8-2011

Improving Safety in High-Speed Work Zones: A Super 70 Study

Andrew P. Tarko
Purdue University - Main Campus, aptarko@gmail.com

Mouyid Bin Islam
Purdue University

Jose E. Thomaz
Purdue University

JOINT TRANSPORTATION RESEARCH PROGRAM

INDIANA DEPARTMENT OF TRANSPORTATION &
PURDUE UNIVERSITY

JTRP

IMPROVING SAFETY IN HIGH-SPEED WORK ZONES: A SUPER 70 STUDY

Andrew P. Tarko
Professor of Civil Engineering
Purdue University
Corresponding Author

Mouyid B. Islam
Graduate Research Assistant
School of Civil Engineering
Purdue University

Jose E. Thomaz
Data Analyst
School of Civil Engineering
Purdue University

SPR-3249
Report Number: FHWA/IN/JTRP-2011/09
DOI: 10.5703/1288284314622

RECOMMENDED CITATION

Tarko, A. P., M. Islam, and J. E. Thomaz. *Improving Safety in High-Speed Work Zones: A Super 70 Study*. Publication FHWA/IN/JTRP-2011/09. Joint Transportation Research Program, Indiana Department of Transportation and Purdue University, West Lafayette, Indiana, 2011.
DOI: 10.5703/1288284314622

CORRESPONDING AUTHOR

Andrew P. Tarko
School of Civil Engineering
Purdue University
(765) 494-5027
tarko@purdue.edu

JOINT TRANSPORTATION RESEARCH PROGRAM

The Joint Transportation Research Program serves as a vehicle for INDOT collaboration with higher education institutions and industry in Indiana to facilitate innovation that results in continuous improvement in the planning, design, construction, operation, management and economic efficiency of the Indiana transportation infrastructure.
https://engineering.purdue.edu/JTRP/index_html

Published reports of the Joint Transportation Research Program are available at: http://docs.lib.purdue.edu/jtrp/

NOTICE

TECHNICAL REPORT STANDARD TITLE PAGE

1. Report No. FHWA/IN/JTRP-2011/09	2. Government Accession No.	3. Recipient's Catalog No.
4. Title and Subtitle Improving Safety in High-Speed Work Zones: A Super 70 Study		5. Report Date 2011
		6. Performing Organization Code
7. Author(s) Andrew Tarko, Mouyid Bin Islam, Jose Thomaz		8. Performing Organization Report No. FHWA/IN/JTRP-2011/09
9. Performing Organization Name and Address Joint Transportation Research Program Purdue University 550 Stadium Mall Drive West Lafayette, IN 47907-2051		10. Work Unit No.
		11. Contract or Grant No. SPR-3249
12. Sponsoring Agency Name and Address Indiana Department of Transportation State Office Building 100 North Senate Avenue Indianapolis, IN 46204		13. Type of Report and Period Covered Final Report
		14. Sponsoring Agency Code

15. Supplementary Notes

Prepared in cooperation with the Indiana Department of Transportation and Federal Highway Administration.

16. Abstract

Super 70 was an urban reconstruction project (March-November 2007) along I-70 in the central part of Indianapolis. INDOT applied in that project several innovative and traditional solutions. This study investigates the safety effect of the solutions. Advanced econometric models were applied to study both the spatial differences in the risk of crash on different roads inside and outside of the construction zone and short-term fluctuations in response to changes in traffic, weather, and traffic management. The single most successful management strategy was rerouting heavy vehicles (13+ tons) on alternative interstate routes. The second significant source of safety benefit was jointly generated by police enforcement, reduced speed, and other traffic management strategies. The safety benefit generated by the two sources was estimated to be 100 crashes saved inside the work zone during the nine months of the road construction. Widening shoulders was indicated as an additional means of improving work zone safety. The study could not confirm that the moveable barriers and consequently adjusting the number of traffic lanes to traffic volumes brought any direct safety benefits inside the work zone. The recommendations could be incorporated to the INDOT supporting materials for traffic management in high-speed urban work zones. The risk prediction equations can be applied to real-time detector data and weather information to assess the risk and identify high-risk conditions. Adequate warning messages could be displayed via VMS placed in advance of work zones.

17. Key Words urban work zone, traffic safety, high-speed road, statistical analysis, logit model, disaggregate data analysis	18. Distribution Statement No restrictions. This document is available to the public through the National Technical Information Service, Springfield, VA 22161

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price
Unclassified	Unclassified	37	

Form DOT F 1700.7 (8-69)

CONTENTS

LIST OF TABLES . . . ii

LIST OF FIGURES . . . iii

EXECUTIVE SUMMARY . . . 1

1 INTRODUCTION . . . 2

2 PAST RESEARCH . . . 2

3 RESEARCH METHODOLOGY . . . 5
3.1 Modeling Crash Severity and Likelihood . . . 6
3.2 Safety Impact Outside of Super 70 . . . 7

4 SOURCE DATA . . . 8
4.1 Sources . . . 8
4.2 Crash Data . . . 9
4.3 Geometry Data . . . 9
4.4. Traffic Data . . . 11
4.5 Weather Data . . . 11
4.6 Traffic Maintenance Data . . . 11
4.7 Enforcement Data . . . 11

5 DATA CONVERSION . . . 11
5.1 Crash Dataset . . . 11
5.2 Geometry Dataset . . . 15
5.3 Traffic Dataset . . . 16
5.4 Weather Dataset . . . 17
5.5 Traffic Maintenance Dataset . . . 18
5.6 Enforcement Dataset . . . 18

6 DATA LINKAGE . . . 18
6.1 Crash-Segment Assignment . . . 18
6.2 Weather Data Linkage . . . 19
6.3 Traffic Dataset Linkage . . . 19
6.4 Traffic Maintenance Dataset Linkage . . . 20
6.5 Enforcement Dataset Linkage . . . 20

7 CRASH LIKELIHOOD MODEL $P(C)$ . . . 20
7.1 Crash Likelihood Sample . . . 20
7.2 Model Estimation . . . 22
7.3 Discussion of the Crash Likelihood Factors . . . 24

8 SINGLE-VEHICLE CRASH LIKELIHOOD MODEL $P(C1|C)$ . . . 28
8.1 Model Estimation . . . 28
8.2 Discussion of the Single-vehicle Crash Factors . . . 28

9 JOINT CRASH SEVERITY MODEL FOR $P(SC1|C1)$ AND $P(SC2|C2)$ . . . 29

10 SUPER 70 SAFETY SIMULATION . . . 30

11 FINDINGS AND RECOMMENDATIONS . . . 35
11.1 Findings . . . 35
11.2 Recommendations . . . 36
11.3 Remarks . . . 37

REFERENCES . . . 37

LIST OF TABLES

Table	Page
Table 4.1 Data Collection Time Frame and Conditions	8
Table 4.2 Sources for Datasets	8
Table 4.3 Crash Data with Variables, Definition, and Formats	9
Table 4.4 Geometry Data with Variables, Definitions, and Formats	10
Table 4.5 Traffic Data with Variables, Definitions, and Formats	11
Table 4.6 Weather Data with Variables, Definitions, and Formats	12
Table 4.7 Traffic Maintenance Data with Variables, Definition and Formats	14
Table 4.8 Enforcement Data with Variables, Definitions, and Formats	14
Table 5.1 Injury Severity in the Converted Crash Dataset	14
Table 5.2 Periods Identified for Crash Date	15
Table 5.3 Summary of Number of Segments for NWZ	16
Table 5.4 Summary of Number of Segments for WZ	16
Table 5.5 Weather Manual Data WX-M Conversion	18
Table 7.1 Entire Sample Size	21
Table 7.2 Number of Observations Initially Generated for Modeling	22
Table 7.3 Number of Observations for Modeling and in the Entire Sample	23
Table 7.4 Number of Crashes Included in the Sample used for Modeling	23
Table 7.5 Intercepts Adjustment for Distorted Proportion of Crash Observations	24
Table 7.6 Crash Likelihood Model P(C) Results	25
Table 7.7 Marginal Effects of Variables on Safety Inside the Super 70 Work Zone	26
Table 8.1 Single-vehicle Crash Likelihood Model P(C1|C) Results	29
Table 9.1 Severe Crash Likelihood Model P(SCn|Cn) Results	30
Table 10.1 Safety in the Super 70 Period under Various Traffic Diversion Scenarios	31
Table 10.2 Ramps Effect	32
Table 10.3 Joint Effect of Other Factors	33
Table 10.4 Interstate Segments and Location	33
Table 10.5 Negative Binomial Test for Different Injury Severity in Different Segments	34
Table 10.6 Two Proportion Test for Different Injury Severity in Different Segments for Interstate	35
Table 10.7 Two Proportion Test for Involvement of Different Injury Severity in Different Segments	36

LIST OF FIGURES

Figure	Page
Figure 1.1 Super 70 Work zone in the Indianapolis Area	2
Figure 3.1 Tree of Safety-related Events in a Time Interval	5
Figure 3.2 Datasets for Safety Modeling	6
Figure 3.3 Research Framework	6
Figure 5.1 Interstate Sections and Traffic Detectors Included in the Study	16
Figure 6.1 Crash-Segment Assignment Process	19
Figure 6.2 Illustrations of Crash Location (LDIST) and Travel Direction (TRADIR)	19
Figure 6.3 Linkages with Weather Dataset	20
Figure 6.4 Linkage with Detector and Traffic Dataset	20
Figure 6.5 Linkages with Maintenance Dataset	21
Figure 6.6 Linkage with Enforcement Dataset	21
Figure 7.1 Sampling Process of Work Zone and Non-work-Zone Dataset	22
Figure 10.1 Simulated Impact of Traffic Volume on the Number of Crashes in the Super 70 Period	31
Figure 10.2 Simulated Impact of Traffic Volume on the Number of Injury Crashes in the Super 70 Period	31
Figure 10.3 Simulated Impact of Shoulder Width on the Number of Crashes in the Super 70 Period	32
Figure 10.4 Simulated Impact of Shoulder Width on the Number of Crashes in the Super 70 Period	32
Figure 10.5 Studied interstate segments	34

EXECUTIVE SUMMARY

Improving Safety in High-Speed Work Zones: A Super 70 Study

Introduction

Highway work zones, particularly those on urban high-speed roads, require special attention and adequate traffic management to reduce the adverse impact of altered geometry and traffic that differ from typical conditions. Super 70 was an urban reconstruction project (March-November 2007) along I-70 in the central part of Indianapolis. In that project, INDOT applied several innovative and traditional solutions such as movable barriers, prohibiting heavy vehicles, closing selected ramps, reducing the speed limit, and aggressive enforcement of traffic restrictions. Unlike many past work zones, the frequency of crashes in the Super 70 work zone segment was reported to be less during its construction period of nine months. The question addressed by the presented study was whether this reduction was merely a result of reduction in traffic volume passing the work zone, or did the novel safety countermeasures that were applied also play a role in enhancing the work zone's safety? Another important safety aspect was the potential migration of the risk of crashes away from the work zone to other roads that received the diverted traffic.

Findings

The presented study addressed these questions by applying advanced econometric models to study both the spatial differences in the risk of crash on different roads inside and outside of the construction zone and short-term fluctuations in response to changes in traffic, weather, and traffic management. The impact of the work zone on the entire Indianapolis interstate system was investigated by using before and after studies. The safety management during the Super 70 project was found to be successful. The single most successful management strategy was rerouting heavy vehicles (13+ tons) on alternative interstate routes. The second significant safety benefit was jointly generated by police enforcement, reduced speed, and other traffic management strategies. The safety benefit generated by the two sources was estimated to be 100 crashes saved inside the work zone during the nine months of the road construction. Widening shoulders was indicated as an additional means of improving work zone safety. The study could not confirm that the moveable barriers and the consequent adjustment of the number of traffic lanes to traffic volumes brought any direct safety benefits inside the work zone. The moveable barriers may have had a positive impact on traffic safety on local roads, however, by providing additional capacity that reduced traffic disturbances on the local system. Closing interchange ramps had only a limited safety effect for the work zone safety. Overall, safety on the affected interstates was higher during the Super 70 project than before.

In light of the research results, the following recommendations for future work zones on urban interstates were derived:

(1) Reroute heavy vehicles (13+ tons) on alternative interstate routes.
(2) Reduce the speed limit and apply shoulders as wide as possible inside the work zone. Use of additional traffic lanes instead of wide shoulders should be considered where shortage of capacity is expected, although this may lead to traffic spillover to surface roads.
(3) Avoid redirecting the through traffic on surface roads if possible by providing a sufficient number of lanes inside the work zone and/or using changeable barriers to adjust capacity to demand.
(4) Reduce the impact of the work zone on local traffic by using the changeable barriers to adjust the number of lanes to current traffic demand and by maintaining as many open ramps as possible.

Consider warning drivers about crash danger via variable message signs displaying adequate messages based on the real-time assessment of the risk of crash. The models developed in this study can facilitate the risk estimation in 30-minute intervals.

Implementation

The recommendations of the research study should be incorporated in the INDOT supporting materials for traffic management in high-speed urban work zones. The risk prediction equations can be applied to real-time detector data and weather information to assess the risk and identify high-risk conditions. Adequate warning messages could be displayed via VMS placed in advance of work zones.

1 INTRODUCTION

Highway work zone conditions require increased attention by motorists since the traffic flow, geometry within the work zone, and maintenance of traffic are significantly different from normal daily driving (Tarko, 2007). Super 70 was an urban reconstruction project completed in March-November 2007 by the Indiana Department of Transportation (INDOT) along the I-70 stretch between two interchanges, I-65 at the west end and I-465 at the east end (shown in Figure 1.1). This is a six-mile section that carries almost 180,000 vehicles daily (I-70 Reconstruction). During the construction period, INDOT applied several innovative and traditional solutions such as movable barriers, prohibiting heavy vehicles, closing selected ramps, reducing speed limit, and aggressive enforcement of traffic restrictions. The Indiana State Police policy from the beginning of this enforcement effort was to issue citations instead of warnings for almost all violations. The rationale for removing truck traffic from the work zone was to reduce both the risk of lane blockage by damaged heavy vehicles and the risk of secondary crashes caused by the traffic backup. Indeed, the Super 70 work zone was never closed due to a truck crash during the entire project.

Most of the past studies have found that the crash frequency is expected to increase in work zone segments due to the conditions different from the typical ones to which drivers are accustomed. Promisingly, there was a reduction in the frequency of crashes in the Super 70 work zone segment during its construction period of nine months according to the crash statistics of the Indiana State Police. The number of crashes reported during the construction period of February 23 – November 16, 2007, was 161. The corresponding period in 2006 had 233 crashes, which is equivalent to a 31 percent reduction.

The interesting question is whether this reduction was caused only by the reduction in the traffic volume passing through the work zone or if the applied novel safety countermeasures also played an additional role in enhancing the work zone's safety. Another important safety aspect was migration of the risk of crash away from the work zone to other interstate roads that received the diverted traffic. Thus, the second question was whether the overall safety in the in Indianapolis interstate system affected by the traffic diversions changed and in what manner.

This study addressed the first research question by applying advanced econometric models to study both the spatial differences in the risk of crash on different roads inside and outside of the construction zone and short-term fluctuations in response to changes in traffic, weather, and traffic management. The impact of the work zone on the entire Indianapolis interstate system is investigated by using the before and after studies. The traffic safety is measured with crash likelihood in short intervals, with crash frequency in nine-month periods, and with likelihood of a severe outcome given that crash happens.

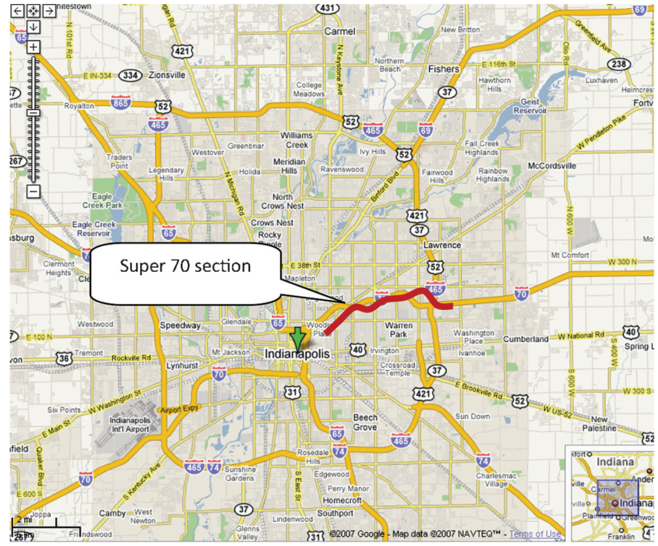


Figure 1.1 Super 70 Work zone in the Indianapolis Area

This report discusses the research objectives and past research relevant to the objectives. Next, the available data, their conversion, and linking are presented. The statistical modeling of short-term safety is then presented and discussed. The second part of the analysis includes the before-and-after study of the interstate and other roads in the Indianapolis area potentially influenced by the work zone. Finally, the results of the modeling and before-and-after study are discussed and final conclusions and recommendations are made.

2 PAST RESEARCH

Several studies have attempted to establish a relationship between work zone crashes and related influencing factors for safety as found in the literature. The important and related studies are mentioned here.

A study by Pal and Sinha (1996) mentioned that crash rates at work zones could be higher at most of the sites, but in some cases they could be lower. However, this study also indicated a counter-intuitive effect of work zone safety, stating that it may be possible that at the sites with high crash rates, motorists were more attentive within work zones. This study pointed out that the lower crash rates could be because of various work zone safety programs, including enhanced enforcement levels (e.g., police patrols).This study focused on the two lane closures strategies, cross-over and partial lane closure, and developed a Negative Binomial and Poisson regression model and normal regression models. This study indicated that Negative Binomial and Poisson regression models could be better predictive models for work zones if larger datasets were to be considered.

Vanogopal and Tarko (2000) reviewed some of the studies on work zone safety with the evidence of higher crash rates during road construction than during periods of regular traffic operations. This study also tried to attribute the conditions behind the increase of crashes: (a) general disruption to the flowing traffic due to sudden discontinuities caused by closed lanes, (b) improper lane merging maneuvers, (c) the presence of heavy construction equipment within the work area, (d) inappropriate use of traffic control devices, and (e) poor traffic management. This study identified some potential factors, such as cost of work, average daily traffic volume, ramps, work zone length, duration of work, and type of work for predicting the number of crashes at rural freeway work zones when INDOT was planning to implement the Indiana Lane Merge System (ILMS) at the entries of the urban freeway work zones to mitigate the aggressive lane changing maneuvers. Traffic volume, length of work zone, and duration of work were found to be significant in the model. Moreover, the cost of the work and the type of work were critical factors of safety inside work zones.

Another study by Khattak et al. (2000) clearly indicated the increase of crash rates and injury and non-injury crash rates during work zone time than other regular time. Wok zone information from the California Department of Transportation was merged with California traffic and crash data files for 1992 and 1993, resulting in 36 observations each on pre-work zone and during-work zone periods in California. This study revealed that the total crash rate during the work zone period was 21.5% higher (0.79 crashes per million vehicle km) than the pre-work zone period (0.65 crashes per million vehicle km). When compared with the pre-work zone period, the non-injury and injury crash rates during the work zone period was 23.8% and 17.3% higher, respectively. This study focused on a Negative Binomial model for estimating the frequency of crashes and found that the frequency increased with an increase in work zone duration, length, and average daily traffic. One important finding of this study is that after controlling for various factors, longer work zone durations significantly increase the frequency of both injury and non-injury crashes.

Khattak and Targa (2004) summarized previous studies on multi-vehicle collisions inside work zones and found that multi-vehicle collisions were consistently overrepresented in work zones and were also significant for large truck-involved collisions in work zone crashes. This study focused on an Ordered Probit model for injury severity and ordinary least squares for the cost associated with each injury category from the data extracted from the Highway Safety Information System (HSIS) for 2000 for North Carolina. The findings from the model indicate that the most injurious/harmful work zone crashes were a) truck-involved collisions when the roadway was closed and a detour was required on the opposite side, which experienced a 38.5% greater chance of injury, b) before the actual work area, where traffic moves out of its normal path, compared with crashes in the advance warning area or adjacent to the actual activity/work area, and c) on two-way undivided roads, which experienced a 19.1% greater chance of injury on such roadways.

Spainhour and Mishra (2008) applied logistic regression to analyze the factors affecting overcorrection in fatal run-off-the-road crashes in the year 2000 in the state of Florida, which account for more than 25 percent of run-off-the-road crashes, using 23 explanatory variables to identify the predictive variables. Backward stepwise regression with a significance level of 0.2 removed all of the original variables except gender, presence of rumble strips, speeding, second shoulder type, and vehicle movement. The results of this study indicated that female drivers are at a 40 percent higher risk than male drivers to overcorrect in fatal run-off-the-road crashes. Also, while fewer than 20 percent of fatal crashes occurred where rumble strips were present, drivers were more than 50 percent more likely to overcorrect when they were present.

Lu et al. (2006) applied ordinal logistic regression procedures to investigate potential associations between the severity of median crossover crashes (property-damage-only, injury, and fatality) and various predictors for 631 median crossover crashes in Wisconsin from 2001 to 2003. The results indicated that the seasons or time of the year may play a critical role in this type of crash severity and the winter months and correlating road conditions with the time of the year when they occurred were responsible factors here.

Li and Bai (2006) developed a logistic regression model based on the fatal crash database from the Kansas Department of Transportation from 1992 to 2004. This study demonstrated the effect of flagger usage and the effect of stop sign/signal usage separately. The results indicated that the presence of flagger controls in work zones could reduce the probability of fatal crashes for male drivers by 15%; flaggers directing the traffic could reduce the conditional probability of heavy truck involvement in fatal crashes by 27%; and the presence of a stop sign or signal is effective in reducing multi-vehicle fatal crashes by 13%.

Dissanayake and Lu (2002) conducted a study using a crash database derived from the Florida Traffic Crash Database from 1997 to 1998. Their study focused on young drivers (16- 25 years) involved in a single vehicle collision with fixed objects. A set of sequential binary logistic regression models was developed to identify the roadway, driver, environmental, and vehicle-related factors influencing crash severity (e.g., fatal, incapacitating, non-incapacitating, and possible injury). Use of alcohol or drugs, ejection in the crash, gender, impact point of the vehicle, restraint device usage, urban/rural nature, grade/curve existence at the crash location, lighting condition, and speed were found to be the most important factors affecting the severity of young driver single-vehicle fixed-object crashes involving passenger cars.

Mercier et al. (1997) applied logistic regression to determine the factors causing injury severity using the Iowa Department of Transportation crash database for the period 1986–1993. Age and gender were found to be two important factors in injury severity in head-on crashes on rural highways. Injury severity was defined as fatal, major, and minor in the dependent variables. Fourteen explanatory variables and their interactions were first considered as forward stepwise selection. The independent variables were age, age squared, position in the vehicle, and protection.

A study by Daniel et al. (2000) examined the influence of work zone activity on the frequency of fatal crashes based on three work zone locations in Georgia during 1995–1997. This study indicated the manner of collision, lighting conditions, truck involvement, and roadway functional classification for fatal crashes in work zones. In this study, construction activities, such as resurfacing and widening projects, were identified for determination of the potential causes for the fatal crashes that occurred. The predominant type of collision within work zones was determined to be single-vehicle collisions (48.6 percent), compared with 56.3 percent for non-work zone locations. Fatal crashes in work zones were more likely to involve a collision with another vehicle than non-work zone fatal crashes. In this regard, a significantly higher proportion of rear-end collisions occurred in active work zones than in idle work zones. The fatal crashes within the study's work zones primarily involved passenger cars, which account for 80 percent of the vehicles involved in fatal crashes. Heavier vehicle involvement, such as trucks, in fatal crashes in work zones, was significantly higher (20 percent) than in non-work zone locations (13 percent). A significantly higher proportion of fatal crashes occurred during darkened conditions in work zones compared with non-work zone locations. Rural highways were riskier (59 percent) than urban roadways (41 percent). The contributing factors attributed by the investigating police officers for work zone fatal crashes were "driver lost control," "driver failed to yield," and "too fast for conditions," which account for 38 percent of fatal crashes in work zones. This study also suggested that speeding is one of the contributing factors for crashes in active work zones, and rear-end collisions are predominant in active work zones than in idle work zones.

A study by Wang et al. (1996) indicates that there are mixed findings on accident severity reported in the literature. A good example indicated that one Texas study of 1978 accident data showed accidents in construction zones tend to be more severe than accidents elsewhere, which was contradicted by another study that used 1977 accident data for Texas. This study indicates that, for all three states (A, B, and C) during 1991–1992, there was an overwhelmingly large percentage of rear-end collisions. The accident data also showed work zone accidents occurred most frequently on major highway facilities in urban areas because a great deal of road work was being conducted in large metropolitan areas with heavy traffic.

Bedar et al. (2002) conducted logistic regression on fatal accident cases involving single vehicles that collided with fixed objects using data for 1975–1998 from a FARS database to estimate the drivers' fatality risk. Their multivariate logistic regression revealed that the odds ratio for fatal injury increased with age. The analysis also revealed that the proportion of sober drivers fatally injured increased from 64.2% for drivers younger than 20 years to 88.5% for drivers aged 80 years and more. The "U" shaped relationship between the blood alcohol content (BAC) and the fatality risk indicated that the risk decreased with an increasing BAC, reaching its lowest point at a BAC of 0.05–0.09 and again increased with a maximum BAC of 0.30 or greater. Frontal impact of roughly 65% of crashes showed the largest source of fatalities followed by these most frequent types, right side with 17.5% (far-side) and left side with 13.5% (driver-side). Driver-side impacts have twice the odds of fatality compared with frontal impact. Speeds in excess of 69 mph prior to or at impact had a higher fatality odds ratio (OR = 2.64) compared with speeds less than 35 mph. Restraint systems, such as three-point seatbelts, were compared with no seatbelt usage. The overall analysis suggested that higher usage of seatbelt, reduced speed, and a decreased number and severity of driver-side impact could potentially prevent fatalities.

A study by Al-Ghamdi (2002) indicated the influence of accident factors on injury severity based on traffic police data at Riyadh in Abu Dhabi during August 1997 to November 1998. The location and cause of the accidents were found to be highly associated with the injury severity in the accidents. Non-intersection accidents exhibited higher odd than an intersection by 2.6 times. Wrong way-related accidents have higher odds than any other causes for injury.

Tay et al. (2008) identified the factors associated with hit-and-run crashes with driver characteristics, vehicle types, crash characteristics, roadway features, and environmental characteristics. Using logistic regression to model the likelihood of hit-and-run crashes, this study revealed that drivers are more likely to leave the scene when an accident occurred at night; on a bridge and flyover, bend, straight road, and near shop houses; involved two vehicles, two-wheelers and imported vehicles; and were male drivers and between the ages of 45 to 69.

The weather conditions have an influential impact on highway safety by directly affecting the dynamic driving conditions of the traffic. Adverse weather can be a hazard for drivers passing the work zone because it is characteristically different from the typical roadway section. Visibility is highly impaired by snow, fog, and rain whereas the vehicle operation is much more complicated on the wet surface caused by precipitation.

A study on traffic safety in New Mexico construction zones by Hall and Lorenz (1989) indicated the influence of adverse weather conditions on increased construction zone crashes while considering typical accident characteristics for construction zones. Another study by Pal and Sinha (1996) on work zone safety in Indiana clearly emphasized the importance of weather along with other factors in road safety by stating that there are many factors that may influence the crash rate on a given highway section, including geometry, pavement condition, traffic volume, length of work zone, weather, lane closure strategy, etc.

While investigated the effect of weather on freeway capacity in Ontario, Canada, the findings of another study by Kaisy et al (2000) suggested that adverse weather affected work zone capacity differently in each direction. This finding could be further interpreted that bad weather imposed restrictions on traffic flow in both directions, which could have had a dominant impact on the queue discharge flow.

Li and Bai (2008) developed a crash-severity-index applying logistic regression analysis for a work zone safety evaluation based on 85 fatal and 604 injury crashes from 1998 to 2004 in Kansas highway work zones. In this study, the crash severity index (CSI) is a numerical value ranging from zero to one and interpreted as the likelihood of having fatalities when a severe crash occurs in a given work zone. They developed two groups of CSI models: driver-independent (DI-CSI) and driver dependent (DD-CSI) models. Each of the groups represents two models each; one is comprehensive and the other is simplified based on the significant variables from the comprehensive models. Typically, risk factors such as poor light condition, truck involvement, having two travel lanes, and high speed limit may lead to high CSI values and, equivalently, high risk levels.

Other than just statistical modeling on the safety of construction zones, different types of statistical tests have been conducted in other studies. A study by Jin et al. (2008) utilized the paired t-test, two-way ANOVA, and Tukey test to assess the trend between mean crash frequencies during construction and non-construction periods in Utah using 202 construction project records between 2002 and 2005. The research shed some light on why crash frequencies during active construction periods on the same highway section. The paired t-test shows that crash rates are statistically the same between construction periods and non-construction periods on urban interstates, urban non-interstates, and rural non-interstates except at lower severity levels for rural interstate highways. The two-way ANOVA and Tukey test indicated that the effect of highway class was not statistically significant for the difference in mean crash rates during construction and non-construction periods. The concluding part of this study clearly mentioned that the trend of higher crash rates during construction periods reported by past work zone studies was not statistically supported by Utah's crash records. However, the authors also mentioned that there was a shortcoming for this study in that it did not consider the work zone traffic control strategies in this analysis.

A work zone is a risky segment on a highway, and most of the previous studies mentioned above found that the crash frequency increased with associated factors. However, in the Super 70 work zone segment for the present study, fewer crashes were reported during its nine-month construction period. According to the crash statistics reported by INDOT, the number of crashes declined from 233 in 2006 during the same time period of the construction (February 23 to November 16) to 161 in 2007 during the actual construction period (February 23 to November 16), which represents a 31% reduction from the previous year. This study will attempt to investigate the factors related to such crash reduction in terms of crash frequency. In addition, crash severity will be also studied.

Although it was determined that the crash frequency increased in work zone areas, the work zone characteristics that influenced the injury severity were not identified in many of the previous studies. Since the reported crashes in the work zone area (I-70 section) decreased compared to previous years, the factors influencing the crash frequency and injury severity will be investigated.

3 RESEARCH METHODOLOGY

Answering the two research questions about the effectiveness of applied safety countermeasures and the overall safety impact of the Super 70 work zone requires two distinct research methods:

(1) Advanced modeling of safety in short time intervals and on relatively short road segments with logistic regression to estimate the impacts of individual safety countermeasures and other safety variables that need to be accounted for to avoid estimation bias, and
(2) Before-and-after study to estimate the overall change in safety in the work zone impact area and its components over time measured with longer periods of several months.

Figure 3.1 shows the tree of events including the crash occurrence, its type (single-vehicle and multiple-vehicle), and the severity of the outcome. A crash is severe if at least one person is injured. The likelihood of these events for a time interval and under specific traffic, road, and weather conditions are estimated using statistical models. The events whose likelihood is estimated include crash, P(C), single-vehicle crash given that crash occurs, P(C1|C), a severe outcome given that a single-vehicle crash occurs P(SC1|C1), and a severe outcome give that

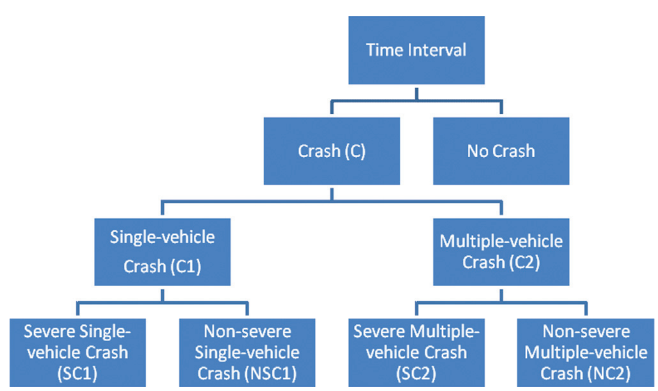


Figure 3.1 Tree of Safety-related Events in a Time Interval

multiple-vehicle crash occurs, P(SC2|C2). The likelihood values obtained from the models are then used to calculate the likelihood of other events in the tree using the following set of equations.

$$\begin{aligned} P(C2|C) &= 1 - P(C1|C) \\ P(NC1|C1) &= 1 - P(SC1|C1) \\ P(NC2|C2) &= 1 - P(SC2|C2) \end{aligned} \quad (1)$$

$$\begin{aligned} P(SC1) &= P(C)P(C1|C)P(SC1|C1) \\ P(SC2) &= P(C)P(C2|C)P(SC2|C2) \\ P(NC1) &= P(C)P(C1|C)P(NC1|C1) \\ P(NC2) &= P(C)P(C2|C)P(NC2|C2) \end{aligned} \quad (2)$$

where P(X) is the likelihood of X while P(Y|X) is the likelihood of Y given that X happens.

The likelihood values estimated either in the developed statistical models or calculated in Eq. 1 and Eq. 2 apply to 30-minute intervals. The expected number of crashes in long periods can be calculated by summing up the likelihood values in all the intervals included in the period as expressed in Eq. 3.

$$\begin{aligned} ESC1 &= \sum_{t \in T}^{\infty} P_t(SC1) \\ ENC1 &= \sum_{t \in T}^{\infty} P_t(NC1) \\ ESC2 &= \sum_{t \in T}^{\infty} P_t(SC2) \\ ENC2 &= \sum_{t \in T}^{\infty} P_t(NC2) \end{aligned} \quad (3)$$

where t is the interval index and T is the period.

To model crash frequency and severity, crash data are needed jointly with other data that reflect the temporal and spatial characteristics, such as road geometry, traffic volume and composition, weather, traffic management, and others.

Data collection and management were the most time- and labor-consuming components of this research and it has been given due attention in this report. The potential model variables included in various datasets are described in the following subsections. Figure 3.2 shows the schematics of data components of development of safety model.

Multivariate analysis of various multiple factors of crash likelihood in 30-minute intervals and crash severity are applied to three distinct periods: before, during, and after the construction activities with different geometry and traffic conditions. On the other hand, the impact of the work zone on the adjacent roads is investigated by comparing safety data aggregated in nine-month periods before and during the construction activities and on prolonged road segments with and without the construction activities.

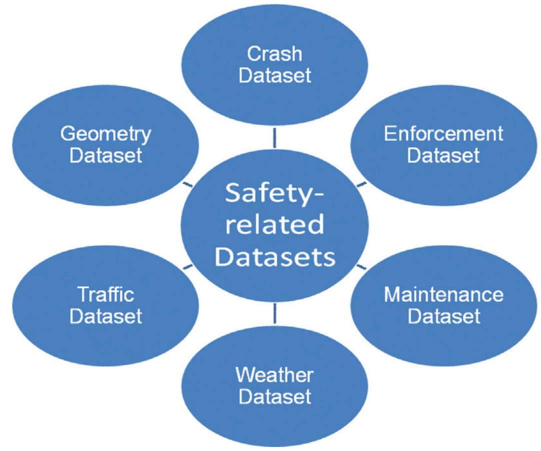


Figure 3.2 Datasets for Safety Modeling

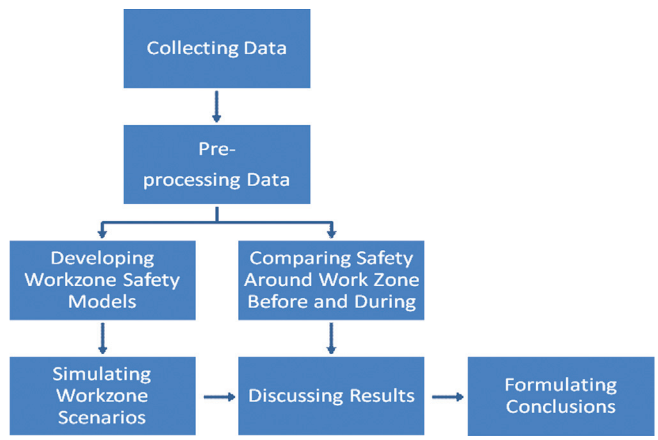


Figure 3.3 Research Framework

Modeling the safety of work zones is conducted by sampling the data for six major components: crashes, traffic, weather, geometry, traffic maintenance, and police enforcement. Since a disaggregate dataset is developed for the statistical model, sampling of the data was undertaken in the modeling segment, and general adjustments associated with the modeling are discussed later in the report. The before-after study for alternative routes (interstates and local routes) focuses on the routes that are primarily considered parallel to the work zone segment of I-70 and diverting routes.

3.1 Modeling Crash Severity and Likelihood

The severity of a crash is defined as the most severe injury experienced by persons involved in the crash. Indiana crash data, similar to many other states, uses the KABCO scale to assess the injury severity at the crash scene by investigating police officers. The KABCO scale was introduced by the National Safety Council in 1966, where K stands for “fatal,” A for “incapacitating,” B for “non-incapacitating,” C for "possible injuries" or "complaint pain," and O for "not-injured." Although there are five levels of severity ranging from "no injury" to "fatality," some of the levels are combined to obtain more observations in the pooled severity categories.

The statistical model most often used to model the crash severity is the logit model, which belongs to the family of discrete choice models. Although there is no choice involved in the severity outcome of a crash, these models are considerable in this case. The logit model applied to crash severity calculates the likelihood of the crash severity level *i* given that crash *n* happened. Thus, the data used in modeling consist of observations that are crash characteristics together with variables that describe the conditions present when the crash happened. The modeling is based on a family of linear functions T of contributing factors Xn for severity level i and a given crash n as shown in Eq. 4:

$$T_{in} = \beta_i X_{in} + \varepsilon_{in} \tag{4}$$

where,

X_{in} is a vector of explanatory variables (crash, traffic, geometry, weather and enforcement and maintenance variables);

β_i is the vector of estimable parameters, and

ε_{in} is the error term.

The likelihood of outcome i (severity level i) has the likelihood of the event that the value T_i for this level is higher than for the other levels:

$$\begin{aligned} P_n(i) &= P(T_{in} \geq T_{in} \forall \neq i \\ P_n(i) &= P(\beta_i X_{in} - \beta_I X_n \geq \varepsilon_{In} - \varepsilon_{in}) \forall I \neq i \end{aligned} \tag{5}$$

where, P_n (i) being the probability of crash n having a discrete outcome *i* ($i \in$ I) and I denotes all possible outcome (Washington et al., 2003).

For analyzing the likelihood of crash, the period of analysis is divided into 30-minute intervals and the studied road network is divided into approximately 0.25-mile segments. An observation in this modeling is a set of traffic, geometry, and other conditions on a 0.25-mile segment during 30 minutes, together with the binary variable indicating that a crash occurred. To estimate the crash likelihood, binary logistic regression was used, which is a special case of the previous model applied to the case of only two outcomes (crash, no crash).

3.2 Safety Impact Outside of Super 70

The safety changes after the I-70 work zone's onset are measured with differences in crash statistics applied to the before and after periods. Safety aspects such as crash frequency, severity, and involvement of different types of vehicles are investigated. Two tests are used here:

1. The Negative Binomial test to investigate the change in the crash frequency of various types of crashes and
2. The Two Proportions test to investigate the change in the proportion of certain types of crashes in the total number of crashes.

These two tests are described in more details in the next two sections in this report.

3.2.1 Negative Binomial Test for Changes in Crash Frequency

The Negative Binomial test for the reduced number of crashes estimates the likelihood p of the observed or lower number of crashes in the after period based on the assumption that the distribution of crashes is not affected by the work zone. The p likelihood is estimated using the cumulative Negative Binomial distribution or the equivalent distribution Beta, which is more convenient as shown in Eq. 6:

$$\begin{aligned} p &= \Pr(C < c_A) = \sum\nolimits_{c=0 \ldots c_A} NegBinomDist[c,r,p] \\ &= BetaDist(p,r,c_A+1) = BetaDist\left[\frac{n_B}{n_B+1}, c_B, c_A+1\right] \end{aligned} \tag{6}$$

where:

c_A = number of crashes in the after nine-month period,

c_B = total number of crashes in the before period (three nine-month sub-periods),

n_B = number of nine-month sub-periods in the before period (n_B=3).

A low value of p indicates that the number of crashes in the after period is too low; thus, the assumption of the lack of the work zone effect is questionable. The conclusion opposite to the assumption is made. The test provides statistical evidence that the work zone has reduced the number of crashes. If p is not sufficiently low, then it can be stated that there is no statistical basis to reject the assumption about the lack of the work zone impact.

To decide if p is sufficiently low to claim the effect of the work zone, a small value α, called level of significance, is used. The estimated likelihood is considered sufficiently low if it is at most equal α. We have used α = 0.10.

The Negative Binomial test is also used for the increased number of crashes. In this case, p is the likelihood of the observed or higher number of crashes in the after period.

$$\begin{aligned} p &= \Pr(C \geq c_A) = 1 - \Pr(C \leq c_A - 1) \\ &= 1 - \sum\nolimits_{c=0 \ldots c_A} NegBinomDist[c,r,p] \\ &= 1 - BetaDist(p,r,c_A) = 1 - BetaDist\left[\frac{n_B}{n_B+1}, c_B, c_A\right] \end{aligned} \tag{7}$$

where,

c_A = number of crashes in the after nine-month period,

c_B = total number of crashes in the before period (three nine-month sub-periods),

n_B = number of nine-month sub-periods in the before period (n_B=3).

A value of p equal to or lower than α indicates that the assumption of the lack of the work zone effect is questionable and the opposite conclusion can be made that the work zone has increased the number of crashes significantly.

3.2.2. Two Proportions Test for Changes in Crash Proportion

The Two Proportions test, based on Z statistic for normal distributions, was used to check if the work zone changed the proportion of certain types of crashes in the total number of crashes. The expression for z-statistic is given in Equation Eq. 8:

$$z=\frac{p_A-p_B}{\sqrt{p.(1-p).\left(\frac{1}{n_B}+\frac{1}{n_A}\right)}} \tag{8}$$

where:

p = "true" proportion of a certain type of crashes x in total number c, $p=\frac{x_A+x_B}{c_B+c_A}$,

p_B = before estimate of the true proportion, $p_B=\frac{x_b}{c_B}$,

p_A = after estimate of the true proportion, $p_A=\frac{x_A}{c_A}$.

The Z statistic is distributed according to the standardized normal distribution and has a corresponding cumulative distribution Pr(Z<z), where Z is the value calculated in Equation 8. A negative value of Z and the corresponding small value of likelihood p = Pr(Z<z) may indicate that the after proportion is lower than the before proportion. It may be concluded that p is equal or lower than a small value α. As for the crash frequency test, we used the level of significance α = 0.10.

On the other hand, a positive value of z and a corresponding small value of p = Pr(Z>z) = 1-Pr(Z<z) may indicate that the after proportion is higher than the before proportion. The significance level α is again used to decide if the difference in the proportions is significant.

TABLE 4.2
Sources for Datasets

Dataset	Source
Crash dataset	Indiana State Police Crash Data Records
Traffic dataset	Detectors set up by INDOT
Geometry dataset	Google Earth and Super 70 work zone drawing
Weather dataset	National Climatic Data Center
Maintenance dataset	Super 70 work zone drawing
Enforcement dataset	Super 70 work zone activity log

4 SOURCE DATA

Construction in the I-70 work zone began in March 2007, but traffic restrictions were activated on February 23, 2007. The Super 70 work zone had two phases – Phase-I had two sub-phases–IA and IB, with Sub-phase-IA conducted from February 23 through July 12 and Sub-phase-IB from July 13 through July 17. Phase-II was conducted from July 17 through November 16. The Super 70 work zone was officially opened to traffic on November 16, 2007.

A different time line was adopted for data collection based on the normal flow of traffic with heavy vehicles (non-work zone segments, before the construction); changeable concrete barrier, number of lane closures, restrictions on heavy vehicles, and closed ramps (during the construction); and improved highway geometry (after the construction). Table 1 shows the data collection time frame under different roadway conditions and vehicle restrictions.

4.1 Sources

Since this study is data-intensive, statistical modeling, different datasets have been developed or converted from the source files. Table 4.1 shows the data collection time frame and the corresponding traffic and geometry condition over time. The source of these datasets is presented in Table 4.2.

TABLE 4.1
Data Collection Time Frame and Conditions

Period	Time Frame	Consideration
1: All Non-work zone segment	Jan 2006 – Jun 2008	Standard concrete barrier and heavy vehicles
2: Before I-70 work zone	Jan 2006 – Feb 2007	Standard concrete barrier and heavy vehicles
3: During the construction of Super 70 (Phase-IA)	Feb – Jul 2007	Eastbound direction of interstate was shared by east and west bound traffic, changeable concrete barrier and no heavy vehicles, lane closure, ramps closure, and level of enforcement
4: During the construction of Super 70 (Phase-IB)	Jul 2007	Newly completed pavement used by westbound traffic, standard drum and no heavy vehicles, lane closure, ramps closure, and level of enforcement
5: During the construction of Super 70 (Phase-II)	Jul – Nov 2007	Westbound direction interstate was shared by east and westbound traffic, changeable concrete barrier and new barrier (permanent) and no heavy vehicles, lane closure, ramps closure, and level of enforcement
6: After completion of Super 70	Dec 2007 – Jun 2008	Standard concrete barrier and heavy vehicles, changes in highway geometry

4.2 Crash Data

The source crash dataset in the comma-delimited-values text format was imported to the Microsoft Access database. In the Access file each row represents a crash and each column represents all of the information collected by the police officer who investigated the crash. These data are organized in columns (i.e., variables). There are many variables in the original Automated Reporting Information Exchange System (ARIES), which is the updated official name in 2009 from VCRS Database. Each crash record starts with a unique master record number that serves as the crash ID. Table 4.3 shows the variables and their definitions and format included in the original crash dataset and imported to the Access database to be used in the project.

4.3 Geometry Data

Geometry data were extracted from the Google Earth geographic database and from the INDOT work zone drawings. The geometry data include the dimensions of the freeway components such as cross-

TABLE 4.3
Crash Data with Variables, Definition, and Formats

Variable Name	Definition	Format
Crash related Information		
MstrRecNbrTxt	This is an investigating agency's locally assigned crash identification number. This number should be a minimum of four (4) digits in length and unique in that it identifies the investigating agency and the individual report.	Text
CollDte	Month, calendar day, and year of crash using numeric symbols. For month use 01=January, 02=February, etc.; for the day use 01, 02, etc.; and for the year use all four digits 2003, 2004, etc.	Date/Time
CollDayWeekCde	Day that the crash occurred. [1-Sunday, 2-Monday, 3-Tuesday, 4-Wednesday, 5-Thursday, 6-Friday, 7-Saturday]	Text
CollTimeTxt	The local time that the crash occurred.	Text
MotorVehInvolvedNmb	The number of motor vehicles involved in the crash.	Number
InjuredNmb	The total number of people injured (including drivers, passengers, and non-motorists). A person shall be counted as injured if they have any injury listed in the Nature of Most Severe Injury category.	Number
DeadNmb	The total number of people fatally injured (including drivers, passengers, and non-motorists).	Number
InterMilemarkNmb	The number of the nearest mile marker to where the crash occurred.	Number
FeetFromPointNmb	The number of feet (or tenths of a mile) from the location identified in the nearest mile marker.	Number
DirFromPointCde	The direction (N, S, E, W, NE, SE, SW, or NW) from the location identified in the nearest/ Intersecting road,mile marker, or interchange.	Text
LatDecimalNmb	Crash latitude, using degrees, minutes and seconds.	Number
LongDecimalNmb	Crash longitude, using degrees, minutes and seconds.	Number
LightCondCde	Light condition at the time and place of crash [01-Daylight, 02-Dawn/Dask, 03-Dark (Lighted), 04-Dark (Not Lighted), 05-Unknown]	Text
WeatherCde	Weather condition at the time and place of crash. [01-clear, 02-Cloudy, 03-Rain, 04-Snow, 05-Sleet/Hail/Freezing Rain, 06-Severe Cross Wind, 07-Blowing Sand/Soil/Snow]	Text
SurfaceCondCde	Road surface conditions at the time and place of the crash. [01-Dry, 02-Wet, 03-Muddy, 04-Snow/Slush, 05-Ice, 06-Loose Material on Road, 07-Water (Standing and Moving)]	Text
MannerCollCde	The type of crashes. [01-Rear End, 02-Head On, 03-Rear to Rear, 04-Same Direction Sideswipe,05-Opposite Direction Sideswipe, 06-Ran-off Road, 07-Right Angle, 08-Left Turn, 09-Right Turn, 10-Left/Right Turn, 11-Backing, 12-Other (explain in narrative), 13-Non-Collision]	Text
RdwyClassCde	The highest classification for the road the crash occurred on [01-Interstate, 02-US Route, 03-State Road, 04-Couty Road, 05-Local/City Road, 06-Unknown]	Text
Person related Information		
InjStatusCde	Injury status of occupants. [01-Fatal, 02-Incaopacitating, 03-Non-incapacitating, 04-Possible, 05-not Reported, 06-Unknown, 07-Refused]	Text
Vehicle related information		
UnitTypeCde	Type of vehicle involved. [01-Passenger Car/Station Wagon, 02-Pickup, 03-Van, 04-Sport Utility Vehicle, 05-Truck (Single 2 Axle, 6 Tires), 06-Truck(Single 3 or more Axles), 07-Truck/Trailer (not semi), 08-Tractor/One Semi Trailer, 09-Tractor/Double Trailer, 10-Tractor/Triple Trailer, 11-Tractor (Cab Only, No Trailer), 12-Motor Home/Recreational Vehicle, 13-Motorcycle, 14-Bus/Seats 9–15 Persons with Driver, 15-Bus/ Seats 15+ Persons with Driver, 16-School Bus, 17-Unknow Type, 18-Farm Type, 19-Combination Vehicle, 20-Pedestrian, 21-Bicycle]	Number
OccupsNmb	Number of occupants, including the driver, that were riding in or on the vehicle at the time of the crash.	Number

sections, ramp intersections, and horizontal curvature of road segments inside the work zone and in the non-work zone area of Indianapolis. The geometry data for all non–work zone segments were extracted with the help of the Google Earth measurement tools. The geometry data for the work zone were measured or read from the INDOT work-zone drawings. Table 4.4 shows the geometry data, definitions, and formats of the data.

Additional information was added to the segment geometry dataset in the form of comments such as the different phases of construction and the entrance and exit ramp name/ID information. Other variables included in the segment geometry dataset are as follows:

- TMP_BRR: Temporary barrier, particularly includes movable barrier (Type – 2), and temporary fixed barrier (Type – 4) in the construction zone during the construction period (Period – 2, 3 and 4). See Appendix C, Figure C.2.
- RLATCL_IN: the average horizontal clearance from the edge of the inside shoulder line of the road to the physical obstruction (i.e., median barrier).
- RLATCL_OUT: the average horizontal clearance from the edge of the outside shoulder line of the road to the physical obstruction (i.e., roadside barriers).
- CD_RMP: It usually happens at the collector-distributor (CD) road connecting the main line. See Appendix C Figure C.1.
- CD_MERGE: when CD road merges to main line. See Appendix C Figure C.1.
- CD_DIVERGE: when CD road diverges from main line. See Appendix C Figure C.1.
- ONE_WDIV: when there is concrete barrier between CD road and mainline. See Appendix C Figure C.1.

Appendix A, Table A.2 gives a complete list of converted variables for geometry dataset.

TABLE 4.4
Geometry Data with Variables, Definitions, and Formats

Variable/p>	Definition	Format
RDNAME	Road name	Text
TRADIR	Traffic direction (F/B)	Text
RSEGLEN	Segment length (mi)	Number
STMP	Start of Mile Post	Number
ENDMP	End of Mile Post	Number
LN_N	Number of lanes in one direction	Number
IN_SHLDR	Presence of inside shoulder	Binary number
OUT_SHLDR	Presence of outside shoulder (1 = yes, 0 = no)	Binary number
LN_W	Average lane width (ft)	Number
IN_SHLDR_W	Average inside shoulder width (ft)	Number
OUT_SHLDR_W	Average outside shoulder width (ft)	Number
SHLDR_USE	Using shoulder as travel lane (1 = yes, 0 = no)	Binary number
UP_RMP	Presence of upstream ramp (1 = yes, 0 = no)	Binary number
UP_RMP_ID	Upstream ramp ID	Text
UP_RMP_EFF	Upstream ramp effect (1-0 influence function based on 0.25 mile)	Number
UP_ON_RMP	On ramp/off ramp (1 = on, 0 = off)	Binary number
UP_RHS_RMP	Right/left hand side (1 = right, 0 = left)	Binary number
UP_OPEN_RMP	Open/closed (1 = open, 0 = closed)	Binary number
UP_TAP_RMP	Tapered/parallel (1 = tapered, 0 = parallel)	Binary number
UP_ONELN_RMP	One lane/two lane (1 = one lane, 0 = two lanes)	Binary number
DWN_RAMP	Presence of downstream ramp (1 = yes, 0 = no)	Binary number
DWN_RMP_ID	Downstream ramp ID	Text
DWN_RMP_EFF	Downstream ramp effect (1-0 influence function based on 0.25 mile)	Number
DWN_ON_RMP	On ramp/off ramp (1 = on, 0 = off)	Binary number
DWN_RHS_RMP	Right/left hand side (1 = right, 0 = left)	Binary number
DWN_OPEN_RMP	Open/closed (1 = open, 0 = closed)	Binary number
DWN_TAP_RMP	Tapered/parallel (1 = tapered, 0 = parallel)	Binary number
DWN_ONELN_RMP	One lane/two lane (1 = one lane, 0 = two lanes)	Binary number
TMP_BRR	Temporary barrier (1 = median/roadside barrier, 0 = outside)	Binary number
MED_BRR	Presence of median barrier (1 = yes, 0 = no)	Binary number
RDSD_BRR	Presence of roadside barrier (1 = yes, 0 = no)	Binary number
LN_DRP	Number of lane(s) dropped	Number
H_CURVE	Average horizontal curve (1/ft)	Number
RLATCL_IN	Average inside lateral clearance (ft)	Number
RLATCL_OUT	Average outside lateral clearance (ft)	Number
CURV_SEG_N	Number of curve per segment	Number
CD_RMP	CD road at gore-point (1 = yes, 0 = no)	Binary number
CD_MERGE	Merging (1 = yes, 0 = no)	Binary number
CD_DIVERGE	Diverging (1 = yes, 0 = no)	Binary number
ONE_WDIV	One way divider (1 = yes, 0 = no)	Binary number

4.4. Traffic Data

Traffic data were obtained from INDOT in the binary form as collected with permanent loop detectors in the Indianapolis area. These binary files were downloaded from the INDOT ftp site and converted with the Traffic Data Management Software (Centurion CC) to ASCII format for further processing. Table 4.5 shows the variables, their meanings, and their formats extracted from the original binary files.

4.5 Weather Data

Weather data were obtained from the National Climatic Data Center (NCDC), Indianapolis International Airport Weather station, having five-digit Weather-Bureau-Army-Navy (WBAN) identifier 93819 and six-digit USAF master station identifier 724380 at the NCDC was selected as source of the weather data for this research work. The NCDC's data explanation was used to determine the desired weather parameters to cover further. Table 4.6 shows the variables, their meanings, and their formats directly extracted from the NCDC database.

4.6 Traffic Maintenance Data

Work zone traffic maintenance data (e.g., cross-over location, temporary barrier movement, lane addition and lane drop, the presence of a standard drum) were extracted from the INDOT work zone drawings. Table 4.7 shows the variables, their meanings, and their formats directly extracted from the INDOT drawings.

4.7 Enforcement Data

Enforcement data include the total worked hours of the highway patrol, traffic citations issued, moving citations, speeding citations, all truck citations, and truck (moving citations). These data were date-aggregated over the duration of the Super 70 work zone. Table 4.8 shows the variables, their definitions, and their formats extracted from the INDOT activity log.

5 DATA CONVERSION

The sources data had to be converted to a format more suitable for statistical modeling. The conversion included but was not limited to changing units, dropping variables not needed, creating variables from other source variables, and aggregation. This chapter describes the conversion process for all the types of collected and processed data.

5.1 Crash Dataset

The obtained variables in the converted crash dataset are shown in Table A.1 in Appendix A. Since the needed data were supposed to apply to individual crashes, it was necessary to aggregate the selected pieces of the source data applying to individual vehicles and the involved persons to represent a crash. The source crash data were processed as follows:

- Severity of injury [SEVINJ] for individuals involved in a crash: The conversion is explained in Table 5.1.
- The severity of a crash is defined by the most severe injury of individuals involved in the crash.
- Number of vehicle involved [VEHNUM]: information is extracted from the vehicle dataset of original database based on Crash_ID.
- Number of occupants involved [OCCNUM]: information is extracted from the person dataset of original database based on Crash_ID.
- Dry surface [RDRY]: SurfaceCondCde variable in the source data was converted based on its value [01-Dry] and the obtained variable is RDRY (1 or 0).
- Wet surface [RWET]: SurfaceCondCde variable in the source data was converted based on its value [02-Wet] and the obtained variable is RWET, which is binary in nature (1 or 0).
- Muddy surface [RMUD]: SurfaceCondCde variable in the source data was converted based on its value [03-Muddy] and the obtained variable is RMUD, which is binary in nature (1 or 0).
- Snow/Slush on surface [RSNOW]: SurfaceCondCde variable in the source data was converted based on its value [04-Snow/Slush] and the obtained variable is RSNOW, which is binary in nature (1 or 0).
- Ice on surface [RICE]: SurfaceCondCde variable in the source data was converted based on its value [05-Ice] and the obtained variable is RICE, which is binary in nature (1 or 0).

TABLE 4.5
Traffic Data with Variables, Definitions, and Formats

Variable	Definition	Format
Lane	Lane designation (Left, middle, right)	Number
Direction	Direction of Traffic	Text
Day	Day of data collection	Date (MM/DD/YY)
Time	Time of traffic data collection	Time (24-hr Clock)
Volume	16 bin traffic count	Number
Speed	13 bin speed count	Number

TABLE 4.6
Weather Data with Variables, Definitions, and Formats

Variable	Definition	Format
Time	GEOPHYSICAL-POINT-OBSERVATION time	Time (HRMM)
Date	The date of a GEOPHYSICAL-POINT-OBSERVATION	Number
VISBY	The horizontal distance at which an object can be seen and identified	Number
TEMP	AIR-TEMPERATURE-OBSERVATION air temperature (deg cel.)	Number
PRECIP – AMT	LIQUID-PRECIPITATION depth/amount (mm)	Number
CLOUD SUMM	The code that denotes the portion of the total celestial dome covered by all layers of clouds and other obscuring phenomena at or below a given height Clear - No coverage: 0 FEW - 2/8 or less coverage (not including zero): 1 SCATTERED - 3/8-4/8 coverage: 2 BROKEN - 5/8-7/8 coverage: 3 OVERCAST - 8/8 coverage: 4	Number
WX-M	Manual present weather: 00: Cloud development not observed or not observable 01: Clouds generally dissolving or becoming less developed 02: State of sky on the whole unchanged 03: Clouds generally forming or developing 04: Visibility reduced by smoke, e.g. veldt or forest fires, industrial smoke or volcanic ashes 05: Haze 06: Widespread dust in suspension in the air, not raised by wind at or near the station at the time of observation 07: Dust or sand raised by wind at or near the station at the time of observation, but no well-developed dust whirl(s) or sand whirl(s), and no duststorm or sandstorm seen or, in the case of ships, blowing spray at the station 08: Well developed dust whirl(s) or sand whirl(s) seen at or near the station during the preceding hour or at the time of observation, but no duststorm or sandstorm 09: Duststorm or sandstorm within sight at the time of observation, or at the station during the preceding hour 10: Mist 11: Patches of shallow fog or ice fog at the station, whether on land or sea, not deeper than about 2 meters on land or 10 meters at sea 12: More or less continuous shallow fog or ice fog at the station, whether on land or sea, not deeper than about 2 meters on land or 10 meters at sea 13: Lightning visible, no thunder heard 14: Precipitation within sight, not reaching the ground or the surface of the sea 15: Precipitation within sight, reaching the ground or the surface of the sea, but distant, i.e., estimated to be more than 5 km from the station 16: Precipitation within sight, reaching the ground or the surface of the sea, near to, but not at the station 17: Thunderstorm, but no precipitation at the time of observation 18: Squalls at or within sight of the station during the preceding hour or at the time of observation 19: Funnel cloud(s) (Tornado cloud or waterspout) at or within sight of the station during the preceding hour or at the time of observation 20: Drizzle (not freezing) or snow grains not falling as shower(s) 21: Rain (not freezing) not falling as shower(s) 22: Snow not falling as shower(s) 23: Rain and snow or ice pellets not falling as shower(s) 24: Freezing drizzle or freezing rain not falling as shower(s) 25: Shower(s) of rain 26: Shower(s) of snow or of rain and snow 27: Shower(s) of hail (Hail, small hail, snow pellets), or rain and hail 28: Fog or ice fog 29: Thunderstorm (with or without precipitation) 30: Slight or moderate duststorm or sandstorm has decreased during the preceding hour 31: Slight or moderate duststorm or sandstorm no appreciable change during the preceding hour 32: Slight or moderate duststorm or sandstorm has begun or has increased during the preceding hour 33: Severe duststorm or sandstorm has decreased during the preceding hour 34: Severe duststorm or sandstorm no appreciable change during the preceding hour 35: Severe duststorm or sandstorm has begun or has increased during the preceding hour 36: Slight or moderate drifting snow generally low (below eye level) 37: Heavy drifting snow generally low (below eye level) 38: Slight or moderate blowing snow generally high (above eye level)	

TABLE 4.6
(Continued)

Variable	Definition	Format
	39: Heavy blowing snow generally high (above eye level) 40: Fog or ice fog at a distance at the time of observation, but not at the station during the preceding hour, the fog or ice fog extending to a level above that of the observer 41: Fog or ice fog in patches 42: Fog or ice fog, sky visible, has become thinner during the preceding hour 43: Fog or ice fog, sky invisible, has become thinner during the preceding hour 44: Fog or ice fog, sky visible, no appreciable change during the preceding hour 45: Fog or ice fog, sky invisible, no appreciable change during the preceding hour 46: Fog or ice fog, sky invisible, has begun or has become thicker during the preceding hour 47: Fog or ice fog, sky invisible, has begun or has become thicker during the preceding hour 48: Fog, depositing rime, sky visible 49: Fog, depositing rime, sky invisible 50: Drizzle, not freezing, intermittent, slight at time of observation 51: Drizzle, not freezing, continuous, slight at time of observation 52: Drizzle, not freezing, intermittent, moderate at time of observation 53: Drizzle, not freezing, continuous, moderate at time of observation 54: Drizzle, not freezing, intermittent, heavy (dense) at time of observation 55: Drizzle, not freezing, continuous, heavy (dense) at time of observation 56: Drizzle, freezing, slight 57: Drizzle, freezing, moderate or heavy (dense) 58: Drizzle and rain, slight 59: Drizzle and rain, moderate or heavy 60: Rain, not freezing, intermittent, slight at time of observation 61: Rain, not freezing, continuous, slight at time of observation 62: Rain, not freezing, intermittent, moderate at time of observation 63: Rain, not freezing, continuous, moderate at time of observation 64: Rain, not freezing, intermittent, heavy at time of observation 65: Rain, not freezing, continuous, heavy at time of observation 66: Rain, freezing, slight 67: Rain, freezing, moderate or heavy 68: Rain or drizzle and snow, slight 69: Rain or drizzle and snow, moderate or heavy 70: Intermittent fall of snowflakes, slight at time of observation 71: Continuous fall of snowflakes, slight at time of observation 72: Intermittent fall of snowflakes, moderate at time of observation 73: Continuous fall of snowflakes, moderate at time of observation 74: Intermittent fall of snowflakes, heavy at time of observation 75: Continuous fall of snowflakes, heavy at time of observation 76: Diamond dust (with or without fog) 77: Snow grains (with or without fog) 78: Isolated star-like snow crystals (with or without fog) 79: Ice pellets 80: Rain shower(s), slight 81: Rain shower(s), moderate or heavy 82: Rain shower(s), violent 83: Shower(s) of rain and snow mixed, slight 84: Shower(s) of rain and snow mixed, moderate or heavy 85: Snow shower(s), slight 86: Snow shower(s), moderate or heavy 87: Shower(s) of snow pellets or small hail, with or without rain or rain and snow mixed, slight 88: Shower(s) of snow pellets or small hail, with or without rain or rain and snow mixed, moderate or heavy 89: Shower(s) of hail (hail, small hail, snow pellets) , with or without rain or rain and snow mixed, not associated with thunder, slight 90: Shower(s) of hail (hail, small hail, snow pellets), with or without rain or rain and snow mixed, not associated with thunder, moderate or heavy 91: Slight rain at time of observation, thunderstorm during the preceding hour but not at time of observation 92: Moderate or heavy rain at time of observation, thunderstorm during the preceding hour but not at time of observation 93: Slight snow, or rain and snow mixed or hail (Hail, small hail, snow pellets), at time of observation, thunderstorm during the preceding hour but not at time of observation	

TABLE 4.6
(Continued)

Variable	Definition	Format
	94: Moderate or heavy snow, or rain and snow mixed or hail(Hail, small hail, snow pellets) at time of observation, thunderstorm during the preceding hour but not at time of observation 95: Thunderstorm, slight or moderate, without hail (Hail, small hail, snow pellets), but with rain and/or snow at time of observation, thunderstorm at time of observation 96: Thunderstorm, slight or moderate, with hail (hail, small hail, snow pellets) at time of observation, thunderstorm at time of observation 97: Thunderstorm, heavy, without hail (Hail, small hail, snow pellets), but with rain and/or snow at time of observation, thunderstorm at time of observation 98: Thunderstorm combined with duststorm or sandstorm at time of observation, thunderstorm at time of observation 99: Thunderstorm, heavy, with hail (Hail, small hail, snow pellets) at time of observation, thunderstorm at time of observation	

TABLE 4.7
Traffic Maintenance Data with Variables, Definition and Formats

Variable	Definition	Format
RDNAME	Road name	Text
TRADIR	Traffic direction (F/B)	Text
RSEGLEN	Segment length (mi)	Number
STMP	Start of Mile Post	Number
ENDMP	End of Mile Post	Number
TIME	Time	Time
DATE	Date	Date
CBARIN	Barrier moved inside (1 = inside, 0 = outside)	Binary Number
CBAROUT	Barrier moved outside (1 = inside, 0 = outside)	Binary Number
CXOVER	Cross-over (1 = yes, 0 = no)	Binary Number
CSHOULDERN	Shoulder narrow (1 = yes, 0 = no)	Binary Number
CSTDDRUM	Standard drum (1 = yes, 0 = no)	Binary Number
LNADD	Lane added (1 = yes, 0 = no)	Binary Number
LNDRP	Lane dropped (1 = yes, 0 = no)	Binary Number

TABLE 4.8
Enforcement Data with Variables, Definitions, and Formats

Variable	Definition	Format
Time	Month	Text
Work Hour	Total hours worked	Number
Traffic Citation	Number of traffic citations	Number
Moving Citation	Number of moving citations	Number
Speed Citation	Number of speed citations	Number
All Truck Citations	Number of all truck citations	Number
Truck (moving) Citations	Number of truck (moving citations)	Number

TABLE 5.1
Injury Severity in the Converted Crash Dataset

Person Injury Record Presence	Original Entry	Converted Value	Remark
Yes	Fatal	Fatal	None
	Incapacitating	Incapacitating	None
	Non-incapacitating	Non-incapacitating	None
	Possible Unknown Refused	Possible	Possible category is the combination of Possible, Unknown, and Refused categories
No	NA	No injury	Absence of the injury record implies no injury damage crash was added in this category

- Loose material on surface [RLOOSE]: SurfaceCondCde variable in the source data was converted based on its value [06-Loose] and the obtained variable is RLOOSE, which is binary in nature (1 or 0).
- Standing water on surface [RWATER]: SurfaceCondCde variable in the source data was converted based on its value [07-Water] and the obtained variable is RWATER, which is binary in nature (1 or 0).
- Head-on collision [HAEDON]: MannerCollCde variable in the source data was converted based on its value [02-headon] and the obtained variable is HEADON, which is binary in nature (1 or 0).
- Rear-end collision [REAREND]: MannerCollCde variable in the source data was converted based on its value [01-rearend] and the obtained variable is REAREND, which is binary in nature (1 or 0).
- Sideswipe collision [SIDESWIPE]: MannerCollCde variable in the source data was converted based on its value [04-sideswipe same direction, 05-sideswipe opposite direction] and the obtained variable is SIDESWIPE, which is binary in nature (1 = same direction or 0 = opposite direction).
- Right angle collision [RANGLE]: MannerCollCde variable in the source data was converted based on its value [07-right angle] and the obtained variable is RANGLE, which is binary in nature (1 or 0).
- Turning collision [TURN]: MannerCollCde variable in the source data was converted based on its value [08-left turn, 09-right angle] and the obtained variable is TURN, which is binary in nature (1 = right or 0 = left).
- 13+ton truck involved [HVEH]: UnitTypeCde variable in the source data was converted based on its value [07-truck/trailer (not semi), 08-tractor/One semi trailer, 09-tractor double trailer, 10-tractor/triple trailer, 12-motor home/recreational vehicle, 15-bus/seats 15+ persons with driver, 16-school bus, 19-combination vehicle] and the obtained variable is HVEH, which is binary in nature (1 or 0).
- Single unit truck involved [SUT]: UnitTypeCde variable in the source data was converted based on its value [05-truck (single 2 axle, 6 tires), 06-truck (single 3 or more axles)] and the obtained variable is SUT, which is binary in nature (1 or 0).
- Non-heavy vehicle involved [NHVEH]: UnitTypeCde variable in the source data was converted based on its value [01-passenger car/station wagon, 02-pickup, 03-van, 04-SUV, 11-tractor (cab only, no trailer), 13-motorcycle, 14-bus/seats 9–15 persons with driver, 18-farm vehicle] and the obtained variable is NHVEH, which is binary in nature (1 or 0).
- Standardized road name [SRDNAME]: space between in the name of road (text format) in source crash dataset was standardized removing all space between words.
- Corrected MM [LDIST]: Linear distance of crash location with respect to linear mile marker was corrected based on the Direction of crash from MM [DIRCMM] and Distance from point (feet) [taken from original source] [DISTMM].
- Single vehicle involved [SINGVEH]: MotorVehInvolvedNmb variable in the source data was filtered into one vehicle involved and the obtained variable is SINGVEH, which is binary in nature (1 or 0).
- Multiple vehicles involved [SINGVEH]: MotorVehInvolvedNmb variable in the source data was filtered into more than one vehicle involved and the obtained variable is MULTVEH, which is binary in nature (1 or 0).
- Crash period according to period defined [CPERIOD]: Crash period was categorized based on the date defined in Table 5.2.
- Standardized travel direction of traffic [STDDIR]: Direction of travel is based on the mile marker of the interstate system as found in the GIS map (see Appendix D). An increasing mile marker is defined as "Forward" or "F" and a decreasing mile marker as "Backward" or "B". See Appendix F for data management for converted crash dataset.

5.2 Geometry Dataset

The length of a segment has been targeted as 0.25 mile and all segments are kept to their end or start from the ramp location (on and off-ramp) in this study. However, due to the short spacing between ramps and some other practical considerations of geometry in the vicinity of the ramp area, the length of a segment could be less or more than 0.25 mile. However, on average, the length of segment is approximately 0.25 mile.

Non-work zone segments were selected based on the location of the detectors and the traffic data availability. There are total 467 segments: 201 are classified as non-work zone (NWZ); and 266 are classified as work zone (WZ), which are further classified as follows:

- Before construction: 49
- During construction (3 phases): 168
- After construction: 49

The WZ segment selection was based on the following criteria:

1. Open and closed ramps within segments identified
2. Change of geometry due to the traffic maintenance by INDOT

In order to clearly understand the INDOT drawings and the traffic operations, correspondence with those parties involved in the work zone construction, such as

TABLE 5.2
Periods Identified for Crash Date

Date	Period	Segment under consideration	Phase
1/1/2006 – 2/22/2007	1	Before construction (I70)	-
2/23/2007 – 7/12/2007	2	During construction (Phase-IA)	Phase-IA
7/13/2007 – 7/17/2007	3	During construction (Phase-IB)	Phase-IB
7/18/2007 –11/16/2007	4	During construction (Phase-II)	Phase-II
11/17/07 – 6/30/2008	5	After construction (I70)	-

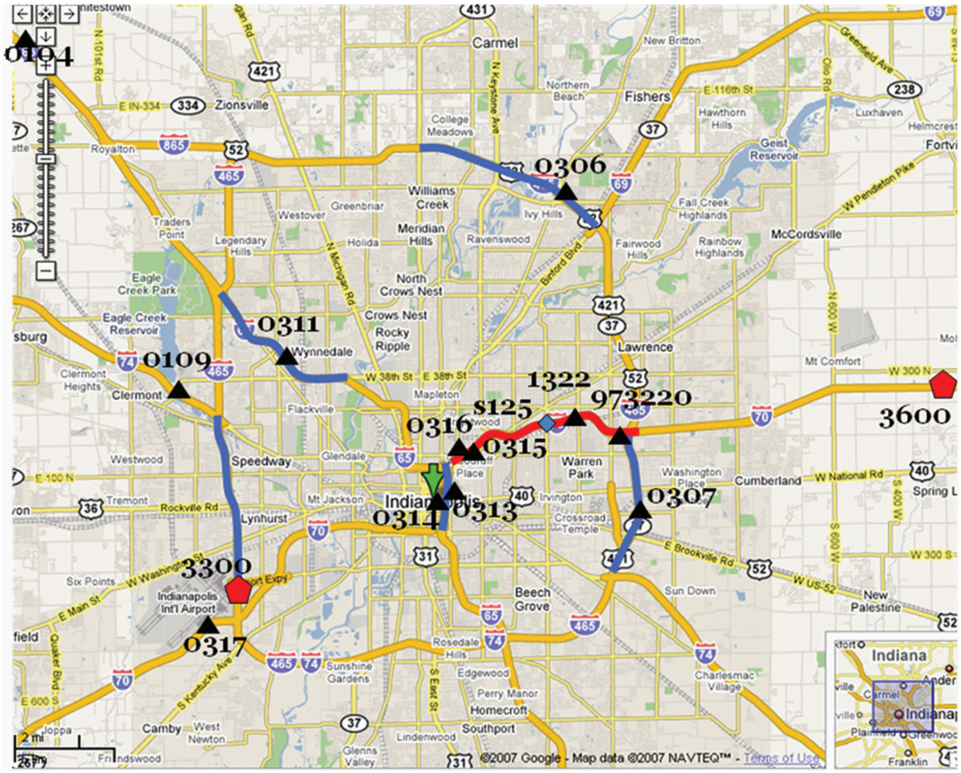


Figure 5.1 Interstate Sections and Traffic Detectors Included in the Study Legend: ATR, WIM, Sidefire unit

Walsh Construction Company and United Consulting Company, were maintained during the analysis and report writing period.

Tables 5.3 and 5.4 define direction-wise the number of segments selected on the Interstates with different mile markers.

The components for highway geometry were collected by using Google Earth measurement tools and the INDOT drawings for the work zone segments.

TABLE 5.3
Summary of Number of Segments for NWZ

Direction	Mile Marker	Interstate	No. of Segments
NB – F	117 – 123	I-65	19
SB – B		I-65	18
EB – F	30 – 37	I-465	25
WB – B		I-465	25
NB – B	44 – 49	I-465	18
SB – F		I-465	19
NB – F	109 – 113	I-65	11
SB – B		I-65	10
NB – F	10 – 16	I-465	29
SB – B		I-465	27
Total	28	I-65 + I-465	201

Here, NB–northbound, SB–southbound, EB–eastbound, WB–westbound, F–forward, B–backward

5.3 Traffic Dataset

The traffic data were extracted from different detectors inside and outside the work zone. Figure 5.1 shows the location of detectors. Table A.3 in Appendix A shows the potential variables considered in the traffic dataset, which is 30-minute interval aggregated from the original 60-minute, 15-minute and 5-minute counts of bins from different detectors surrounding the study area. In addition, volume data were available for all detectors, but speed data were missing for several detectors during different time periods of data collection.

The vehicle classification is based on the Federal Highway Administration (FHWA) classification, which

TABLE 5.4
Summary of Number of Segments for WZ

Traffic Direction	No. of Segments	Phases	
Eastbound – F	26	Phase-IA	Phase-I
	26	Phase-IB	
	27	Phase-II	Phase-II
Sub total	*79*		
Westbound – B	29	Phase-IA	Phase-I
	30	Phase-IB	
	30	Phase-II	Phase-II
Subtotal	*89*		
Total	168		

was followed in the bins setup for all the detectors (i.e., Automatic Traffic Recorder (ATR)) under consideration. The Vehicles Classification (i.e., HEVVEH, MEDVEH, LIGHTVEH) was followed, which was based on the figure shown in Appendix B. During the data acquisition period from INDOT from their ftp source, two sets of traffic data obtained:

1) Individual detectors data as obtained periodically (binary format)
2) Converted traffic data (binary to ASCII) for all detectors (ATRs, WIMs)

In order to main consistency in the data over time from 2006 to 2007, all of the traffic data were converted based on source one (from binary to ASCII). And only 2008 traffic data were converted based on source two as mentioned earlier. Original traffic dataset was converted and aggregated as follows:

- Volume rate [VOLRATE]: Traffic volume was obtained in 30-minute interval on an hourly basis from the source data of five-minute (i.e., side fire radar unit) and 15-minute and 60-minute (ATRs, WIM) interval readings. A difference in time intervals for different detectors was found throughout the study period so readings for different intervals were all processed in MS Access for the uniform 30-minute interval.
- Average Speed [AVGSPEED]: Travel speed was obtained in 30-minute intervals on an hourly basis from the source data of five-minute (i.e., side fire radar unit) and 15-minute and 60-minute (ATRs, WIM) interval readings. A difference in time intervals for different detectors was found throughout the study period so readings for different intervals were all processed in MS Access for the uniform 30-minute interval.
- Percent of 13+ ton trucks [HEVVEH]: According to FWHA vehicle classification, classes 8–13 were combined to represent this class of vehicle in the sample. Bins 8–13 were aggregated and converted to the variable HEVVEH in percentage.
- Percent of less than 13-ton trucks [MEDVEH]: According to FWHA vehicle classification, classes 4-7 were combined to represent this class of vehicle in the sample. Bins 4–7 were aggregated and converted to the variable MEDVEH in percentage.
- Percent of 4-and 2-wheelers [LIGHTVEH]: According to FWHA vehicle classification, classes 1–3 were combined to represent this class of vehicle in the sample. Bins 1–3 were aggregated and converted to the variable LIGHTVEH in percentage.

Since there are four detectors (i.e., 1315, S125, 1322, and 973220) closely located inside the work zone, the detectors were assigned to the corresponding segments based on the distances between them. The segments and the corresponding detector information are given in Table E.1 in Appendix E.

Due to operational issues (detector shutdown or malfunction) during the construction period inside and outside the work zone for ATRs, WIM, and the side-fire radar unit, there was no data collection occasionally during 2006 and 2007. Some important detectors where traffic data were considered crucial for this study are mentioned here:

Detector related –

- ATR 0315 and 0316 were shut down from May 2006 to March 2008.
- Battery–operated ATR 973220 was not working during WZ but started at the end of project (November 2007 until end of project).
- ATR 0306 was not working mid-2007 until the end of the study period (August 2007 – June 2008)
- WIM 3300 collected no traffic data during January-November 2006.

Once the needed traffic data are obtained from all the detectors of interest, count models for imputation of traffic data were run based on the available traffic data from adjacent detectors. The time line for imputation of traffic data is shown in Figure F.7 in Appendix F. The simple multiple regression models were developed and run in a SAS program and processed in MS Excel. Table F.1 in Appendix F shows the 12 count models for traffic data – volume rate and heavy vehicles for respective detectors. In the count models, the time interval of a day (24-hour) (Appendix F, Table F.2) was considered (i.e., B1, B2, B3, B4, B5, B6) and interaction of these time interval with different detectors used in model development.

5.4 Weather Dataset

The Indianapolis International Airport weather station was chosen to provide the detailed weather parameters to represent the weather condition throughout the study area (work zone and non-work zone areas). In the source data, the time interval for observation was not exactly hourly basis so in the converted weather dataset, every interval of observation was shifted to the closet hour in order to maintain uniform hourly observations for the weather dataset. These generating intervals were conducted in the SAS program. The original traffic dataset was converted and aggregated as follows:

- The [CV] code denotes the portion of the total celestial dome covered by all layers of clouds and other obscuring phenomena at or below a given height. Missing values (app. 25%) have been imputed by assuming observations form adjacent hours. The codes and meaning are as follows:
  00: None, SKC or CLR
  01: One okta - 1/10 or less but not zero
  02: Two oktas - 2/10 - 3/10, or FEW
  03: Three oktas - 4/10
  04: Four oktas - 5/10, or SCT
  05: Five oktas - 6/10
  06: Six oktas - 7/10 - 8/10
  07: Seven oktas - 9/10 or more but not 10/10, or BKN
  08: Eight oktas - 10/10, or OVC
  09: Sky obscured, or cloud amount cannot be estimated
  10: Partial obscuration
- Visibility [WVISIB]: VISBY in the source data was in meters and so unit was converted in miles and missing

observations imputed by applying values from the adjacent hours. The new variable is named WVISIB.

- Temperature [BTEMP]: TEMP in the source data was in degree Celsius. The missing values (app. 25%) have been imputed by applying the observations from adjacent hours. Then, the variable has been converted to a binary variable representing the freezing conditions taking value 1 when TEMP<0 and 1, otherwise.
- Dawn, Dusk, and Dark interval [EDAWN and EDUSK]: EDAWN, EDUSK and DARK_HRS time for Indiana has been followed based on website weather information of Indiana [http://www.gaisma.com/en/location/indianapolis-indiana.html]. The chart of dawn, dusk, sunrise, sunset, and dark hours of the day for different months of the calendar year has been converted based on the time interval of the whole day. If certain intervals of the day corresponded to dawn, dusk, and dark periods of the day, they were converted to 1 or 0 based on the time interval.
- The manual observations coded in variable WX-M has been converted to binary variables representing the type of precipitation and its intensity following conversion rules summarized in Table 5.5. Due to a large amount of missing data, these values were not imputed and instead, the missing values were coded as 1 in binary variable BWXMISS in order to properly account for them in the models.

5.5 Traffic Maintenance Dataset

Traffic maintenance data were derived from the INDOT work zone drawings. Different traffic maintenance strategies adopted by INDOT were considered as potential traffic maintenance. The original traffic maintenance data were interpreted as follows:

- Barrier moved inside/outside [CBARIN and CBAROUT]: A barrier, particularly a median barrier, which is Type–4, was moved inside or outside during certain periods of time along certain segments within the work zone. However, there were certain time periods when a median barrier (Type-4) was not moved inside or outside, but rather was kept fixed during construction (i.e., Period 3). See Appendix C Figure C.2.
- Presence of cross-over [CXOVER]: Presence of a cross-over along certain segments. See Appendix C Figure C.3.
- Narrow shoulder [CSHOULDERN]: Presence of the “SHOULDER NARROWS” sign posted ahead warning about the narrow shoulder in certain segments during the construction. See Appendix C Figure C.4.
- Standard drum [CSTDDRUM]: Presence of a standard drum (2 feet diameter) along the segments in difference phases of construction. See Appendix C Figure C.5.
- Lane add and Lane drop [LNADD and LNDRP]: The transition interval for a movable barrier in different segment of the I-70 work zone section is considered as lane added (addition of one lane) and lane dropped (dropping of one lane). See Table G.3 in Appendix G.

5.6 Enforcement Dataset

Source enforcement data were aggregated over a month; all of the enforcement variables having monthly data were assigned to 30-minute intervals for both directions for the work zone segment. Appendix A Table A.6 shows the enforcement dataset. The converted enforcement dataset contains the following variables:

- Day: original monthly data were expanded throughout the day of corresponding month.
- Hour: original monthly data were attached to all hours of the day of the respective months.
- Minute: original monthly data were attached to every 30 minutes of the day of the respective months.
- HRWRK: obtained variable for total work hours of highway patrol.
- TRACT: obtained variable for traffic citations.
- MVCT: obtained variable for total moving citations.
- SPDCT: obtained variable for total speed citations.
- TRUCKCT: obtained variable for truck citations.
- TRUCKMCT: obtained variable for truck moving citations.

6 DATA LINKAGE

Different datasets were linked based on their common variables or linking variables between two datasets. This linkage process was completed for the crash severity and crash likelihood models. The following subsection introduces the process of linkage and the linking variables used in different steps.

6.1 Crash-Segment Assignment

In this phase, crash data with the linear location of crash (i.e., the nearest mile marker of the crash location) has been processed for whole crash dataset.

TABLE 5.5
Weather Manual Data WX-M Conversion

Binary Variable	Name	= 1 if WX-M code is:
Fog	BFOG	11,12,28,40-49,76-78
Rain	BRAIN	21,23,24-27,58-69, 81-84,87-95,97
Snow	BSNOW	20,22,23,26,27,36-39,68-75,77,78,83-90,93-97,99
Thunderstorm	BTSTRM	17,29,91-99
Slight Intensity	BSLIGHT	50,51,56,58,60,61,66,68,70,71,80,83,85,87,89,91,93
Slight or moderate intensity	BSLMOD	30-32,36,38,95,96
Moderate intensity	BMODER	52,53,62,63,72,73
Moderate or heavy intensity	BMODHV	57,59,67,69,81,84,86,88,90,92,94
Heavy intensity	BHEAVY	37,39,54,55,64,65,74,75,97,99

LDIST is a variable in the converted crash dataset used for crash assignment in the segment dataset. In the original crash data, the variables of Interstate mile marker (InterMilemarkNmb), direction from the mile marker (DirFromPointCde), and distance (feet) from the mile marker (FeetFromPointNmb) were used to determine the LDIST variable in the converted crash dataset. In the segment dataset, there are starting mile post (STMP) and ending mile post (ENDMP) variables. LDIST in the crash dataset is compared with STMP and ENDMP under a criterion STMP<LDIST<ENDMP. If this condition satisfied, Crash IDs were assigned to segment datasets where the Segment ID, Travel Direction (TRADIR), and Road Name (RDNAME) were used in the linking process in MS Access. In the assignment process, the segment dataset with Segment ID receives Crash ID. The crash-segment assignment is flow is shown in Figure 6.1.

Figure 6.2 explains the process for designating the location of the crash and direction of travel in crash assignment, and Segment ID set up with a schematic diagram. If the location of the crash within the identified segment (i.e., bounded by STMP and ENDMP) is found to be east of the mile marker which is increasing in the direction of the mile marker (MM 86 as shown in Figure 6.2), then LDIST = InterMildemarkNmb + FeetFromPointNmb (converted to miles). If the travel direction is found to be eastbound, then standardized travel direction (i.e., STDDIR) is F which is applicable for Crash A (shown in Figure 6.2). The other scenario, shown as Crash B, which is in a decreasing direction of the mile marker (MM 87 as shown in Figure 6.2), then LDIST = InterMildemarkNmb – FeetFromPointNmb (converted to miles) and travel direction is westbound and STDDIR is B.

6.2 Weather Data Linkage

After crash-segment assignment, the weather dataset is linked based on the date and time. In this dataset, the observation time for weather parameters were not uniform as the intended time interval (i.e., 0, 30 minutes). So this time of observation was processed to make it uniform considering the closest 30-minute interval. Then the weather dataset was linked with the crash-segment dataset as shown in Figure 6.3.

6.3 Traffic Dataset Linkage

After linking the weather dataset with the crash-segment dataset, proper linking variables were selected (e.g., Segment ID, Road Name, Travel Direction, and Period) with the detector dataset. In the detector dataset, the detectors were assigned to the segments (i.e., Segment IDs). In the traffic dataset, the detectors

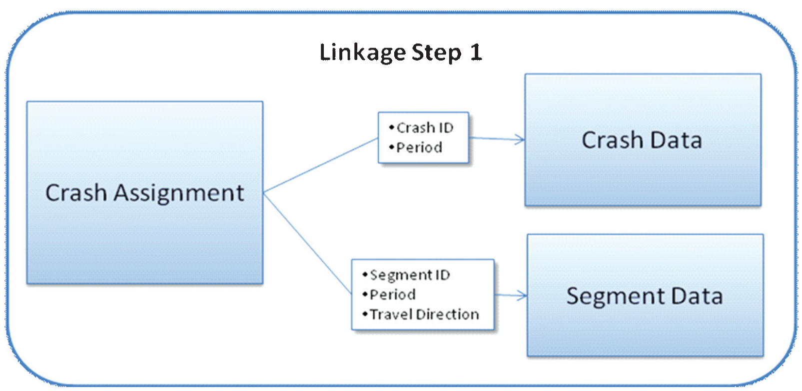


Figure 6.1 Crash-Segment Assignment Process

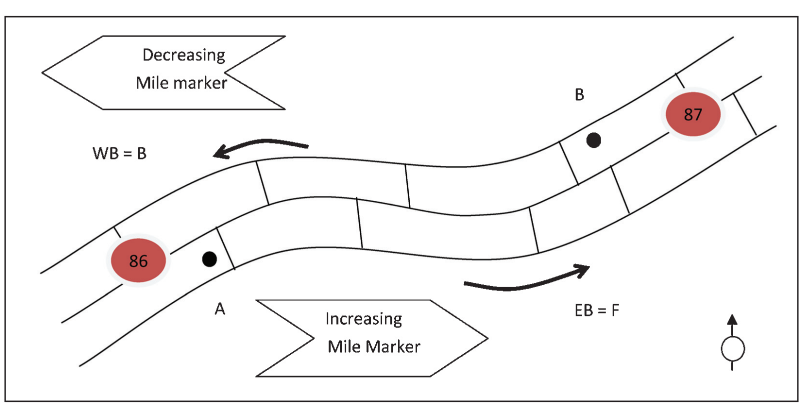


Figure 6.2 Illustrations of Crash Location (LDIST) and Travel Direction (TRADIR)

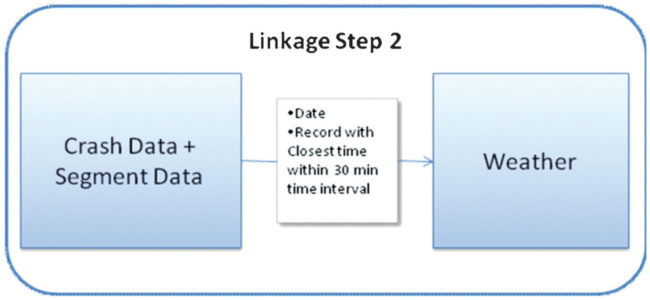


Figure 6.3 Linkages with Weather Dataset

were assigned so the first linking of the crash-segment-weather dataset with the detector dataset produced the Segment IDs. This new combined dataset received the linking variables, the Detector IDs from the traffic dataset. Linkage was then completed with the traffic dataset based on the Detector ID, the Date, the Travel Direction and the Time Interval as shown in Figure 6.4.

6.4 Traffic Maintenance Dataset Linkage

After the linking of the crash-segment-weather-detector-traffic, the traffic maintenance dataset was linked based on linking variables (Segment ID, Road Name, Travel Direction, and Period). All traffic maintenance related data were assigned to segments (i.e., Segment IDs) which were used in linking with the remaining linked dataset as shown in Figure 6.5.

6.5 Enforcement Dataset Linkage

The enforcement dataset was linked at the final stage. Date and Time were the linking variables to the prior linked variables of the dataset as shown in Figure 6.6.

7 CRASH LIKELIHOOD MODEL *P*(*C*)

This chapter describes development of the crash likelihood model P(C) mentioned in Chapter 3. This model is a key component of investigating the safety factors. The following sections describe the sampling procedure necessary to handle a large number of observations; presents the obtained model; and discusses the obtained results.

7.1 Crash Likelihood Sample

The number of observations is the product of the number of 30-minute intervals in the studied 2.5-year period and the number of studied road segments.

Table 7.1 estimates the total number of observations. There are 201 road segments outside of the Super 70 work zone during the entire studied period, while the number of road segments in the Super 70 section was different in different parts of the study period due to the differences in road geometry before, during, and after the construction project. The study period was divided into five sub-periods with a corresponding numbers of road segments (Table 7.1).

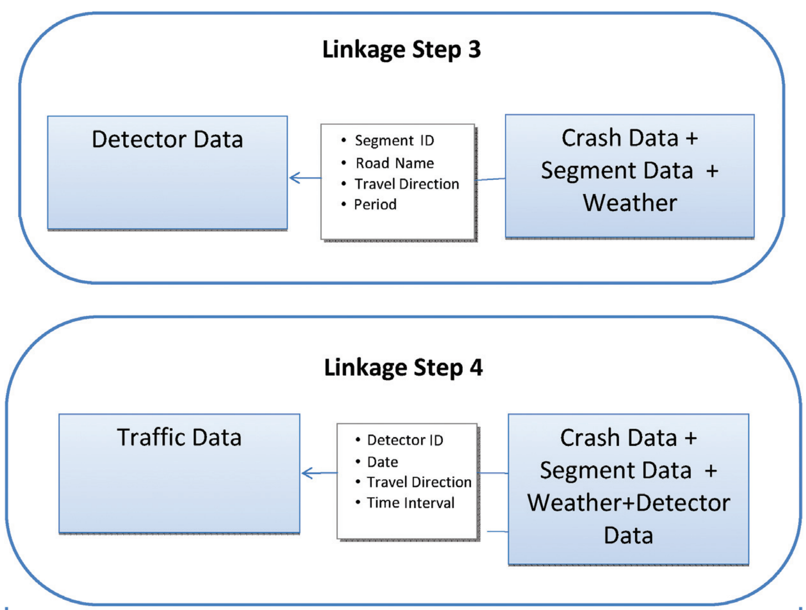


Figure 6.4 Linkage with Detector and Traffic Dataset

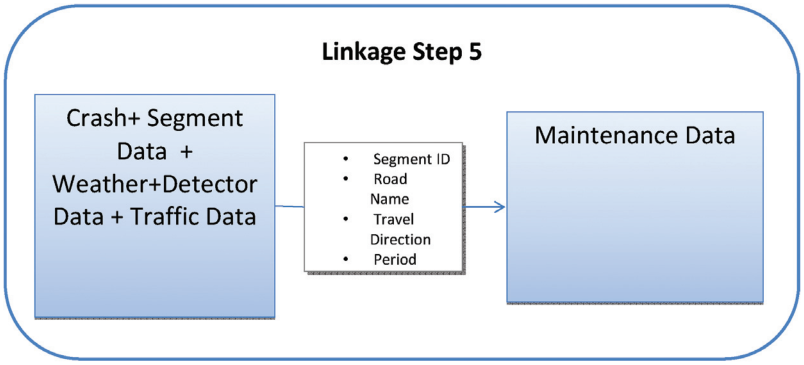


Figure 6.5 Linkages with Maintenance Dataset

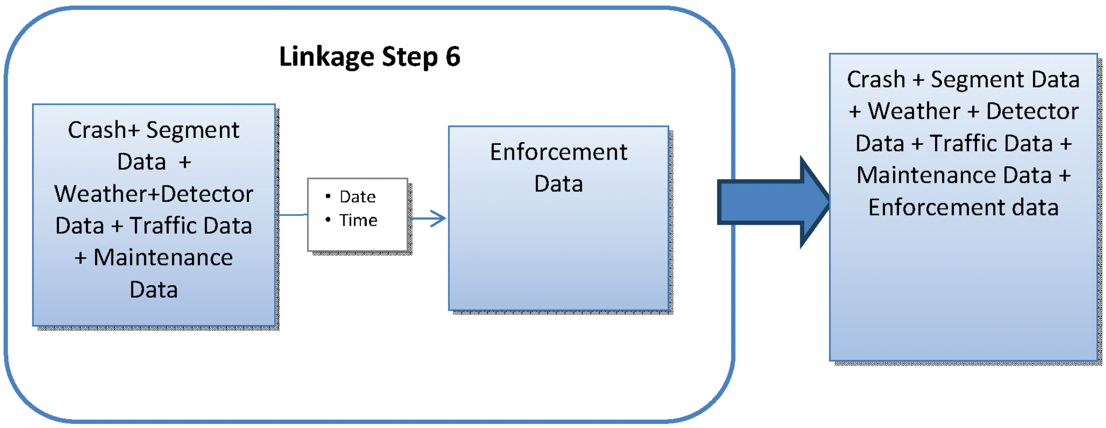


Figure 6.6 Linkage with Enforcement Dataset

TABLE 7.1
Entire Sample Size

Subsample	Sub-period	Work Zone	No. of Segments	No. of days	No. of Observations = No. of Segments x No. of Days x 48
Non-Super 70 section	1 – 5	No	201	899	8,673,552
Super 70 section	1	No	49	411	966,672
Super 70 section	2	Yes	110	140	739,200
Super 70 section	3	Yes	56	5	13,440
Super 70 section	4	Yes	114	119	651,168
Super 70 section	5	No	49	224	526,848
	Entire sample size				11,570,880

The sample size was over 11 million observations. It was not possible to use all the observations in the model estimation due to the large sample size. Instead, we adopted a sampling method to randomly select a manageable number of observations from the entire sample. All observations with crashes and about 850,000 randomly selected observations without crashes were selected. Approximately half of the sample comes from the I-70 section and half from the other sections. The process of selecting non-crash observations is presented in Figure 7.1. Consequently, the sub-sample for model estimation included all crash observations and around one percent of the non-crash observations.

Table 7.2 provides the numbers of observations after the random selection. The sample did not include observation duplicates and observations with missing traffic data. This initial sample for modeling included missing values of other variables because we did not know at the time of sampling which of these variables would be included in the final model, except traffic volume. During the modeling phase, we eliminated some observations that had missing values of variables included in the model.

Table 7.3 shows the number of observations used to develop a final model.

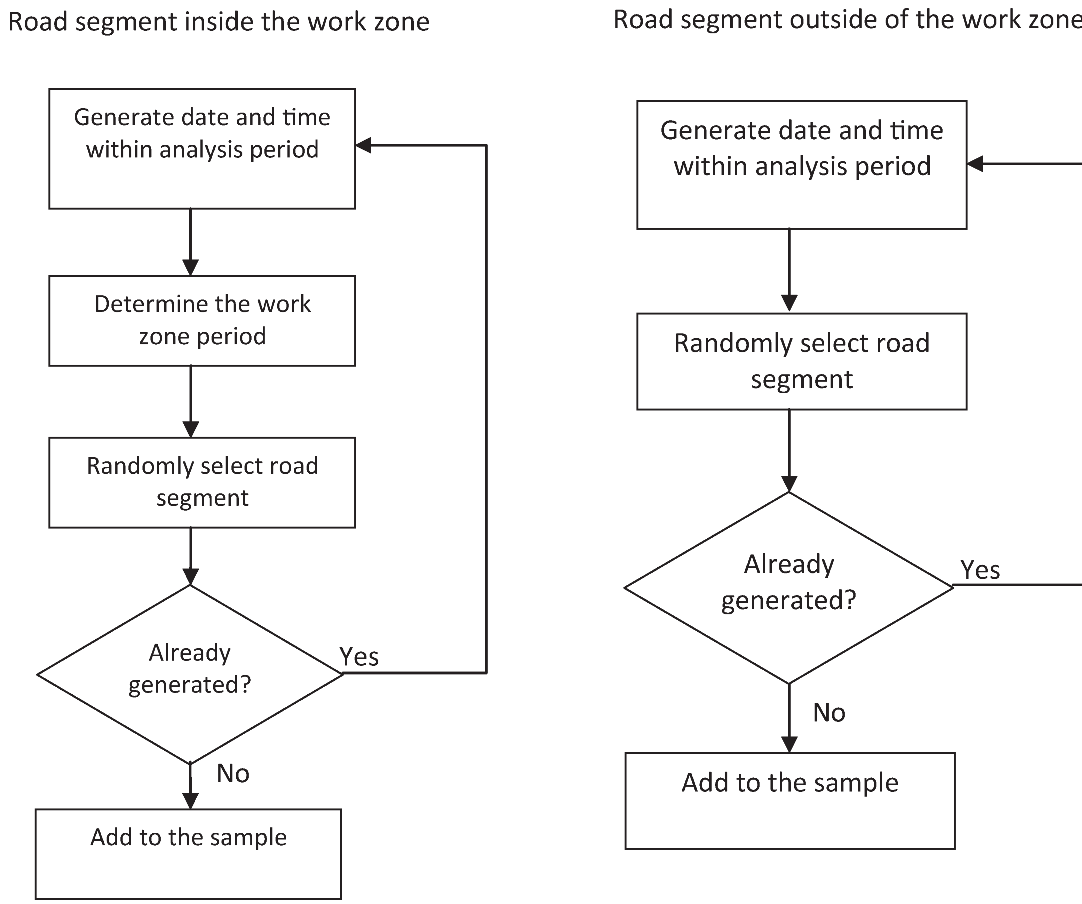


Figure 7.1 Sampling Process of Work Zone and Non-work-Zone Dataset

TABLE 7.2
Number of Observations Initially Generated for Modeling

Segments	Crash Observations	Non-crash Observations	Total
Work zone	167	195,695	195,862
Non-work zone	2,568	654,154	565,722
All segments	2,735	849,849	852,584

7.2 Model Estimation

Out of the initial 852,584 observations, 466,483 time intervals with 1,403 crashes were useable to estimate the model (Table 7.3). The omitted observations had missing values in some of the variables included in the final model. The numbers of crashes used for estimating the models presented in the following sections of this report are presented in Table 7.4. The model estimation was conducted with the SAS software package, ver. 9.2, developed by the SAS Institute Inc (SAS, 2007). The variables included in the model are significant at the 10% significance level. A complete set of the variables investigated in this study can be found in Appendix A, which also presents the basic statistics of these variables including the minimum, maximum, and average values, and the number of values 1 for the binary variables. These statistics are shown for the work zone and non-work zone observations separately.

The range of an important continuous variable – traffic volume - was divided into narrow intervals and each interval is represented by a separate binary variable. These binary variables were included in the initial models to check whether a single model coefficient is applicable to the entire range of traffic volume. The analysis of the obtained coefficients indicated that a single parameter was acceptable so the traffic volume was used in the model instead of the mentioned binary variables. The same test was applied to the volume of heavy vehicles (13+ tons) and the results again confirmed that a single coefficient was appropriate. Through a sequence of attempts, the most efficient combination of the traffic volume and other variables was obtained and was included in the final model (see Appendix H for the SAS report). The AIC criterion was used to compare the models among themselves to identify the most efficient one.

The obtained model is reported with statistical details in Appendix H and its variables with associated model parameters (betas) are shown below:

$$P(C) = \frac{\exp(\boldsymbol{\beta X})}{1 + \exp(\boldsymbol{\beta X})} \tag{9}$$

TABLE 7.3
Number of Observations for Modeling and in the Entire Sample

Segments	Sample Used for Modeling			Entire Sample Representing Population		
	Crash Observations	Non-crash Observations	Total	Crash Observations	Non-crash Observations	Total
Work zone	132	156,514	156,646	167	1,403,643	1,403,808
Non-work zone	1,271	308,566	309,837	2,568	10,164,500	10,167,072
All segments	1,403	465,080	466,483	2,735	11,568,143	11,570,880

TABLE 7.4
Number of Crashes Included in the Sample used for Modeling

Freeway Segments	Single-vehicle Crashes	Multiple-vehicle Crashes	All Crashes	Crashes with Injuries
Non-work-zone	353	918	1271	192
Work-zone	20	112	132	22
All	373	1030	1403	214

where:

$$
\begin{aligned}
\boldsymbol{\beta X} = {} & -8.0593^{\mathrm{a}} + 0.000239\ \text{VOLUME} - 0.00114\ \text{HVVOL} + \\
& 0.00533\ \text{HVVOL WORKZONE} + 3.2997\ \text{RSEGLEN} - \\
& 1.7903\ \text{WORKZONE RSEGLEN} + 0.0848\ \text{CURVSEGN} + 0.5458\ \text{L5} - \\
& 0.0300\ \text{INSHLDW} - 0.0534\ \text{OUTSHLDW} + 0.1326\ \text{DONRAMP} + \\
& 0.1503\ \text{DOFFRAMP} + 0.3003\ \text{BTEMP} - 0.4363\ \text{BTEMPMISS} - \\
& 0.0498\ \text{VISIBILITY} + 0.2766\ \text{BPRECIP} - 1.2529\ \text{BHEAVY} - 0.8039\ \text{T2} \\
& - 0.8394\ \text{T3} - 0.6299\ \text{T4} - 0.3756\ \text{T5} + 0.3442\ \text{T7} + 0.6271\ \text{T8} + \\
& 0.4592\ \text{T9} + 0.3778\ \text{T16} + 0.4701\ \text{T17} + 0.6656\ \text{T18} - 0.5086\ \text{DW1} - \\
& 0.3084\ \text{DW4} - 0.2111\ \text{DW5} - 0.5911\ \text{DW6} - 0.8197\ \text{DW7} + \\
& 0.5218\ \text{WORKZONE DW4} + 0.6222\ \text{WORKZONE DW5} - 0.4452\ \text{I65} - \\
& 0.9776\ \text{I70} - 0.27294^{\mathrm{b}}\ \text{WORKZONE}
\end{aligned}
\tag{10}
$$

where:
BHEAVY = 1 if an atmospheric event of a heavy intensity (rain, snow, storm, etc.) as reported in the weather database,
BPRECIP = 1 if rain or snow precipitation, = 0 otherwise;
BTEMP = 1 if temperature below the freezing point (32^{o}F), = 0 otherwise;
BTEMPMISS = 1 if temperature information is missing, = 0 otherwise;
CURVSEGN = number of horizontal curves on the segment;
DOFFRAMP = 1 if an off-ramp terminal is at the downstream end of the segment, = 0 otherwise;
DONRAMP = 1 if an on-ramp terminal is at the downstream end of the segment, = 0 otherwise;
DW1 = 1 if Monday, = 0 otherwise;
DW4 = 1 if Thursday, = 0 otherwise;
DW5 = 1 if Friday, = 0 otherwise;
DW6 = 1 if Saturday, = 0 otherwise;
DW7 = 1 if Sunday, = 0 otherwise;
HVVOL = volume of trucks 13T or heavier, veh/h;
I65 = 1 if I65 interstate, = 0 otherwise;
I70 = 1 if I70 interstate, = 0 otherwise;
INSHLDW = inside shoulder width, ft;
L5 = = 1 if segment has 5 traffic lanes (one direction), = 0 otherwise
OUTSHLDW = outside shoulder width, ft;
RSEGLEN = segment length - should be as close to 0.25 mile as possible, mi;
T2 = 1 if between 1 AM and 2 AM, = 0 otherwise;
T3 = 1 if between 2 AM and 3 AM, = 0 otherwise;
T4 = 1 if between 3 AM and 4 AM, = 0 otherwise;
T5 = 1 if between 4 AM and 5 AM, = 0 otherwise;
T7 = 1 if between 6 AM and 7 AM, = 0 otherwise;
T8 = 1 if between 7 AM and 8 AM, = 0 otherwise;
T9 = 1 if between 8 AM and 9 AM, = 0 otherwise;
T16 = 1 if between 3 PM and 4 PM, = 0 otherwise;
T17 = 1 if between 4 PM and 5 PM, = 0 otherwise;
T18 = 1 if between 5 PM and 6 PM, = 0 otherwise;
VOLUME = volume rate (veh/h);
VISIBILITY = visibility distance reported by a weather agency, mi;
WORKZONE = 1 if the segment is inside a work zone, = 0 otherwise.

The results are summarized in Table 7.6 and further discussed in the next section of this report. The original SAS results with additional statistical results are included in Appendix H. The intercept shown in Eq. 10 is an adjusted original value reported in the SAS results. This adjustment in accordance with Washington et al. (2003) was needed because of the

Notes:
[a]This intercept has been adjusted by -2.7898 (see Table 7.5) to incorporate the sampling scheme which reduces the number of intervals without crashes. This adjustment is explained in the following part of this section. Appendix H reports among other results also the original value of the intercept equal to -5.2695.
[b]This parameter has been adjusted for the difference between the adjustments of work-zone and non-work-zone predictions as presented in Table 7.5. The original value reported in Appendix H has been modified by adding (-1.9714 + 2.7898).

TABLE 7.5
Intercepts Adjustment for Distorted Proportion of Crash Observations in the Sample Used to Develop the P(C) Model

	Sample Used for Modeling		Entire Sample Representing Population		
Segments	Crash Observations	Total	Crash Observations	Total	Intercept Adjustment
Work zone	132	156,514	165	1,403,808	-1.9714
Non-work zone	1,271	309,837	2,572	10,167,072	-2.7898
All segments	1,403	466,351	2,737	11,570,880	-

reduction in the number of observations described in the previous section. To adjust the SAS-estimated intercept, a value calculated as shown below has to be added to the original intercept.

$$-LN\left(\frac{SR_i}{PR_i}\right)$$

where, SR_i = ratio of observations having outcome i to other observations in the sample

PR_i = ratio of observations having outcome i to other observations in the total population.

Following the assumption that the weather and traffic variables may not fully grasp the variability of safety in time, the impacts of the time of day, day of week, and month were analyzed representing these temporary effects. In fact, we found that early morning, rush periods, and week of day should be included in the models while other periods of the day and months do not have to be included.

7.3 Discussion of the Crash Likelihood Factors

Traffic factors, roadway (geometry) factors, weather-related factors, special driving circumstances in specific roadway sections, and time of the day were identified as significant in the crash likelihood model. Some of them are discussed below in more detail.

Traffic Volume (VOLUME)

Traffic volume is considered to be one of the most prominent factors of safety. After all, no crashes can happen without vehicles on the road. The effect of traffic volume, however, is more complex than doubling the number of crashes when the traffic volume doubles. The interaction between vehicles increases with traffic volume but also does the drivers' attention. The results in Table 7.6 confirm the results obtained in most previous studies, namely, that that the likelihood of crash grows with an increase in traffic volume. This effect has been estimated as an additional 0.35 crash/year/0.24 mi in one direction per 1,000 vehicles added in this direction (Table 7.7). This is approximately equivalent to an additional 14 crashes annually along a 10-mile long work zone of characteristics similar to the Super 70. This number doubles if the volume increase applies to both directions of traffic. We could not confirm that the traffic volume impact is any different inside and outside the work zone as long as the additional volume does not include heavy vehicles.

Heavy Vehicles (13+ ton, HVVOL)

The presence of heavy vehicles (13+ tons) on roads outside of the Super 70 project was associated with the reduced risk of crash (Table 7.6). The positive impact of one added vehicle is stronger than the negative effect of adding one non-heavy vehicle. There are two possible explanations of this effect: (1) truck drivers are trained professionals who expose others and themselves to a smaller risk; and (2) the presence of heavy trucks increases the alertness and vigilance of other drivers as their perception of risk increases and it promotes safety very effectively.

At the same time, the safety impact of heavy vehicles on safety inside the Super 70 project was found to be negative. It seems that much more challenging conditions inside the construction zone, particularly for large and heavy trucks, could have a strong reverse impact that exceeded the added caution of other drivers. Another option was that other drivers, themselves more challenged with the altered geometry and controlled conditions, did not pay as much attention as on regular road segments and did not take sufficient precautions around heavy vehicles. Finally, large trucks obstruct visibility of other vehicles adding more difficulty to the already challenging conditions. An additional 10 heavy vehicles per hour in one direction increased the number of annual crashes by 0.06 per a quarter-mile one-way segment (Table 7.7). This number amounts to almost five crashes a year in a ten-mile work zone.

This result indicates that rerouting heavy vehicles on alternative roads might have a positive double impact of removing dangerous vehicles from the work zone and "calming" other drivers on the alternative routes. This effect is further analyzed in the next section.

Moveable Barrier

A moveable barrier was applied inside the Super 70 work zone to add a lane to the traffic direction during the time of day when high traffic volumes were experienced. This lane adjustment strategy was expected to improve traffic safety by dispersing traffic across more traffic lanes and reducing the frequency and intensity of risky interactions between vehicles. To

TABLE 7.6
Crash Likelihood Model P(C) Results

Variable	Parameter	Standard Error	P-Value	Interpretation
Intercept	-8.0593	0.2778	<.0001	Intercept adjusted for the distorted proportion of 1s in the non-work-zone sample.
VOLUME	0.000239	0.000027	<.0001	Higher traffic volume per lane is associated with more crashes.
HVVOL	-0.00114	0.000281	<.0001	Presence of heavy vehicles (+13 ton) is associated with lower crash frequency on segments with regular traffic conditions (no work zone). This interesting result is further discussed in the next chapter. Inside the work zone, though, the effect is reversed.
HVVOL*WORKZONE	0.00533	0.00227	0.0189	
RSEGLEN	3.2997	0.3416	<.0001	The risk of crash is higher at longer road segments. This effect is weaker inside work zones.
WORKZONE*RSEGLEN	-1.7903	0.9312	0.0545	
CURVSEGN	0.0848	0.0463	0.0671	Horizontal curves increase the risk of crash.
L5	0.5458	0.1000	<.0001	Segments with five lanes in one direction are more dangerous than other segments.
INSHLDW	-0.0300	0.0102	0.0035	The risk of crash is lower on segments with wider inside (median) shoulders
OUTSHLDW	-0.0534	0.0108	<.0001	The risk of crash is lower on segments with wider outside shoulders.
DONRAMP	0.1326	0.0736	0.0716	Downstream ramps, on and off, increase the risk of crash.
DOFFRAMP	0.1503	0.0736	0.0411	
BTEMP	0.3003	0.0773	0.0001	Temperature below the freezing point increases the risk of crash. The intercept adjustment factor BTEMPMISS should be used if the temperature information is not known.
BTEMPMISS	-0.4363	0.1402	0.0019	
VISIBILITY	-0.0498	0.0138	0.0003	
BPRECIP	0.2766	0.0970	0.0044	Rain or snow precipitation increases the risk of crash. On the other hand, if this precipitation or another atmospheric event such as storm or fog are of heavy intensity, then the risk is even lower than without the precipitation.
BHEAVY	-1.2529	0.7137	0.0792	
T2	-0.8039	0.2571	0.0018	Very early morning hours exhibit lower risk of crash.
T3	-0.8394	0.2652	0.0016	
T4	-0.6299	0.2375	0.008	
T5	-0.3756	0.2081	0.0711	
T7	0.3442	0.1208	0.0044	The effect of traffic volume represented by the coefficient associated with the VOLUME variable does not fully reflect the rush hour safety impact. These coefficients indicate that during these hours the risk is even higher and needs to be adjusted up.
T8	0.6271	0.1089	<.0001	
T9	0.4592	0.1137	<.0001	
T16	0.3778	0.1161	0.0011	
T17	0.4701	0.1137	<.0001	
T18	0.6656	0.1068	<.0001	
DW1	-0.5086	0.0926	<.0001	It seems that there is an additional risk associated with Tuesdays and Wednesdays. Sunday needs the strongest adjustment down.
DW4	-0.3084	0.0855	0.0003	
DW5	-0.2111	0.0815	0.0096	
DW6	-0.5911	0.1019	<.0001	
DW7	-0.8197	0.1156	<.0001	
WORKZONE*DW4	0.5218	0.2472	0.0348	Thursday and Friday inside the work zone have additional risk associated with Thursdays and Fridays.
WORKZONE*DW5	0.6222	0.2250	0.0057	
I65	-0.4452	0.0825	<.0001	Both, I-65 and I-70 segments have slightly reduced risk of crash than I-465.
I70	-0.9776	0.1460	<.0001	
WORKZONE	-0.27294	0.3192	0.0006	This coefficient indicates additional unexplained by the model factors that make the Super 70 work zone safer. The presented value has been adjusted to account for the distorted proportions of 1s in the work-zone and non-work-zone samples.

confirm this assumption, the impact of the traffic volume per lane was checked. In spite of numerous attempts, we could not confirm any effect of the traffic volume per lane in addition to what was already captured by the total traffic volume. Also, the explicit effect of the number of lanes, except five lanes, could not be confirmed. The last attempt involved modeling the effect of an additional lane in certain time intervals inside the Super 70 section. Also this attempt failed providing plausible indication of any positive effect.

The only potentially positive safety effect of varying the number of lanes was indirect and was produced by

TABLE 7.7
Marginal Effects of Variables on Safety Inside the Super 70 Work Zone

Variable	Parameter	Average	Initial value	Change	Unit	Effect (crash/int)	Effect[1] (crash/year)
VOLUME	0.000239	1153	1150	1000	veh/h	1.98E-05	0.346
HVVOL	0.00419	18.7	18.7	10	veh/h	3.46E-06	0.061
RSEGLEN	1.5094	0.24	0.24	1	mi	1.25E-04	2.187
CURVSEGN	0.0848	0.55	0	1		6.69E-06	0.117
L5	0.5458	0.003	0	1		4.51E-05	0.789
INSHLDW	-0.03	1.97	2.0	1	ft	-2.48E-06	-0.043
OUTSHLDW	-0.0534	4.09	4.1	1	ft	-4.42E-06	-0.077
DONRAMP	0.1326	0.059	0	1		1.09E-05	0.191
DOFFRAMP	0.1503	0.142	0	1		1.22E-05	0.213
BTEMP	0.3003	0.024	0	1		2.47E-05	0.432
VISIBILITY	-0.0498	6.079	6.1	1	mi	-4.12E-06	-0.072
BPRECIP	0.2766	0.044	0	1		2.26E-05	0.396
BHEAVY	-1.2529	0.002	0	1		-1.04E-04	-1.820
T2	-0.8039	0.041	0	1		-6.87E-05	-1.204
T3	-0.8394	0.04	0	1		-7.18E-05	-1.258
T4	-0.6299	0.041	0	1		-5.34E-05	-0.936
T5	-0.3756	0.042	0	1		-3.16E-05	-0.553
T7	0.3442	0.042	0	1		2.81E-05	0.491
T8	0.6271	0.042	0	1		5.05E-05	0.885
T9	0.4592	0.042	0	1		3.72E-05	0.653
T16	0.3778	0.042	0	1		3.07E-05	0.539
T17	0.4701	0.042	0	1		3.81E-05	0.668
T18	0.6656	0.041	0	1		5.36E-05	0.938
DW1	-0.5086	0.149	0	1		-4.54E-05	-0.795
DW4	0.2134	0.137	0	1		1.36E-05	0.238
DW5	0.4111	0.152	0	1		2.88E-05	0.504
DW6	-0.5911	0.149	0	1		-5.34E-05	-0.935
DW7	-0.8197	0.143	0	1		-7.62E-05	-1.335

[1](Crash/year) = (Crash/int) • 365 days • 48 intervals

increasing the throughput of the work zone and reducing the number of non-heavy vehicles that had to divert to alternative longer routes with potentially higher exposure to crash.

Shoulder Width (INSHLDW, OUTSHLDW)

Segments with wider inside and outside shoulders tend to be safer than other segments (Table 7.6). Wide shoulders provide a last-second chance to regain control of a vehicle and return to traffic. Wide shoulders also offer an opportunity to swerve towards the shoulder to avoid collision with another vehicle in the traffic lane. We could not confirm that a shoulder width had different impacts inside and outside of the Super 70 work zone. The marginal effect analysis presented in Table 7.7 indicates that increasing the inside and outside shoulder widths by one foot may reduce the annual number of crashes by 0.043 and 0.077, respectively. The corresponding reduction for a 10-mile work zone is 3.4 and 6.2.

Segment Length (RSEGLEN)

As expected, longer segments tend to have more crashes than shorter segments due to the higher exposure to risk. This effect does not have to be proportional (twice longer segments having twice as many crashes) because the segment length may also represent other factor such as the frequency of ramps if the gore point of the ramp is selected as the segment end as in our project.

Presence of Horizontal Curves (CURVSEGN)

Segments with curves experience more crashes than straight segments (Table 7.6). It is quite plausible that a change in the road direction creates an “opportunity” for human error. Not paying attention to the road may surprise a driver who may leave the lane and collide with a vehicle in the adjacent lane or initiate a chain of events leading to a collision between other vehicles. One curve adds 0.1 crashes annually under average Super 70 conditions (Table 7.7).

Presence of Ramp Intersections (DONRAMP, DOFFRAMP)

The presence of a downstream ramp intersection increases the risk of crash (Table 7.6). Traffic merging or leaving the freeway involves lane-change activities that add more risk. Drivers wanting to leave the freeway aggressively change lanes to position themselves in the rightmost lane. Drivers who enter the freeway try to leave the rightmost lane and take typically faster left lanes. The Highway Capacity Manual (HCM) assumes that considerable perturbation caused by the ramps is present along a distance of 1,500 ft. upstream of off-ramps and downstream of on-ramps. Indeed, the results have confirmed that along such segments approximately 0.25 long, the risk of crash is higher than somewhere else. The marginal effect analysis presented in Table 7.7 indicates that one downstream ramp causes 0.2 crashes a year under the average Super 70 conditions.

Interestingly, the weak deterioration of safety was also detected upstream of on-ramps and downstream of off-ramps. Freeway drivers approaching an on-ramp anticipate the perturbation past the ramp gore point and try to move to the left in advance to avoid this. The off-ramp reduces traffic in the rightmost lane and some drivers take advantage of this by moving to the right, including those who change lanes upstream of the ramp and then return to this lane after passing the off-ramp. The safety effect of these maneuvers is much weaker than in the previous two cases and could not be confirmed statistically as significant.

Adverse Weather (BTEMP, BPRECIP, BHEAVY)

Adverse winter conditions are represented by the temperature below the freezing point (BTEMP), the rain/snow precipitation (BPRECIP), and the intensity of atmospheric events (BHEAVY). All three variables are binary (Table 7.6). As expected, the freezing temperature and precipitation were found to significantly increase the risk of crash. The freezing temperature creates conditions for slippery roads while rain and snow precipitation reduces the pavement-tire friction coefficient and worsens the visibility through the windshield with poorly performing wipers. The impacts of the two conditions are similar to each other and twice as strong as the impact of a downstream ramp. Each of the two conditions generates approximately 0.4 crash per year (conditions assumed continuous) per direction and along a quarter-mile segment.

Any atmospheric conditions characterized by the weather service agencies as heavy (snow, rain, fog, storm, wind, etc.) reduce the likelihood of crash quite considerably. This result might be the effect of drivers' overcompensation of the risk by driving slower and sometimes even abandoning the trip and staying off the road. The above conditions exclude unusually severe events, such as hurricanes and flood, that cause significant loss and cannot be controlled by usual compensation behavior.

Air Opacity (VISIBILITY)

The air opacity measured by the visibility distance affects the risk of crash (Table 7.6). This effect needs additional explanations. The air opacity is caused by small particles suspended in the air. These particles can be of the industrial or natural origin. They are considered pollutants and most likely affect human performance. In the considered case, the air pollution may reduce safety through affecting drivers' psychological performance. It is highly unlikely that the visibility distance directly affects the risk of collision as it is typically much longer than the sight distance required for safe driving. This effect magnitude associated with the increase in the visibility by one mile is comparable to increasing the outside shoulder width everywhere by one foot (Table 7.7).

Time of Day (T2-T5, T7-T9, T16-T18)

Traffic and weather are those safety factors that change with time and their effects have been captured through several variables. Since none of these temporal factors can be fully included in any model, the effect of the time of day, day of the week, month, and longer periods have been investigated. The time of day was represented through 24 binary variables representing hourly intervals. The risk of crash during three periods: early morning (1–5 AM), morning rush period (6–9 AM), and afternoon rush period (3–6 PM) needed to be adjusted even after accounting for the traffic volume. The early morning adjustment reduced the risk of crash while the two rush periods needed to be adjusted up. These adjustments might indicate that these periods contribute to safety with unknown factors, but it is also possible that the logistic curve does not fit well the traffic impact relationship and it needs to be adjusted for low and high traffic volumes.

Day of Week (DW1, DW4-DW7)

The results indicate that weekends are safer than weekdays even after adjusting for lower traffic. Drivers are more relaxed and rested, and the value of time is lower than on weekdays when business-related travel prevails. This difference may lead to lower risk-taking behavior, less aggression, and fewer crashes.

The Super 70 segments tended to experience higher crash risk on Thursdays and Fridays than other segments. More challenging geometry conditions combined with the buildup of time pressure and fatigue towards the end of week might contribute to this result.

Neither seasonal effect nor longer-term safety fluctuations were detected.

Other Work Zone Factors (WORKZONE)

The effect of the work zone on safety is complex. Altered geometry, visual distractions, reduced number of lanes causing dense traffic, reduced speed, closed ramps attributing to confusion, police enforcement, and special traffic management are among the work zone factors. The factors explicitly investigated include altered traffic volumes, reduced presence of heavy vehicles, reduced lateral clearance, and presence of horizontal curves and ramps.

The Super 70 work zone was unique in respect to police enforcement. Intensive enforcement included almost continuous presence of patrol cars, with constant monitoring of drivers of heavy vehicles who violated the access restriction to the work zone and those who violated the 45 mi/h speed limit. Although the enforcement level varied to some extend during the nine months of the work zone, this variation was too weak to detect is impact on safety. The factors that could not be separated from the police enforcement were the reduced speed, the traffic of construction trucks, and the impact of detailed traffic management solutions (drums, crossovers, warning signs, etc). The joint positive effect of all these variables is represented by the work zone variable in Table 7.6. This effect is also simulated in the next section of this report.

8 SINGLE-VEHICLE CRASH LIKELIHOOD MODEL $P(C1|C)$

8.1 Model Estimation

A model for single-vehicle crashes estimates the probability of an involvement of only a single vehicle given that a crash happened. Several traffic, roadway, and weather factors were found to be significant in the Super 70 work zone. The model is presented in Eq. 8 and Eq. 9 and the original SAS report in Appendix I. Table 16 lists the variables included in the model and it explains their meaning. These variables are further discussed in the next section of the report.

$$P(C1) = \frac{\exp(\boldsymbol{\beta X})}{1 + \exp(\boldsymbol{\beta X})} \tag{8}$$

with

$$\begin{aligned} &-0.3434 - 0.00050\ \text{VOLUME} - 0.0562\ \text{WORKZONE HVVOL} + \\ &0.1769\ \text{CURVSEGN} + 0.0424 - \text{OUTSHLDW} + 0.4623\ \text{BPRECIP} + \\ &0.3734\ \text{BTEMP} + 1.197\ \text{I70} - 0.3294\ \text{PERIOD1} - 0.997\ \text{WORKZONE}, \end{aligned} \tag{9}$$

where:

VOLUME = hourly volume rate representing the studied half-an-hour interval, veh/h;
WORKZONE = 1 if segments is in the active work zone section, =0 otherwise;
HVVOL = volume of trucks 13T or heavier, veh/h;
CURVSEGN = number of horizontal curves on the segment;
OUTSHLDW = outside shoulder width, ft;
BPRECIP = 1 if rain or snow precipitation, = 0 otherwise;
BTEMP = 1 if the temperature is below the freezing point, =0 otherwise;
I70 = 1 if the segment is on I-70, =0 otherwise;
PERIOD1 = 1 if period before the Super 70 project, = 0 otherwise.

8.2 Discussion of the Single-vehicle Crash Factors

Modeling single-vehicle involvement in a crash is justified with possibly different factors of crash probability and severity effective in a single-vehicle crash than in a multiple-vehicle crash. With some simplification, one may say that single vehicle crashes are caused by a driver's incorrect judgment about the road and speed or by the lack of attention and when the interaction between vehicles is weak or non-existent. Nighttime, with more challenging conditions and low traffic volumes, is expected to have more single-vehicle crashes than daytime conditions. There are also commonalities between the two types of crashes. Some crashes initiated by an interaction between two vehicles end up being a single vehicle crash when one of the drivers avoids collision with the other vehicle, leaves the road, and rolls over or hits an obstruction. It is also possible that a driver's error triggering a collision is not affected by other vehicles. The driver loses control over the vehicle and he or she collides with another vehicle before leaving the road.

Traffic Volume (VOLUME)

As expected, sizeable traffic volumes reduce the proportion of crashes that end up as a single vehicle crash. Less space between vehicles increases the risk of faulty interactions between vehicles and also the risk of collision with another vehicle even when the triggering event was not caused by a vehicle interaction. This effect of volume is twice as strong in the work zone due to the higher concentration of traffic on a segment with a reduced number of lanes. The limited lateral clearance also contributes to this effect because a vehicle colliding with a barrier is not fully off the road and can be hit by another vehicle from behind.

Volume of Heavy Vehicles Inside the Super 70 Work Zone (HVVOL*WORKZONE)

The crash likelihood model indicates that the frequency of crashes inside the Super 70 work zone increased with the increased presence of heavy vehicles (13+ tons). The single-vehicle crash model points out those heavy vehicles increased the risk of multivehicle crashes, which is manifested through the reduced likelihood of the single-vehicle crash inside the work zone. This effect was not detected outside of the Super 70 work zone.

Shoulder Width (OUTSHLDW)

The above comment about lateral clearance is reinforced by this effect. Segments with wider shoulders tend to have a higher percentage of single-vehicle crashes than other segments. Wide shoulders not only contain the crashed vehicles, but also reduce the risk of returning these vehicles back to the traffic stream and colliding with other vehicles. In addition to this physical

TABLE 8.1
Single-vehicle Crash Likelihood Model P(C1|C) Results

Variable	Parameter	St. Dev.	P-Value	Interpretation
Intercept	-0.3434	0.2887	0.2343	-
VOLUME	-0.0005	0.000052	<.0001	Volume increases reduces the proportion of single-vehicle crashes
WORKZONE*HVVOL	-0.0562	0.0265	0.0341	Presence of heavy vehicles inside the work zone increases the proportion of multiple-vehicle crashes
CURVSEGN	0.1769	0.1049	0.0918	Horizontal curves increases the proportion of single-vehicle crashes
OUTSHLDW	0.0424	0.0216	0.0497	Wide outside shoulders reduce the proportion of multiple-vehicle crashes
BPRECIP	0.4623	0.1753	0.0084	Adverse weather conditions increase the proportion of single-vehicle crashes
BTEMP	0.3734	0.1702	0.0282	
I70	1.197	0.2922	<.0001	I70 segments experience higher proportion of single-vehicle crashes than other segments. This overrepresentation is weaker inside the I70 work zone.
WORKZONE	-0.997	0.5964	0.0946	
PERIOD1	-0.3294	0.1433	0.0215	Lower percent of single-vehicle crashes in the period before the Super 70 project

interpretation, there is also a possible psychological effect on drivers who, in the case of facing the danger of hitting another vehicle, may more frequently decide to use a shoulder to avoid the collision if the shoulder is wide.

Horizontal Curve (CURVSEGN)

The presence of a horizontal curve adds difficulty in remaining in the travelled way and creates a risk of leaving the roadway and hitting a barrier. This added risk is captured by the higher proportion of single vehicles on road segments with a horizontal curve.

Snow, Rain, and Freezing Temperature

Adverse weather conditions and slippery roads increase the driving difficulty, which increases the number of road departures with a single-vehicle outcome.

Work Zone

The work zone experiences a lower proportion of single-vehicle crashes due to the denser traffic in a reduced number of lanes and limited lateral clearance due to barriers inside the work zone. The work zone does not provide sufficient protection from other moving vehicles when a vehicle hits the barrier.

Other Factors

Other significant variables whose interpretation is not obvious include the increased proportion of single-vehicle crashes on I-70 outside of the work zone and a reduced proportion of single vehicle crashes during the period before the Super 70 project began.

9 JOINT CRASH SEVERITY MODEL FOR *P(SC1|C1) AND P(SC2|C2)*

Crash severity models estimate the probability of a severe outcome (fatality or injury) given that the crash has happened. A single joint model was developed with several interacting variables whose parameters depend on whether the crash is single-vehicle or multiple-vehicle. Developing a single model increases the model's performance by utilizing a larger sample to estimate a single model rather than two separate models. Adding interactions with the type of crash is justified by the fact that the two types of crashes may be associated with different outcomes. In a single vehicle collision, a vehicle rolls over or hits a fixed obstruction. Such crashes tend to generate more severe outcome than multiple vehicle crashes. It is also possible that most single-vehicle crashes are not reported if the outcome is not severe. On the other hand, even fender benders in multiple-vehicle crashes are most likely reported. Drivers who did not cause the crash are willing to report the crash regardless of its severity in order to be covered by their insurance. The developed models can be used in conjunction with the other models to calculate the number of severe and non-severe crashes of single vehicle and multiple-vehicle crashes.

In this study, a binary logit model is used for crash severity. In this model, crashes with fatal, incapacitating, and non-incapacitating injuries were considered as severe crashes while other crashes where only possible injury and property-damage-only were recorded, were considered as non-severe crashes. The variable INJURY represents the severity outcome. INJURY=1 indicates a severe crash while INJURY=0 indicates a non-severe crash. The initial 2,703 crashes in the original sample were reduced to 1,403 crashes due to missing data, mainly traffic volumes. Out of these crashes, 214 were crashes with injuries, 373 were single-vehicle crashes, and 1,030 were multiple-vehicle crashes. The modeling was facilitated with SAS. The form of severity model is shown in Eq. 11. The output from SAS is attached in Appendix J.

$$P(SC_n|C_n) = \frac{\exp(\boldsymbol{\beta X})}{1 + \exp(\boldsymbol{\beta X})} \tag{11}$$

where:

$$\begin{aligned} &-1.3431 - 0.395 \text{ SINGVEH} - 0.00023 \text{ MULTVEH VOLUME} + \\ &0.0323 \text{ INSHLDW} - 0.0419 \text{ OUTSHLDW } + \\ &0.5923 \text{ SINGVEH PERIOD2} + 1.1209 \text{ SINGVEH DW1} \\ &+ 0.5088 \text{ DW5} + 0.9667 \text{ SINGVEH DW6}, \end{aligned} \tag{12}$$

and:
SINGVEH = 1 if single-vehicle crash, = 0 otherwise;
MULTVEH = 1 if multivehicle crash, = 0 otherwise;
INSHLDW = inside shoulder width, ft;
OUTSHLDW = outside shoulder width, ft;
PERIOD2 = 1 if the first phase of the Super 70 project, = 0 otherwise;
DW1 = 1 if Monday, = 0 otherwise;
DW5 = 1 if Friday, = 0 otherwise;
DW6 = 1 if Saturday, = 0 otherwise;
VOLUME = hourly volume rate per lane, veh/h/lane;

The main difference between modeling crash events and severity outcome is that the latter is also affected by factors effective during the crash and not known beforehand. The factors of crash severity include the characteristics of the impact the characteristics of the vehicles, and even the positions of individuals inside the involved vehicles. Some of these data were not available and, if they are included in crash reports, they cannot be used to model the frequency of severe and non-severe crashes. That is why the variables included in the model represent not only themselves but also other variables omitted in the model but correlated with the ones in the model. The limited number of observations was another modeling problem. Several hundred crashes with around 100 of them severe carry limited information and many variables could not be added to the model because they were insignificant. Their statistical insignificance does not mean that they do not affect severity, rather the sample was too small to prove their importance.

That is why the obtained models were viewed not so much a means of investigating the impact of the road and the weather, but rather tools to predict the proportion of severe crashes once the number of crashes were known from the other models. Such an approach is much better than relying only on the average proportion of severe crashes obtained from the crash statistics. We will not attempt to interpret all the results of the above models.

Only two effects seem to be easy to explain: the number of vehicles involved and the volume. Single-vehicle crashes tend to be less severe than crashes that involve more vehicles. There is a quite simple statistical explanation of this result. The crash severity is defined by the most severe injury among people involved in the crash. It is less likely to have someone injured in a crash with fewer people involved – an obvious case in a single-vehicle crash.

Growing traffic volume reduces the severity of multivehicle crashes due to a reduced speed during busy periods. It is also possible that vehicles during periods with lower traffic are occupied by more people and the mechanism explained for single vehicles may also be a plausible explanation of this result.

10 SUPER 70 SAFETY SIMULATION

The set of equations explained in Chapter 3, Eq. 1-Eq. 3 and the sequence of calculations shown in Figure 3.1 were used to predict the number of crashes expected in prolonged periods and under certain traffic, weather, and geometry conditions. It should be noted that all the factors that affect the crash frequency also affect the number of severe crashes. Thus, any analysis of severity factors should be based on the sequential prediction of various types of crashes, their splitting into severe and not severe, and on aggregation of the

TABLE 9.1
Severe Crash Likelihood Model P(SCn|Cn) Results

Variable	Parameter	St. Dev.	P-Value	Interpretation
Intercept	-1.3431	0.2915	<.0001	-
SINGVEH	-0.395	0.2651	0.1362	Single-vehicle crashes tend to have lower crash severity than multiple-vehicle crashes.
MULTVEH* VOLUME	-0.00023	0.000067	0.0007	Severity of multiple-vehicle crashes tend to be less severe under heavy traffic.
INSHLDW	0.0323	0.0172	0.0609	Severity of crashes on segments with wide inside
OUTSHLDW	-0.0419	0.0224	0.0618	shoulders tend to be more severe while on segments with wide outside shoulders tend to be less severe.
SINGVEH* PERIOD2	0.5923	0.3203	0.0644	Single-vehicle crashes tend to be more severe during the first phase if the Super 70 project.
SINGVEH*DW1	1.1209	0.3765	0.0029	Single-vehicle crashes tend to be more severe on
SINGVEH*DW6	0.9667	0.3425	0.0048	Mondays and Fridays than on other days of week.
DW5	0.5088	0.1862	0.0063	All crashes tend to be more severe on Fridays.

TABLE 10.1
Safety in the Super 70 Period under Various Traffic Diversion Scenarios

Scenario	Super 70 Work Zone Traffic	All Crashes	Injury Crashes
1	Traffic level as before the Super 70 project	235.8	47.5
2	Traffic expected if rerouting of heavy vehicles not enforced	190.5	41.9
3	Actual traffic experienced during the Super 70 project	**161**	**36**
4	Traffic of non-heavy vehicles only during the Super 70 project (all heavy vehicles diverted)	142.0	33.7

obtained predictions by severity or by other criterion suitable for the analysis (Figure 3.1).

The sample of 156,646 30-minute intervals with 132 crashes was selected from the total 1,403,808 intervals with 161 crashes (official INDOT data) during the Super 70 period. This sample reflects the historical geometric, traffic, and weather conditions during the Super 70 period, and it was used to simulate selected safety effects. Safety under altered conditions was simulated by calculating the crash likelihood values for each 30-minute interval and aggregating them according to Figure 3.1 and Eq. 13. The resulting number of crashes was then converted from the sample intervals to the entire Super 70 period by proper scaling. The results obtained for various scenarios are discussed below.

Traffic Volume Effect

Figure 10.1 presents the simulated safety impact of changes in traffic volume. A twice larger traffic volume is associated with approximately twice as many crashes. This relationship is not linear, however, and further increases in traffic volume may have a disproportionally stronger effect on safety. The results indicate that heavy vehicles contribute more to the increase in crash frequency than other vehicles. The increase in the number of injury crashes shown in Figure 10.2 is weaker than for the total number of crashes, which also indicates that with the increase in the traffic volume, the proportion of injury crashes may decrease. In this case, heavy vehicles have a weaker impact on crash severity than other vehicles.

Given the estimated relationships between traffic volume and the safety of the Super 70 work zone, it was interesting to estimate safety in the work zone under various traffic diversion scenarios. Table 10.1 presents the simulated results. Under Scenario 1 where the work zone traffic remains at the level before the construction period, the number of expected (simulated) crashes was 236, including 48 injury crashes. This level of safety is close to the number of crashes officially reported by INDOT for the corresponding nine-month period in 2006: 233 and 45, respectively. This result itself indicates the success of the Super 70 traffic management strategies that would allow keeping the safety level almost unchanged under the challenging traffic and geometry conditions of the work zone even if the traffic level had not been reduced. This result will be further discussed in the remaining part of this chapter.

Scenario 2 in Table 10.1 assumes the volume of non-heavy vehicles as observed during the construction period, which was lower than during the corresponding before period in the previous year. This reduction was not enforced and was rather the results of drivers' personal decisions to avoid the work zone. In Scenario 2, this "natural" diversion rate was applied to heavy vehicles to represent the case where police enforcement is not present. In Scenario 2, 191 crashes with 42 injury crashes are simulated. This result indicates that enforcing heavy vehicles on alternative routes saved 30 crashes including six injury crashes inside the work zone. It should be noted that the positive effect of the presence of heavy vehicles on the alternative routes brought additional positive safety effects of rerouting these vehicles.

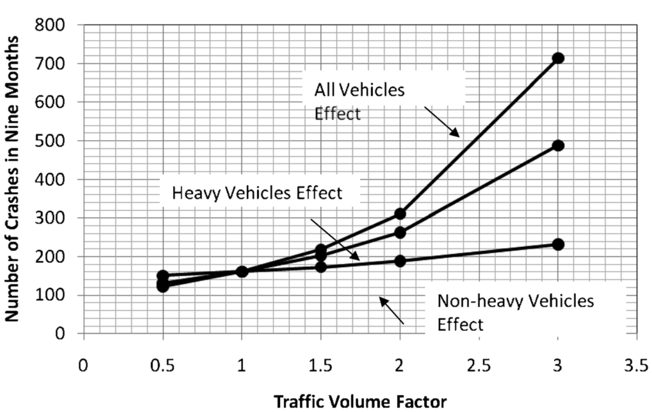


Figure 10.1 Simulated Impact of Traffic Volume on the Number of Crashes in the Super 70 Period

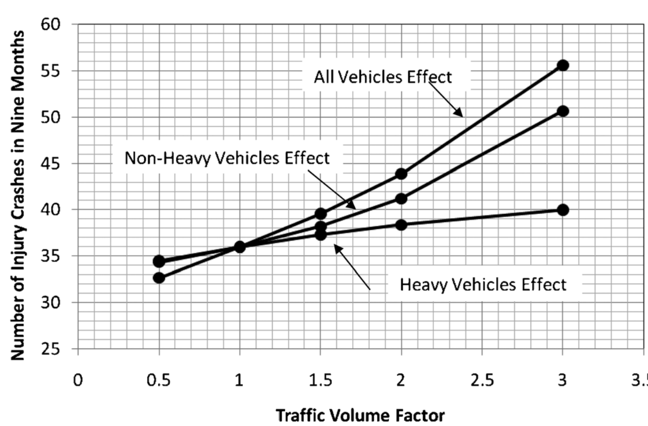


Figure 10.2 Simulated Impact of Traffic Volume on the Number of Injury Crashes in the Super 70 Period

Scenario 3 represents the actual case of the Super 70 work zone with the officially reported 161 crashes, including 36 injury crashes. These crash numbers could be further reduced inside the work zone if all the heavy vehicles were rerouted. The reductions were 19 and two, respectively.

Shoulder Width

Among several geometry variables, shoulder width is one that can to some extent be controlled by engineers who design work zones. We could not detect any significant difference of this effect between inside and outside the work zone. Figure 10.3 and Figure 10.4 present the effect of widening the inside and outside shoulders by three and six feet and under various traffic volumes. Widening the inside and outside shoulders by six feet along the entire work zone and in both directions might save up to 60 crashes during the nine-month construction period and also reduce the injury crashes by 16 crashes. These reductions, particularly the injury crashes, are considerably higher than the effect of traffic volumes. The mechanism of reducing the injury crashes is three-fold: (1) Wider shoulders save some crashes because of the added opportunity of avoiding a crash by using the shoulder for evasive maneuvers. (2) Widening shoulders "converts" multiple-vehicle crashes into single-vehicle crashes by reducing the risk of a secondary event of hitting an already crashed single vehicle if the vehicle is retained by the roadside. Single-vehicle crashes tend to be less severe. (3) There is also a direct and positive, although weak, effect of widening both the inside and outside shoulders on crash severity.

In the case of adding an additional 12 ft. for widening both the shoulders, a careful consideration should be given to an alternative solution of adding an additional traffic lane and keeping the shoulders narrow. This decision depends on the capacity of the work zone, the traffic demand, and the safety effect of rerouting vehicles to alternative routes when the work zone capacity is insufficient.

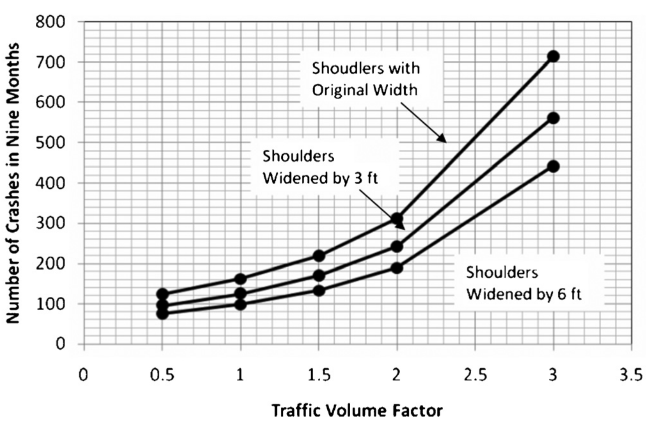


Figure 10.3 Simulated Impact of Shoulder Width on the Number of Crashes in the Super 70 Period

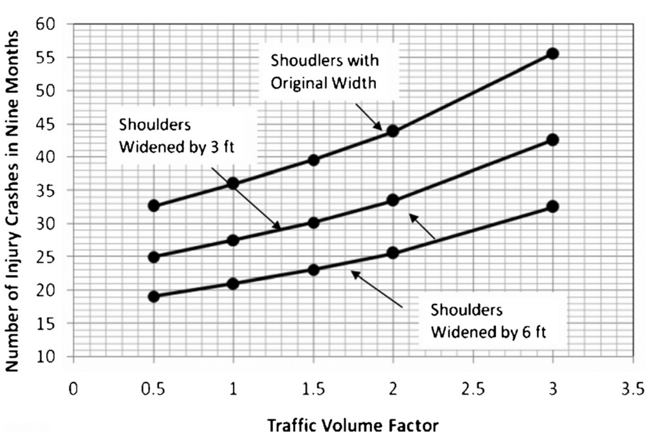


Figure 10.4 Simulated Impact of Shoulder Width on the Number of Crashes in the Super 70 Period

Interchange Ramps

Table 10.2 presents the effect of closing all ramps during the Super 70 construction period. The effect on the ramps on work zone safety was found to be rather limited. The scenario of closing the ramps is rather theoretical because ramps are typically closed out of necessity when construction of the ramp is needed. Also, ramps are typically kept open as long as possible to reduce the disturbance to the origin-destination traffic that may access the freeway and the abutting areas through these ramps. Otherwise, the rerouting would be longer and transferred to the surface streets. The effect of the diverted traffic on the safety of local roads is presented in the next chapter.

Other Factors

A rather low number of crashes inside the work zone and the simultaneous presence of various additional factors made estimation of the individual effects of police enforcement intensity, speed reduction, and other traffic management techniques difficult. Instead, a single estimation of the joint effect of these factors was performed. Other factors, such as individual drivers' adjustments to new conditions, possibly contributed to the results as well. A simulation scenario was run with the assumption that this joint effect was not present. Table 10.3 shows that this single assumption increased the number of crashes by 50 including 11 injury crashes. These 50 crashes can be attributed to all

TABLE 10.2
Ramps Effect

	Ramps Opened		All Ramps Closed	
Traffic Volume Factor	All Crashes	Injury Crashes	All Crashes	Injury Crashes
1	**161.0**	**36.0**	156.3	35.0
2	310.8	43.9	301.6	42.6
3	713.9	55.6	693.5	54.0

TABLE 10.3
Joint Effect of Other Factors

Scenario	All Crashes	Injury Crashes
Super 70 project	**161**	**36**
Super 70 project without other effects remaining unidentified by the statistical analysis (lower speed, enforcement-related behavioral changes, drivers' adaptation to the work zone, etc.)	211.4	47.3

TABLE 10.4
Interstate Segments and Location

Group number	Geographical Location	Road segments identified
1	Inside I-465 ring	Segment: 3,4,5,22
2	I-465 ring	Segment: 8,9,10,11,12,13,14,15
3	Outside I-465 ring	Segment: 1,2,7,16,17,18,20
4	I-70 work zone segment	Segment: 21

Note: Segment 19 has been removed because it has experienced major geometry changes during the study period due to the relocation of the Indianapolis International Airport terminal. It is unlikely that this event has affected other sections of the interstate system in the study area.

other traffic management strategies including speed reduction and police enforcement. In other words, the earlier estimate in Scenario 1 (Table 10.1), of 236 crashes, including 48 injury crashes, would have to be increased by at least 50 and 11, respectively. We believe that these additional precautions offset the added work zone difficulty, thereby bringing about the safety in Scenario 1 to a level comparable with the regular traffic conditions.

This effect also indicates that rerouting the non-heavy vehicles around the work zone may cause additional crash increases along these routes if they experience a similar level of risk by a typically longer increase to the crash exposure.

Before-and-During Study

A before-and-during study was conducted to estimate the safety change on other roads in the I-70 work zone area after the work zone onset on February 22, 2007. The I-70 work zone was present between February 22 and November 16, 2007. Thus, the period March-November, 2007 was named the during period and all data and results that apply to this period were labeled as during. The March-November periods during 2004, 2005, and 2006 were combined into a single before period and all data and results that apply to this period were labeled as before. Using the same months for the before and during periods eliminates from the study the effect of seasonal changes observed in traffic and safety.

The studied interstate roads were divided into four groups:

(1) Inside I-465 ring – expected lower traffic during the construction period
(2) I-465 ring – expected higher traffic during the construction period
(3) Outside I-465 ring – expected limited impact of the construction on traffic
(4) I-70 work zone – major impact

The studied interstate segments are listed in Table 10.4 and shown in Figure 10.5.

Crash Frequency

Different types of injury severities, such as fatal, incapacitating, non-incapacitating, possible injury, and property-damage-only crashes, and different types of vehicles involved in the crashes were investigated for negative binomial tests for different segments.

Due to INDOT's policy of rerouting of traffic around I-465, traffic reduced substantially on the I-70 work zone stretch as evidenced from the detector data of ATR 0313 and ATR 0314 (see Appendix A and Appendix B). Increases in traffic around I-465 were confirmed by the detector data, particularly ATR 0307 and WIM 3300, for all traffic and heavy vehicles (Appendix A and Appendix B).

All types of crashes decreased inside the I-465 ring and inside the I-70 work zone stretch, and the overall interstate crashes decreased. However, all crashes increased in the I-465 ring and outside the I-465 ring as explained earlier because of increased traffic in the I-465 ring and outside of the I-465 ring as people tried to avoid the I-70 work zone. Another cause could be a construction zone on the west section of I-465 which could contribute to the occurrence of additional crashes.

Table 10.5 indicates that non-incapacitating and possible injury crashes decreased significantly inside the I-465 ring, outside of I-465, and in the I-70 work zone. There was also some reduction of fatal and incapacitating injuries inside the I-465 ring.

Property-damage-only crashes decreased inside the I-465 ring. I-70 work zone and overall interstate crashes decreased as well whereas there was increase in the I-465 ring and outside the I-465 ring.

From Table 10.5, it is clear that moderately severe crashes decreased significantly. This shift caused property-damage-only crashes to increase in the I-465 ring and outside of the I-465 ring. In addition, more severe injuries, such as fatal and incapacitating injuries, increased outside of the I-465 ring.

Negative Binomial tests for different types of crashes are presented in Table 10.5. The P-value for Negative Binomial tests is shown in the last column. A P-value higher than 0.10 is considered the cut-off point and indicates that the change in crashes was significantly lower and vice versa.

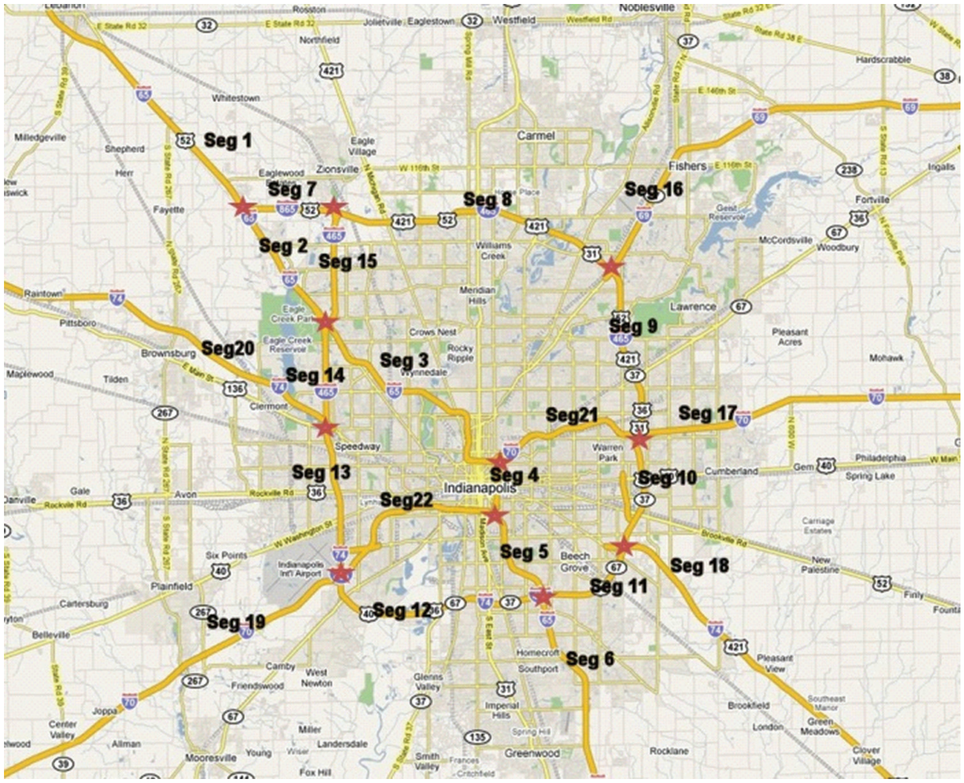


Figure 10.5 Studied interstate segments

TABLE 10.5
Negative Binomial Test for Different Injury Severity in Different Segments

Crash Severity	Roads	Crashes *Before* (2004–2006) Total	Crashes *Before* (2004–2006) Average	Crashes *During* 2007	Change in Crashes per nine months	Negative Binomial Test (P-value)
All crashes	Inside I-465 ring	1282	427.3	303	**-124.3**	0.000
	I-465 ring	2992	997.3	1064	**66.7**	0.036
	Outside I-465 ring	768	256	280	**24.0**	0.103
	I-70 work zone	664	221.3	151	**-70.3**	0.000
	All interstates	6205	2068.3	1901	**-167.3**	0.001
Fatal and incapacitating injury crashes	Inside I-465 ring	22	7.3	4	-3.3	0.184
	I-465 ring	58	19.3	18	-1.3	0.456
	Outside I-465 ring	20	6.7	15	**8.3**	0.011
	I-70 work zone	6	2.0	2	0.0	0.555
	All interstates	106	35.3	39	3.7	0.262
Non-incapacitating and possible injury crashes	Inside I-465 ring	232	77.3	51	**-26.3**	0.003
	I-465 ring	565	188.3	149	**-39.3**	0.005
	Outside I-465 ring	255	85.0	47	**-38.0**	0.000
	I-70 work zone	133	44.3	31	**-13.3**	0.040
	All interstates	1198	399.3	298	**-101.3**	0.000
Property-damage-only crashes	Inside I-465 ring	1033	344.3	248	**-96.3**	0.000
	I-465 ring	2369	789.7	897	**107.3**	0.001
	Outside I-465 ring	597	199.0	223	24.0	0.077
	I-70 work zone	525	175.0	118	**-57.0**	0.000
	All interstates	4524	1508	1486	-22.0	0.317

Crash Proportion for Interstate

Different types of injury severities – fatal, incapacitating, non-incapacitating, possible injury, and property-damage-only crashes - were investigated for the two proportion test for different segments considered.

As shown in Table 10.5, non-injury crashes–property-damage-only crashes increased. whereas moderate injuries reduced significantly. Given the crashes occurred, there is an increase tendency in non-injury crashes than more severe injuries.

The different types of vehicles involved in crashes were also investigated for a crash proportion test (Table 10.6). These types of vehicles are:

- Heavy vehicle [HVEH]: it includes truck/trailer (not semi), tractor/one semi trailer, tractor/double trailer, tractor/triple trailer, motor home/recreational vehicle, bus/seats 15+ persons with drivers, school bus, combination vehicle.
- Single Unit Truck [SUT]: It includes trucks (single 2 axle, 6 tires), truck (single 3 or more axles).
- Non-heavy vehicle [NHVEH]: It includes passenger car/station wagon, pickup, van and sport utility vehicles.

As mentioned earlier, heavy vehicles were restricted inside the I-70 work zone stretch, and a significant reduction of crashes involving heavy vehicles was observed inside the I-465 ring and the I-70 work zone stretch (Table 10.6). Nonetheless, crashes involving heavy vehicles significantly increased in the I-465 ring.

Crashes involving non-heavy vehicles increased significantly inside the I-465 ring, and the I-70 work zone stretch. In addition, single unit truck crashes decreased in the I-465 ring.

From Table 10.6, it is clear that the I-465 ring experienced an increase in heavy vehicle crashes and a decrease in single unit truck and non-heavy vehicle crashes. The I-70 work zone experienced an exactly opposite pattern – an increase in non-heavy vehicles and single unit trucks and a decrease in heavy vehicles.

11 FINDINGS AND RECOMMENDATIONS

This research study addressed the question of whether the Super 70 construction project was a successful example of traffic management from the viewpoint of safety. Another important objective was to identify the components of the work zone management that were shown effective in increasing safety. Finally, recommendations for future high-speed urban work zones supported by the conducted research are provided. The research objectives of this study were addressed in two complementary ways: (1) an advanced detailed statistical analysis and simulation of the work zone, and (2) an aggregated before-and-during study of the work zone's impact area.

The following findings and recommendations are a summary of the detailed results discussed in Chapters 7–10. This summary reflects only the results that are relevant to safety management in urban high-speed work zones. The reader should refer to the previous chapters for the discussion of other safety factors including weather.

11.1 Findings

The safety management during the Super 70 project can be considered highly successful in the light of our

TABLE 10.6
Two Proportion Test for Different Injury Severity in Different Segments for Interstate

Crash Severity	Roads	Crashes Before (2004–2006)	Crashes During 2007	Prop. *Before*	Prop. During	Diff. in Prop.	Z- value	P-value
All crashes	Inside I-465	1282	303	na	na	Na	na	na
	I-465 ring	2992	1064	na	na	Na	na	na
	Outside I-465	768	280	na	na	Na	na	na
	I-70 work zone	664	151	na	na	Na	na	na
	All interstates	5706	1798	na	na	Na	na	na
Fatal and incapacitating injury crashes	Inside I-465	22	4	0.017	0.013	-0.004	-0.488	0.313
	I-465 ring	58	18	0.019	0.017	-0.002	-0.510	0.305
	Outside I-465	20	15	0.026	0.054	0.027	**2.195**	0.014
	I-70 work zone	6	2	0.009	0.013	0.004	0.474	0.318
	All interstates	106	39	0.018	0.022	0.003	0.836	0.201
Non-incapacitating and possible injury crashes	Inside I-465	232	51	0.181	0.168	-0.013	-0.517	0.303
	I-465 ring	565	149	0.189	0.140	-0.049	**-3.590**	0.000
	Outside I-465	255	47	0.332	0.167	-0.164	**-5.192**	0.000
	I-70 work zone	133	31	0.200	0.205	0.005	0.138	0.445
	All interstates	1185	278	0.208	0.155	-0.053	**-4.453**	0.000
Property- damage-only crashes	Inside I-465	1033	248	0.806	0.818	0.013	0.505	0.307
	I-465 ring	2369	897	0.792	0.843	0.051	**3.627**	0.000
	Outside I-465	597	223	0.777	0.796	0.019	0.663	0.746
	I-70 work zone	525	118	0.791	0.781	-0.009	-0.250	0.401
	All interstates	4524	1486	0.793	0.826	0.034	**1.399**	0.081

TABLE 10.7
Two Proportion Test for Involvement of Different Injury Severity in Different Segments

Vehicle Type	Roads	Crashes *Before* (2004–2006)	Crashes *During* 2007	Prop. *Before*	Prop. During	Diff. in Prop.	Z-value	P-value
All vehicles	Inside I-465	2362	548	na	na	Na	na	na
	I-465 ring	5801	2114	na	na	Na	na	na
	Outside I-465	1373	513	na	na	Na	na	na
	I-70 work zone	1293	349	na	na	Na	na	na
	All interstates	10829	3524	na	na	Na	na	na
Heavy Vehicles	Inside I-465	259	33	0.110	0.060	-0.049	**-3.470**	0.000
	I-465 ring	565	295	0.097	0.140	0.042	**5.331**	0.000
	Outside I-465	178	75	0.130	0.146	0.017	0.939	0.174
	I-70 work zone	126	8	0.097	0.023	-0.075	**-4.513**	0.000
	All interstates	1128	411	0.104	0.117	0.012	**2.077**	0.019
Single Unit Truck	Inside I-465	58	16	0.025	0.029	0.005	0.622	0.267
	I-465 ring	189	51	0.033	0.024	-0.008	**-1.941**	0.026
	Outside I-465	30	9	0.022	0.018	-0.004	**-0.585**	0.279
	I-70 work zone	25	9	0.019	0.026	0.006	0.751	0.226
	All interstates	302	85	0.028	0.024	-0.004	-1.199	0.115
Non-Heavy Vehicles	Inside I-465	2003	486	0.848	0.887	0.039	**2.329**	0.010
	I-465 ring	4979	1731	0.858	0.819	-0.039	**-4.325**	0.000
	Outside I-465	1147	425	0.835	0.828	-0.007	-0.360	0.359
	I-70 work zone	1112	328	0.860	0.940	0.080	**4.028**	0.000
	All interstates	9241	2970	0.853	0.843	-0.011	**-1.529**	0.063

Note: Heavy Vehicles: truck/trailer (not semi), tractor/one semi trailer, tractor/double trailer, tractor/triple trailer, motor home/recreational vehicle, bus/seats 15+ persons with drivers, school bus, combination vehicle. Single-Unit-Truck: trucks (single 2 axle, 6 tires), truck (single 3 or more axles). Non-Heavy Vehicles: passenger car/station wagon, pickup, van and sport utility vehicles, tractor (cab only, no trailer), motorcycle, bus/seats 9–15 persons with driver, farm vehicle.

research results. The single most successful management strategy was rerouting heavy vehicles (13+ tons) on alternative interstate routes. The safety benefit was produced by two phenomena: (1) The reduced presence of heavy vehicles inside the work zone prevented a considerable number of crashes. (2) The subsequent increased number of the same heavy vehicles on alternative interstate roads had a "calming" effect on other vehicles and contributed to the absence of negative safety impacts on these roads.

The second significant source of safety benefit was jointly generated by police enforcement, reduced speed limits, and other traffic management strategies. The magnitude of this joint effect was estimated at 50 work zone crashes during the nine construction months and was approximately equal to the work zone safety effect of rerouting heavy vehicles on adjacent interstate roads. Thus, the safety benefit generated by the two sources was around 100 work zone crashes.

Widening shoulders has been indicated as an additional means of improving work zone safety. Widening shoulders may reduce the number and severity of work zone crashes. The possible limitation is typical shortage of the right of way in urban conditions amplified by the road construction.

The presented research cannot confirm that utilizing moveable barriers and consequently adjusting the number of traffic lanes to the traffic volumes provided any direct safety benefits inside the work zone. However, a positive impact on traffic safety on local roads may have occurred by providing additional capacity that reduced traffic disturbances on the local road system.

Closing interchange ramps had not only a limited safety effect for the work zone safety, but might be offset or exceeded by additional crashes on surface roads due to disturbance to the origin-destination and local traffic. Analyzing the safety effect of the Super 70 project on the local road network was outside of the research scope.

Comparing safety on the interstates affected by the Super 70 project before and during the construction period indicated an expected safety pattern. Safety was higher on interstate roads inside the I-465 rings where the traffic volumes were reduced by the rerouting and safety was reduced on the I-465 rings where the traffic volumes increased. The overall safety on the affected interstates was higher during the Super 70 project than before. This is considered a very positive outcome.

11.2 Recommendations

The following recommendations for urban high-speed construction zones were derived from the results of this study:

1. Reroute heavy vehicles (13+ tons) on alternative interstate routes. Advanced information about the restriction, accompanied with aggressive police enforcement, has proven to be an efficient in maintaining a high diversion rate among heavy vehicles.

2. Implement 24-hour State Police enforcement of the heavy vehicles restriction in the construction zone during the entire construction period.
3. Reduce speed limits and apply as wide as possible shoulders inside the work zone given the local conditions. Police enforcement should include speed enforcement. Use of additional traffic lanes instead of wide shoulders should be considered where shortage of capacity is expected, which may lead to traffic spillover to surface roads.
4. Avoid redirecting interstate through traffic on surface roads if possible. Consider providing a sufficient number of lanes inside the work zone. Using moveable barriers to adjust capacity to demand during a day will provide additional capacity without negative safety impact inside the work zone.
5. Reduce to a minimum the impact of the work zone on the local traffic by using the moveable barriers to adjust the number of lanes to current traffic demand and by maintaining as many ramps open as possible.
6. Consider developing a real time crash risk assessment tool based on the developed short-term safety model to evaluate the risk in real time and warn drivers entering the work zone about the heightened risk level.

11.3 Remarks

The magnitude of the research effort was defined mainly by the size of the dataset, which included nearly 14 million time intervals and the scope of the 150 variable values for each time interval including traffic, weather, geometry, enforcement, crashes, and traffic management. The total number of two billion data values posed a formidable data management problem. Even minor issues, normally dealt with easily, multiplied in this project to major time-consuming tasks. Source data cleaning, conversion, linking into a single dataset, and imputing some missing values took more than one year of extensive effort on processing, quality testing, and corrections. To make the modeling task manageable, we decided to use around four percent of the records without crashes and were able to use 60 percent of the observations with crashes.

The above remarks are meant to help others who may in the future consider using massive disaggregate data for similar safety analysis. Although this research approach is powerful and offers insight unmatched by more traditional analysis, the effort required managing and model data is considerable and must be included in the research planning.

The crash model developed in this research indicates the strong variability of the risk of crash in 30-minute intervals and, even stronger, per vehicle. This variability prompts an attractive new safety management tool warning drivers in real time about the crash danger. Variable message signs displaying adequate messages seem to be a promising means of implementation. The input needed to calculate the risk includes available today input: the geometry characteristics of the road, weather data easy to obtain on-line from local weather stations, and nearby traffic detector counts.

REFERENCES

Al-Ghamdi, A. S. (2002). "Using Logistic Regression to Estimate the Influence of Accident Factors on Accident Severity." *Accident Analysis and Prevention*, 34(6), 729–741, doi: 10.1016/S0001-4575(01)00073-2.

Bedard, M., Guyatt, G. H., Stones, M. J., and Hirdes, J. P. (2002). "The Independent Contribution of Driver, Crash, and Vehicle Characteristics to Driver Fatalities." Accident Analysis and Prevention, 34(6), 717–727, doi: 10.1016/S0001-4575(01)00072-0.

Compton, C. P. (2005). "Injury severity codes: A comparison of police injury codes and medical outcomes as determined by NASS CDS Investigators." Journal of Safety Research – Traffic Records Forum Proceedings, 36(5), 483–484.

Dissanayake, S. and Lu, J. (2006). "Analysis of Severity of Young Driver Crashes, Sequential Binary Logistic Regression Modeling." *Transportation Research Record: Journal of the Transportation Research Board*, No. 1784, TRB, National Research Council, Washington D.C., 108 – 114.

Hall, J. W., and Lorenz, V. M. (1989). "Characteristics of Construction-Zone Accidents." Transportation Research Record, 1230, 20–27.

Interstate 70 Reconstruction Project, [http://www.in.gov/indot/div/projects/i70/projfact.html].

Jin, T. G., Saito, M., and Eggett, D. L. (2008). "Statistical Comparison of the Crash Characteristics on Highways between Construction Time and Non-construction Time." Accident Analysis and Prevention, 40, 2015–2023, doi: 10.1016/j.aap.2008.08.024.

Kaisy, A. A., Zhou, M., and Hall, F. (2000). "New Insights into Freeway Capacity at Work Zones: Empirical Case Study." Transportation Research Record, 1710, 154–160, doi: 10.3141/1710-18.

Khattak, A. J., Khattak, A. J., and Council, F. M. (2002). "Effects of Work Zone Presence on Injury and Non-injury Crashes." Accident Analysis and Prevention, 34(1), 19–29, doi: 10.1016/S0001-4575(00)00099-3.

Khattak, A. J., and Targa, F. (2004). "Injury Severity and Total Harm in Truck-involved Work Zone Crashes." Transportation Research Record, 1877, 106–116, doi: 10.3141/1877-12.

Li, Y., and Bai, Y. (2006). "Investigating the Characteristics of Fatal Crashes in Highway Construction Zones." *Proceedings of CIB W99 International Conference on Global Unity for Safety & Health in Construction*, Beijing, China.

Li, Y., and Bai, Y. (2008). "Development of Crash-Seveirty-Index Models for Measurement of Work Zone Risk Levels." Accident Analysis and Prevention, 40, 1724–1731, doi: 10.1016/j.aap.2008.06.012.

Lu, G., Noyce, D. A., and McKendry, R. J. (2006). "Analysis of the Magnitude and Predictability of Median Crossover Crashes Utilizing Logistic Regression." *Proceedings of TRB 85th Annual Meeting* (CD-ROM) TRB, National Research Council, Washington D.C., Paper #06-2681.

Mercier, C. R., Shelley, M. C., Rimkus, J. B. and Mercier, J. M. (1997). "Age and Gender as Predictors of Injury Severity in Head-on Highway vehicular Collisions." *Transportation Research Record: Journal of the Transportation Research Board*, No. 1581, TRB, National Research Council, Washington D.C., 37 – 46.

Pal, R., and Sinha, K. C. (1996). "Analysis of Crash Rates at Interstate Work Zones in Indiana." *Transportation Research Record*, No. 1529, 45–53.

SAS Institute Inc. (2007). SAS OnlineDoc® 9.2**.,** Cary, North Carolina.

Spainhour, A. K., and Mishra, A. (2008). "Analysis of Fatal Run Off the Road Crashes Involving Overcorrection." *Transportation Research Record: Journal of the Transportation Research Board*, No. 2069, National Research Council, Washington, D.C., 1–8.

Tay, R., and Rifaat, S. M., and Chin, H. C. (2008). "A Logistics Model of the Effects of Roadway, Environmental, Vehicle, Crash and Driver Characteristics on hit-and-run Crashes." Accident Analysis and Prevention, 40, 1330–1336, doi: 10.1016/j.aap.2008.02.003.

Tarko, A. P., (2007). "Improving Safety in High-Speed Work Zones A Super 70 Study." Proposal for First-Track Research Study.

Wang, J., Hughes, W. E., Council, F. M., and Paniati, J. F. (1996). "Investigation of Highway Work Zone Crashes: What We Know and What We Don't Know." *Transportation Research Record: Journal of the Transportation Research Board*, No. 1529, TRB, National Research Council, Washington D.C., 54 – 62.

Washington, S. P., Karlaftis, M. G., and Mannering, F. L, (2003). "Statistical and econometric Methods for Transportation Data Analysis." CRC press company.

Venugopal, S. and Tarko, A. (2000). "Safety Models for Rural Freeway Work Zones." Transportation Research Record, 1715, 1–9, doi: 10.3141/1715-01.

Pairs 254446 c 0.639

APPENDIX A

TABLE A.1:
Identified and Investigated Variables

Variable	Description
BFOG	= 1 if fog, =0 otherwise
BHEAVY	= 1 if heavy intensity of the atmospheric event, =0 otherwise
BLIGHT	= 1 if light intensity of the atmospheric event, =0 otherwise
BLTMOD	= 1 if light/moderate intensity of the atmospheric event, =0 otherwise
BMDHEAV	= 1 if moderate/heavy intensity of the atmospheric event, =0 otherwise
BMODER	= 1 if moderate intensity of the atmospheric event, =0 otherwise
BPRECIP	= 1 if rain or snow precipitation, = 0 otherwise
BRAIN	= 1 if rain precipitation, = 0 otherwise
BSNOW	= 1 if snow precipitation, = 0 otherwise
BTSTRM	= 1 if thunderstorm, =0 otherwise
BTEMP	= 1 if temperature below the freezing point, = 0 otherwise
CRASH	= 1 if crash occurs, = 0 otherwise
CURVSEGN	Number of horizontal curves on the segment
DOFFRAMP	= 1 if an off-ramp terminal is at the downstream end of the segment, = 0 otherwise
DONRAMP	= 1 if an on-ramp terminal is at the downstream end of the segment, = 0 otherwise
DW1	= 1 if Monday, = 0 otherwise
DW2	= 1 if Tuesday, = 0 otherwise
DW3	= 1 if Wednesday, = 0 otherwise
DW4	= 1 if Thursday, = 0 otherwise
DW5	= 1 if Friday, = 0 otherwise
DW6	= 1 if Saturday, = 0 otherwise
DW7	= 1 if Sunday, = 0 otherwise
HEVVEH	Percent of 13+ ton trucks, percent
HRWRK	Total hours of enforcement in a month (men-hrs/month)
I465	= 1 if segment on I465 interstate, = 0 otherwise
I65	= 1 if segment on I65 interstate, = 0 otherwise
I70	= 1 if segment on I70 interstate, = 0 otherwise
INJURY	= 1 if crash with at least one person injured (any type of injuries in KABCO), = 0 otherwise
INSHLDW	Inside shoulder width, ft
L1	= 1 if segment has 1 traffic lanes (one direction), = 0 otherwise
L2	= 1 if segment has 2 traffic lanes (one direction), = 0 otherwise
L3	= 1 if segment has 3 traffic lanes (one direction), = 0 otherwise
L4	= 1 if segment has 4 traffic lanes (one direction), = 0 otherwise
L5	= 1 if segment has 5 traffic lanes (one direction), = 0 otherwise
L67	= 1 if segment has 6 or 7 traffic lanes (one direction), = 0 otherwise
LANESNUM	Number of lanes in one direction
M1	= 1 of January, = 0 otherwise
M2	= 1 if February, = 0 otherwise
M3	= 1 if March, = 0 otherwise
M4	= 1 if April, = 0 otherwise
M5	= 1 if May, = 0 otherwise
M6	= 1 if June, = 0 otherwise
M7	= 1 if July, = 0 otherwise
M8	= 1 if August, = 0 otherwise
M9	= 1 if September, = 0 otherwise
M10	= 1 if October, = 0 otherwise
M11	= 1 if November, = 0 otherwise
M12	= 1 if December, = 0 otherwise
MEDBRR	= 1 if median barrier, = 0 otherwise
MULTVEH	= if crash with involved multiple vehicles, = 0 otherwise
MVCT	Number of movement violation citations in a month,
NWZ	= 1 if non-work-zone segment, = 0 otherwise
OUTSHLDW	Outside shoulder width, ft
PERIOD1	= 1 if before the Super 70 period, = 0 otherwise
PERIOD2	= 1 if during the first Super 70 phase, = 0 otherwise
PERIOD3	= 1 if during transition between the first and second phases of Super 70, = 0 otherwise
PERIOD4	= 1 if during the second Super 70 phase (2), = 0 otherwise
PERIOD5	= 1 if after the Super 70 period, = 0 otherwise
RDSDBRR	= 1 if roadside barrier present, = 0 otherwise
RSEGLEN	Segment length - should be as close to 0.25 mile as possible, mi

TABLE A.1
(Continued)

Variable	Description
SINGVEH	= 1 if crash with involved single vehicle, = 0 otherwise
SPDCT	Number of speeding violations citations in a month
SPEEDLMT	Speed limit, mi/h
TRADIR	= F if mileposts grow in the direction of traffic, =B if mileposts decrease in the direction of traffic
T1	=1 if between midnight and 1 AM
T2	= 1 if between 1 AM and 2 AM, = 0 otherwise
T3	= 1 if between 2 AM and 3 AM, = 0 otherwise
….	….
T23	= 1 if between 10 PM and 11 PM, = 0 otherwise
T24	= 1 if between 11 PM and midnight, = 0 otherwise
TRACT	Number of traffic violations citations in a month
TRUCKCT	Number of truck violations citation in a month
TRUCKMCT	Number of moving violations citations by trucks in a month
UOFFRAMP	= 1 if an off-ramp starts at the upstream end of the segment, = 0 otherwise
UONRAMP	= 1 if an on-ramp merges at the upstream end of the segment, = 0 otherwise
VOLPLN	Traffic volume per lane, veh/h/lane
VOLUME	Traffic volume in one direction, veh/h
VISIBILITY	= air visibility as reported by a weather agency, mi
WORKZONE	= 1 if the segment is inside the work zone, = 0 otherwise

TABLE A.2
Summary of the non-binary sample variables

WORK ZONE SEGMENTS (156,646)											
Variable	Mean	St. Dev.	Min.	1st Prcntle	5st Prcntle	25st Prcntle	50st Prcntle	75st Prcntle	95st Prcntle	99th Prcntle	Max.
RSEGLEN (ft)	0.24	0.08	0.00	0.05	0.09	0.23	0.25	0.25	0.35	0.41	1.25
LN_W (ft)	11.00	0.00	11.0	11.0	11.0	11.0	11.0	11.0	11.0	11.0	11.0
IN_SHLD_W (ft)	1.97	0.25	0.0	0.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0
OUT_SHLD_W (ft)	4.09	4.10	0.0	0.0	1.0	2.0	2.0	3.0	15.0	15.0	15.0
CURV_SEG_N	0.55	0.50	0	0	0	0	1	1	1	1	1
VISIBILITY (FT)	9726	7331	0	0	0	0	16000	16093	16093	16093	16093
TEMP (°C)	19.0	8.7	-9	-4	1	14	21	26	31	33	36
VOLUME (veh/h)	1153	979	0	94	144	444	888	1629	3128	4910	6412
HVVOL (veh/h)	18.7	20.5	0	2	4	7	14	23	47	97	362
LANESNUM	2.63	0.75	1	1	1	2	3	3	4	4	5
VOLPLN (veh/h/ln)	455	386	0	37	56	184	359	607	1186	1886	3674
HRWRK (hrs/ month)	1526	664	0	0	0	1654	1655	1929	2055	2055	2055
NON-WORK-ZONE SEGMENTS (309,837)											
Variable	Average	Stand. Dev.	Minimum	1st Prcntle	5st Prcntle	25st Prcntle	50st Prcntle	75st Prcntle	95st Prcntle	99th Prcntle	Max.
RSEGLEN (ft)	0.26	0.08	0.06	0.06	0.16	0.25	0.25	0.25	0.44	0.48	0.49
LN_W (ft)	11.91	0.44	11.0	11.5	11.5	11.7	11.7	12.0	12.5	14.0	14.1
IN_SHLD_W (ft)	12.23	3.42	0.0	4.0	6.7	11.0	12.0	16.0	17.0	17.0	18.7
OUT_SHLD_W (ft)	10.96	2.99	0.0	0.0	8.0	10.0	11.0	12.0	16.0	23.0	26.0
CURV_SEG_N	0.56	0.62	0	0	0	0	0	1	2	2	2
VISIBILITY (FT)	9448	7273	0	0	0	0	14000	16093	16093	16093	16093
TEMP (°C)	10.0	11.2	-21	-14	-7	1	9	19	28	32	36
VOLUME (veh/h)	1969	1373	0	170	285	761	1770	2856	4571	5537	9048
HVVOL (veh/h)	189.2	127.5	0	0	28	95	161	257	442	552	861
LANESNUM	3.47	0.96	2	2	3	3	3	4	6	7	7
VOLPLN (veh/h/ln)	602	445	0	41	83	221	526	870	1441	1847	3444
HRWRK (hrs/ month)	583	838	0	0	0	0	0	1654	1929	2055	2055
ALL SEGMENTS (466,483)											
Variable	Average	Std. Dev.	Minimum	1st Prcntle	5st Prcntle	25st Prcntle	50st Prcntle	75st Prcntle	95st Prcntle	99th Prcntle	Max.
RSEGLEN (ft)	0.26	0.08	0.00	0.06	0.13	0.25	0.25	0.25	0.40	0.48	1.25
LN_W (ft)	11.60	0.56	11.0	11.0	11.0	11.0	11.7	12.0	12.3	14.0	14.1
IN_SHLD_W (ft)	8.79	5.59	0.0	0.0	2.0	2.0	11.0	12.5	17.0	17.0	18.7
OUT_SHLD_W (ft)	8.65	4.70	0.0	0.0	2.0	2.0	10.0	11.7	15.0	21.3	26.0
CURV_SEG_N	0.55	0.58	0	0	0	0	1	1	1	2	2
VISIBILITY (FT)	9541	7294	0	0	0	0	14484	16093	16093	16093	16093
TEMP (°C)	13.0	11.3	-21	-13	-6	3	14	22	29	32	36
VOLUME (veh/h)	1695	1313	0	115	197	626	1421	2458	4296	5359	9048
HVVOL (veh/h)	131.9	132.0	0	2	5	20	96	201	404	525	861
LANESNUM	3.19	0.98	1	1	2	3	3	3	5	7	7
VOLPLN (veh/h/ln)	552	432	0	39	70	206	464	784	1385	1855	3674
HRWRK (hrs/ month)	900	902	0	0	0	0	1536	1812	2055	2055	2055

TABLE A.3
Summary of the binary sample variables (proportion of observations with value 1)

Variable	Work zone	No-work zone	All segments	Variable	Work zone	No-work zone	All segments	Variable	Work zone	No-work zone	All segments
M1	0.000	0.130	0.086	T1	0.042	0.042	0.042	L1	0.081	0.000	0.027
M2	0.025	0.129	0.094	T2	0.041	0.042	0.042	L2	0.289	0.046	0.127
M3	0.120	0.150	0.140	T3	0.040	0.041	0.041	L3	0.555	0.638	0.610
M4	0.004	0.057	0.039	T4	0.041	0.041	0.041	L4	0.073	0.194	0.154
M5	0.037	0.084	0.068	T5	0.042	0.042	0.042	L5	0.003	0.068	0.046
M6	0.123	0.089	0.101	T6	0.041	0.042	0.042	L67	0.000	0.054	0.036
M7	0.202	0.063	0.110	T7	0.042	0.042	0.042	I65	0.000	0.315	0.209
M8	0.130	0.047	0.075	T8	0.042	0.042	0.042	I70	1.000	0.065	0.379
M9	0.145	0.041	0.076	T9	0.042	0.041	0.042	UONRAMP	0.041	0.395	0.276
M10	0.138	0.069	0.092	T10	0.042	0.042	0.042	UOFFRAMP	0.156	0.309	0.258
M11	0.076	0.068	0.070	T11	0.041	0.041	0.041	DONRAMP	0.059	0.353	0.254
M12	0.000	0.073	0.049	T12	0.042	0.042	0.042	DOFFRAMP	0.142	0.363	0.289
DW1	0.149	0.138	0.142	T13	0.042	0.042	0.042	BSLIGHT	0.035	0.088	0.070
DW2	0.141	0.141	0.141	T14	0.042	0.042	0.042	BSLMOD	0.005	0.006	0.005
DW3	0.129	0.143	0.138	T15	0.042	0.042	0.042	BMODER	0.005	0.008	0.007
DW4	0.137	0.148	0.144	T16	0.042	0.042	0.042	BMODHV	0.001	0.001	0.001
DW5	0.152	0.149	0.150	T17	0.042	0.041	0.042	BHEAVY	0.002	0.003	0.003
DW6	0.149	0.143	0.145	T18	0.041	0.041	0.041	BFOG	0.001	0.003	0.002
DW7	0.143	0.138	0.140	T19	0.042	0.042	0.042	BRAIN	0.041	0.068	0.059
PERIOD1	0.000	0.422	0.280	T20	0.042	0.042	0.042	BSNOW	0.009	0.039	0.029
PERIOD2	0.359	0.196	0.251	T21	0.042	0.041	0.042	BTSTRM	0.008	0.007	0.007
PERIOD3	0.086	0.009	0.035	T22	0.042	0.041	0.041	BTEMP	0.024	0.133	0.096
PERIOD4	0.555	0.153	0.288	T23	0.042	0.042	0.042	BTEMPMISS	0.335	0.321	0.326
PERIOD5	0.000	0.221	0.147	T24	0.041	0.041	0.041	VISIBMISS	0.335	0.321	0.326
INJURY	0.167	0.151	0.153	MULTVEH	0.848	0.722	0.734	SINGVEH	0.152	0.278	0.266

APPENDIX B

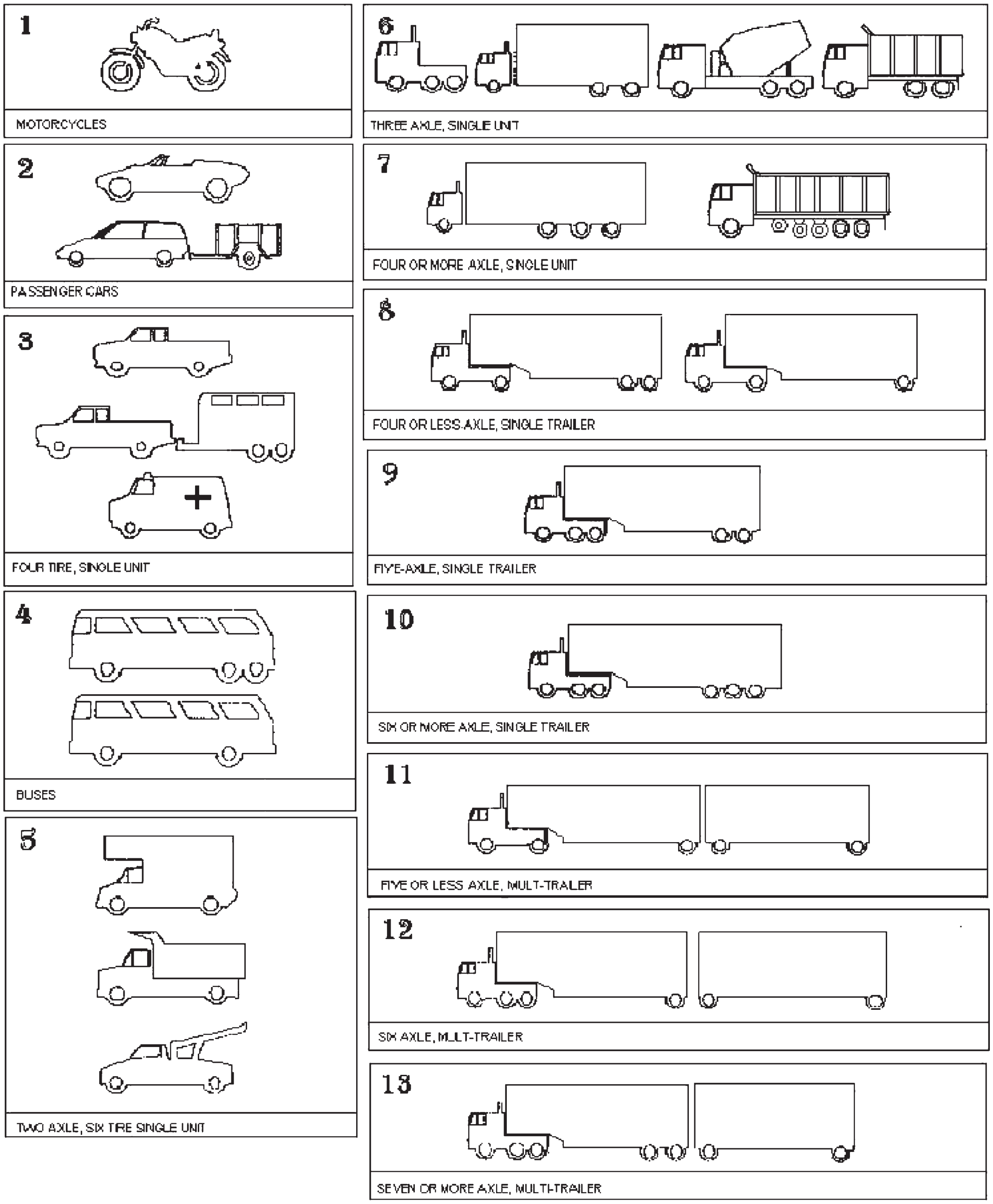


Figure B.1: FHWA Vehicle Classification

Figure C.1: Geometry showing CD connecting to main line

TMP_BRR: Movable barrier – Type 4, and Temporary barrier – Type 2 inside the work zone

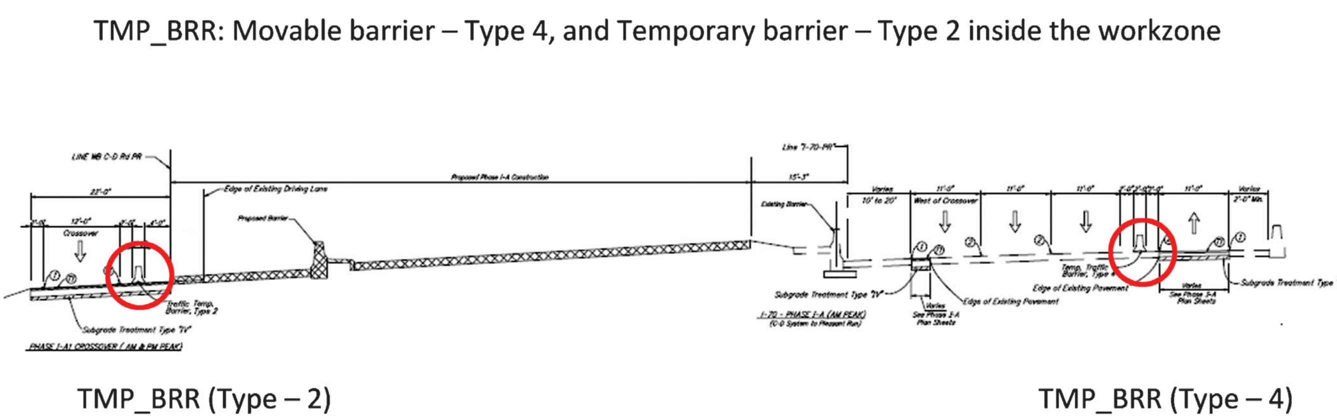


Figure C.2: Movable barrier and temporary barrier

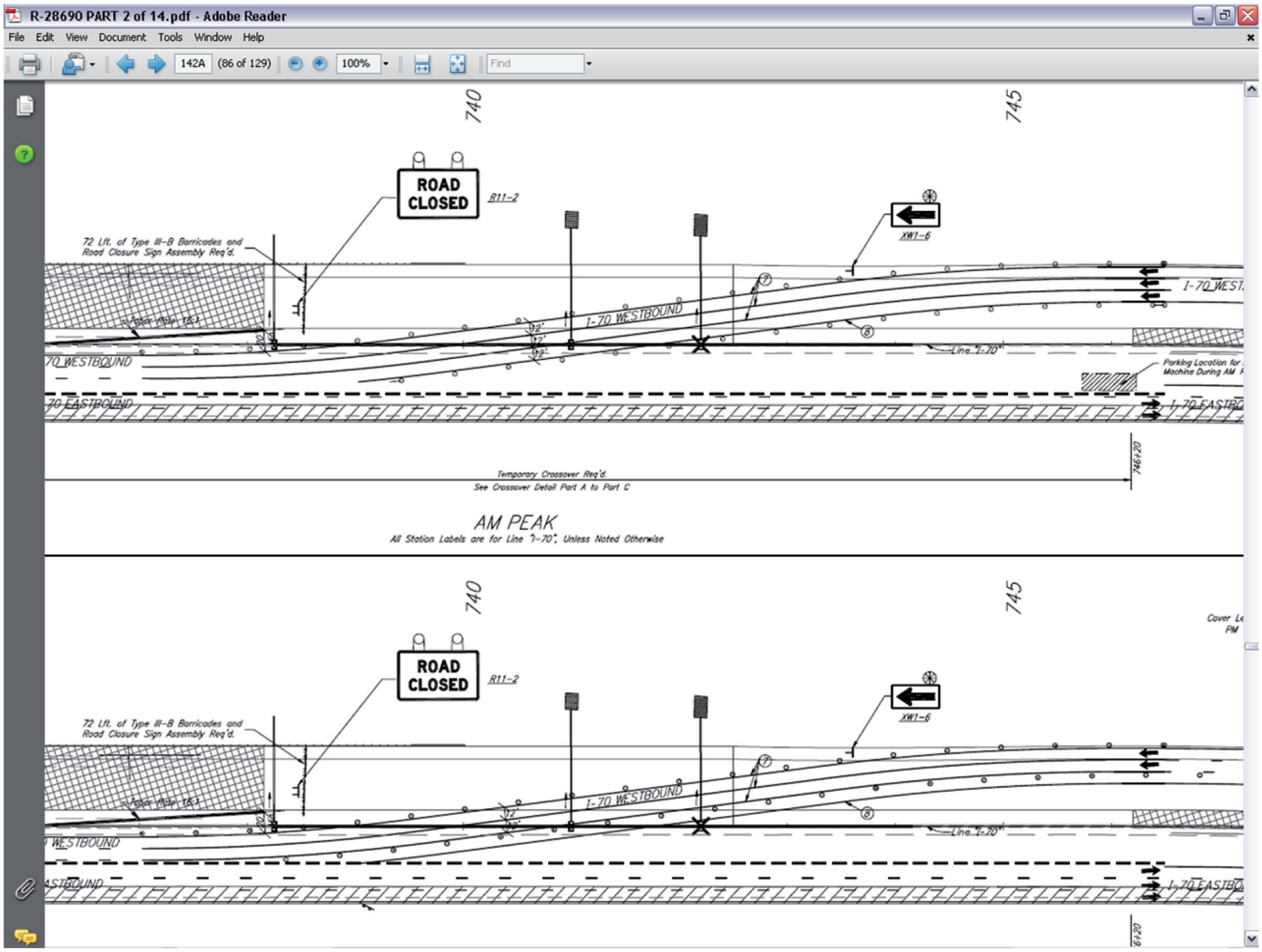


Figure C.3: Presence of Crossover

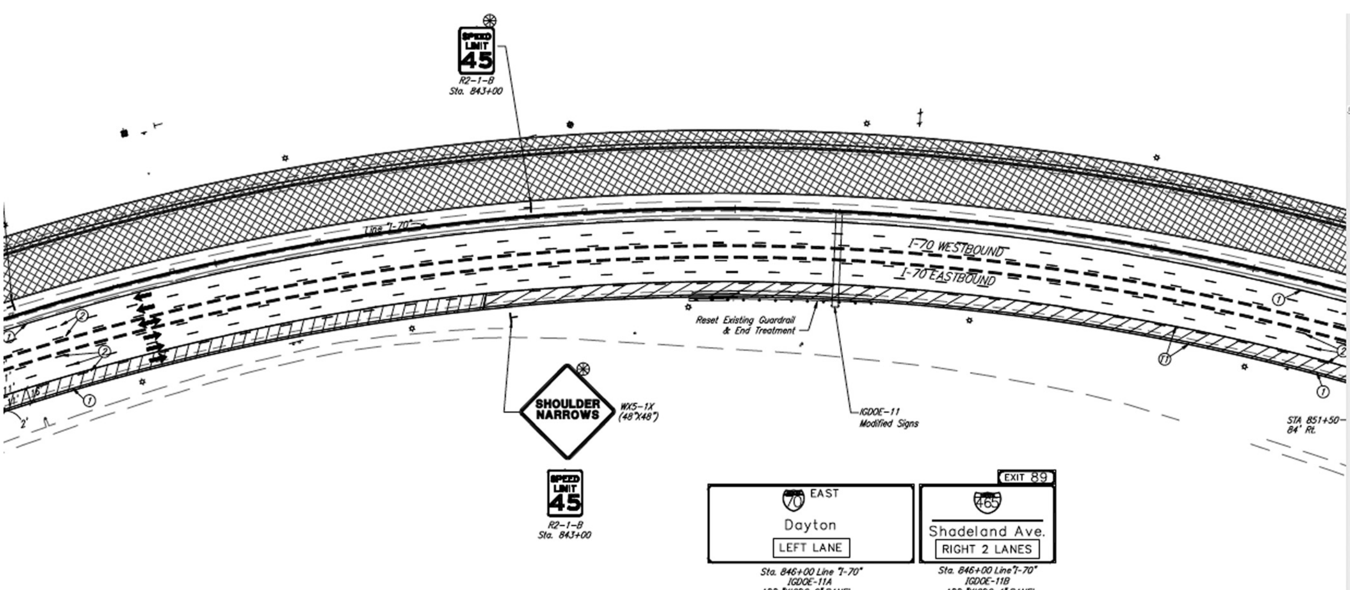


Figure C.4: Presence of “Shoulder Narrow” sign

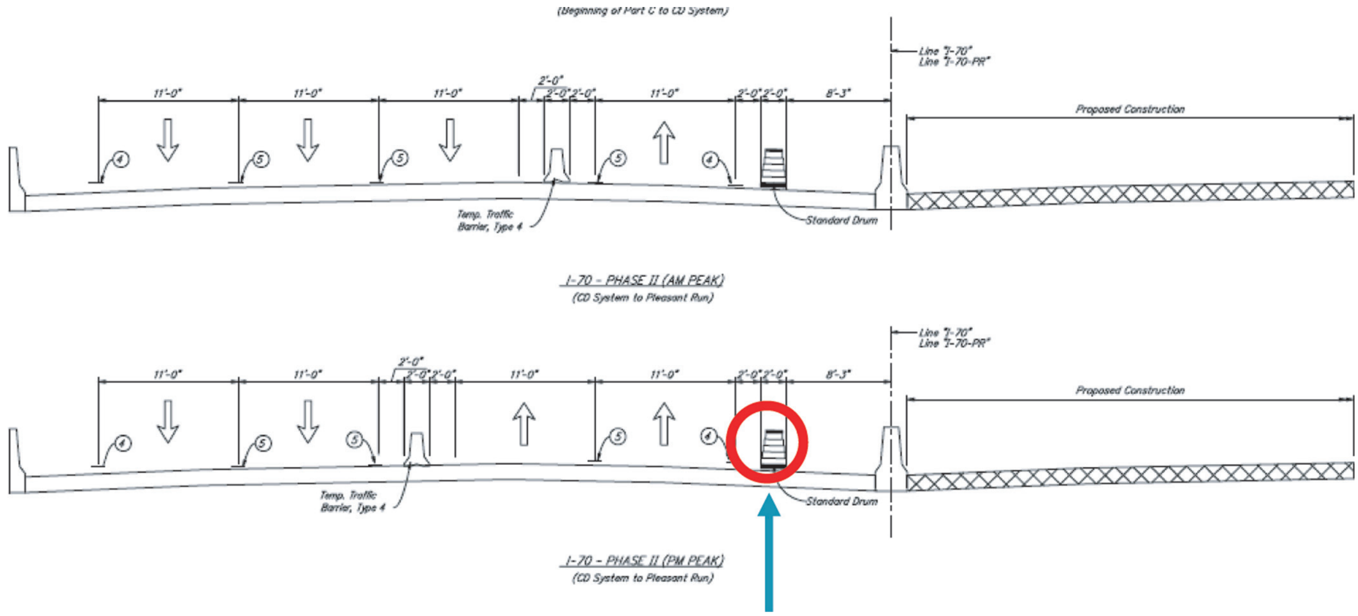


Figure C.5: Presence of different type of barrier inside WZ

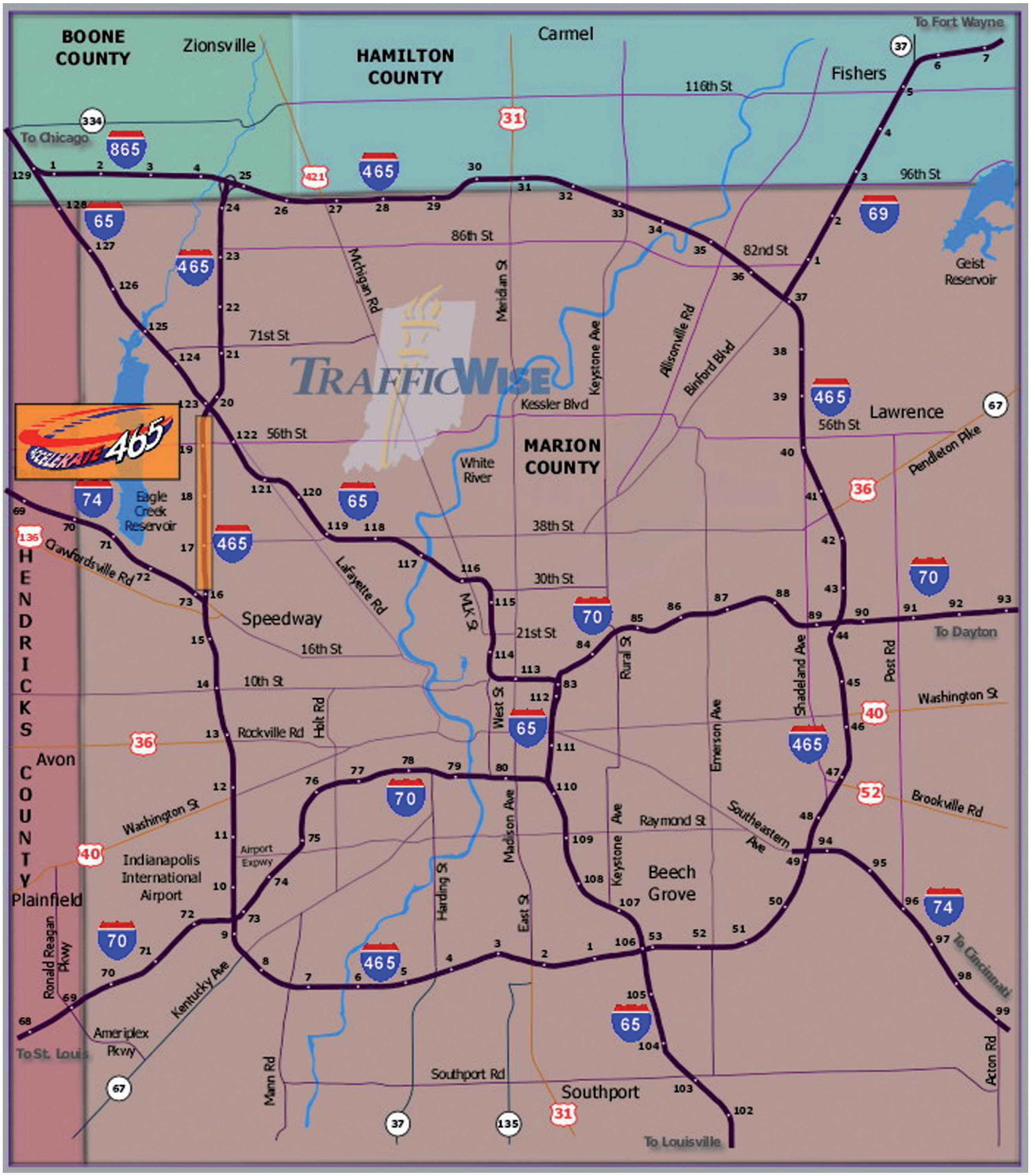


Figure D.1: Map showing mile marker as followed in the direction of traffic

APPENDIX E

TABLE E.1:
Detector and assigned Segment Information

Period	Segment range	Direction	Detectors
1	83.40 – 85.57	F	0315F
1	85.57 – 86.67	F	S125F
1	86.67 – 87.80	F	S125F
1	87.80 – 90.22	F	973220F
1	83.46 – 85.56	B	0316B
1	85.56 – 86.68	B	S125B
1	87.73 – 90.08	B	973220B
2	83.40 – 85.44	F	0315F
2	85.44 – 86.82	F	S125F
2	86.82 – 87.80	F	S125F
2	87.80 – 90.22	F	973220F
2	85.44 – 83.40	B	0316B
2	87.80 – 85.44	B	S125B
2	90.10 – 87.80	B	973220B
3	83.40 – 85.44	F	0315F
3	85.44 – 86.57	F	S125F
3	86.57 – 87.80	F	S125F
3	87.80 – 90.22	F	973220F
3	85.44 – 83.40	B	0316B
3	87.80 – 85.44	B	S125B
3	87.80 – 90.10	B	973220B
4	83.40 – 85.44	F	0315F
4	85.44 – 86.70	F	S125F
4	86.70 – 87.70	F	S125F
4	87.70 – 90.22	F	973220F
4	85.44 – 83.40	B	0316B
4	87.80 – 85.44	B	S125B
4	90.10 – 87.80	B	973220B
5	83.40 – 85.57	F	0315F
5	85.57 – 86.67	F	S125F
5	86.67 – 87.80	F	0322F
5	87.80 – 89.60	F	973220F
5	85.56 – 83.46	B	0316B
5	87.73 – 85.56	B	S125B
5	90.08 – 87.73	B	973220B

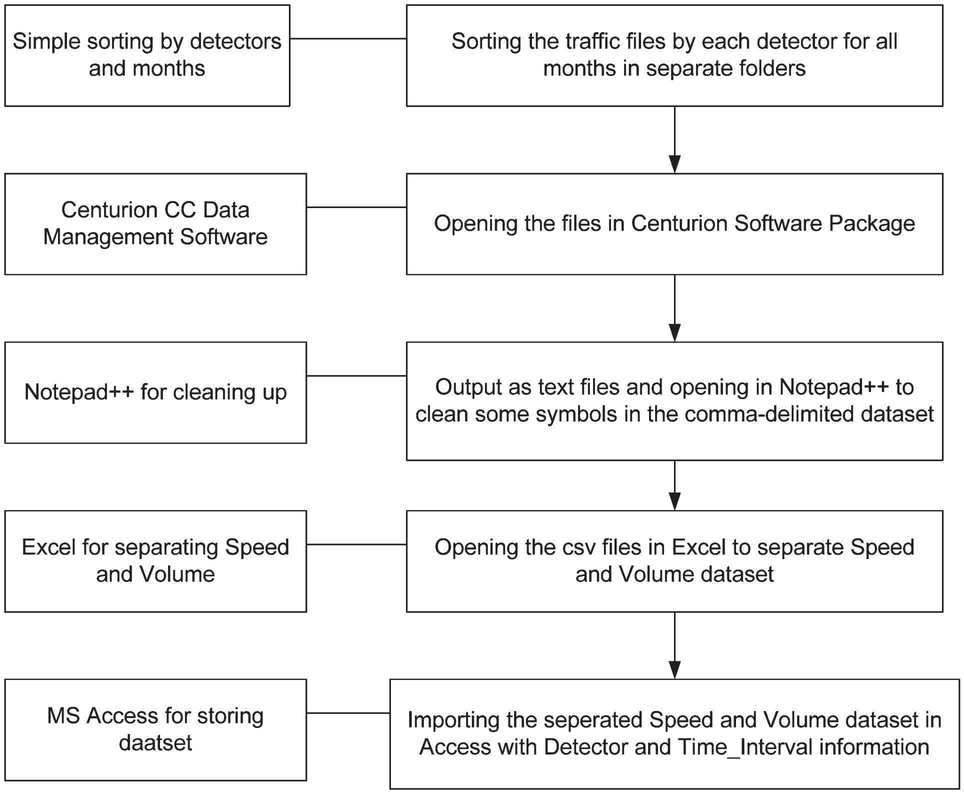
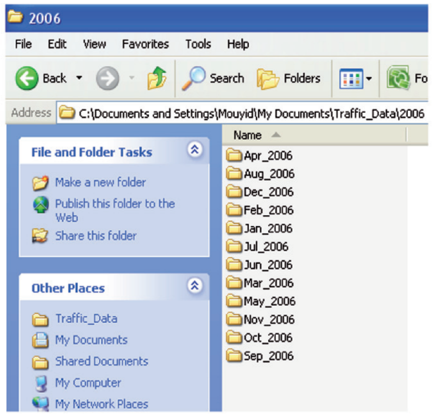


Figure E.1: Stepwise Process for Traffic Data Conversion to Storage

Source File Format: Month by month folder The original files are arranged month by month of all detectors that had been received from INDOT ftp.



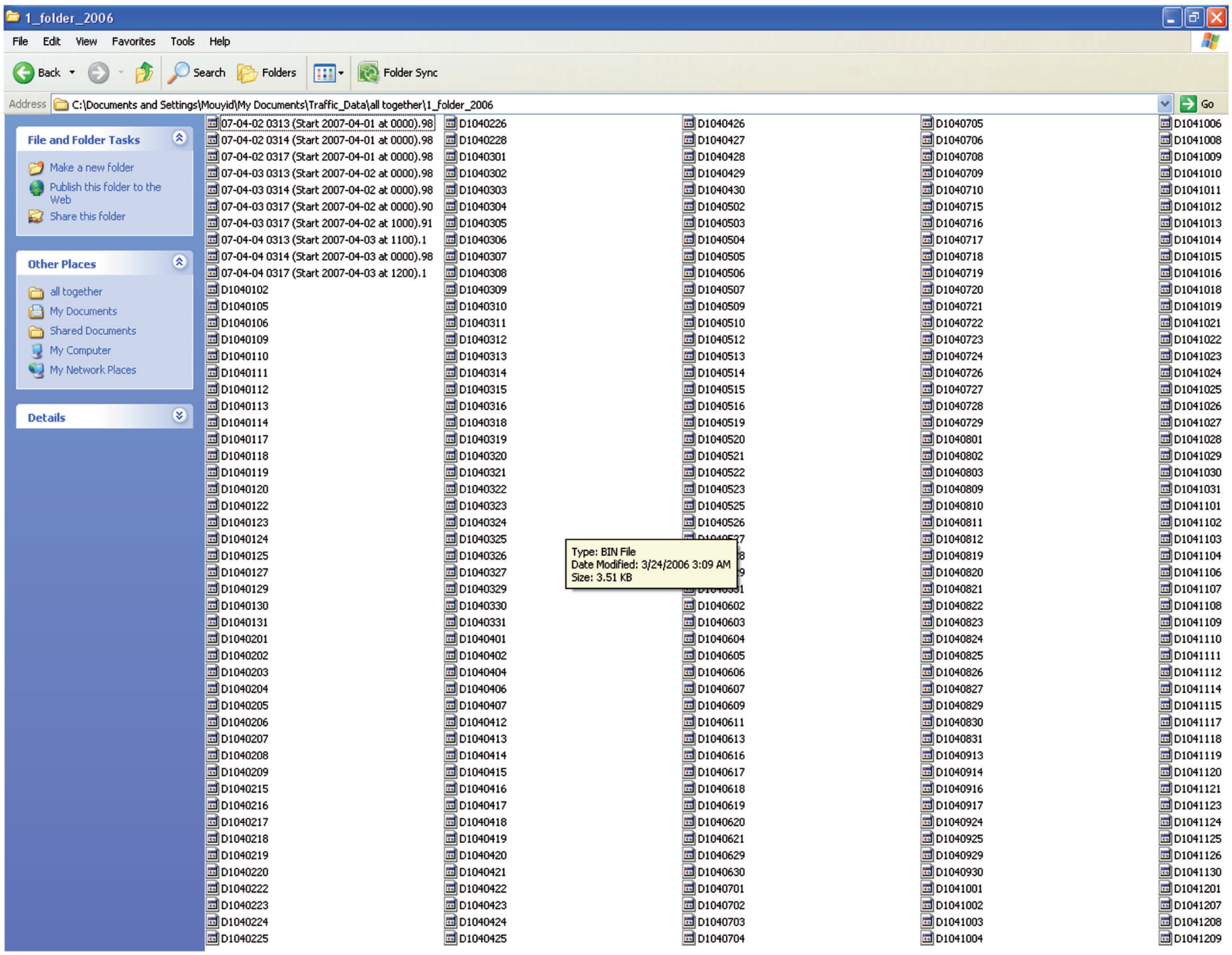
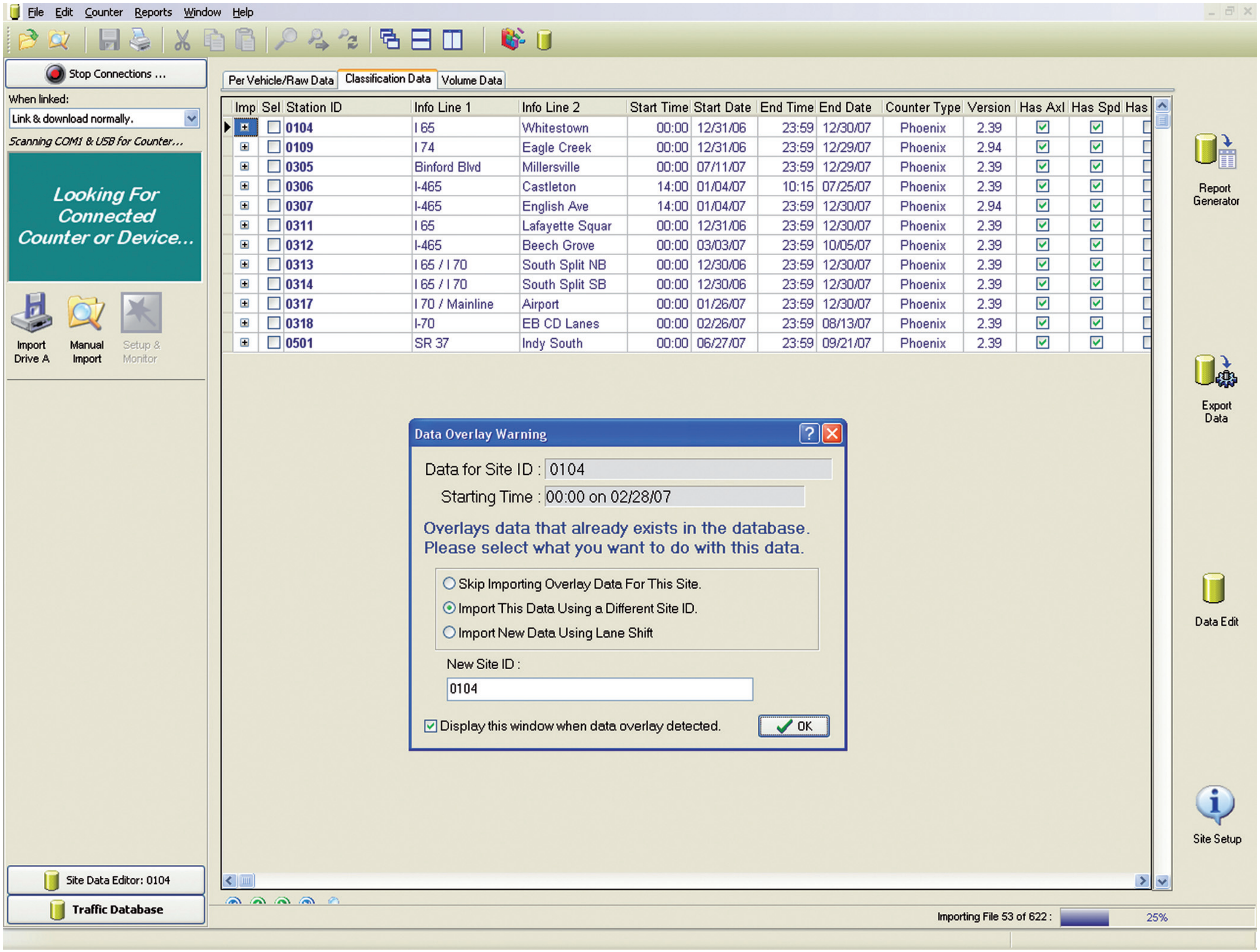


Figure E.2: All the files of different detectors are kept in the one folder before inputting for conversion from binary format to ASCII format. So there are no more month by month folders. It is kept in one single folder. Opening the Centurion Data Management Software, the files of the single folder are selected and input to its database, then output as set as Old ASCII format no labels. And the output files are text format and saved as different detectors.



Imp	Sel	Station ID	Info Line 1	Info Line 2	Start Time	Start Date	End Time	End Date	Counter Type	Version	Has Axl	Has Spd	Has
		0104	I 65	Whitestown	00:00	12/31/06	23:59	12/30/07	Phoenix	2.39			
		0109	I 74	Eagle Creek	00:00	12/31/06	23:59	12/29/07	Phoenix	2.94			
		0305	Binford Blvd	Millersville	00:00	07/11/07	23:59	12/29/07	Phoenix	2.39			
		0306	I-465	Castleton	14:00	01/04/07	10:15	07/25/07	Phoenix	2.39			
		0307	I-465	English Ave	14:00	01/04/07	23:59	12/30/07	Phoenix	2.94			
		0311	I 65	Lafayette Squar	00:00	12/31/06	23:59	12/30/07	Phoenix	2.39			
		0312	I-465	Beech Grove	00:00	03/03/07	23:59	10/05/07	Phoenix	2.39			
		0313	I 65 / I 70	South Split NB	00:00	12/30/06	23:59	12/30/07	Phoenix	2.39			
		0314	I 65 / I 70	South Split SB	00:00	12/30/06	23:59	12/30/07	Phoenix	2.39			
		0317	I 70 / Mainline	Airport	00:00	01/26/07	23:59	12/30/07	Phoenix	2.39			
		0318	I-70	EB CD Lanes	00:00	02/26/07	23:59	08/13/07	Phoenix	2.39			
		0501	SR 37	Indy South	00:00	06/27/07	23:59	09/21/07	Phoenix	2.39			

Figure E.3: Opening the Centurion Data Management Software, the files of the single folder are selected and input to its database

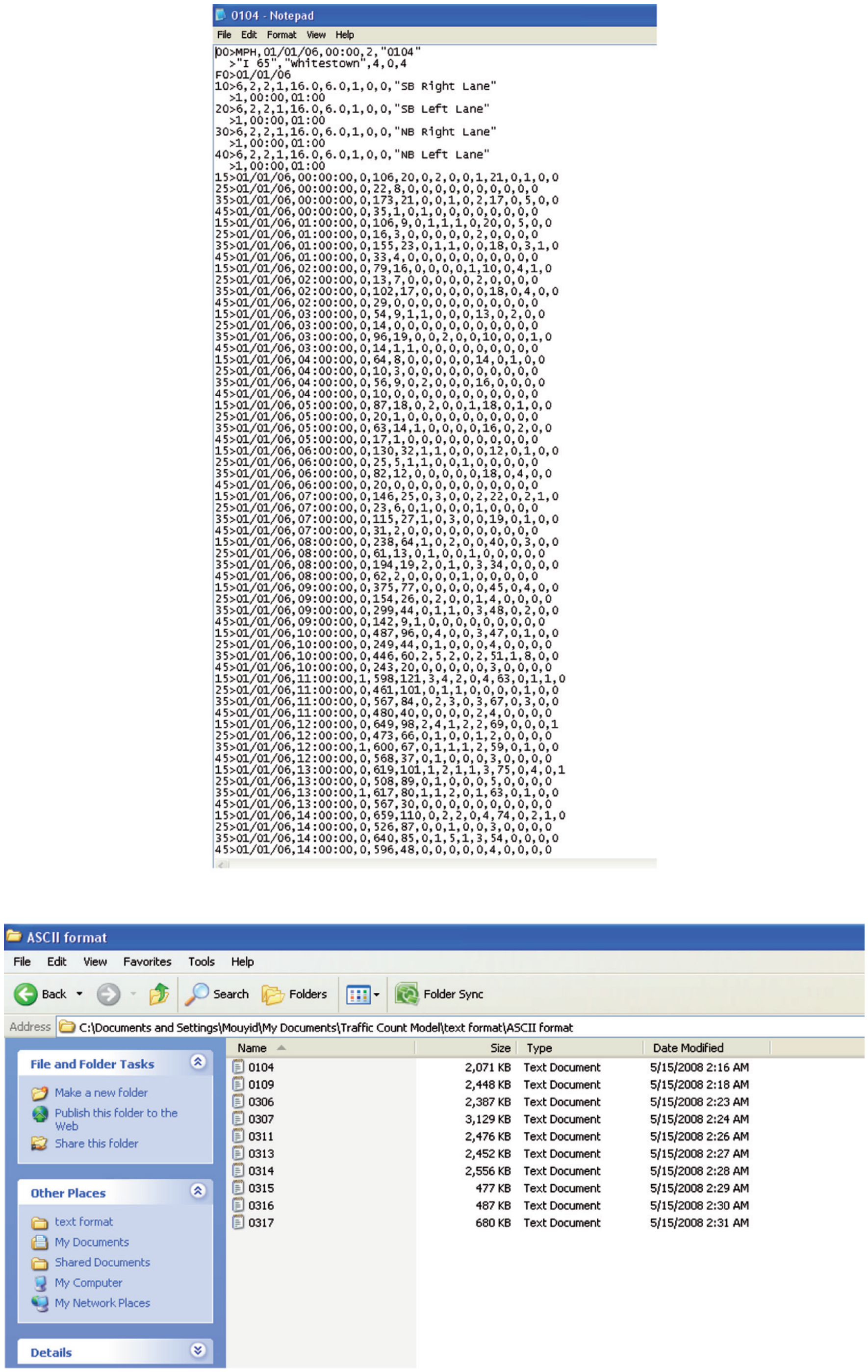


Figure E.4: Output files are set as Old ASCII format no labels

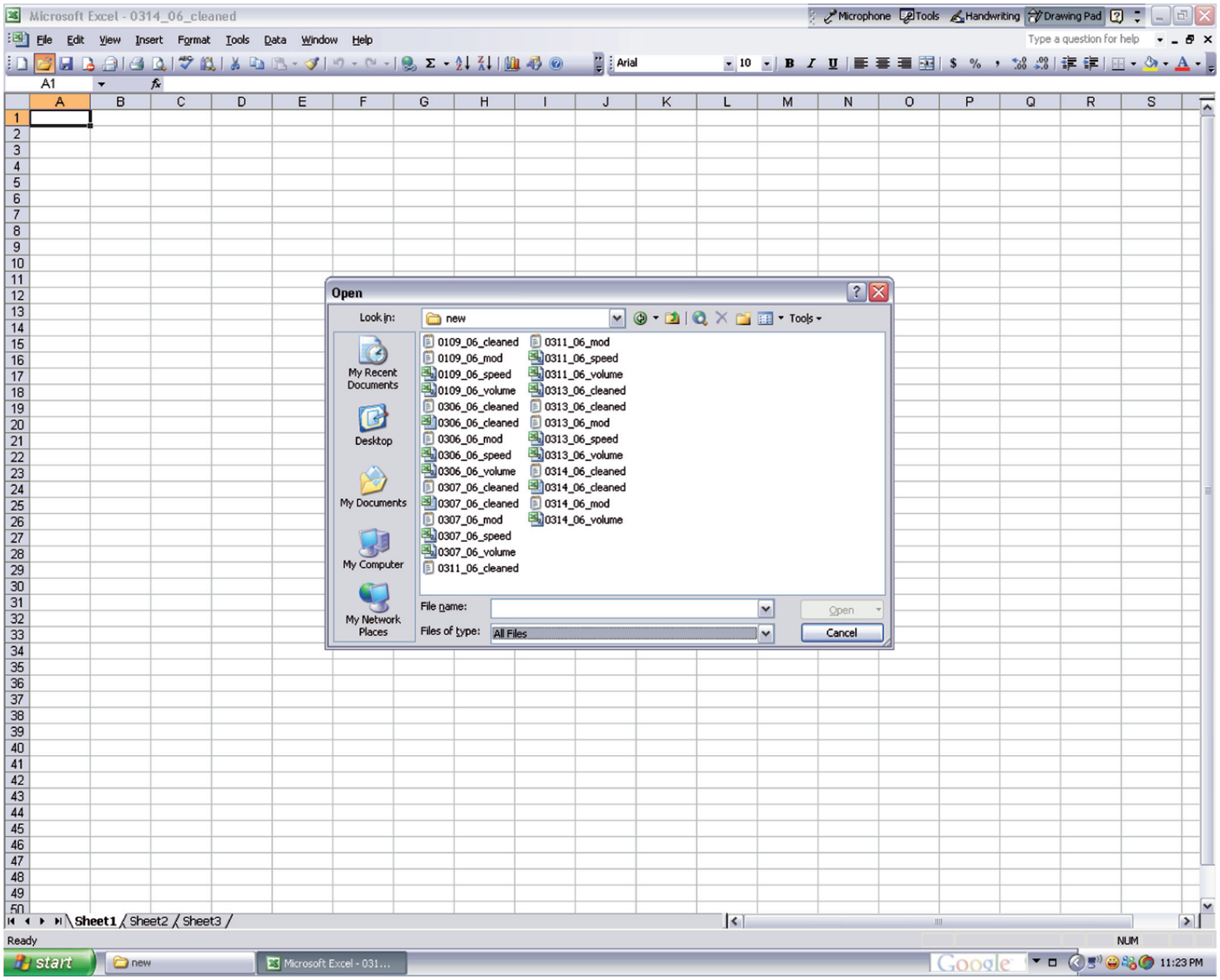


Figure E.5: Text Files are input in Excel

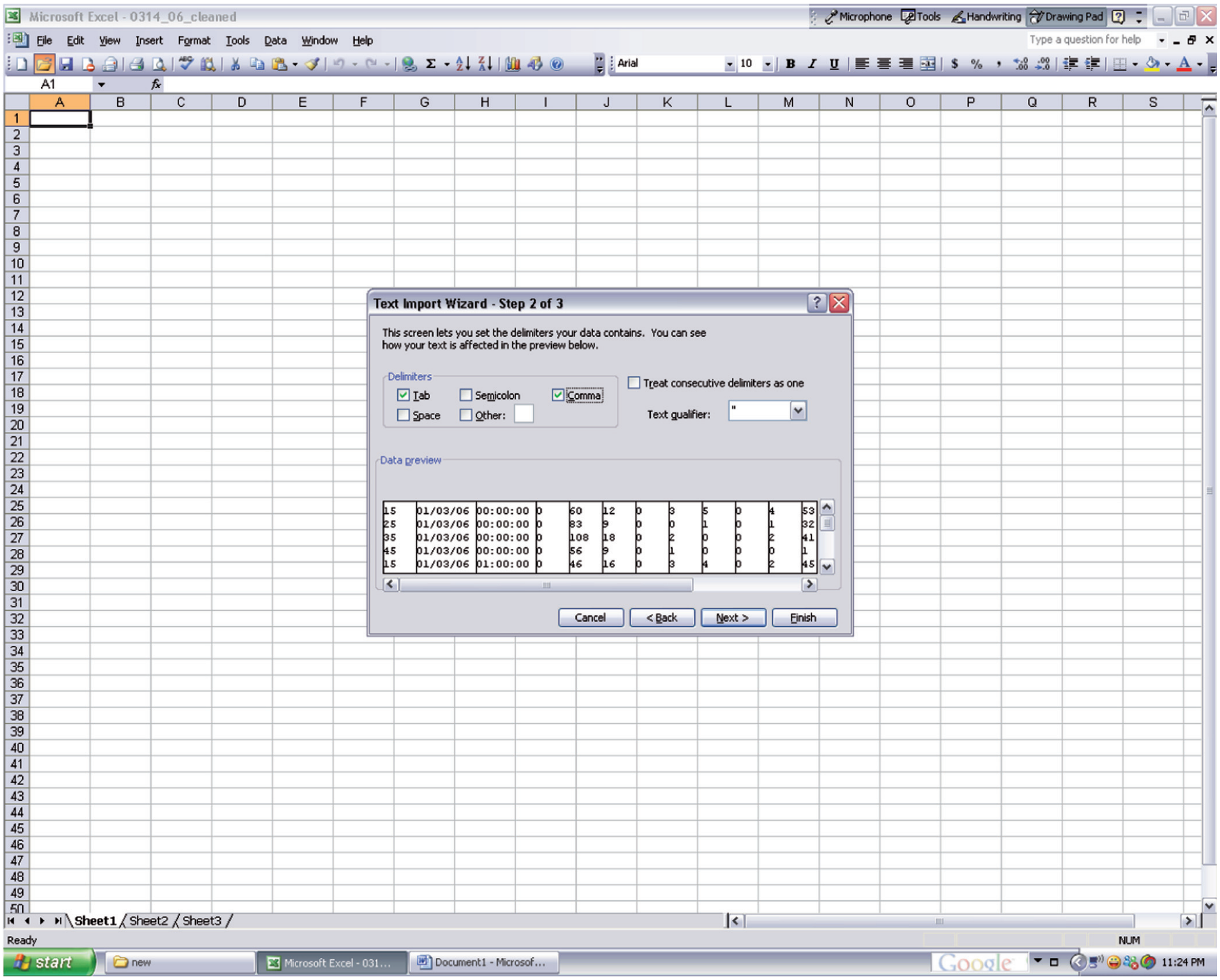


Figure E.6: Text Files are opened in Excel in “tab” and “comma”

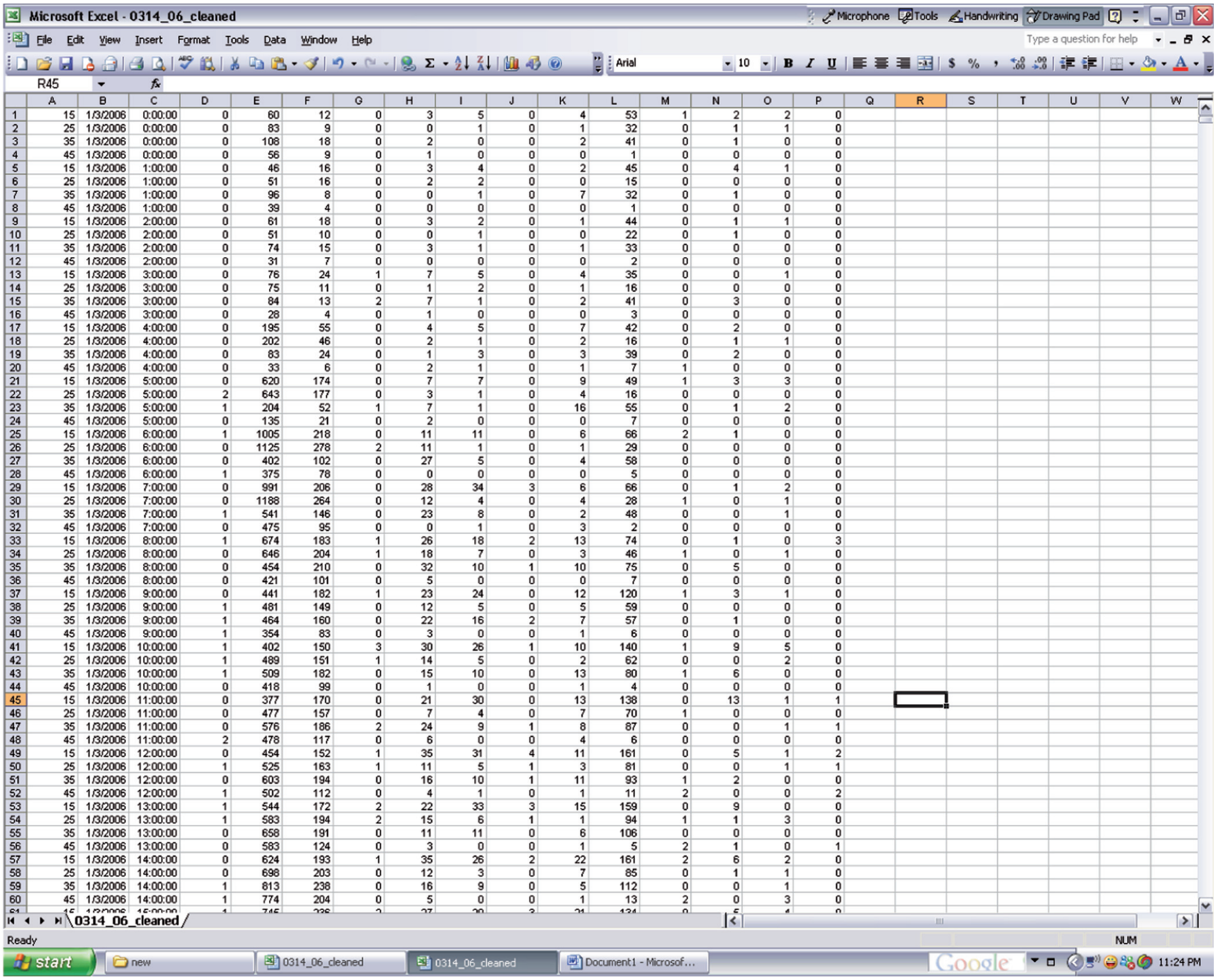
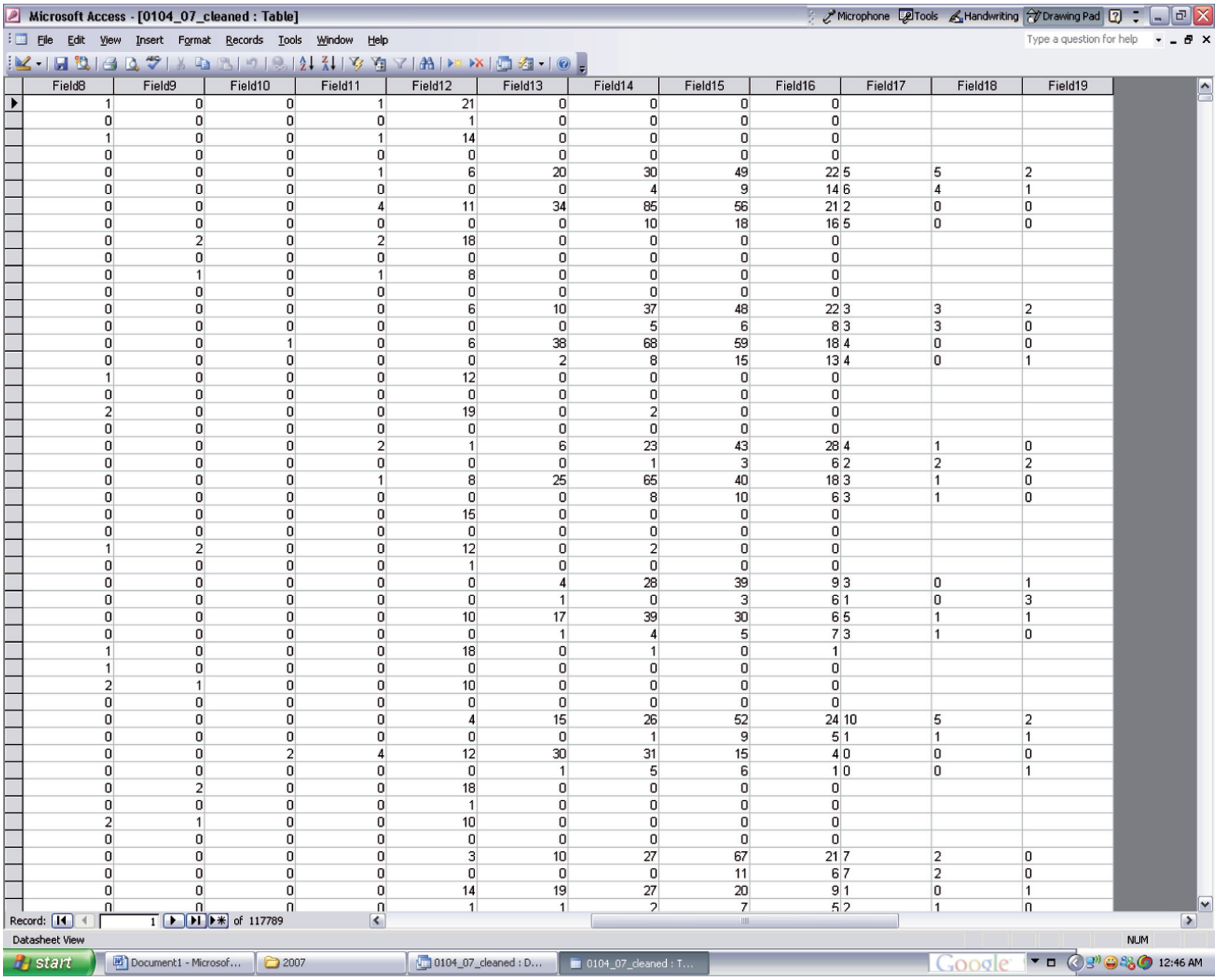


Figure E.7: Opening in Excel after cleaning some symbols on Notepad++



Field8	Field9	Field10	Field11	Field12	Field13	Field14	Field15	Field16	Field17	Field18	Field19
1	0	0	1	21	0	0	0	0			
0	0	0	0	1	0	0	0	0			
1	0	0	1	14	0	0	0	0			
0	0	0	0	0	0	0	0	0			
0	0	0	1	6	20	30	49	22	5	5	2
0	0	0	0	0	0	4	9	14	6	4	1
0	0	0	4	11	34	85	56	21	2	0	0
0	0	0	0	0	0	10	18	16	5	0	0
0	2	0	2	18	0	0	0	0			
0	0	0	0	0	0	0	0	0			
0	1	0	1	8	0	0	0	0			
0	0	0	0	0	0	0	0	0			
0	0	0	0	6	10	37	48	22	3	3	2
0	0	0	0	0	0	5	6	8	3	3	0
0	0	1	0	6	38	68	59	18	4	0	0
0	0	0	0	0	2	8	15	13	4	0	1
1	0	0	0	12	0	0	0	0			
0	0	0	0	0	0	0	0	0			
2	0	0	0	19	0	2	0	0			
0	0	0	0	0	0	0	0	0			
0	0	0	2	1	6	23	43	28	4	1	0
0	0	0	0	0	0	1	3	6	2	2	2
0	0	0	1	8	25	65	40	18	3	1	0
0	0	0	0	0	0	8	10	6	3	1	0
0	0	0	0	15	0	0	0	0			
0	0	0	0	0	0	0	0	0			
1	2	0	0	12	0	2	0	0			
0	0	0	0	1	0	0	0	0			
0	0	0	0	0	4	28	39	9	3	0	1
0	0	0	0	0	1	0	3	6	1	0	3
0	0	0	0	10	17	39	30	6	5	1	1
0	0	0	0	0	1	4	5	7	3	1	0
1	0	0	0	18	0	1	0	1			
1	0	0	0	0	0	0	0	0			
2	1	0	0	10	0	0	0	0			
0	0	0	0	0	0	0	0	0			
0	0	0	0	4	15	26	52	24	10	5	2
0	0	0	0	0	0	1	9	5	1	1	1
0	0	2	4	12	30	31	15	4	0	0	0
0	0	0	0	0	1	5	6	1	0	0	1
0	2	0	0	18	0	0	0	0			
0	0	0	0	1	0	0	0	0			
2	1	0	0	10	0	0	0	0			
0	0	0	0	0	0	0	0	0			
0	0	0	0	3	10	27	67	21	7	2	0
0	0	0	0	0	0	0	11	6	7	2	0
0	0	0	0	14	19	27	20	9	1	0	1
0	0	0	0	1	1	2	7	5	2	1	0

Figure E.8: Separating Speed and Volume Data in Excel

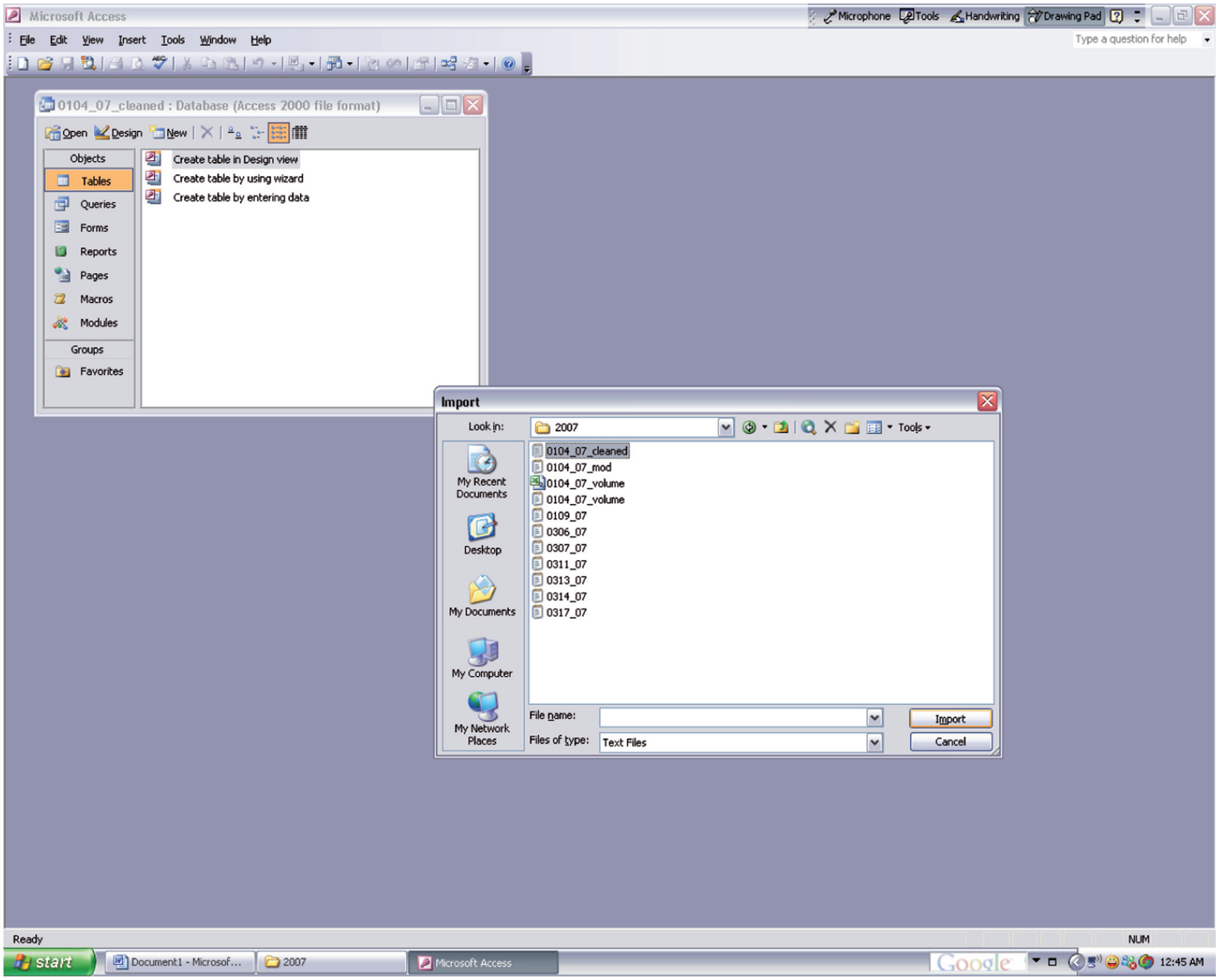


Figure E.9: Storing Data in MS Access

APPENDIX F

Crash Dataset Management:

Source: State Police Database

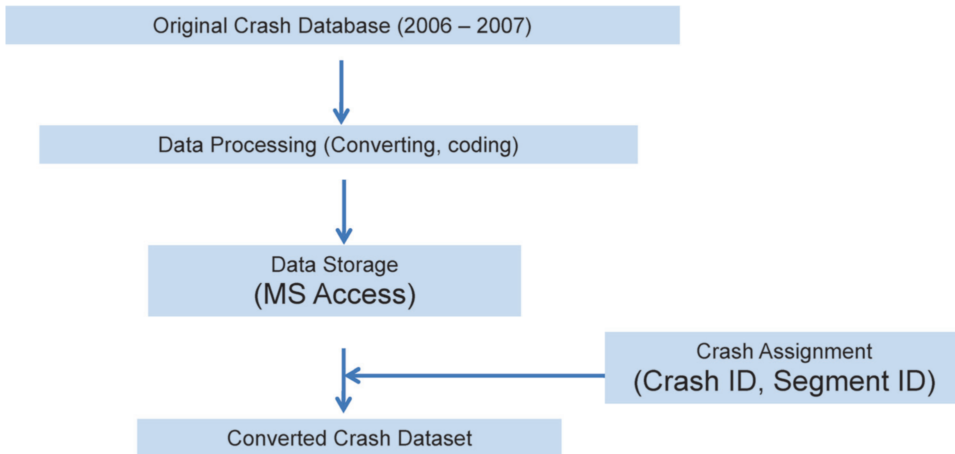


Figure F.1: Crash Assignment

Crash Table	Segment Table
SRDNAME	RDNAME
CPERIOD	Period
STDDIR	TRADIR

Criterion: STMP < LDIST < ENDMP

Figure F.2: Criterion and Linking of Crash and Segment

Traffic Dataset Management:

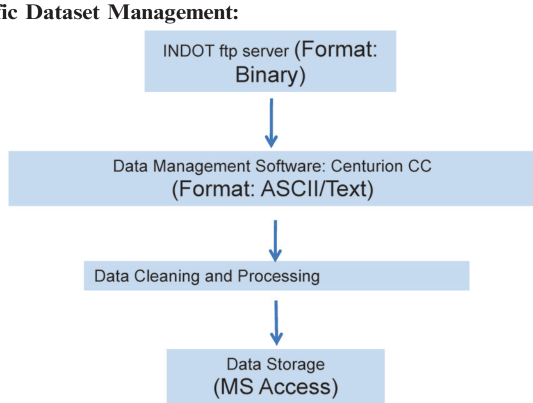


Figure F.3: Traffic Data Management

Segment Geometry Dataset Management:

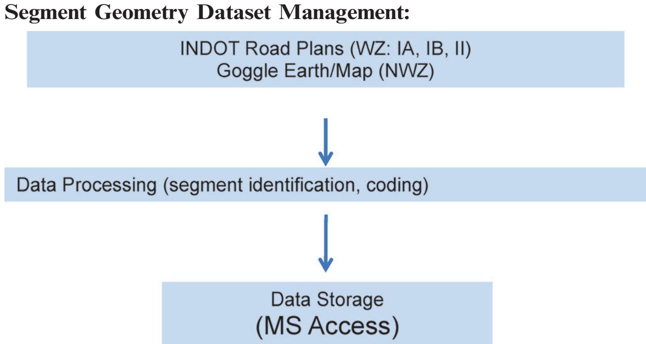


Figure F.4: Geometry Data Management

Traffic Maintenance Dataset Management:

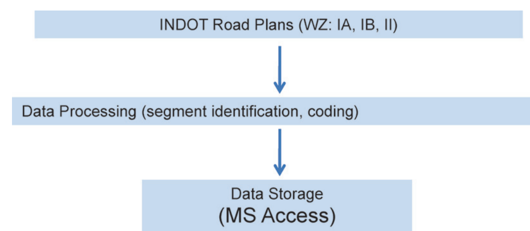


Figure F.5: Maintenance Data Management

Weather Dataset Management:

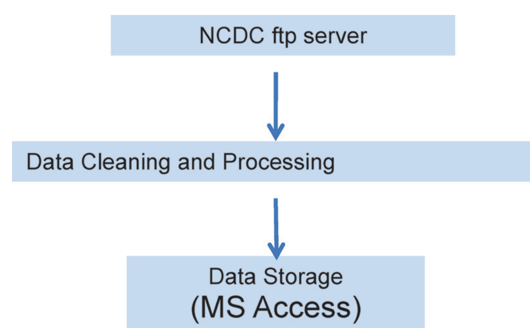


Figure F.6: Weather Data Management

TABLE F.1:
Count Models Development and Imputation

INPUT		Output		
Period	Detector	Detector	Model	Action
Jan – Apr 2006	0311B, 0313	0315	1	Development
Jan – Apr 2006	0311F, 0314	0316	2	Development
May 2006 – Feb 2007 Dec 2007 – Feb 2008	0311B, 0313	0315	1	***Imputation***
May 2006 – Feb 2007 Dec 2007 – Feb 2008	0311F, 0314	0316	2	***Imputation***
Jan – Mar 2008	0311B, 0313	973220F	3	Development
Jan – Mar 2008	0311F, 0314	973220B	4	Development
Jan 2006 – Feb 2007	0311B, 0313	973220F	3	***Imputation***
Jan 2006 – Feb 2007	0311F, 0314	973220B	4	***Imputation***
Sept – Nov 2007	0311F, 0314	973220B	5	Development
Sept – Nov 2007	0311B, 0313	973220F	6	Development
Mar – Aug 2007	0311F, 0314	973220B	5	***Imputation***
Mar – Aug 2007	0311B, 0313	973200F	6	***Imputation***
Sept 2007	973220F	SF/F	7	Development
Sept 2007	973320B	SF/B	8	Development
Jan 2006 – Aug 2007 Oct 2007 – Mar 2008	973320F	SF/F	7	***Imputation***
Jan 2006 – Aug 2007 Oct 2007 – Mar 2008	973220B	SF/B	8	***Imputation***
Mar 2008	973220F	0315	9	Development
Mar 2008	973220B	0316	10	Development
Mar – Nov 2007	973220F	0316	9	***Imputation***
Mar – Nov 2007	973220B	0315	10	***Imputation***
Jan – Dec 2007	0109B, 0104F	3300F	11	Development
Jan – Dec 2007	0104B, 0109F	3300B	12	Development
Jan – Nov 2006	0109B, 0104F	3300F	11	***Imputation***
Jan – Nov 2006	0104B,0109F	3300B	12	***Imputation***

TABLE F.2:
Time Interval used in Count Model

Time Interval (24-hr)	Notation	Logic
00 – 05	B1	>=0 and <6
06 – 09	B2	>=6 and <=9
10 – 12	B3	>=10 and <=12
13 – 15	B4	>=13 and <=15
16 – 19	B5	>=16 and <=19
20 – 24	B6	>=20 and <=24

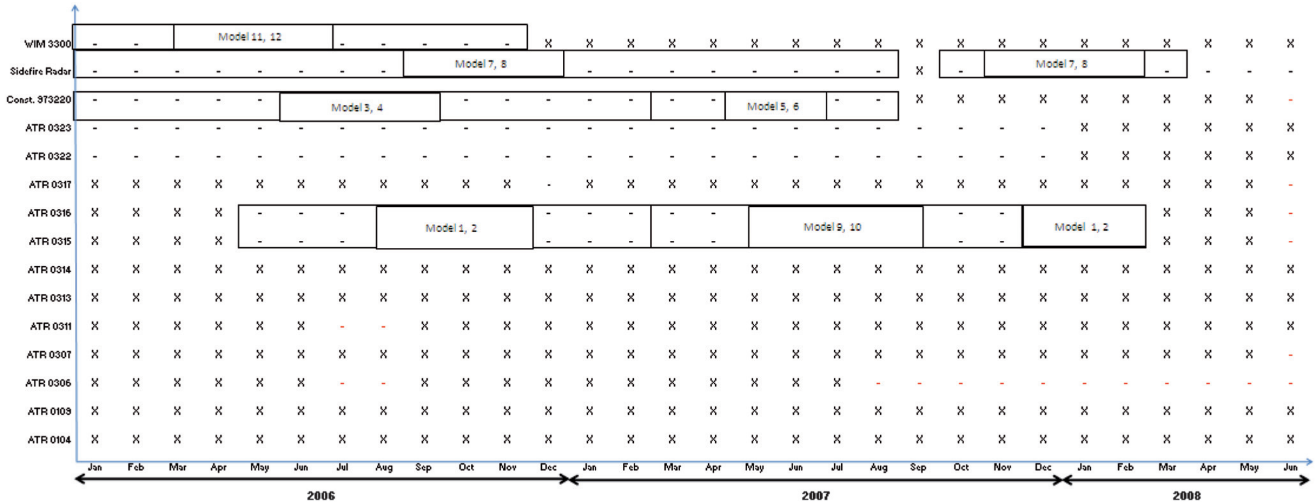


Figure F.7: Count Model for Importation of Missing Traffic Data

TABLE F.3:
Detector Locations with Lane assigned

Detector_ID	Highway	Road Name	Direction	Lane_Designation	Lane_Meaning	Location (MM)	Data
0104	I65	Whitestown	SB	15	Right lane	RP 131 + 90	Volume
			SB	25	Left lane		
			NB	35	Right lane		
			NB	45	Left lane		
0104	I65	Whitestown	SB	13	Right lane	RP 131 + 90	Speed
			SB	23	Left lane		
			NB	33	Right lane		
			NB	43	Left lane		
0109	I74	Eagle Creek	EB	15	Right lane	RP 71 + 80	Volume
			EB	25	Left lane		
			WB	35	Right lane		
			WB	45	Left lane		
0109	I74	Eagle Creek	EB	13	Right lane	RP 71 + 80	Speed
			EB	23	Left lane		
			WB	33	Right lane		
			WB	43	Left lane		
0306	I465	Castleton	SB	15	Right lane	RP 36 + 20	Volume
			SB	25	Middle lane		
			SB	35	Left lane		
			NB	45	Right lane		
			NB	55	Middle lane		
			NB	66	Left lane		
0306	I465	Castleton	SB	13	Right lane	RP 36 + 20	Speed
			SB	23	Middle lane		
			SB	33	Left lane		
			NB	43	Right lane		
			NB	53	Middle lane		
			NB	63	Left lane		
0307	I465	English Avenue	NB	15	Right lane	RP 46 + 30	Volume
			NB	25	Middle lane		
			NB	35	Left lane		
			SB	45	Right lane		
			SB	55	Middle lane		
			SB	66	Left lane		

TABLE F.3
(Continued)

Detector_ID	Highway	Road Name	Direction	Lane_Designation	Lane_Meaning	Location (MM)	Data
0307	I465	English Avenue	NB	13	Right lane	RP 46 + 30	Speed
			NB	23	Middle lane		
			NB	33	Left lane		
			SB	43	Right lane		
			SB	53	Middle lane		
			SB	63	Left lane		
0311	I65	Lafayette	SB	15	Right lane	RP 119 + 70	Volume
			SB	25	Middle lane		
			SB	35	Left lane		
			NB	45	Right lane		
			NB	55	Middle lane		
			NB	65	Left lane		
0311	I65	Lafayette	SB	13	Right lane	RP 119 + 70	Speed
			SB	23	Middle lane		
			SB	33	Left lane		
			NB	43	Right lane		
			NB	53	Middle lane		
			NB	63	Left lane		
0313	I65/I70	South Split	NB	15	Right lane	110 + 60	Volume
			NB	25	Right Middle lane		
			NB	35	Left Middle lane		
			NB	45	Left lane		
0313	I65/I70	South Split	NB	13	Right lane	110 + 60	Speed
			NB	23	Right Middle lane		
			NB	33	Left Middle lane		
			NB	43	Left lane		
0314	I65/I70	South Split	SB	15	Right lane	RP 110 + 50	Volume
			SB	25	Right Center lane		
			SB	35	Left Center lane		
			SB	45	Left lane		
0314	I65/I70	South Split	SB	13	Right lane	RP 110 + 50	Speed
			SB	23	Right Center lane		
			SB	33	Left Center lane		
			SB	43	Left lane		
0315	I70	North Split	EB	15	Right lane	RP 084 + 49	Volume
			EB	25	Right Middle lane		
			EB	35	Middle lane		
			EB	45	Left Middle lane		
			EB	55	Left lane		
0315	I70	North Split	EB	13	Right lane	RP 084 + 49	Speed
			EB	23	Right Middle lane		
			EB	33	Middle lane		
			EB	43	Left Middle lane		
			EB	53	Left lane		
0316	I70	North Split	WB	15	Right lane	RP 084 + 49	Volume
			WB	25	Right Middle lane		
			WB	35	Middle lane		
			WB	45	Left Middle lane		
			WB	55	Left lane		
0316	I70	North Split	WB	13	Right lane	RP 084 + 49	Speed
			WB	23	Right Middle lane		
			WB	33	Middle lane		
			WB	43	Left Middle lane		
			WB	53	Left lane		
0317	I70	Airport	WB	15	Right lane	RP 70 + 60	Volume
			WB	25	Middle lane		
			WB	35	Left lane		
			EB	45	Right		
			EB	55	Middle lane		
			EB	65	Left lane		

TABLE F.3
(Continued)

Detector_ID	Highway	Road Name	Direction	Lane_Designation	Lane_Meaning	Location (MM)	Data
0317	I70	Airport	WB	13	Right lane	RP 70 + 60	Speed
			WB	23	Middle lane		
			WB	33	Left lane		
			EB	43	Right		
			EB	53	Middle lane		
			EB	63	Left lane		
0322	I70	Ritter Avenue	EB	15	Right lane	RP 87 + 00	Volume
			EB	25	Center Right lane		
			EB	35	Center Left lane		
			EB	45	Left lane		
0322	I70	Ritter Avenue	EB	13	Right lane	RP 87 + 00	Speed
			EB	23	Center Right lane		
			EB	33	Center Left lane		
			EB	43	Left lane		
0323	I70	Ritter Avenue	WB	15	Right lane	RP 87 + 00	volume
			WB	25	Right Middle lane		
			WB	35	Left Middle lane		
			WB	45	Left lane		
0323	I70	Ritter Avenue	WB	13	Right lane	RP 87 + 00	Speed
			WB	23	Right Middle lane		
			WB	33	Left Middle lane		
			WB	43	Left lane		
973220	I70	Special	EB	12	EB	RP 88+70	Volume
		Shadeland	EB	22	Changeable EB		
		Avenue	WB	32	Left lane		
			WB	42	Right lane		
			WB	52	Changeable WB		
S125	I70	Emerson Avenue	EB	1	Right	RP 86+50	Volume
			EB	2	Left		
			WB	3	Changeable WB		
			WB	4	Left		
			WB	5	Right		
S125	I70	Emerson Avenue	EB	1	Right	RP 86+50	Speed
			EB	2	Left		
			WB	3	Changeable WB		
			WB	4	Left		
			WB	5	Right		

TABLE F.4:
Transition Interval as Lane Change in Different Segments

Segment	Interval Lane_Added (+1)	Interval Lane_Dropped (-1)
070-F-083.40-083.65-2	10:30	22:00
070-F-083.65-083.90-2	10:30	22:00
070-F-083.90-084.15-2	10:30	22:00
070-F-084.15-084.55-2	10:30	22:00
070-F-084.55-084.72-2	10:30	22:00
070-F-084.72-084.97-2	10:30	22:00
070-F-084.97-085.19-2	10:30	22:00
070-F-085.19-085.44-2	10:30	22:00
070-F-085.44-085.73-2	10:30	22:00
070-F-085.73-085.95-2	10:30	22:00
070-F-085.95-086.32-2	10:30	22:00
070-F-086.32-086.57-2	10:30	22:00
070-F-086.57-086.82-2	10:30	22:00
070-F-086.82-087.07-2	10:00	22:30
070-F-087.07-087.32-2	10:00	22:30
070-F-087.32-087.55-2	10:00	22:30
070-F-087.55-087.80-2	10:00	22:30
070-F-087.80-088.08-2	10:00	22:30
070-F-088.08-088.33-2	10:00	22:30
070-F-088.33-088.58-2	10:00	22:30
070-F-088.58-088.93-2	10:00	22:30
070-F-088.93-089.18-2	10:00	22:30
070-F-089.18-089.35-2	10:00	22:30
070-F-089.35-089.60-2	10:00	22:30
070-F-089.60-089.97-2	10:00	22:30
070-F-089.97-090.22-2	10:00	22:30
070-F-083.40-083.65-4	10:30	22:00
070-F-083.65-083.90-4	10:30	22:00
070-F-083.90-084.15-4	10:30	22:00
070-F-084.15-084.55-4	10:30	22:00
070-F-084.55-084.72-4	10:30	22:00
070-F-084.72-084.97-4	10:30	22:00
070-F-084.97-085.19-4	10:30	22:00
070-F-085.19-085.44-4	10:30	22:00
070-F-085.44-085.73-4	10:30	22:00
070-F-085.73-085.95-4	10:30	22:00
070-F-085.95-086.20-4	10:30	22:00
070-F-086.20-086.45-4	10:30	22:00
070-F-086.45-086.70-4	10:30	22:00
070-F-086.70-086.95-4	10:30	22:00
070-F-086.95-087.20-4	10:00	22:30
070-F-087.20-087.45-4	10:00	22:30
070-F-087.45-087.70-4	10:00	22:30
070-F-087.70-087.95-4	10:00	22:30
070-F-087.95-088.20-4	10:00	22:30
070-F-088.20-088.37-4	10:00	22:30
070-F-088.37-088.62-4	10:00	22:30
070-F-088.62-088.87-4	10:00	22:30
070-F-088.93-089.18-4	10:00	22:30
070-F-089.18-089.47-4	10:00	22:30
070-F-089.47-089.60-4	10:00	22:30
070-F-089.60-089.85-4	10:00	22:30
070-F-089.85-090.22-4	10:00	22:30
070-B-083.65-083.40-2	22:00	10:30
070-B-083.90-083.65-2	22:00	10:30
070-B-084.15-083.90-2	22:00	10:30
070-B-084.40-084.15-2	22:00	10:30
070-B-084.72-084.40-2	22:00	10:30
070-B-085.02-084.72-2	22:00	10:30
070-B-085.19-085.02-2	22:00	10:30

TABLE F.4
(Continued)

Segment	Interval Lane_Added (+1)	Interval Lane_Dropped (-1)
070-B-085.44-085.19-2	22:00	10:30
070-B-085.72-085.44-2	22:00	10:30
070-B-085.95-085.72-2	22:00	10:30
070-B-086.20-085.95-2	22:00	10:30
070-B-086.45-086.20-2	22:00	10:30
070-B-086.70-086.45-2	22:00	10:30
070-B-086.87-086.70-2	22:00	10:30
070-B-087.12-086.87-2	22:30	10:00
070-B-087.37-087.12-2	22:30	10:00
070-B-087.55-087.37-2	22:30	10:00
070-B-087.80-087.55-2	22:30	10:00
070-B-088.08-087.80-2	22:30	10:00
070-B-088.29-088.08-2	22:30	10:00
070-B-088.54-088.29-2	22:30	10:00
070-B-088.79-088.54-2	22:30	10:00
070-B-088.87-088.79-2	22:30	10:00
070-B-088.92-088.87-2	22:30	10:00
070-B-089.20-088.92-2	22:30	10:00
070-B-089.45-089.20-2	22:30	10:00
070-B-089.60-089.45-2	22:30	10:00
070-B-089.85-089.60-2	22:30	10:00
070-B-090.10-089.85-2	22:30	10:00
070-B-083.65-083.40-4	22:00	10:30
070-B-083.90-083.65-4	22:00	10:30
070-B-084.15-083.90-4	22:00	10:30
070-B-084.40-084.15-4	22:00	10:30
070-B-084.72-084.40-4	22:00	10:30
070-B-085.02-084.72-4	22:00	10:30
070-B-085.19-085.02-4	22:00	10:30
070-B-085.44-085.19-4	22:00	10:30
070-B-085.72-085.44-4	22:00	10:30
070-B-085.95-085.72-4	22:00	10:30
070-B-086.20-085.95-4	22:00	10:30
070-B-086.45-086.20-4	22:00	10:30
070-B-086.70-086.45-4	22:00	10:30
070-B-086.87-086.70-4	22:30	10:00
070-B-087.12-086.87-4	22:30	10:00
070-B-087.37-087.12-4	22:30	10:00
070-B-087.55-087.37-4	22:30	10:00
070-B-087.80-087.55-4	22:30	10:00
070-B-088.08-087.80-4	22:30	10:00
070-B-088.36-088.08-4	22:30	10:00
070-B-088.54-088.36-4	22:30	10:00
070-B-088.79-088.54-4	22:30	10:00
070-B-088.87-088.79-4	22:30	10:00
070-B-088.92-088.87-4	22:30	10:00
070-B-089.02-088.92-4	22:30	10:00
070-B-089.20-089.02-4	22:30	10:00
070-B-089.45-089.20-4	22:30	10:00
070-B-089.60-089.45-4	22:30	10:00
070-B-089.85-089.60-4	22:30	10:00
070-B-090.10-089.85-4	22:30	10:00

APPENDIX G

TABLE G.1:
Average Daily Traffic for Detectors and Segments

Detector ID	Segment	Group	Average Daily Traffic (ADT) (Time Range of Traffic Observation)		
			2006 (12 months)	2007 (12 months)	2008 (6 months)
0306	Segment: 8	I-465 ring	2340	2330	-
0307	Segment: 10	I-465 ring	1832	2010	1880
3300	Segment:13	I-465 ring	1796	2033	1850
0311	Segment:3	Inside I-465 ring	1173	1187	1306
0313, 0314	Segment:4	Inside I-465 ring	2534	2209	2607
0104	Segment:1	Outside I-465 ring	1179	1143	1144
0109	Segment:20	Outside I-465 ring	680	643	636

TABLE G.2:
Average Daily Traffic for Detectors and Segments

Detector ID	Segment	Group	Average Daily Traffic (ADT) for Heavy Vehicles [HEVVEH] (Time Range of Traffic Observation)		
			2006 (12 months)	2007 (12 months)	2008 (6 months)
0306	Segment: 8	I-465 ring	133	133	-
0307	Segment: 10	I-465 ring	196	298	226
3300	Segment:13	I-465 ring	147	252	210
0311	Segment:3	Inside I-465 ring	130	103	166
0313, 0314	Segment:4	Inside I-465 ring	275	155	622
0104	Segment:1	Outside I-465 ring	263	234	274
0109	Segment:20	Outside I-465 ring	120	114	87

APPENDIX H

Crash Probability Model P(C)

The SAS System 15:30 Monday, November 9, 2009 134

The LOGISTIC Procedure

Model Information

Data Set	ALL.SAMPLE2B
Response Variable	CRASH
Number of Response Levels	2
Model	binary logit
Optimization Technique	Fisher's scoring

Number of Observations Read	466483
Number of Observations Used	466483

Response Profile

Ordered Value	CRASH	Total Frequency
1	1	1403
2	0	465080

Probability modeled is CRASH='1'.

Model Convergence Status

Convergence criterion (GCONV=1E-8) satisfied.

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	19097.120	17772.179
SC	19108.173	18181.139
-2 Log L	19095.120	17698.179

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	1396.9414	36	<.0001
Score	1531.8066	36	<.0001
Wald	1242.7278	36	<.0001

The LOGISTIC Procedure

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-5.2695	0.2778	359.7679	<.0001
Volume	1	0.000239	0.000027	77.8800	<.0001
HVVol	1	-0.00114	0.000281	16.3476	<.0001
HVVol*WORKZONE	1	0.00533	0.00227	5.5071	0.0189
RSEGLEN	1	3.2997	0.3416	93.2791	<.0001
WORKZONE*RSEGLEN	1	-1.7903	0.9312	3.6963	0.0545
CURV_SEG_N	1	0.0848	0.0463	3.3530	0.0671
L5	1	0.5458	0.1000	29.7747	<.0001
IN_SHLD_W	1	-0.0300	0.0102	8.5477	0.0035
OUT_SHLD_W	1	-0.0534	0.0108	24.2293	<.0001
DONRAMP	1	0.1326	0.0736	3.2463	0.0716
DOFFRAMP	1	0.1503	0.0736	4.1722	0.0411
BTEMP	1	0.3003	0.0773	15.0775	0.0001
BTEMPMISS	1	-0.4363	0.1402	9.6848	0.0019
VISIBILITY	1	-0.00003	8.604E-6	13.0776	0.0003
BPRECIP	1	0.2766	0.0970	8.1303	0.0044
BHEAVY	1	-1.2529	0.7137	3.0823	0.0792
T2	1	-0.8039	0.2571	9.7764	0.0018
T3	1	-0.8394	0.2652	10.0138	0.0016
T4	1	-0.6299	0.2375	7.0325	0.0080
T5	1	-0.3756	0.2081	3.2564	0.0711
T7	1	0.3442	0.1208	8.1144	0.0044
T8	1	0.6271	0.1089	33.1476	<.0001
T9	1	0.4592	0.1137	16.3059	<.0001
T16	1	0.3778	0.1161	10.5944	0.0011
T17	1	0.4701	0.1137	17.0820	<.0001
T18	1	0.6656	0.1068	38.8214	<.0001
DW1	1	-0.5086	0.0926	30.1520	<.0001
DW4	1	-0.3084	0.0855	13.0161	0.0003
DW5	1	-0.2111	0.0815	6.7149	0.0096
DW6	1	-0.5911	0.1019	33.6665	<.0001
DW7	1	-0.8197	0.1156	50.3107	<.0001
WORKZONE*DW4	1	0.5218	0.2472	4.4573	0.0348
WORKZONE*DW5	1	0.6222	0.2250	7.6464	0.0057
I65	1	-0.4452	0.0825	29.0970	<.0001
I70	1	-0.9776	0.1460	44.8669	<.0001
WORKZONE	1	-1.0914	0.3192	11.6934	0.0006

Odds Ratio Estimates

Effect	Point Estimate	95% Wald Confidence Limits	
Volume	1.000	1.000	1.000
CURV_SEG_N	1.088	0.994	1.192

The LOGISTIC Procedure

Odds Ratio Estimates

Effect	Point Estimate	95% Wald Confidence Limits	
L5	1.726	1.419	2.100
IN_SHLD_W	0.970	0.951	0.990
OUT_SHLD_W	0.948	0.928	0.968
DONRAMP	1.142	0.988	1.319
DOFFRAMP	1.162	1.006	1.342
BTEMP	1.350	1.160	1.571
BTEMPMISS	0.646	0.491	0.851
VISIBILITY	1.000	1.000	1.000
BPRECIP	1.319	1.090	1.595
BHEAVY	0.286	0.071	1.157
T2	0.448	0.270	0.741
T3	0.432	0.257	0.727
T4	0.533	0.334	0.848
T5	0.687	0.457	1.033
T7	1.411	1.113	1.788
T8	1.872	1.512	2.318
T9	1.583	1.267	1.978
T16	1.459	1.162	1.832
T17	1.600	1.280	2.000
T18	1.946	1.578	2.399
DW1	0.601	0.501	0.721
DW6	0.554	0.454	0.676
DW7	0.441	0.351	0.553
I65	0.641	0.545	0.753
I70	0.376	0.283	0.501

Association of Predicted Probabilities and Observed Responses

Percent Concordant	67.1	Somers' D	0.507
Percent Discordant	16.4	Gamma	0.608
Percent Tied	16.5	Tau-a	0.003
Pairs	652507240	c	0.754

APPENDIX I

Single-vehicle Crash Probability Model P(C1|C)

The SAS System 15:30 Monday, November 9, 2009 137

The LOGISTIC Procedure

Model Information

Data Set	ALL.SAMPLE2C
Response Variable	SINGVEH1
Number of Response Levels	2
Model	binary logit
Optimization Technique	Fisher's scoring

Number of Observations Read	1403
Number of Observations Used	1403

Response Profile

Ordered Value	SINGVEH1	Total Frequency
1	1	373
2	0	1030

Probability modeled is SINGVEH1=1.

Model Convergence Status

Convergence criterion (GCONV=1E-8) satisfied.

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	1626.944	1448.032
SC	1632.191	1500.496
-2 Log L	1624.944	1428.032

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	196.9124	9	<.0001
Score	179.7519	9	<.0001
Wald	153.0915	9	<.0001

The LOGISTIC Procedure

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-0.3434	0.2887	1.4147	0.2343
Volume	1	-0.00050	0.000052	93.1997	<.0001
WORKZONE*HVVol	1	-0.0562	0.0265	4.4873	0.0341
CURV_SEG_N	1	0.1769	0.1049	2.8417	0.0918
OUT_SHLD_W	1	0.0424	0.0216	3.8505	0.0497
BPRECIP	1	0.4623	0.1753	6.9503	0.0084
BTEMP	1	0.3734	0.1702	4.8135	0.0282
I70	1	1.1970	0.2922	16.7875	<.0001
PERIOD1	1	-0.3294	0.1433	5.2829	0.0215
WORKZONE	1	-0.9970	0.5964	2.7951	0.0946

Odds Ratio Estimates

Effect	Point Estimate	95% Wald Confidence Limits	
Volume	1.000	0.999	1.000
CURV_SEG_N	1.193	0.972	1.466
OUT_SHLD_W	1.043	1.000	1.088
BPRECIP	1.588	1.126	2.239
BTEMP	1.453	1.041	2.028
I70	3.310	1.867	5.869
PERIOD1	0.719	0.543	0.953

Association of Predicted Probabilities and Observed Responses

Percent Concordant	73.7	Somers' D	0.477
Percent Discordant	26.0	Gamma	0.479
Percent Tied	0.3	Tau-a	0.186
Pairs	384190	c	0.739

APPENDIX J

Severe Single-vehicle Crash Probability
Joint Model for P(SC1|C1) and P(SC2|C2

The LOGISTIC Procedure

Model Information

Data Set	ALL.SAMPLE2C
Response Variable	INJURY
Number of Response Levels	2
Model	binary logit
Optimization Technique	Fisher's scoring

Number of Observations Read	1403
Number of Observations Used	1403

Response Profile

Ordered Value	INJURY	Total Frequency
1	1	214
2	0	1189

Probability modeled is INJURY='1'.

Model Convergence Status

Convergence criterion (GCONV=1E-8) satisfied.

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	1200.367	1165.084
SC	1205.614	1212.301
-2 Log L	1198.367	1147.084

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	51.2834	8	<.0001
Score	55.1773	8	<.0001
Wald	50.1748	8	<.0001

The LOGISTIC Procedure

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-1.3431	0.2915	21.2344	<.0001
SINGVEH1	1	-0.3950	0.2651	2.2201	0.1362
MULTVEH1*Volume	1	-0.00023	0.000067	11.5376	0.0007
IN_SHLD_W	1	0.0323	0.0172	3.5124	0.0609
OUT_SHLD_W	1	-0.0419	0.0224	3.4896	0.0618
SINGVEH1*PERIOD2	1	0.5923	0.3203	3.4192	0.0644
SINGVEH1*DW1	1	1.1209	0.3765	8.8625	0.0029
DW5	1	0.5088	0.1862	7.4696	0.0063
SINGVEH1*DW6	1	0.9667	0.3425	7.9675	0.0048

Odds Ratio Estimates

Effect	Point Estimate	95% Wald Confidence Limits	
IN_SHLD_W	1.033	0.999	1.068
OUT_SHLD_W	0.959	0.918	1.002
DW5	1.663	1.155	2.396

Association of Predicted Probabilities and Observed Responses

Percent Concordant	63.3	Somers' D	0.277
Percent Discordant	35.6	Gamma	0.280
Percent Tied	1.0	Tau-a	0.072
Pairs	254446	c	0.639