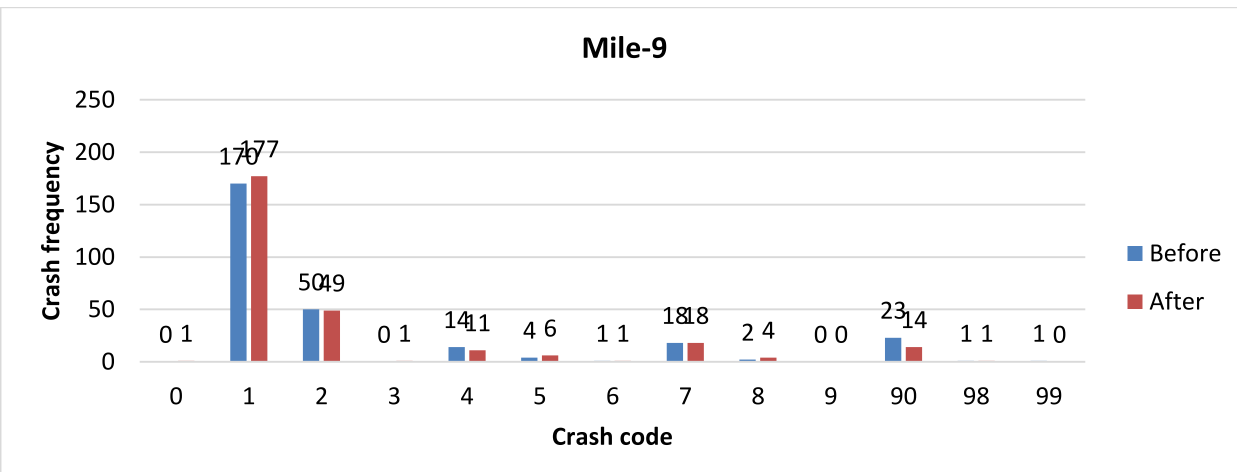

Table A. 10 Crash Summary Table for Section Mile-10

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

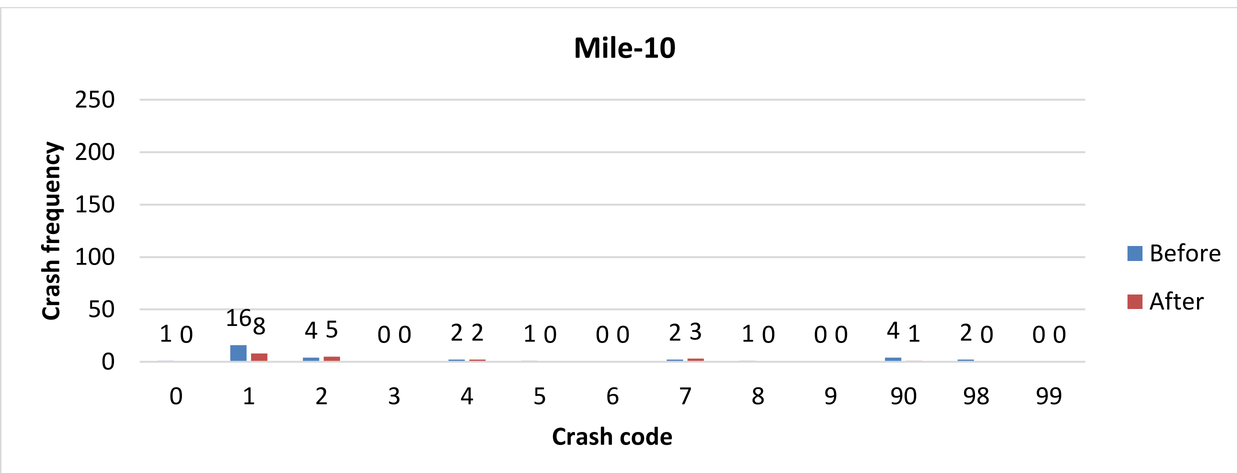

Table A. 11 Crash Summary Table for Section Mile-11

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

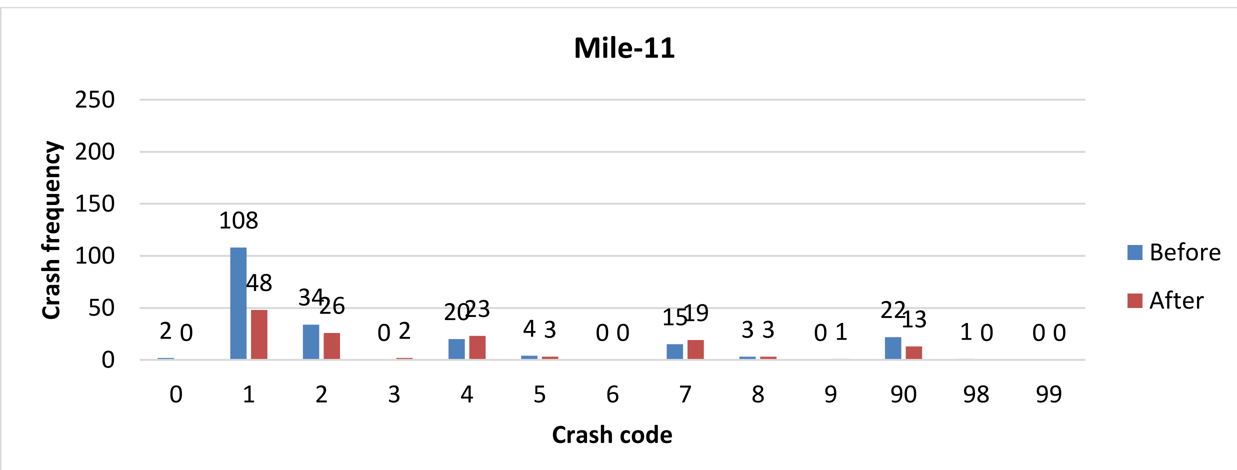

Table A. 12 Crash Summary Table for Section Mile-12

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

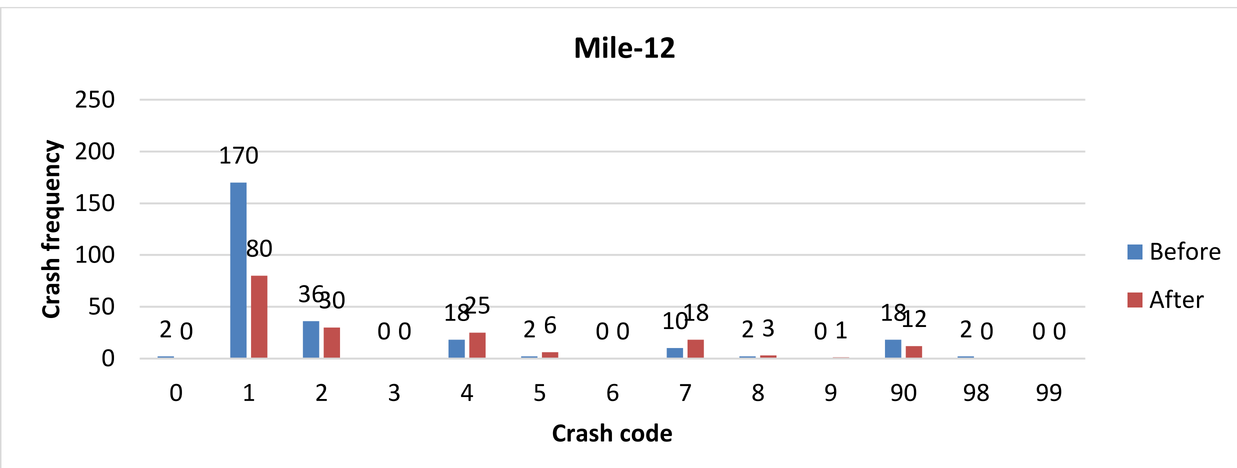

Table A. 13 Crash Summary Table for Section Mile-13

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

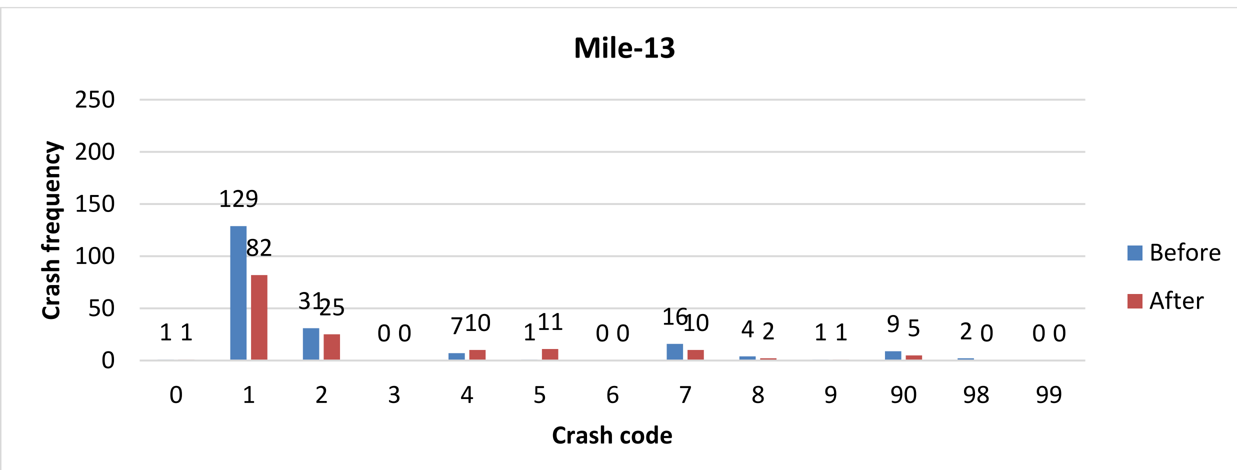

Table A. 14 Crash Summary Table for Section Mile-14

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

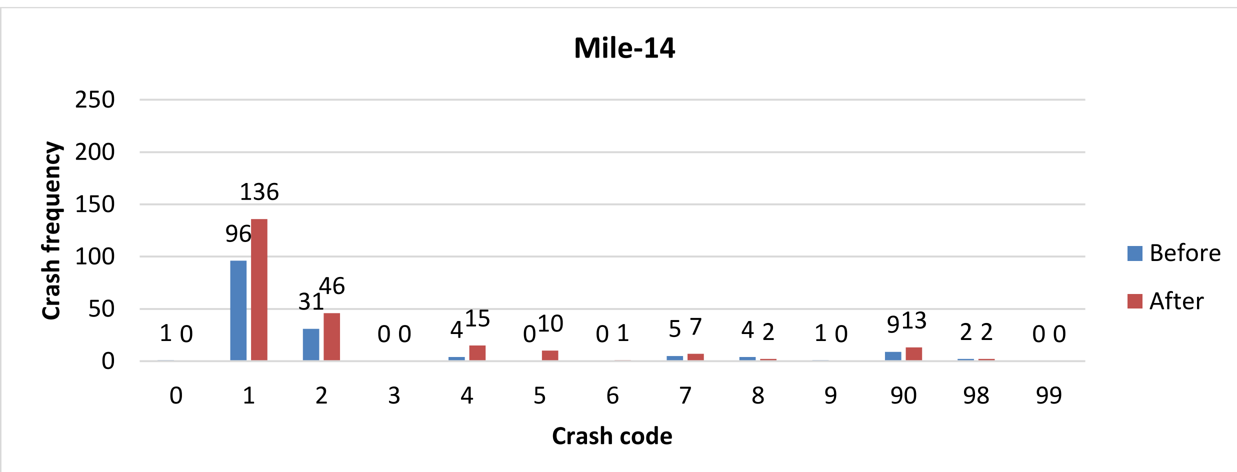

Table A. 15 Crash Summary Table for Section Mile-15

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

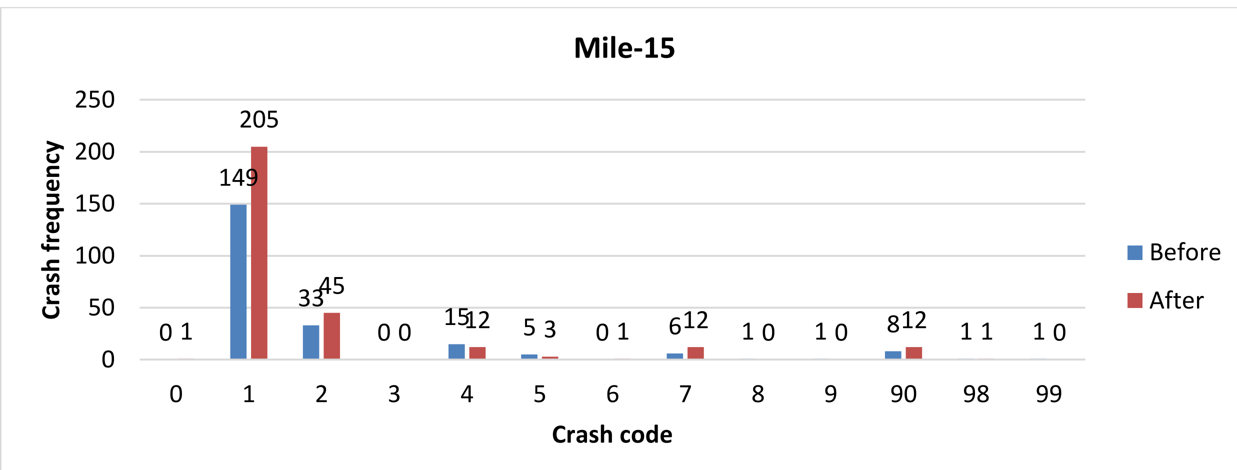

Table A. 16 Crash Summary Table for Section Mile-16

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

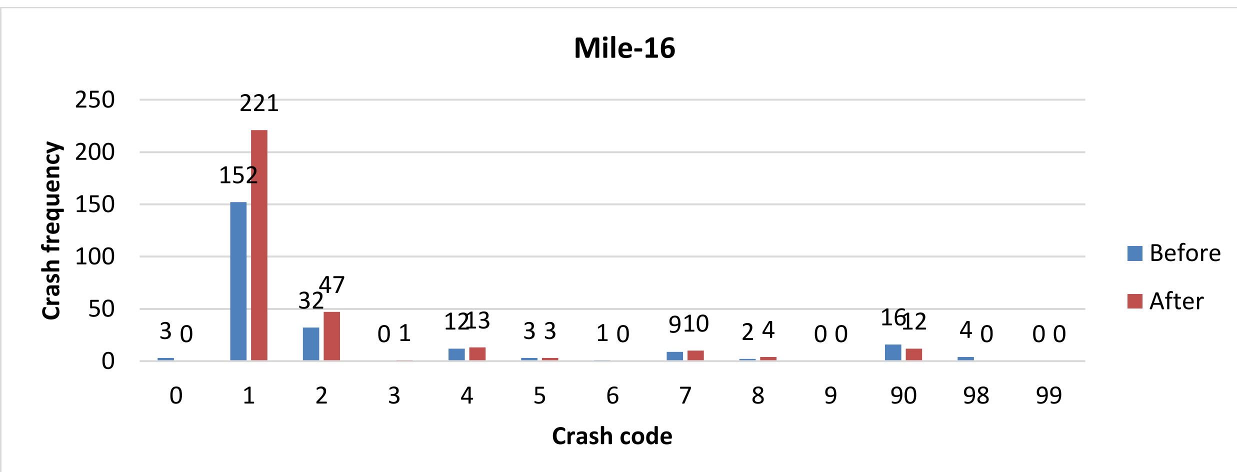

Table A. 17 Crash Summary Table for Section Mile-17

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