# APPLICATION OF MICROSCOPIC SIMULATION TO EVALUATE THE SAFETY PERFORMANCE OF FREEWAY WEAVING SECTIONS

A Thesis

by

THANH QUANG LE

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

December 2009

Major Subject: Civil Engineering

# APPLICATION OF MICROSCOPIC SIMULATION TO EVALUATE THE SAFETY PERFORMANCE OF FREEWAY WEAVING SECTIONS

A Thesis

by

THANH QUANG LE

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

Approved by:

| | |
|---|---|
| Chair of Committee, | Dominique Lord |
| Committee Members, | Thomas Wehrly |
| | Yunlong Zhang |
| Head of Department, | John Niedzwecki |

December 2009

Major Subject: Civil Engineering

## ABSTRACT

Application of Microscopic Simulation to Evaluate the Safety Performance of Freeway Weaving Sections. (December 2009)

Thanh Quang Le, B.S, Hanoi University of Communications and Transport

Chair of Advisory Committee: Dr. Dominique Lord

This study adopted the traffic conflict technique, investigated and applied it for evaluation of freeway weaving section safety performance. Conflicts between vehicles were identified based on the state of interactions between vehicles in the traffic stream at the microscopic level. VISSIM microscopic simulation model was employed to simulate traffic operation. Surrogate safety measures were formulated based on the deceleration rate required to avoid crashes and these simulation-based measures were statistically compared and validated using crash data collected from the same study site. Three study sites located in the Houston and Dallas areas were selected. Geometric and traffic data were collected using various techniques including the use of traffic surveillance cameras and pneumatic tubes. The study revealed the existence of links between actually observed crashes and the surrogate safety measures. The study findings support the possible the use of microscopic simulation to evaluate the safety performance of weaving areas and other transportation facilities.

# DEDICATION

This thesis is dedicated to my late grandfather, who inspired me and taught me the importance of education; to my father, my mother, my sister, the rest of my family, and my friends who have always supported me.

## ACKNOWLEDGEMENTS

I would like to thank my committee chair, Dr. Dominique Lord, for his patience, guidance and support. I also thank my committee members, Dr. Yunlong Zhang and Dr. Thomas Wehrly, for serving on my committee and providing me with valuable suggestions throughout the course of this research.

I would like to extend my special thanks to Dr. Kay Fitzpatrick, Dr. Richard Jon Porter, and Dr. Chi-Leung Chu, of the Texas Transportation Institute, who have always been helpful and supportive. I would not have been able to complete this thesis without them. I also want to thank Dr. Geza Pesti, Dr. Eun Sug Park and Mr. Lance Bullard for their kind support and helpful advice.

Thanks also go to my friends and colleagues at the Texas Transportation Institute and the department faculty and staff for making my time at Texas A&M University a great experience.

Finally, thanks to my mother, my father and my sister for their encouragement and support.

# NOMENCLATURE

| | |
|---|---|
| ADT | Average Daily Traffic |
| AASHTO | American Association of State Highway and Transportation Officials |
| CI | Crash Index |
| CMF | Crash Modification Factor |
| CPI | Crash Potential Index |
| CRIS | Crash Record Information System |
| DR | Deceleration Rate |
| DRAC | Deceleration Rate to Avoid Crash |
| HCM | Highway Capacity Manual |
| HOV | High Occupancy Vehicle (lane) |
| IH | Interstate Highway |
| LOS | Level of Service |
| MADR | Maximum Available Deceleration Rate |
| MPH | Miles per hour |
| MTTC | Modified Time to Collision |
| MUTCD | Manual on Uniform Traffic Control Devices |
| NCHRP | National Cooperative Highway Research Program |
| PHF | Peak Hour Factor |
| SH | (Texas) State Highways |

| | |
|---|---|
| TTC | Time to Collision |
| TTI | Texas Transportation Institute |
| TxDOT | Texas Department of Transportation |
| US | United States Numbered Highways |

## TABLE OF CONTENTS

## LIST OF FIGURES

FIGURE Page

FIGURE Page

FIGURE Page

# LIST OF TABLES

TABLE Page

TABLE Page

TABLE Page

# CHAPTER I

# INTRODUCTION

## 1.1 Background and problem statement

A weaving section is defined as a highway segment where the pattern of traffic entering and leaving at contiguous points of access results in vehicle paths crossing each other. One of the most common types of weaving sections is between an entrance ramp, followed by an exit ramp and connected by an auxiliary lane (HCM, 2000; AASHTO, 2004). On freeways, problems are most likely to happen in the vicinity of interchanges and weaving sections (Rosenbaum, et al., 1982; Lunenfeld, 1993). Traffic conditions at weaving areas are characterized by intense lane-changing maneuvers. Because of complex vehicle interactions, traffic flow turbulences may cause operational problems and potential safety issues at weaving areas (Wynn, 1946; HCM, 1950; Fazio, et al., 1993).

Operational characteristics of weaving sections were first investigated in 1940s and a weaving design and analysis procedure was first incorporated in the 1950 version of the Highway Capacity Manual (Wynn, 1946; HCM, 1950). After this first release, weaving analysis methodologies have been developed and included in the succeeding editions of the Highway Capacity Manual (HCM, 1965; 1985; 2000). The latest version released in 2000 came with a relatively comprehensive analysis procedure for freeway weaving operation. Although safety is always considered as one of the fundamental

---

This thesis follows the style of Accident Analysis & Prevention.

indicators of operational performance of any transportation facility, few studies have focused on the safety performance of weaving sections.

## 1.2 Study objective

This study focuses on investigating a particular type of freeway weaving section with an entrance ramp followed by an exit ramp, connected by an auxiliary lane (as shown in Figure 1-1). This type of weaving section is categorized as type A weaving configuration in the latest and current Highway Capacity Manual (HCM, 2000). By analyzing the safety performance of the freeway weaving areas from microscopic simulation approach, this study provides more understanding on weaving area safety, which reinforces the way leading to an effective, inexpensive and convenient tool for safety analysis and evaluation of the facilities that are going to be built and an alternative for analysis of facilities that have been built and put into operation.

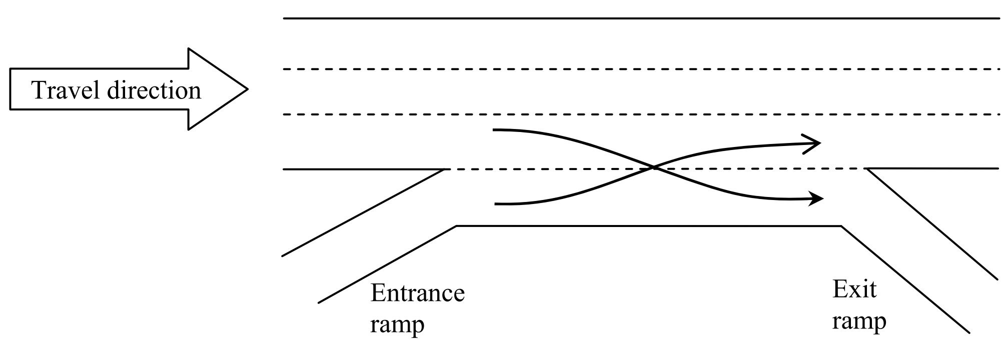



**Fig. 1-1 Layout of type A weaving section (entrance ramp followed by an exit ramp)**

### 1.3 Thesis organization

This thesis is composed of six chapters. After this introduction chapter, the next five chapters are structured as the followings:

Chapter II includes a thorough review of previous studies on weaving sections, and safety of weaving sections. Review of literature on application of simulation in traffic safety is also provided this chapter.

Chapter III contains details about the data collection process, and data reduction. It includes traffic and geometric data collection from the selected study sites using video cameras and pneumatic tubes. Extraction of crash data from the CRIS database is also described in this chapter.

Chapter IV has details of the traffic conflict technique and formulation of the surrogate safety measures. This chapter also provides information on simulation model, descriptions of simulation model calibration and simulation run.

Chapter V has detailed information on data analysis and discussion of the analysis results.

Chapter VI concludes the study findings and provides recommendations for future studies.

# CHAPTER II

# LITERATURE REVIEW

This chapter provides a review of previous studies on weaving sections, weaving section safety and application of simulation in traffic safety analysis. The chapter is organized in three parts. A review of studies on weaving sections, operational characteristics and analysis is presented in the first section. The second section is a review of previous studies on safety of weaving sections. The third part has brief descriptions of studies on the application of simulation to analyze and evaluate the safety performance of roadway facilities. A brief review of previous studies on traffic simulation is also provided in this second section.

## 2.1. Weaving sections, operational characteristics and analysis

One of the first studies on weaving traffic documented was conducted by F. Wynn of Yale University (Wynn, 1946). By using the photographic method to capture vehicle behavior in two weaving areas on Henry Hudson Parkway in New York City, the study revealed operational characteristics of traffic in weaving areas. The author looked into and quantified lengths of weave, speed of weaving vehicles, weaving time, lateral movement, speed differentials, vehicle spacing, and overtaking time. Another early weaving area study was conducted by Fisher and published in 1949 (Fisher, 1949). This study found that the angle of approach, the width, and the length were three main elements of a weaving section that could affect the overall weaving section operation. The study suggested that uniform operating speeds cannot be achieved unless these three main elements are properly proportioned and any change in one of the three would have an effect on the speed, and reduce capacity. However, this study did not have enough data to recommend proper combination of the aforesaid weaving section elements for various traffic conditions. After these early studies, a weaving design and analysis methodology was incorporated in the Highway Capacity Manual, for the first time, in 1950, based on data collected at six weaving areas located in Washington DC and Arlington, Virginia (HCM, 1950).

The next version of Highway Capacity Manual released in 1965 (HCM, 1965) included an improved methodology which was result of a number of published and unpublished research efforts after the 1950 release. The revised version included further graphically developed charts to determine the operational characteristics of weaving sections. This version of the HCM also recognized that performance of a weaving area depends on its geometric characteristics as well as the traffic composition. Furthermore, the 1965 HCM suggested providing one more lane at the weaving areas to satisfactorily serve the combined traffic with complex weaving maneuvers from two different approaches. Analysis methods for multiple weaving sections were also included in this release of the HCM.

Although the analysis methodology for weaving sections incorporated in the 1965 version of the HCM was significantly improved from the one included in the previous HCM, it was found to have problems that led to inadequate results in a study published in 1974 (Roess, et al., 1974). This study concluded that the 1965 HCM procedure was inaccurate in level of service prediction because it did not include lane configuration into the procedure. A modified procedure which took this into account and avoided the aforementioned problems was proposed in this research.

New analysis and design procedures for weaving sections were introduced in an NCHRP report published in 1975 (Pignataro, et al., 1975a). Field data collected at 14 different sites were used to analyze and evaluate the 1965 HCM weaving analysis procedure. This study included geometric characteristics, traffic composition, and traffic volumes on the main line freeway as well as the number of weaving vehicles into the analysis procedures. The study found several fundamentals of the 1965 HCM weaving analysis techniques were not accurate. This study also noted that the geometric configuration of the weaving sections was a vital design factor. Consecutively, the authors proposed a number of fundamental changes to the weaving analysis techniques. This NCHRP report was followed up by a paper published in the same year by the same group of researchers (Pignataro, et al., 1975b). In this paper, an analysis and design procedure for weaving areas was recommended. The procedure was the outcome of the above-mentioned NCHRP study.

In the following years, a number of researchers spent significant efforts in investigating the weaving traffic characteristics and improving techniques for analysis and design of weaving sections (Leisch, 1979; Leisch, 1983; Borchardt, et al., 1984; Reilly, et al., 1984). A revised weaving analysis methodology was developed from these studies and incorporated in the next version of Highway Capacity Manual released in

1985 (HCM, 1985; Roess, 1987). The procedures presented in this version of the HCM had complete definitions of different types of weaving configurations. And such weaving types were defined by number of lane changes that a vehicle had to make to complete a weaving maneuver. This is also the principle of the analysis procedures incorporated in the current version of the HCM (HCM, 2000). Furthermore, in the 1985 HCM, LOS (level of service) criteria were established.

After the release of the 1985 HCM, a lot of effort was devoted to investigating weaving area, evaluating the analysis techniques presented in this HCM and seeking for better solutions for weaving analysis (Fazio and Rouphail., 1986; Skabardonis, et al., 1989; Cassidy, et al., 1989; Cassidy and May., 1991; Ostrom, et al., 1993). The 1985 MCH methodology was found to have shortcomings and one of those was underestimating speeds while the principle of this method relied on predicting weaving and non-weaving vehicle speeds. An update for the HCM weaving analysis procedures was released in 1997. However, the speed underestimation problem still existed. An improved methodology developed by Roess and Ulerio (2000) based on previously existing methods, was published in 2000 and was later incorporated in the 2000 HCM (HCM, 2000).

## 2.2 Safety of weaving section

With the release of the 2000 HCM, a relatively comprehensive method to analyze freeway weaving sections is available, although a number of limitations still existed. It provides prediction models for both weaving capacity and level of service. Despite the fact that weaving sections have long been investigated and a lot of effort has been spent on developing and improving the operational analysis of weaving sections, studies focusing on weaving safety are relatively limited in number. One of the early studies particularly focusing on freeway weaving safety was conducted by Cirillo of the Federal Highway Administration, published on Highway Research Record in 1970 (Cirillo, 1970). Approximately 700 selected weaving sections on Interstate Highways in 20 states extracted from a database of 2288 weaving sections collected in 7 years, from 1959 to 1965, were analyzed in this study. One of the most important findings of this study was the relationship between the crash rates and the lengths of acceleration and deceleration lanes, where weaving areas with shorter acceleration and deceleration lanes had higher crash rates and vice versa.

In addition to the 1970 Cirillo's study, a few other studies focusing on weaving area safety were done by Fazio et al., (1993), Glad et al. (2001), and Golob et al. (2004). While Fazio et al. (1993) took the simulation approach, with the use of traffic conflict as an alternative to crash rate, to analyzing weaving area safety; Golob et al. (2004) conducted their study based on crash data and Glad el al. (2001) adopted both approaches to investigating the issue. Golob et al. (2004) analyzed 55 weaving sections on 5 freeways in Orange County, California and revealed the crash patterns and severity levels associated with different weaving section configurations. While a crash associated with Type A weaving configuration is more likely to occur in an interior lane and is likely to be less severe, a Type B weaving crash is less likely to occur in an interior lane, more likely to be involved with lane changing maneuvers and more likely to have injuries. Type C weaving crashes are more likely to occur in the left lane and more likely to occur during weekdays. From the analysis results, the study also resulted in recommendations for improving the safety of weaving sections of specific configurations (Golob, et al., 2004). By analyzing crash data by collision type and severity, collected from 1994 to 1996 at a weaving area in Olympia, Washington area, Glad et al. (2001) found that most crashes occurring during peak hours were rear-end collisions. These crashes were likely to occur at lower speeds upstream of the weaving section. The ITRAF traffic simulation model was also applied to estimate the safety performances of four alternative designs of this weaving area (Glad, et al., 2001).

## 2.3 Weaving length and safety

As previously mentioned, Cirillo (1970) performed a study on the relationship between weaving length and safety. The author conducted analyses based on a dataset collected from approximately 700 weaving areas from 20 states. The data were aggregated in groups based on weaving length and traffic volume (Average Daily Traffic or ADT), and accident rate (number of accident per 100 million merging and diverging vehicles) was used as the indicator of safety. A part of the study findings is illustrated in Figure 2-1. This figure shows that at a given level of traffic volume, crash rate decreases as length of weaving area increases. The study results also suggest that when the weaving length increases up to a certain point, the crash rate stops increasing and goes flat. This trend was observed for most groups of traffic volume except the lowest one with ADT less than 10,000 vehicles per day. No discernible trend was found for this particular traffic volume group probably because the small sample size in this category or because the for very low traffic volume conditions, the weaving lengths provided were all adequate, discussed the author.

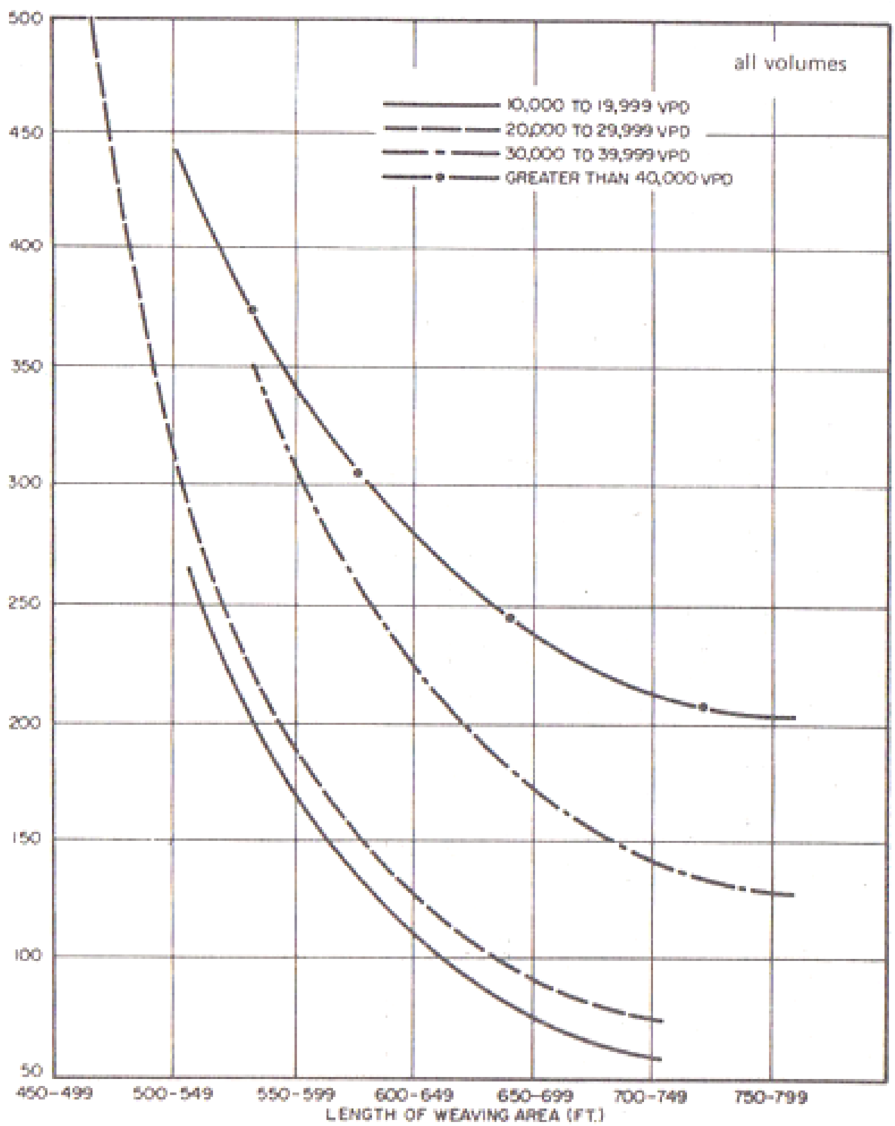



**Fig. 2-1 Crash rate by weaving length and traffic volume (Cirillo, 1970)**

Another study examined the effect of weaving are characteristics on safety was done by Pulugurtha and Bhatt (2008). The study focused on investigating the relationship between crashes and various characteristics, including weaving configuration, weaving length and the percentage of weaving traffic volume. The study used data collected from 25 weaving sections located in Las Vegas, Nevada area. The study result shows that crashes are inversely proportional to weaving length. Weaving sections with length less than 1,320 ft had most crashes (26%) and number of crashes decreases with the increase of weaving length. This trend was also found for rear-end and side-swipe crashes The study also found failure to reduce speed, improper lane-change and following too close were the top 3 contributing factors to crash causation.

## 2.4 Application of simulation in traffic study and traffic safety analysis

Statistical modeling and before and after study analysis are two popularly used approaches to traffic safety analysis in transportation safety research and practice. Although these methods have demonstrated their own advantages, many drawbacks have also been identified and documented: those analyses are based on crash data, as well as geometric and operational characteristics of the facilities, which means existence and operational conditions of roadways, and occurrences of traffic crashes are required to evaluate the safety levels of those roadways. Not to mention that collecting enough crash data for analysis is a long process. Crash reporting and computerizing processes also have a lot of error potentials, which may significantly affect the safety analysis results.

The traffic conflict concept was suggested for the first time by Perkins and Harris (1967) of General Motors. Approaching from this concept of traffic conflict, Fazio et al. (1993) conducted a study specially focusing on weaving section safety with the use of freeway conflict rates as an alternative to crash rates. In this study, although traffic conflict was not clearly recognized as a surrogate measure for safety, a strong correlation between traffic conflicts and crashes in certain types of weaving section configurations and certain traffic conditions was found. It also concluded that traffic conflict rates are predictors of the probability of crash occurrence and could be used as a good indicator of traffic safety.

Since the potential use of a microscopic simulation for safety analysis was first documented in a paper by Darzentas et al. (1980). A number of recent studies have investigated the issue using different measures of safety performance. The most widely discussed measures are time-to-collision (TTC), deceleration rate (DR), and the modified versions of those surrogate safety measures (Sayed, et al., 1994; Gettman and Head, 2003; Cunto and Saccomanno., 2007, 2008; Cunto et al., 2009; Ozbay, et al., 2008).

Initially, TTC was formulated as:

$$\mathrm{TTC} = \frac{\mathrm{D}}{\Delta\mathrm{V}} \qquad (2-1)$$

Where D is the relative distance between two vehicles and ΔV is their relative speed.

Ozbay et al. improved the model by taking speeds of both vehicles in conflict into consideration and created modified time-to-collision (MTTC) as a new surrogate measure of safety performance which includes both conflict severity level (time-to-collision) and crash severity level (vehicle speeds). A new crash index (CI) was proposed to reflect the severity of a potential crash (Ozbay, et al., 2008).

From another approach, by using deceleration rate as the primary factor to reflect the level of conflict, Cunto and Saccomanno (2007, 2008) proposed a crash potential index (CPI) in their recent studies in the following form:

$$\mathrm{CPI_i} = \frac{\sum_{t=ti_i}^{tf_i} P\left(\mathrm{DRAC}_{i,t} > \mathrm{MADR}_{i,t}\right)\Delta t.\,b}{T_i} \qquad (2-2)$$

Where:

- $CPI_i$ is the crash potential index for vehicle i;
- $DRAC_{i,t}$ is the deceleration rate to avoid the crash for vehicle i;
- $MADR_{i,t}$ is the maximum available deceleration rate or maximum breaking capability for vehicle i;
- P is the probability that DRAC of a given vehicle exceeds its MADR;
- $ti_i$ is the initial time interval for vehicle i;
- $tf_i$ is the final time interval for vehicle i;
- $\Delta t$ is the observation time interval;
- b is a binary state variable, and;
- $T_i$ is the total simulated time for vehicle i.

The latest published study on the issue was done by Cunto et al. (2009). In this study, the authors focused on comparing CPI with observed crashes and establishing a link between this simulated safety measure and actual crash occurrence information. This study used data collected from a 19 kilometer limited access freeway segment of Queen Elizabeth Way west of Toronto, Canada. The simulated safety performance measure was compared with crash data collected from the same study site. Three tests were conducted and reported in the study. Although this study has not been able to prove an adequate link between CPI and observed crash data, the study findings provided some evidence that crashes are likely to occur when CPI is high. Statistical evidence of this study suggested that CPI can be a used as an indicator of high risk driving behavior and therefore, potential of crash occurrence.

## 2.5 Chapter summary

This chapter has summarized previous studies the issues related to weaving section operations, weaving section safety performance and the relationship between weaving length and safety, as well as studies on traffic conflict technique and the use of microscopic simulation in safety research. A large number of studies have focused on the operational aspect of weaving sections. Findings of those studies have resulted in operational analysis procedures for weaving sections incorporated in the various versions of the HCM. Although a relatively comprehensive analysis methodology for weaving operation has been available since the release of the 2000 HCM, safety aspect of weaving sections has not been well understood.

With a wide range of application in transportation research and practice, microscopic simulation has begun to find its use in safety research. Several surrogate safety measures have been established with the objectives of using those measures and alternatives to crash data in safety analysis and evaluation. Although no study has successfully achieved that goal, the study findings have reinforced a promising future application of traffic conflict technique and microscopic simulation in traffic safety research.

The next chapter provides descriptions about the selection of study sites, collection and reduction of data which were used for this study.

# CHAPTER III

# DATA COLLECTION, PROCESSING AND ASSEMBLY

This chapter provides a description about the geometric, traffic and crash data used for this study. This chapter also includes details on data collection and reduction process. Section 3.1 covers details on geometric and traffic data. This first section also describes the selection of study sites, collection and reduction of data carried out. Section 3.2 provides details about the crash data collection process.

## 3.1 Geometric and traffic data

Geometric and traffic data used in this study are a part of the data that were collected for the TTI's project 0-5860 funded by TxDOT: Guidelines for Ramp Terminal Spacing for Freeways (Fitzpatrick et al., 2009). Therefore, the data collection, processing and assembly processes described in this section were in fact conducted under scope of the aforementioned TTI's project. For serving the goal of this TTI's project, the target operational measures for the field data collection efforts were identified as:

- Volumes by lane and location;
- Speed magnitudes by lane, location and movement;
- Speed variability by lane, location and movement; and
- Number, direction and location of lane changes.

The level of detail and disaggregation for these measures was limited by practicality and safety issues associated with field data collection. A number of potential factors put into consideration for selecting study sites were ramp spacing, volume, posted speed limit, number of through lanes, area type and truck restrictions.

### *3.1.1 Data collection equipment and tools*

Data collection method initially identified in the proposal of TTI's 0-5860 project (Fitzpatrick et al., 2009) were to use camera trailers, supplemental video recorders and traffic surveillance camera for lane changing and volume data. Traffic sensors and LIDAR guns were also considered as potential options for collecting speed data. After being field-tested on State Highway 6 in College Station, Texas; and evaluated based on data collection experiences in the past, several key points were concluded:

- Ideal camera views would be from upstream of the painted gore of the entrance ramp or downstream of the exit ramp's painted gore. These two camera positions provide the best views and the better camera position depends on the direction of the vertical grade within the weaving section.
- The presence of camera trailer at abovementioned locations could be a possible safety hazard. Sight lines could be blocked and the emergency recovery areas could be occupied.
- Trailer arm stability could be affected by wind. Minimal camera movement could be acceptable for most camera trailer applications but it

was not desirable for counting and identifying locations of weaving activities.

- Using LIDAR guns for speed data collection was difficult and impractical for a large amount of data needed. Because of complex activities, tracking individual vehicles through the entire weaving area was not possible.
- LIDAR guns could easily be conspicuous to drivers.

By taking different issues into consideration, a decision was made to use traffic surveillance cameras and pneumatic tubes as the first choice for data collection. Traffic management cameras, located along major roadway routes in Houston, Dallas and San Antonio, are operated by traffic management centers which known respectively as Transtar, DalTrans and TransGuide. By taking advantage of these traffic surveillance camera systems, several concerns regarding camera trailers could be resolved because these cameras offer optimal height, stability, no safety concern, and ease of video recording. However, a number of disadvantages also arose from the decision made to use of these traffic surveillance cameras. The selection of study sites was controlled more by camera view availability than by the site selection criteria initially considered in the proposal of the project. Camera views were also likely to vary for different locations, thus, the data reduction process would be more difficult and require flexible techniques. More importantly, these cameras are used for traffic and incident management purposes; therefore, the traffic management centers might take control of the cameras whenever needed. Obviously, this would result in extended time periods with the camera being aimed away from the study sites while data were being collected. Three examples of various camera views are showed in Figure 3-1, Figure 3-2 and Figure 3-3.

Google Earth and TxDOT's database were used to identify and select the study sites based on the identified criteria and the traffic surveillance camera availability. Geometric configurations of the weaving sections were also obtained from these sources.

Pneumatic tubes were used to collect volume and speed data at different locations within the weaving section.


**Fig. 3-1 Camera view of SH288 SB between Reed Rd and Airport Blvd**


**Fig. 3-2 Camera view of IH45 NB between FM 2351 and FM 1959**


**Fig. 3-3 Camera view of US67 SB between Red Bird and Camp Wisdom**

Figure 3-4 shows the general layout of the tubes used for data collection. The tube layout configuration was based on the objective of collecting all desired data. It was also considered based on safety concerns as well as the practicality of the data collection process, including installation, durability and removal of the tubes on a major multilane freeway. Two pairs of tubes were placed in the rightmost through travel lane to capture speeds and volumes immediately upstream and downstream of the entrance and exit movements. A single tube was placed on the entrance and exit ramps to collect entering and exiting volumes. The two pairs of tubes located at the ends of the painted solid lines in Figure 3-4 were primarily for speeds, but could also be used for volumes. The tubes at the end of the solid line near the merge tip were meant to capture entering speeds, but they also captured some vehicles that exited the freeway mainline early or that entered the segment from the entrance ramp, remained in the auxiliary lane and exited. Similarly, the tubes at the end of the solid line near the diverging tip were meant to capture exiting speeds, but they also captured some vehicles that entered the freeway mainline late or that entered the segment from the entrance ramp, remained in the auxiliary lane and exited. Freeway volumes in the outer through lanes as well as the numbers and locations of lane changes were counted manually using the recorded video. Additional detail is provided in a subsequent section on data collection and reduction.

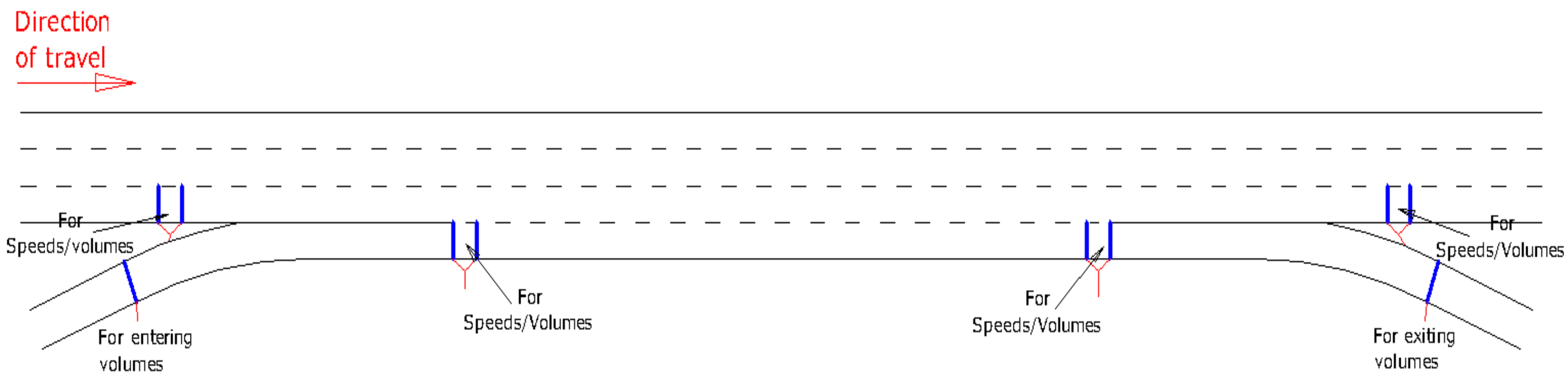



**Fig 3-4: Layout of pneumatic tubes**

### *3.1.2 Study site identification and selection*

The original proposal of TTI's project 0-5860 (Fitzpatrick et al., 2009) listed a number of potential factors to be taken into consideration during study site selection process. Those factors were ramp spacing, traffic volume, posted speed limit, number of through lanes, area type, and truck restrictions. However, the selection of study sites was in fact controlled more by the camera view availability than by the originally identified criteria because of the decision made to use traffic surveillance cameras. Even though, sites with a range in the key variable of interest, ramp spacing, were still desired. Seven study sites on major freeway routes in Houston and Dallas-Fort Worth Metropolitan areas were selected for the TTI's 0-5860 project (Fitzpatrick et al., 2009). Three sites among those seven sites were chosen for this thesis. The characteristics of three selected study sites are described on Table 3-1. The table includes the route designation and travel direction of the freeway, cross streets connecting to the ramps of the weaving sections, posted speed limit on the freeway within the study site, number of lanes on the mainline freeway, number of lanes on the entrance and exit ramps, and three different measures of ramp spacing. The three definitions of ramp spacing, illustrated in Figure 3-5, are based on newer definitions of weaving lengths that are currently being considered for incorporation into the new version of the HCM, expected to be released in 2010. $L_s$ is the length of the skip-line between the auxiliary lane and the main lanes (legal lane-changing area). $L_B$ is the distance between the painted gores and $L_L$ is the distance between the physical gores.

**Table 3-1  Study Site Characteristics**

| Freeway | | Ramp | | Posted Speed (mph) | Number of lanes | | | Spacing[1] (ft) | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Site | Route | Entrance from | Exit to | | Through | Entrance | Exit | $L_S$ | $L_B$ | $L_L$ |
| 1 | SH 288 SB | Reed | Airport | 60 | 3 | 1 | 1 | 490 | 1,100 | 1,600 |
| 2 | IH 45 NB | FM 2351 | FM 1959 | 65 | 3 | 1 | 1 | 3,150 | 3,800 | 4,300 |
| 3 | US 67 SB | Red Bird | Camp Wisdom | 60 | 2[2] | 1 | 1 | 530 | 1150 | 1,800 |

[1] see Figure 3-5 for definitions
[2] does not include adjacent HOV lane in median separated from the traveled way by painted solid lines and rumble strips

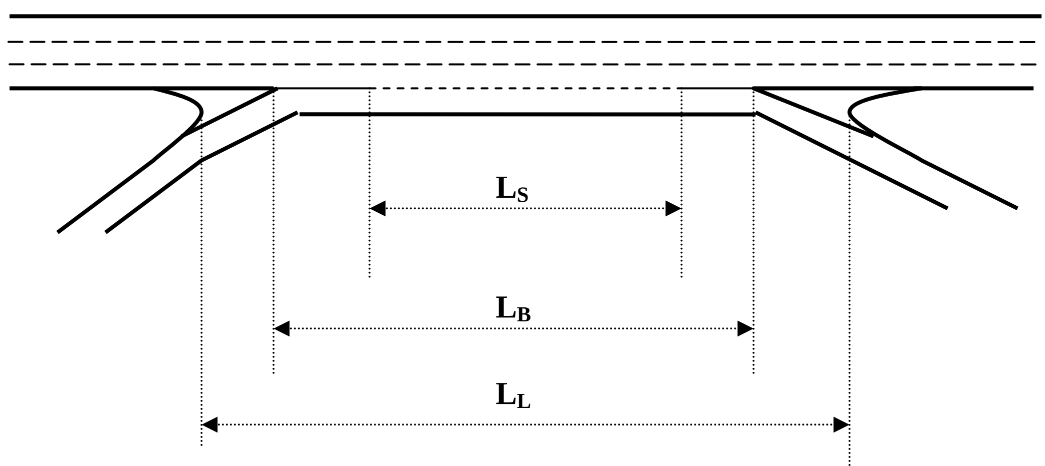



**Fig. 3-5  Measurement of the ramp terminal (Roess et al., 2008)**

Figure 3-6 shows the locations of the three sites selected for this study. Two sites are located on SH 288 and IH 45 in Houston area and the other site is located on US 67 in Dallas.

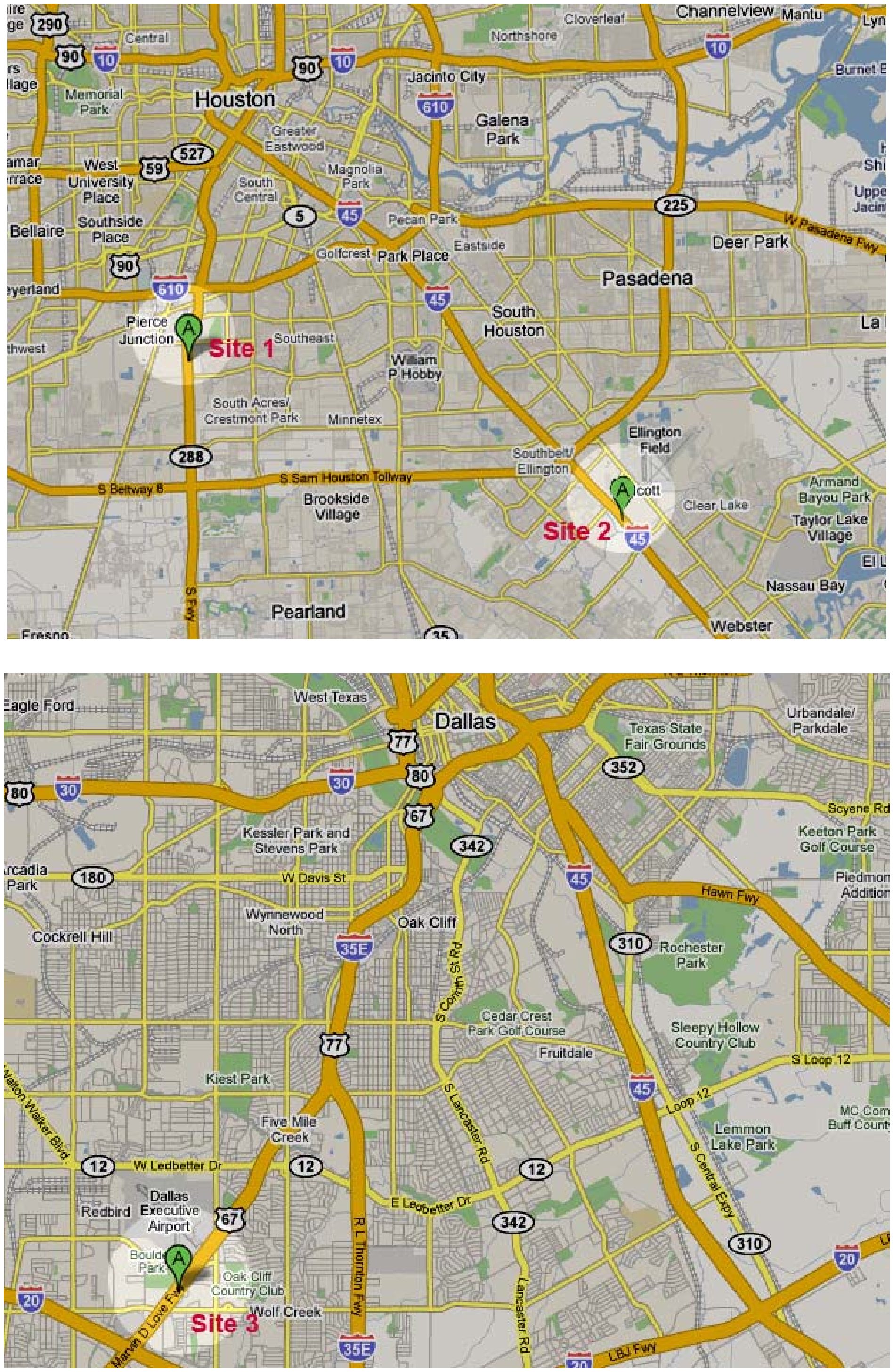



**Fig. 3-6 Study site locations in Houston and Dallas areas (Based on Google Maps)**

### *3.1.3 Collection, reduction and assembly of data*

Pneumatic tubes were placed on the study sites by TTI's researchers on a Monday with support from TxDOT for temporary traffic control. The tubes continuously collected and recorded volume and speed data until being removed on Friday of the same week. Those tubes were closely monitored during the data collection process to ensure their proper working conditions and timely response to any possible malfunction or removal by traffic. The data were collected for at least three consecutive days of data from each site therefore an approximately one work week was needed.

The cameras selected for data collection were aimed at the sites of interest on a Monday and adjusted properly to get the best camera views possible of the study sites. The cameras then recorded in digital video format from Tuesday to Thursday of the same week, dawn to dusk. Because those cameras are used for traffic surveillance, therefore, in some instances, the camera views were moved away from the study sites for monitoring traffic and supporting incident managements.

The objective was to get at least one full day of data, spanning peak periods and lower volume conditions, with the desired camera view and functioning pneumatic tubes for each site.

The video files were saved either directly onto a computer hard drive in a format compatible to most commonly used video players (for study sites in Houston), or onto the hard drive of a digital video recorder and then transferred to computer hard-drives. Tube data were saved in a comma separated value (.csv) format which was compatible with most spreadsheet-based data management and statistical analysis applications. Time stamps on the video and tube data were either synchronized prior to data collection, or the differences were recorded and the tube data were later adjusted during data reduction to match the time stamp on the video files.

Each weaving area was divided into five sections. A numbering and labeling scheme for the data was created to keep the data organized, and facilitate the data reduction and data assembly process. This scheme is illustrated in Figure 3-7 and briefly described as below:

- **Section A:** from painted entrance gore to the downstream end of the solid painted line extending from the painted entrance gore.
- **Section B:** from the downstream end of solid painted line extending from the painted entrance gore to the midpoint of the short weaving section.
- **Section C:** from the midpoint of the short weaving section to the upstream end of the solid painted line extending from the painted exit gore.
- **Section D:** from the upstream end of the solid painted line extending from the painted exit gore to the painted exit gore.
- **Section E:** downstream of the painted exit gore.

Table 3-2 shows the segment length of each section for all three study sites.

**Table 3-2 Lengths of Divided Sections within the Weaving Areas**

| Site | Section length (ft) | | | | |
|---|---|---|---|---|---|
| | ***A*** | ***B*** | ***C*** | ***D*** | ***E*** |
| SH 288 SB | 345 | 245 | 245 | 265 | NA |
| IH 45 NB | 325 | 1575 | 1575 | 325 | NA |
| US 67 SB at Red Bird | 320 | 265 | 265 | 300 | NA |

Lanes were numbered, beginning with the auxiliary lane as *1* and increasing in a direction towards the freeway median. The following measures were then defined using this referencing system:

- Vol_i_j = volume entering section *i,* lane *j*;
- Mean_Sp_i_j = mean speed from tube data in section *i*, lane *j*;
- Std_Sp_i_j = standard deviation of speed from tube data in section *i*, lane *j*; and
- i_LC_jk = number of lane changes from lane *j* to lane *k* in section *i*.

Flexibility in data reduction process was required because camera views and the measures of the above sections at each location varied. Counts of all volumes and lane changes were desired, but not always possible or practical. At a minimum, the following measures were counted:

- Vol_i_A for all *i*;
- Vol_i_E for all *i*; and
- i_LC_12 and i_LC_21 for different combinations of *i* = A, *i* = B, *i* = C and/or *i* = D.

Vol_1_A, Vol_2_A, Vol_1_B, Vol_1_C, Vol_1_E, Vol_2_E, Mean_Sp_A_2, Std_Sp_A_2, Mean_Sp_B_1, Std_Sp_B_1, Mean_Sp_C_1, Std_Sp_C_1, Mean_Sp_E_2 and Std_Sp_E_2 were computed from the pneumatic tube data. Mean_Sp_A_2 and Std_Sp_A_2 are labeled ‘A’ even though their locations are slightly upstream of Section A (see Figure 3-7) in order to simplify the complexity of the notation.

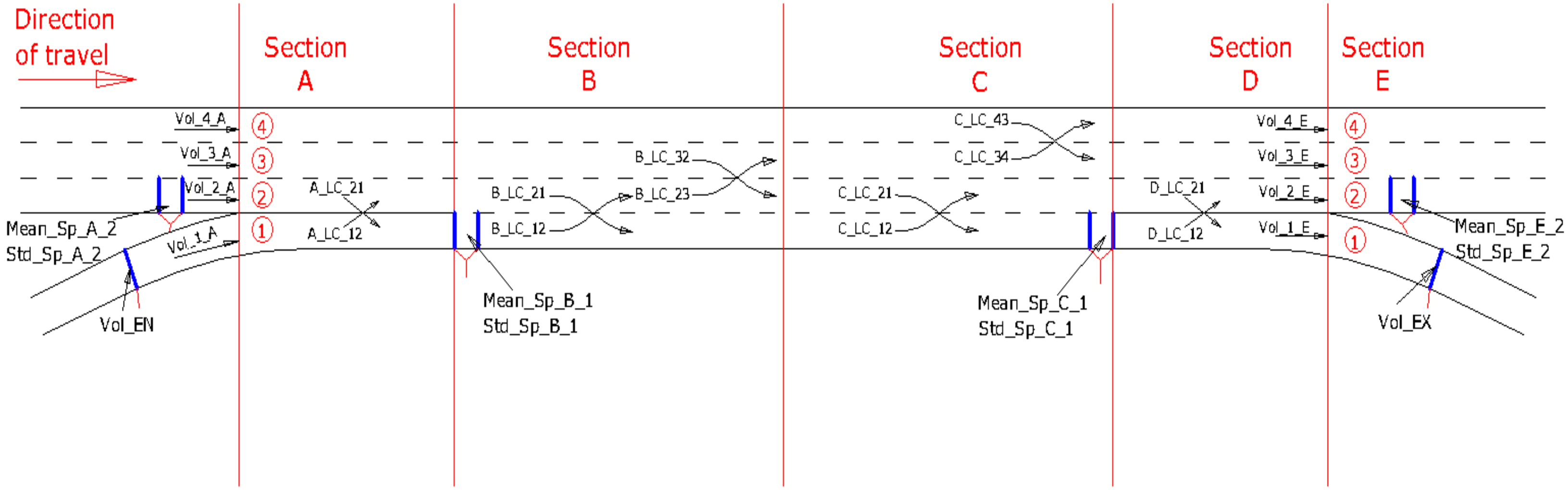



**Fig.3-7 Data reduction numbering and labeling scheme**

In general, the data reduction process followed a five-step procedure as described below:

***Step 1:*** Videos and tube data were scanned to identify day and time periods when the cameras were set at the desired view and the pneumatic tubes were functioning;

***Step 2:*** The usable videos were played on computers using appropriate video playing software. Volumes in the outer lanes where no tubes were placed were counted manually and aggregated in 5-minute and 15-minute intervals.

***Step 3:*** Using the video for the same time periods that volumes were counted, lane changing maneuvers within particular areas of the weaving section were also counted and aggregated for 5-minute and 15-minute intervals;

***Step 4:*** The raw volume and speed data files collected with the pneumatic tubes were processed and the data were aggregated into 5-minute and 15-minute intervals for all available data. Standard deviations of speed data were also calculated and aggregated into 5-minute and 15-minute bins.

***Step 5:*** The processed data including volume (counted and collected from tubes), speed and lane changing data were assembled and merged into a complete database using date and time as linking variables.

The result of the above five-step process created a comprehensive dataset spanning up to three to four days at each site. Volumes and speeds collected from tubes are available for all hours that the corresponding tubes were operational. Number of lane changing maneuvers and volumes in the outer lanes where tubes could not be placed are available for a limited number of selected hours. For lane changing and volume data which manually counting was required, attempts were made to collect at least one hour of relatively high flow rate and one hour of low flow rate.

The data collected from tubes also included acceleration data. These data were used together with lane changing and speed data and volume data for calibration of the microscopic traffic simulation model. The description of simulation calibration is presented in Chapter IV, “Traffic conflict technique and measuring safety performance using microscopic simulation model”. Volume and lane changing data were also used as inputs for simulation run.

Figures 3-8, 3-9 and 3-10 show the variation of speeds over time for study site 1, study site 2 and study site 3 respectively. Each figure shows speeds aggregated in 15-minute intervals collected from pneumatic tubes located at four different locations (shown in Figure 3-4). On each figure, a horizontal bold line is also included to illustrate the posted speed limit at that specific study site. On these graphs, the discontinuation of data is noticeable. These missing data were the consequence of tube failures. The issue is especially significant for data collected from study site 3. However, the data were collected for at least three consecutive days to make sure that data are available at any given time of a day.

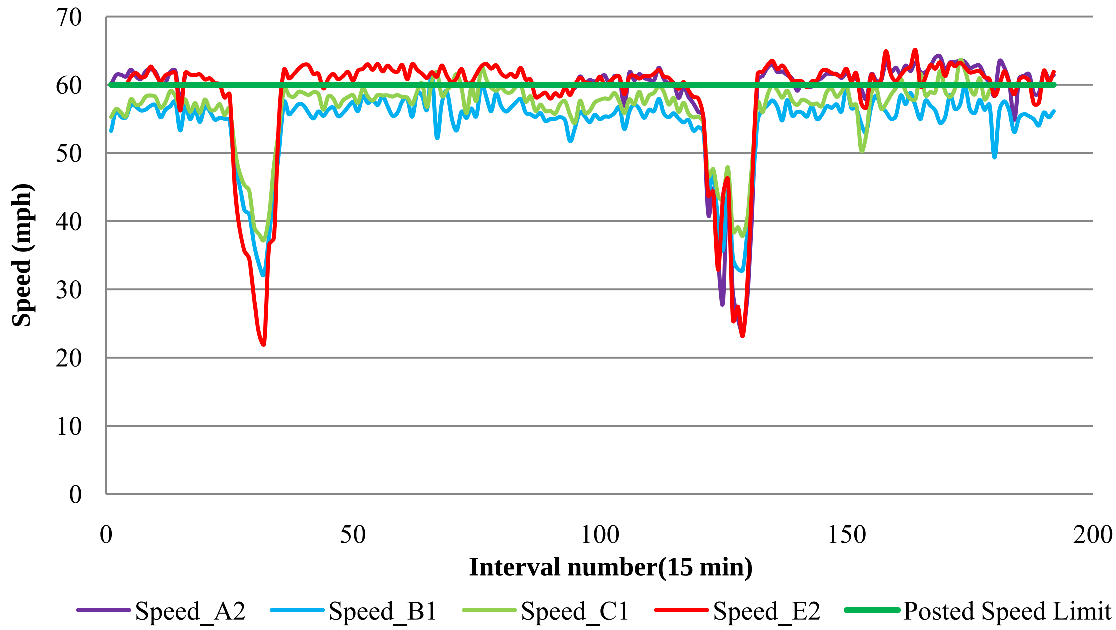



**Fig.3-8 Speeds at different locations over time at study site 1**

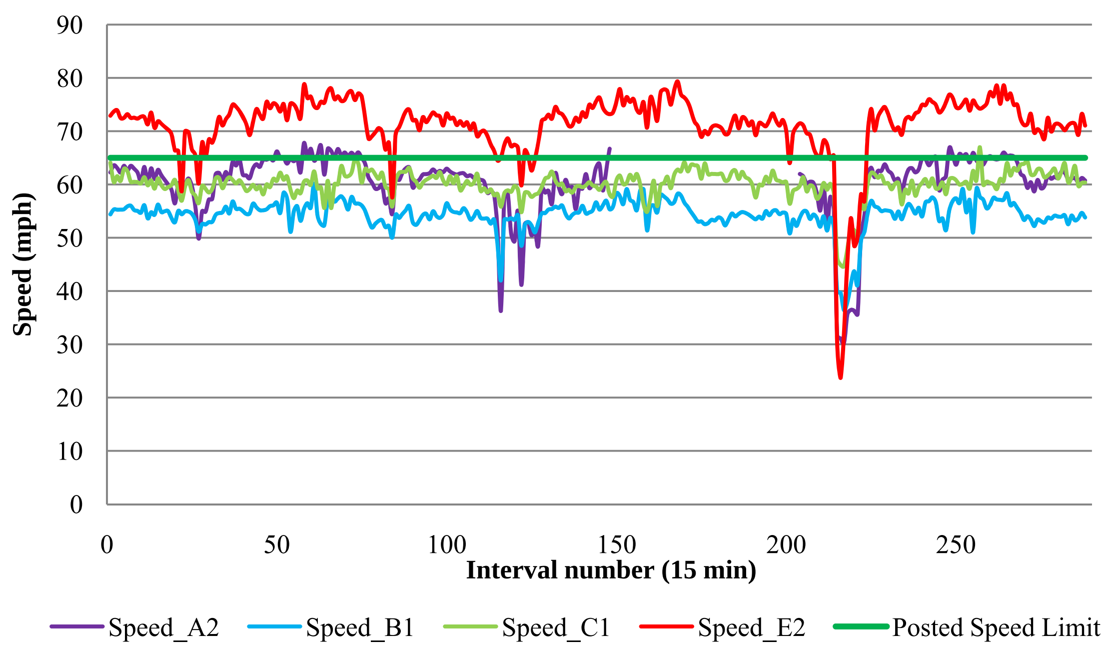



**Fig.3-9 Speeds at different locations over time at study site 2**

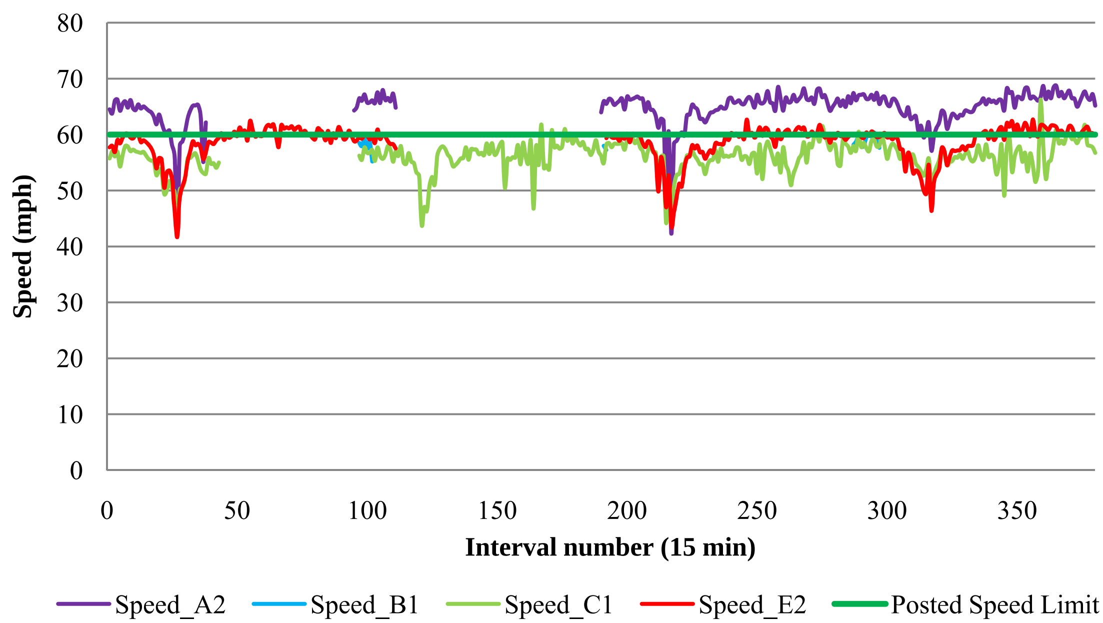
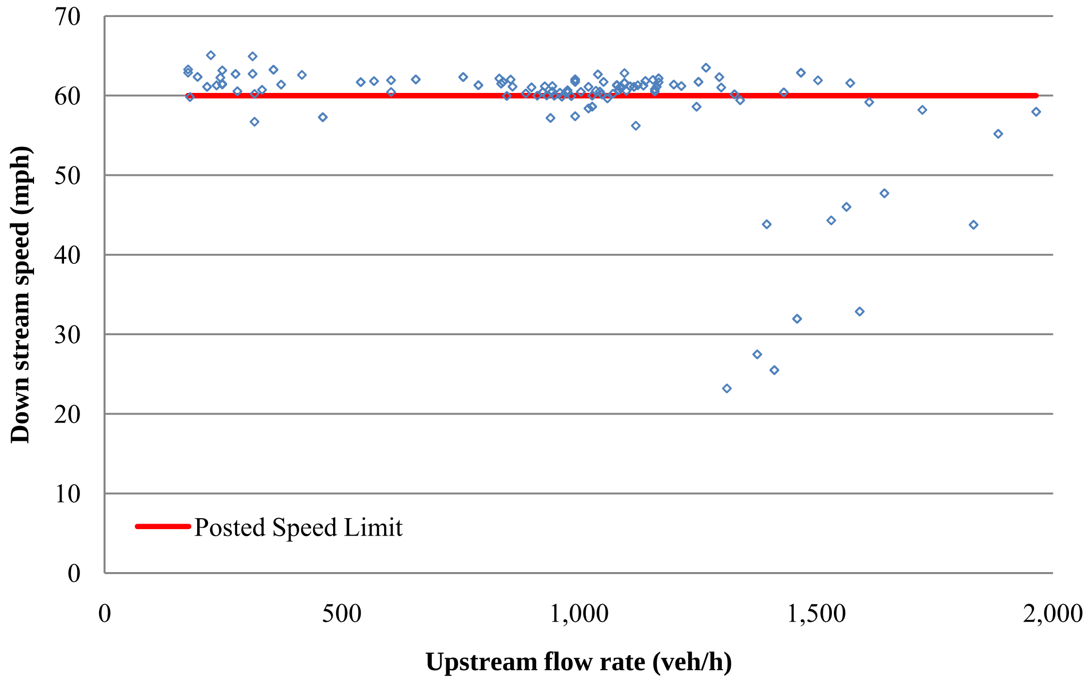



**Fig.3-10 Speeds at different locations over time at study site 3**

Figures 3-11 to 3-13 show scatter plots illustrating the relationships between speed and flow rates for three study sites. On each figure, the upper graph shows the scatter plot of speed in the right lane downstream of the weaving area and the flow rate in the right lane upstream of the weaving area (E2 and A2 in Figure 3-7). The lower graph shows the relationship between speed and flow rate in the right lane downstream of the weaving area (E2 in Figure 3-7).

The relationships between speed and flow rate shown on these graphs are same as expected to be. Although the points where speed starts to decline when flow rate increases may vary, they look very similar to what has been discussed in the May's traffic flow theory text book (May, 1990) and the HCM (HCM, 2000).



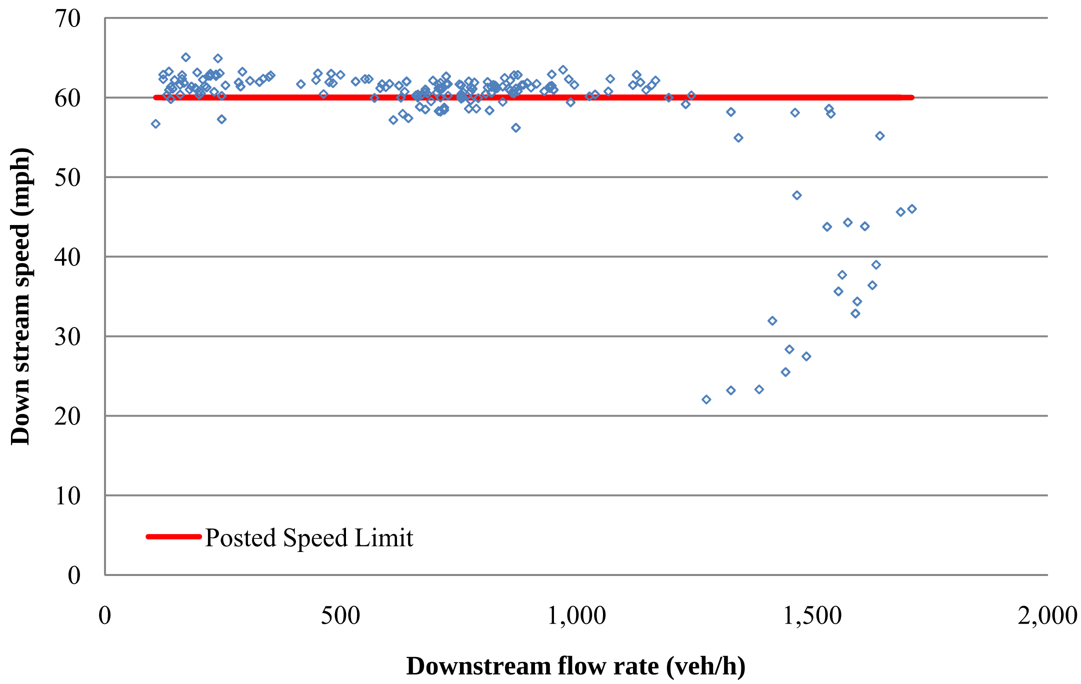
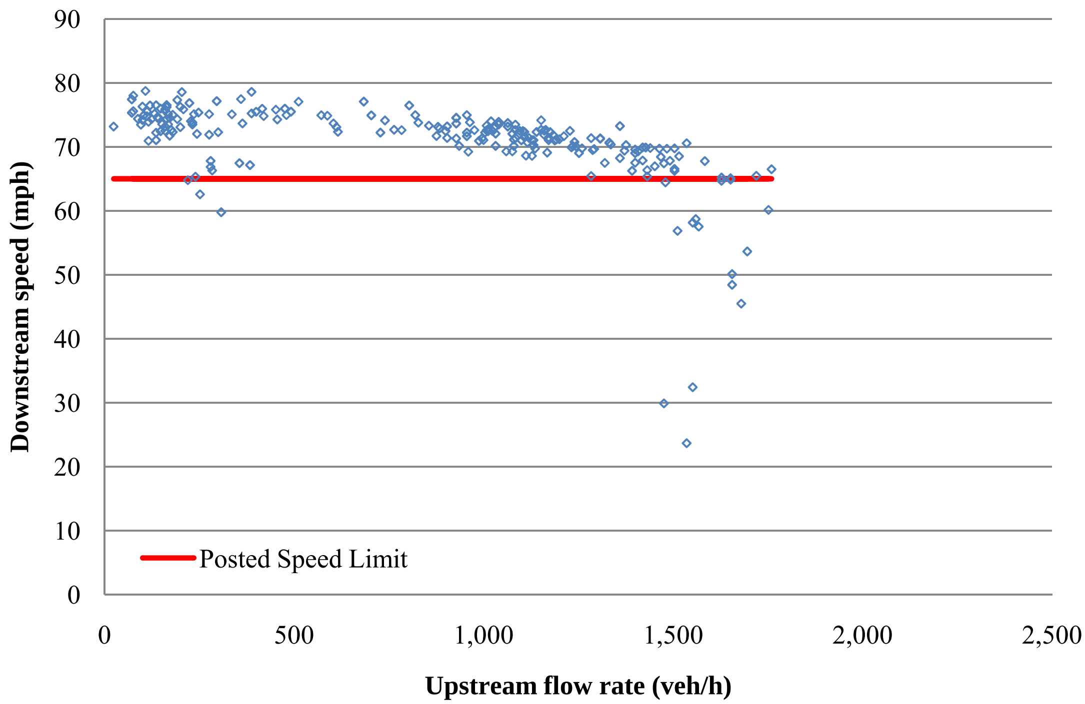




**Fig.3-11 Speed and flow rate relationship for study site 1**



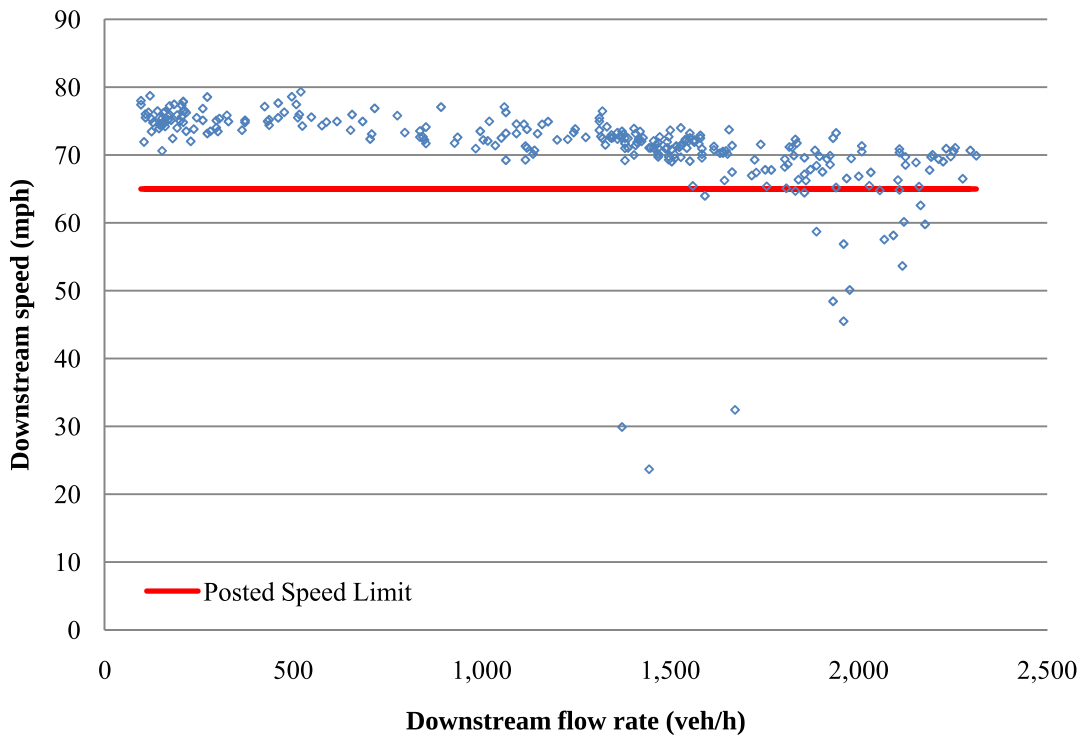




**Fig.3-12 Speed and flow rate relationship for study site 2**

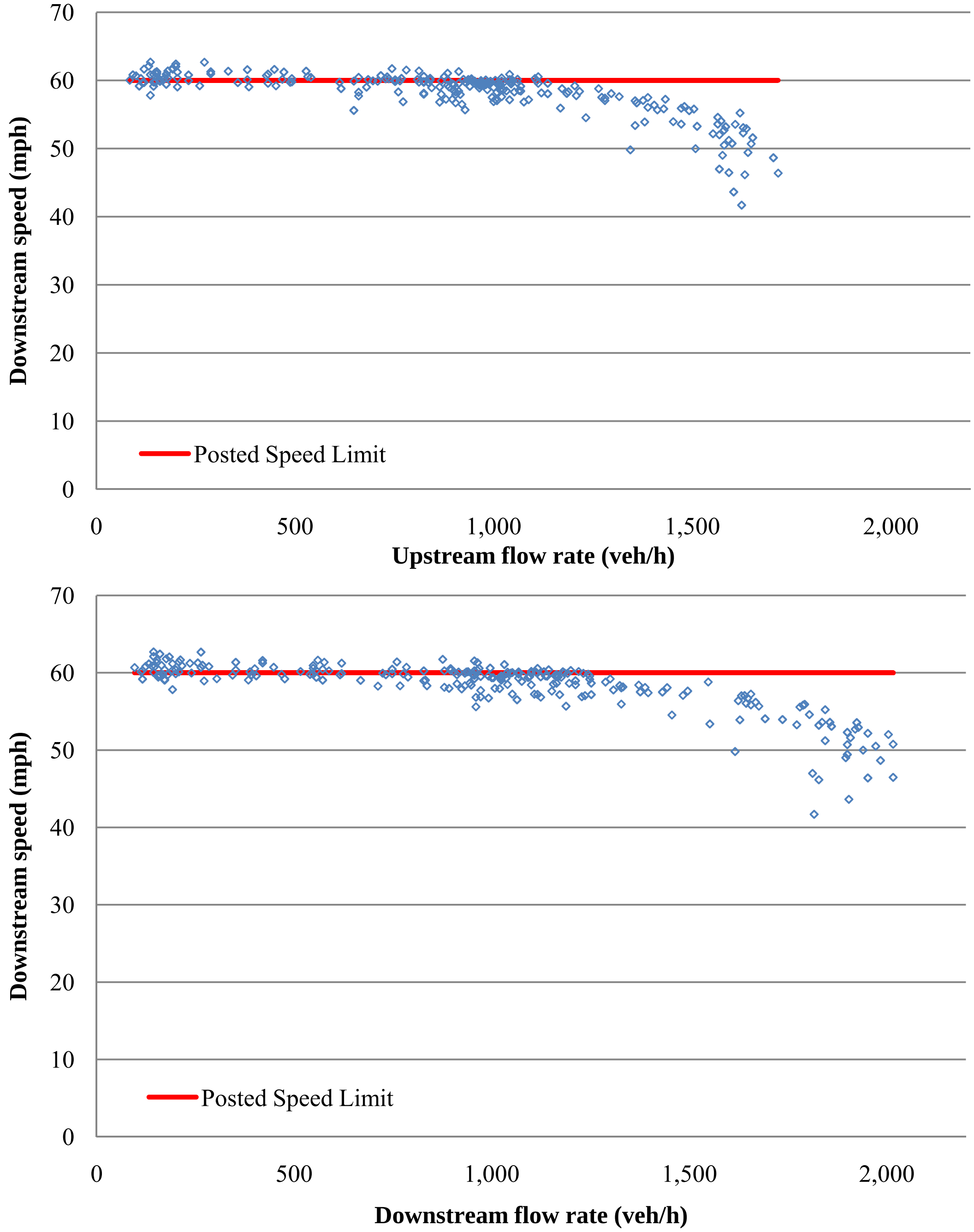



**Fig.3-13 Speed and flow rate relationship for study site 3**

## 3.2 Crash data

Five years of latest crash data, from 2003 to 2007, were acquired from the Texas Department of Transportation Crash Record Information System (CRIS). Crashes occurring at each site were then located and extracted from the database using latitude, longitude, control section, and mile point associated with each particular crash in the database following several steps as below:

- Texas Control Section Map was used to identify particular sections of roadway on which the study sites are located. This information was then used to locate all crashes occurred within these particular sections of roadway, including all crashes on the study sites. After this step, a much smaller subset of data which included crashes occurring at three sections of roadway (control sections) was extracted. This dataset had all crashes on the sites of interest, but it was much smaller and easier to work with.

- Latitude and longitude information was used to create several “bench marks” around the locations of interest in Google Earth, the popular virtual globe, map and geographic information software provided by Google Inc. These bench marks were then used to locate the boundaries of the study sites. These boundaries, associated with mile point information, were then used to extract all crashes occurred within these boundaries, which defined the study sites. In this study, upstream and downstream areas of influence were considered to include in the weaving sections. The traffic condition within these influence areas were actually affected by the activities within the weaving sections and thus, the safety within these influence areas reflected the safety of the entire weaving areas. The influence areas were decided to be 1,000 ft areas upstream and downstream of the mainline freeway and 500 ft areas on the entrance and exit ramps, from the painted gores. This decision was made based on several recommendations and choices made by Chen et al.(2009), Sarhan et al.,(2008) and Batenhorst and Gerken (2000) in their studies on similar issues. The influence areas are illustrated in Figure 3-14.

- The travel direction information of vehicles being involved in the crash was used to identify the crashes on each travel direction. Only crashes within the study sites (including the influence areas) on a particular travel direction were extracted and used for the analyses in this study.

- The information on number of vehicles involved in the crash was used to separate single-vehicle crashes and multi-vehicle crashes. Two datasets were created: one with all crashes (single-vehicle and multi-vehicle crashes) included, and the other one with only multi-vehicle crashes (removed all single-vehicle crashes).

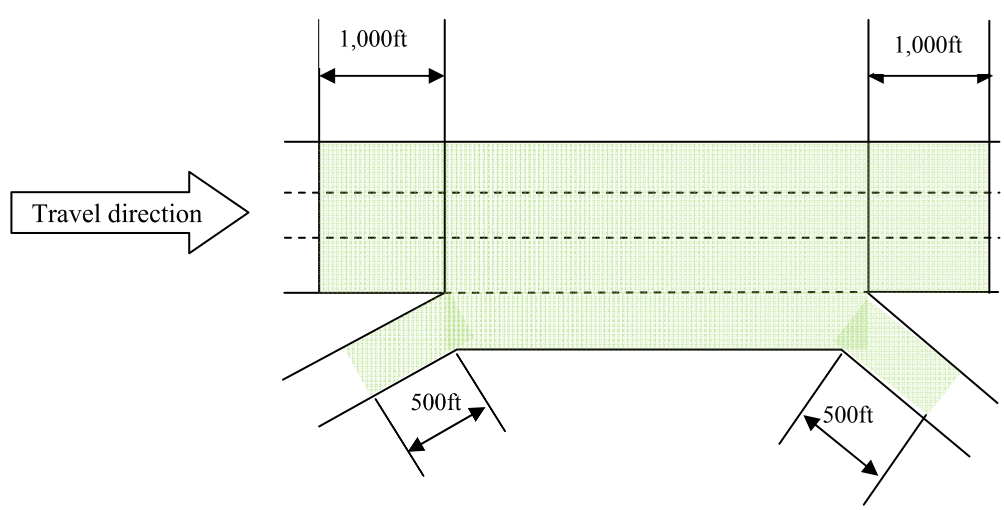



**Fig.3-14 Upstream and downstream influence areas of a weaving section**

The crashes were then grouped by time of occurrence. Tables 3-3 to 3-5 show crash data for three study sites with all five severity levels: Fatal injury (K), Incapacitating injury (A), Non-incapacitating injury (B), Possible injury (C), Property Damage Only (PDO), and Unknown (Un).

**Table 3-3 Crashes by Severity Level in 1-Hour Intervals for Study Site 1 (2003-2007)**

| Hour | All crashes | | | | | | | Multi-vehicle crashes only | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | **K** | **A** | **B** | **C** | **PDO** | **Un** | **Total** | **K** | **A** | **B** | **C** | **PDO** | **Un** | **Total** |
| 0-1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 1-2 | 0 | 0 | 1 | 0 | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 1 |
| 2-3 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 1 |
| 3-4 | 0 | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 4-5 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 5-6 | 0 | 0 | 0 | 0 | 2 | 0 | 2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 6-7 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 7-8 | 0 | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 1 | 0 | 0 | 1 |
| 8-9 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 9-10 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 10-11 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 11-12 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 12-13 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 13-14 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 14-15 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 15-16 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 16-17 | 0 | 0 | 0 | 2 | 1 | 0 | 3 | 0 | 0 | 0 | 2 | 1 | 0 | 3 |
| 17-18 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 18-19 | 0 | 0 | 0 | 1 | 1 | 0 | 2 | 0 | 0 | 0 | 0 | 1 | 0 | 1 |
| 19-20 | 0 | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 1 | 0 | 0 | 1 |
| 20-21 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 21-22 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 22-23 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 23-24 | 0 | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 1 | 0 | 0 | 1 |
| | **0** | **0** | **1** | **7** | **10** | **0** | **18** | **0** | **0** | **1** | **5** | **3** | **0** | **9** |

**Table 3-4 Crashes by Severity Level in 1-Hour Intervals for Study Site 2 (2003-2007)**

| Hour | All crashes | | | | | | | Multi-vehicle crashes | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | **K** | **A** | **B** | **C** | **PDO** | **Un** | **Total** | **K** | **A** | **B** | **C** | **PDO** | **Un** | **Total** |
| 0-1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 1-2 | 0 | 1 | 1 | 2 | 0 | 0 | 4 | 0 | 0 | 0 | 2 | 0 | 0 | 2 |
| 2-3 | 0 | 0 | 1 | 0 | 4 | 0 | 5 | 0 | 0 | 1 | 0 | 2 | 0 | 3 |
| 3-4 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 4-5 | 0 | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 1 | 0 | 0 | 1 |
| 5-6 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 1 |
| 6-7 | 1 | 0 | 0 | 0 | 4 | 0 | 5 | 0 | 0 | 0 | 0 | 4 | 0 | 4 |
| 7-8 | 0 | 0 | 1 | 1 | 1 | 0 | 3 | 0 | 0 | 1 | 1 | 1 | 0 | 3 |
| 8-9 | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 9-10 | 0 | 0 | 0 | 0 | 2 | 0 | 2 | 0 | 0 | 0 | 0 | 1 | 0 | 1 |
| 10-11 | 0 | 0 | 0 | 1 | 2 | 0 | 3 | 0 | 0 | 0 | 1 | 2 | 0 | 3 |
| 11-12 | 0 | 0 | 1 | 1 | 2 | 0 | 4 | 0 | 0 | 0 | 1 | 2 | 0 | 3 |
| 12-13 | 0 | 0 | 2 | 3 | 3 | 0 | 8 | 0 | 0 | 1 | 2 | 2 | 0 | 5 |
| 13-14 | 0 | 1 | 0 | 2 | 3 | 0 | 6 | 0 | 1 | 0 | 1 | 0 | 0 | 2 |
| 14-15 | 0 | 1 | 0 | 1 | 1 | 0 | 3 | 0 | 0 | 0 | 1 | 0 | 0 | 1 |
| 15-16 | 0 | 1 | 0 | 1 | 1 | 0 | 3 | 0 | 1 | 0 | 1 | 0 | 0 | 2 |
| 16-17 | 0 | 0 | 0 | 1 | 2 | 0 | 3 | 0 | 0 | 0 | 1 | 2 | 0 | 3 |
| 17-18 | 0 | 0 | 0 | 4 | 4 | 0 | 8 | 0 | 0 | 0 | 4 | 3 | 0 | 7 |
| 18-19 | 0 | 0 | 0 | 1 | 2 | 1 | 4 | 0 | 0 | 0 | 0 | 2 | 1 | 3 |
| 19-20 | 0 | 1 | 0 | 1 | 1 | 0 | 3 | 0 | 0 | 0 | 1 | 1 | 0 | 2 |
| 20-21 | 0 | 0 | 2 | 2 | 0 | 1 | 5 | 0 | 0 | 0 | 1 | 0 | 0 | 1 |
| 21-22 | 0 | 0 | 1 | 1 | 0 | 0 | 2 | 0 | 0 | 0 | 1 | 0 | 0 | 1 |
| 22-23 | 0 | 0 | 0 | 1 | 1 | 0 | 2 | 0 | 0 | 0 | 1 | 1 | 0 | 2 |
| 23-24 | 0 | 1 | 0 | 0 | 1 | 0 | 2 | 0 | 1 | 0 | 0 | 0 | 0 | 1 |
| | **2** | **6** | **9** | **24** | **36** | **2** | **79** | **0** | **3** | **3** | **20** | **24** | **1** | **51** |

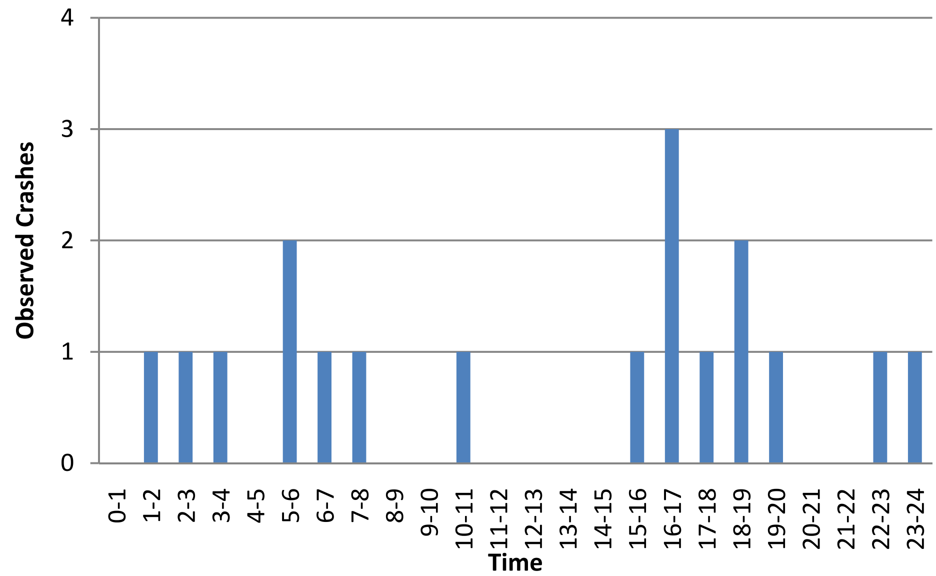
**Table 3-5 Crashes by Severity Level in 1-Hour Intervals for Study Site 3 (2003-2007)**

| Hour | All crashes | | | | | | | Multi-vehicle crashes | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | K | A | B | C | PDO | Un | Total | K | A | B | C | PDO | Un | Total |
| 0-1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 1-2 | 0 | 0 | 0 | 0 | 2 | 0 | 2 | 0 | 0 | 0 | 0 | 2 | 0 | 2 |
| 2-3 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 3-4 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 4-5 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 1 |
| 5-6 | 0 | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 6-7 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 7-8 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 8-9 | 0 | 1 | 0 | 0 | 2 | 0 | 3 | 0 | 0 | 0 | 0 | 1 | 0 | 1 |
| 9-10 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 10-11 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 11-12 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 12-13 | 0 | 0 | 1 | 0 | 1 | 0 | 2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 13-14 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 1 |
| 14-15 | 0 | 0 | 0 | 1 | 1 | 0 | 2 | 0 | 0 | 0 | 1 | 0 | 0 | 1 |
| 15-16 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 16-17 | 0 | 0 | 0 | 1 | 1 | 0 | 2 | 0 | 0 | 0 | 1 | 0 | 0 | 1 |
| 17-18 | 0 | 0 | 0 | 1 | 4 | 0 | 5 | 0 | 0 | 0 | 1 | 4 | 0 | 5 |
| 18-19 | 0 | 0 | 1 | 5 | 1 | 0 | 7 | 0 | 0 | 0 | 5 | 1 | 0 | 6 |
| 19-20 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 20-21 | 0 | 0 | 0 | 1 | 1 | 0 | 2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 21-22 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 22-23 | 0 | 0 | 0 | 1 | 1 | 0 | 2 | 0 | 0 | 0 | 1 | 0 | 0 | 1 |
| 23-24 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| | **0** | **1** | **2** | **11** | **19** | **1** | **34** | **0** | **0** | **0** | **9** | **10** | **0** | **19** |

Data shown on these tables are in 1-hour aggregation based on time of occurrence. The tables show both data sets: one with all crashes included and the other one with multi-vehicle crashes included (removed all single-vehicle crashes). The 8th column of these tables shows the total number of crashes within that period of time for the data set with all crashes (single and multi-vehicle crashes) included. Similarly, the last column of these tables shows total number of crashes for the second data set with all single-vehicle crashes being removed. These total numbers were actually used in all statistical analyses. As described in Chapter IV, the traffic conflict technique is based on 2-vehicle conflict in the traffic stream; therefore all analyses in this study were done with both data sets.

Figures 3-15 to 3-17 show the distributions of crashes in a 24 hour period at study site 1, study site 2 and study site 3 respectively. Each graph shows two graphs for two data sets as previously described: one data set with all crashes included and the other one with all single-vehicle crashes removed. The data shown on these graphs are the total number of crashes in five years aggregated by time of occurrence in 1-hour intervals. Crash severity levels are not distinguished on these graphs.

All crashes




Multi-vehicle crashes only

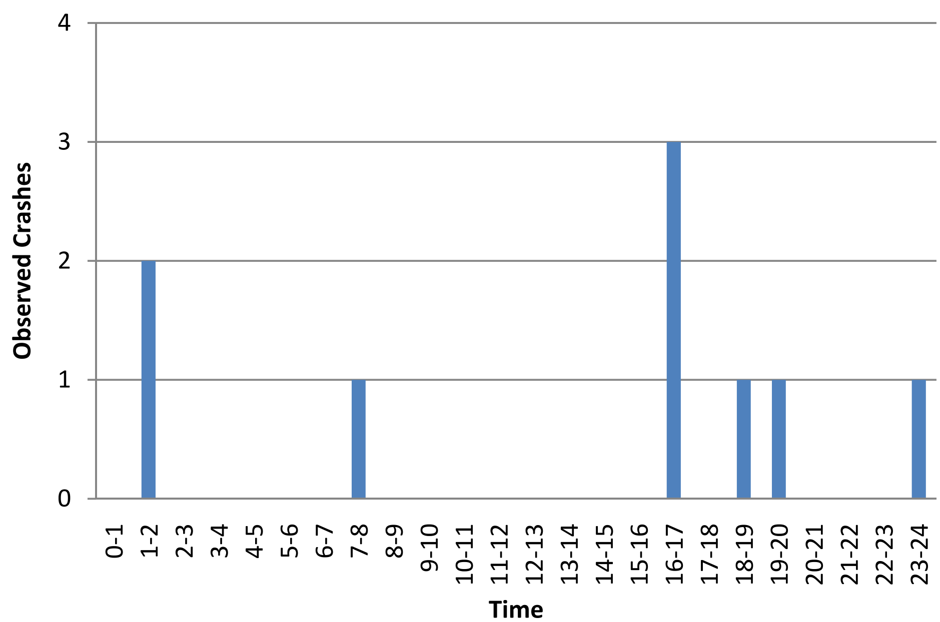



**Fig. 3-15 Crash distribution in 24 hours, aggregated in 1-hour intervals for study site 1 (2003-2007)**

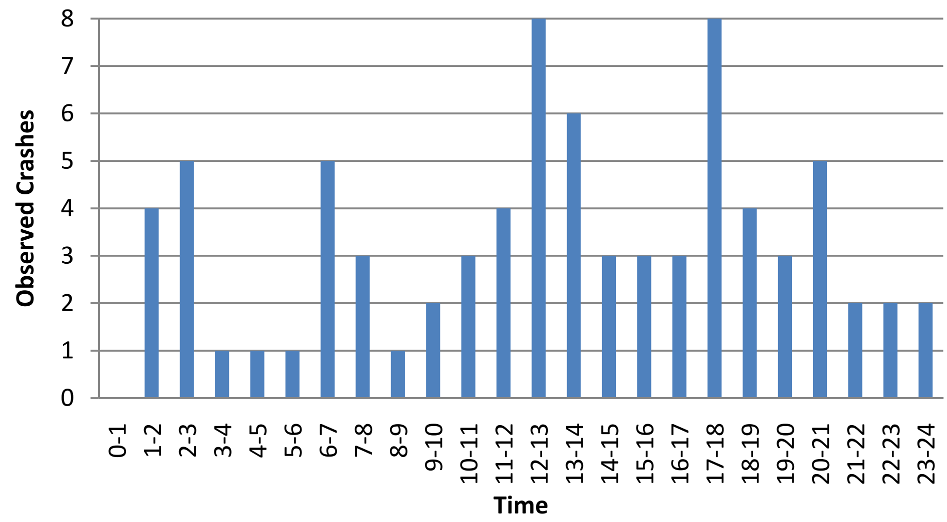
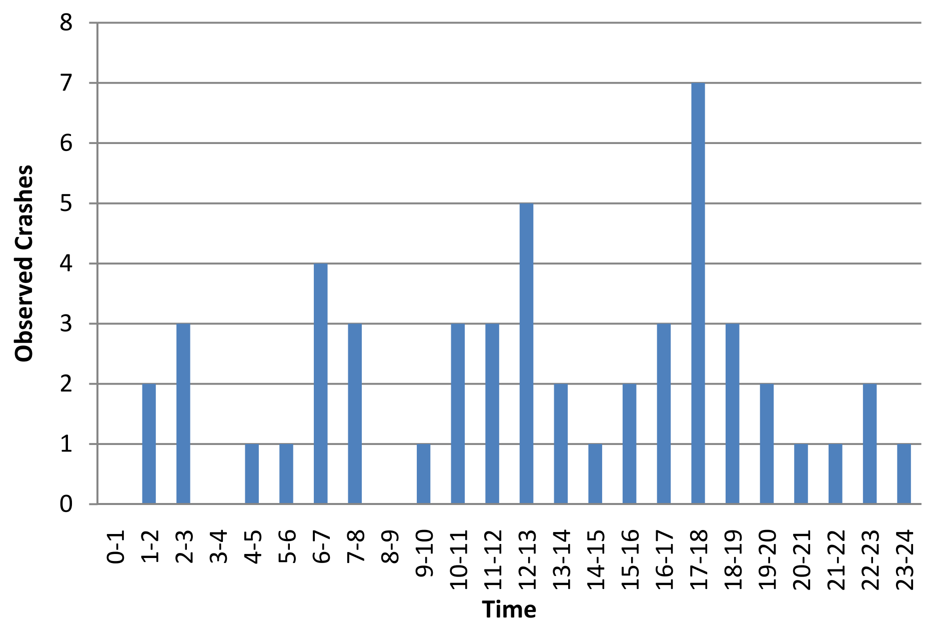
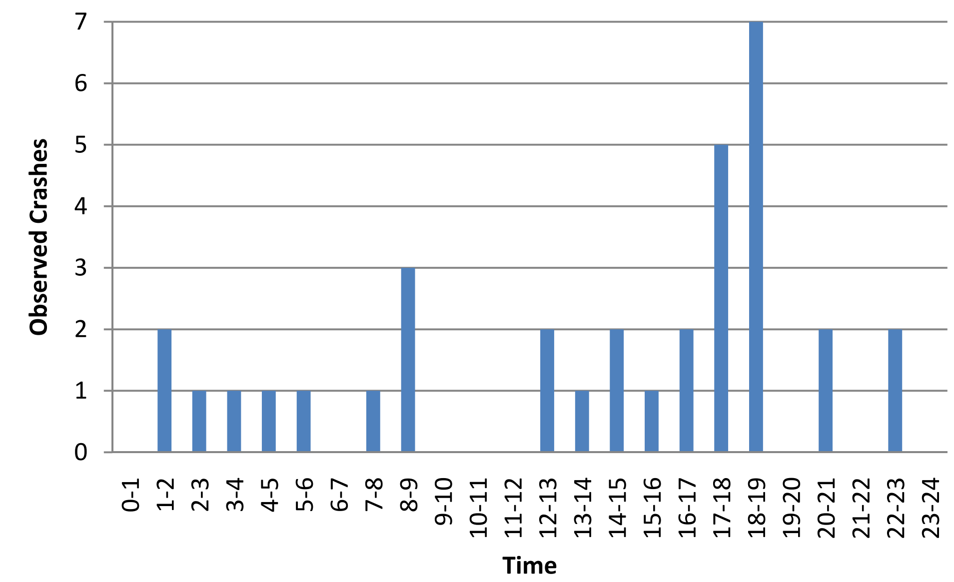
| | |
|---|---|
| All crashes |  |
| Multi-vehicle crashes only |  |

**Fig. 3-16 Crash distribution in 24 hours, aggregated in 1-hour intervals for study site 2 (2003-2007)**

All crashes




Multi-vehicle crashes only

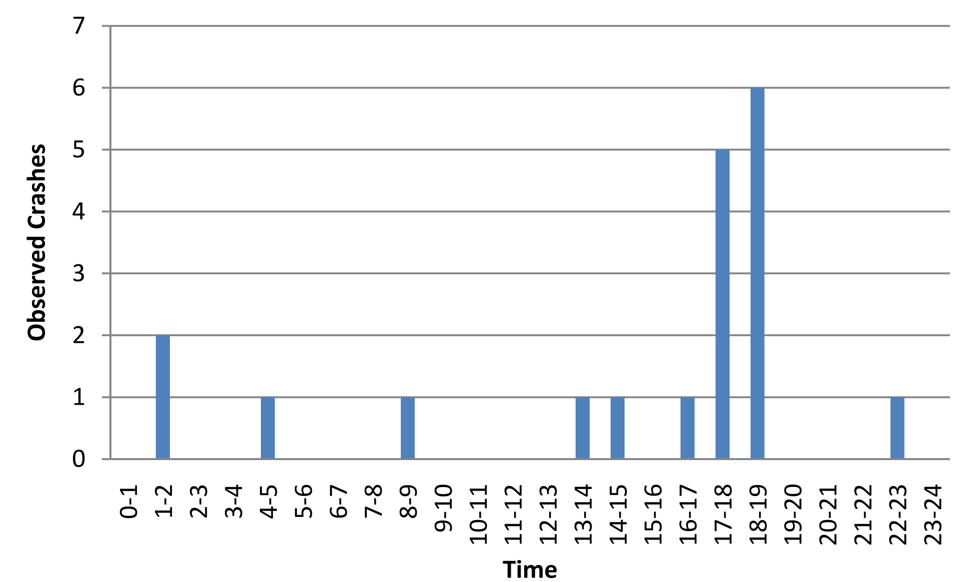



**Fig. 3-17 Crash distribution in 24 hours, aggregated in 1-hour intervals for study site 3 (2003-2007)**

Tables 3-6 to 3-8 show the same subsets of data aggregated by time of day as bellow: Night (from midnight-to 6 AM), Morning peak (from 6 to 9 AM), Off-Peak (from 9 AM to 3 PM), Evening peak (from 3PM to 6 PM) and Evening off-peak (from 6 PM to midnight). The tables include number of crashes for different level of severity (K, A, B, C, PDO, and Un) and the average number of all crashes per hour. The total numbers of crashes were transformed into average values to avoid bias because the data were not aggregated in intervals with different lengths.

**Table 3-6 Crashes by Severity Level Aggregated Based on Time of Day for Study Site 1 (2003-2007)**

| Time | Description | All crashes | | | | | | | Multi-vehicle crashes only | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | K | A | B | C | PDO | UN | All/hr | K | A | B | C | PDO | UN | All/hr |
| 0-6 | Night | 0 | 0 | 1 | 1 | 3 | 0 | 0.83 | 0 | 0 | 1 | 0 | 1 | 0 | 0.33 |
| 6-9 | Morning peak | 0 | 0 | 0 | 1 | 1 | 0 | 0.67 | 0 | 0 | 0 | 1 | 0 | 0 | 0.33 |
| 9-15 | Off-peak | 0 | 0 | 0 | 0 | 1 | 0 | 0.17 | 0 | 0 | 0 | 0 | 0 | 0 | 0.00 |
| 15-18 | Evening peak | 0 | 0 | 0 | 2 | 3 | 0 | 1.67 | 0 | 0 | 0 | 2 | 1 | 0 | 1.00 |
| 18-24 | Evening off-peak | 0 | 0 | 0 | 3 | 2 | 0 | 0.83 | 0 | 0 | 0 | 2 | 1 | 0 | 0.50 |
| | | **0** | **0** | **1** | **7** | **10** | **0** | | **0** | **0** | **1** | **5** | **3** | **0** | |

**Table 3-7 Crashes by Severity Level Aggregated Based on Time of Day for Study Site 2 (2003-2007)**

| Time | Description | All crashes | | | | | | | Multi-vehicle crashes only | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | K | A | B | C | PDO | UN | All/hr | K | A | B | C | PDO | UN | All/hr |
| 0-6 | Night | 0 | 1 | 2 | 3 | 6 | 0 | 2.00 | 0 | 0 | 1 | 3 | 3 | 0 | 1.17 |
| 6-9 | Morning peak | 2 | 0 | 1 | 1 | 5 | 0 | 3.00 | 0 | 0 | 1 | 1 | 5 | 0 | 2.33 |
| 9-15 | Off-peak | 0 | 2 | 3 | 8 | 13 | 0 | 4.33 | 0 | 1 | 1 | 6 | 7 | 0 | 2.50 |
| 15-18 | Evening peak | 0 | 1 | 0 | 6 | 7 | 0 | 4.67 | 0 | 1 | 0 | 6 | 5 | 0 | 4.00 |
| 18-24 | Evening off-peak | 0 | 2 | 3 | 6 | 5 | 2 | 3.00 | 0 | 1 | 0 | 4 | 4 | 1 | 1.67 |
| | | **2** | **6** | **9** | **24** | **36** | **2** | | **0** | **3** | **3** | **20** | **24** | **1** | |

**Table 3-8 Crashes by Severity Level Aggregated Based on Time of Day for Study Site 3 (2003-2007)**

| Time | Description | All crashes | | | | | | | Multi-vehicle crashes only | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | K | A | B | C | PDO | UN | All/hr | K | A | B | C | PDO | UN | All/hr |
| 0-6 | Night | 0 | 0 | 0 | 1 | 4 | 1 | 1.0 | 0 | 0 | 0 | 0 | 3 | 0 | 0.50 |
| 6-9 | Morning peak | 0 | 1 | 0 | 0 | 3 | 0 | 1.3 | 0 | 0 | 0 | 0 | 1 | 0 | 0.33 |
| 9-15 | Off-peak | 0 | 0 | 1 | 1 | 3 | 0 | 0.8 | 0 | 0 | 0 | 1 | 1 | 0 | 0.33 |
| 15-18 | Evening peak | 0 | 0 | 0 | 2 | 6 | 0 | 2.7 | 0 | 0 | 0 | 2 | 4 | 0 | 2.00 |
| 18-24 | Evening off-peak | 0 | 0 | 1 | 7 | 3 | 0 | 1.8 | 0 | 0 | 0 | 6 | 1 | 0 | 1.17 |
| | | **0** | **1** | **2** | **11** | **19** | **2** | | **0** | **0** | **0** | **9** | **10** | **0** | |

Figures 3-18 to 3-20 are visual presentations of the average number of all crashes per hour shown on tables 3-6 to 3-8 for study site 1, study site 2 and study site 3 respectively. These graphs do not show crashes by different severity levels. Only the average number of all crashes per hour, aggregated based on time of occurrence, were presented in the figures. These average values per hour are actually used in the analyses in the next chapter of this thesis.

As shown on these figures, the evening peak hour has the most number of crashes per hour. The figures also show that a significant proportion of single-vehicle crashes occur at night and in the evening. These crashes at night may be contributed by other factors, not simple because of traffic conflicts since the traffic flow is usually very low and most of the vehicles do not have interactions with each other. This supports the idea of including both data sets in the analyses to see the possible differences in the results.

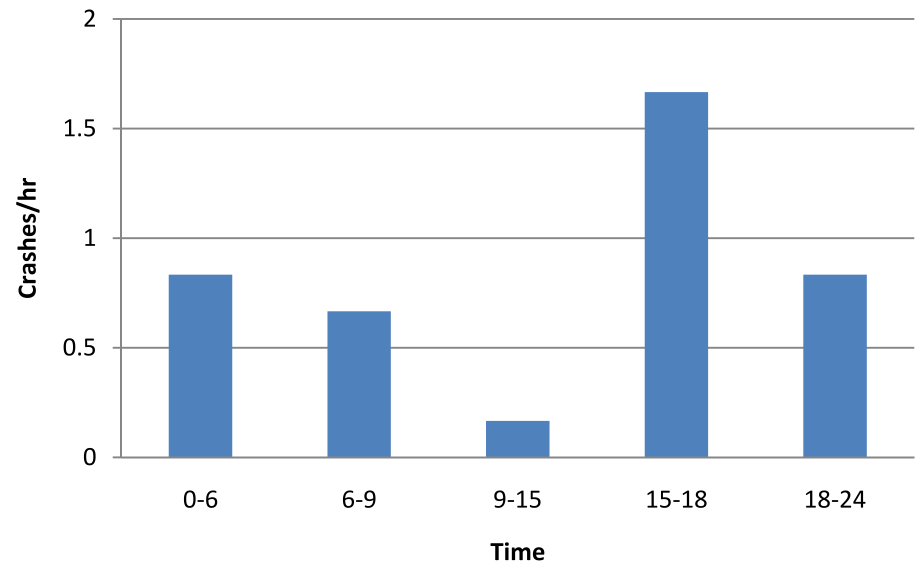
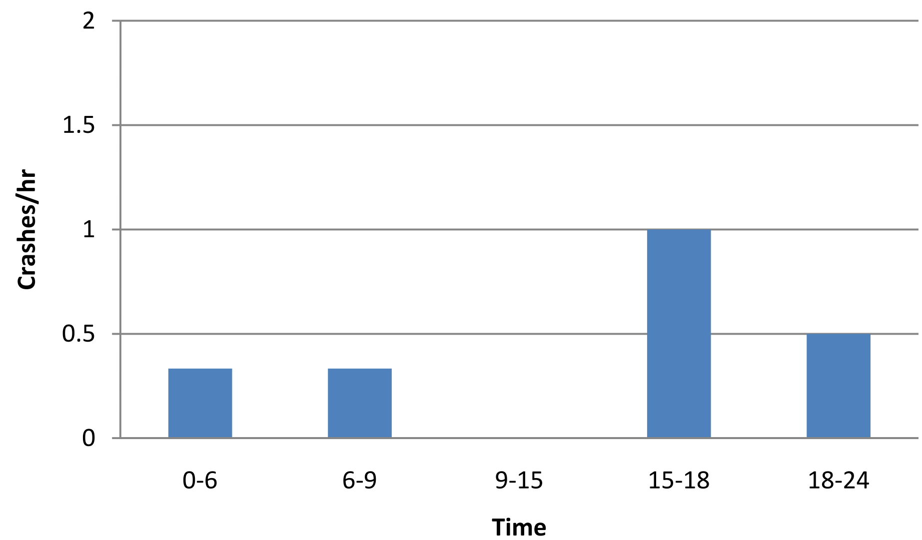
<table>
<tr><td>All crash</td><td>
</td></tr>
<tr><td>Multi-vehicle crash only</td><td>
</td></tr>
</table>

**Fig. 3-18 Crash distribution in 24 hours, aggregated based on time of day for study site 1 (2003-2007)**

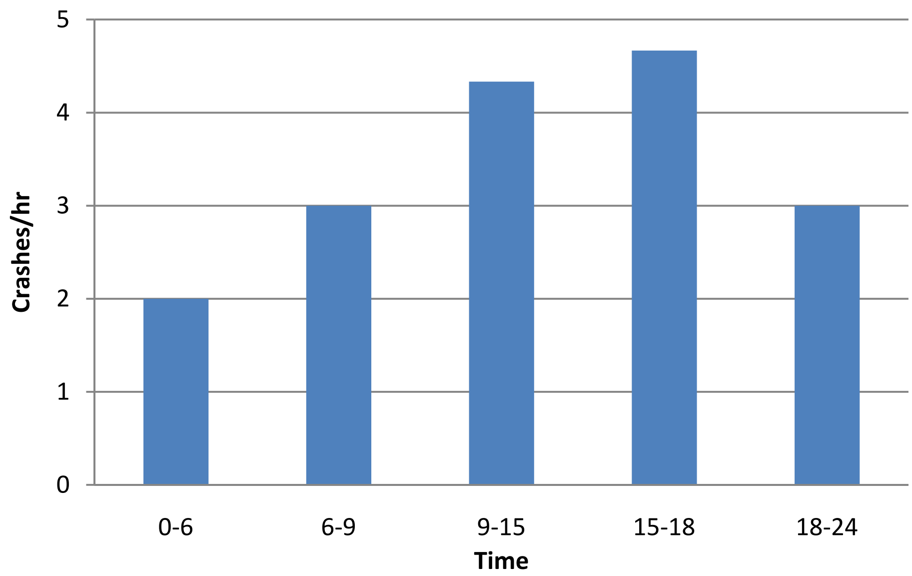
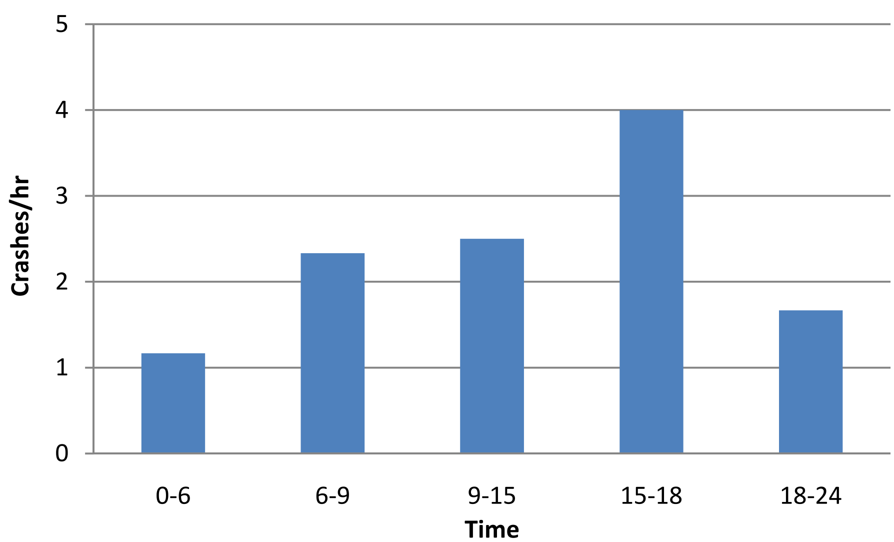
| | |
|---|---|
| All crash |   |
| Multi-vehicle crash only |   |

**Fig. 3-19 Crash distribution in 24 hours, aggregated based on time of day for study site 2 (2003-2007)**

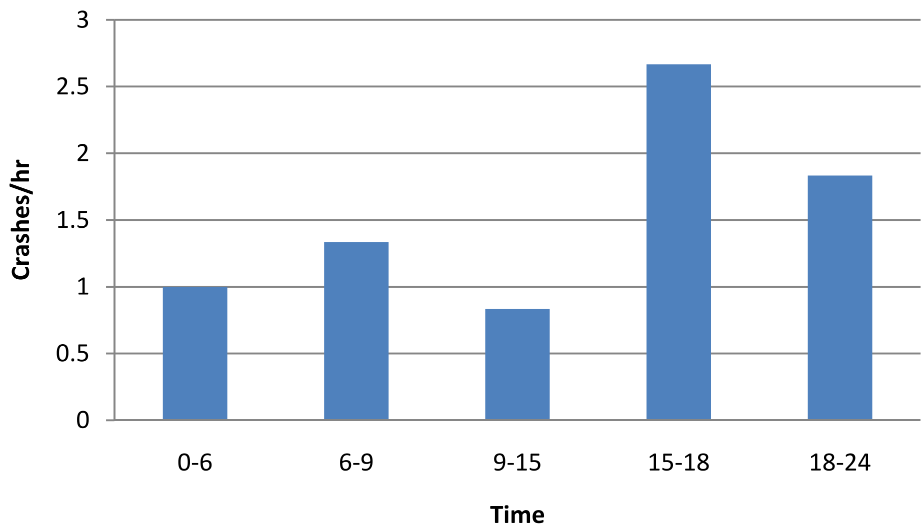
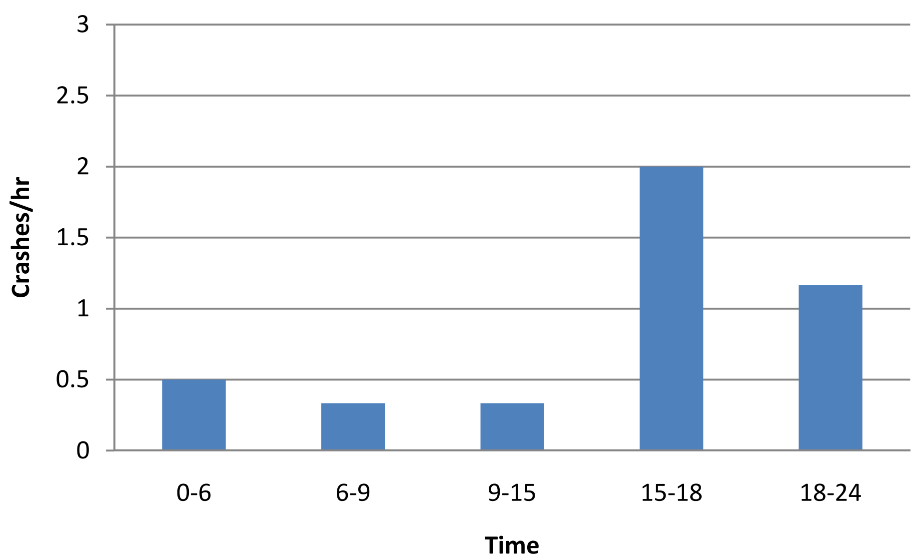
| | |
|---|---|
| All crash |  |
| Multi-vehicle crash only |  |

**Fig. 3-20 Crash distribution in 24 hours, aggregated based on time of day for study site 3 (2003-2007)**

Summaries of the crash data for all three study sites are presented on Table 3-9 and Table 3-10. Table 3-9 shows a summary of crash data for three sites based on the manner of collision. The crashes shown on this table were categorized as rear-end, side-swipe, and others based on the information on the manner of collision included in the CRIS database. The numbers shown on this table are the total numbers of crashes in five years for each study site.

Table 3-10 summarizes that crash data on the basis of exposure. It shows total number of crashes in a 5-year period. Number of crashes per mile was also calculated and shown on this table. This number computed by the total number of crashes divided by the length of each study site. This table also has number of crashes per 1,000 vehicle-miles for each study site. This number is the total crashes at each study site divided by the exposure. The exposure is the total product of number of vehicles travelling through each study site per day and the length of that particular weaving section.

**Table 3-9 Summary of Crash Data by Manner of Crash for 3 Study Sites (2003-2007)**

| Site | | All crashes | | | Multi-vehicle crashes only | | |
|---|---|---|---|---|---|---|---|
| | | Rear-end | Side swipe | Others | Rear-end | Side swipe | Others |
| 1 | SH 288 SB | 7 | 2 | 9 | 7 | 2 | 0 |
| 2 | IH 45 NB | 32 | 19 | 28 | 32 | 19 | 0 |
| 3 | US 67 SB | 13 | 6 | 15 | 13 | 6 | 0 |

**Table 3-10 Summary of Crash Data for 3 Study Sites (2003-2007)**

| Site | | Segment Length (mile) | Daily Traffic (veh/day) | All crashes | | | Multi-vehicle crashes only | | |
|---|---|---|---|---|---|---|---|---|---|
| | | | | Total crashes | Crashes per mile | Crashes per 1,000 vehicle-miles | Total crashes | Crashes per mile | Crashes per 1,000 vehicle-miles |
| 1 | SH 288 SB | 0.59 | 82,000 | 18 | 30.7 | 0.37 | 9 | 15.3 | 0.19 |
| 2 | IH 45 NB | 1.10 | 95,000 | 79 | 71.9 | 0.76 | 51 | 46.4 | 0.49 |
| 3 | US 67 SB | 0.60 | 45,500 | 34 | 57.0 | 1.25 | 19 | 31.8 | 0.70 |

## 3.3 Chapter summary

This chapter describes the data collection process. Geometric and traffic operation data were collected using various tools and equipment. In addition to the TxDOT database, the popular satellite mapping application Google Earth was employed to support the study site identification and selection. Three study sites located in Houston and Dallas areas were selected. Google Earth was also used for collecting the major geometric configuration information of the selected study sites. To collect traffic data, traffic surveillance cameras and pneumatic tubes were employed. The tubes collected speed, volume and acceleration data and the video provided supplemental volumes and lane changes. The video data were reduced by playing the videos and counting manually. Crash data were extracted from the CRIS database.

This chapter also has a summary of data. The graphs and tables provide general ideas about the data characteristics. The next chapter provides details about the traffic conflict technique adopted for this study. The next chapter also describes the application of microscopic simulation model to measure the traffic conflicts between vehicles in the traffic stream. The simulation model calibration process and configuration of simulation outputs are also described in this next chapter.

# CHAPTER IV

# TRAFFIC CONFLICT TECHNIQUE AND MEASURING SAFETY PERFORMANCE USING MICROSCOPIC SIMULATION MODEL

Section 4.1 of this chapter provides details on the application of traffic conflict technique to measure the safety performance in this study. Descriptions of configuration microscopic simulation model, extraction of simulation outputs and computation of traffic conflicts are also included in this first part of this chapter. The second section of this chapter provides brief descriptions on the selection of microscopic traffic simulation model and information on the selected model. Details on calibration of simulation model are presented under Section 4.3; and Section 4.4 describes the simulation process as well as provides a brief summary of simulation outputs and initial processing of the raw data.

## 4.1 Traffic conflict technique and surrogate measure of safety performance

As presented in Chapter II, traffic conflicts have been used as surrogate safety measures and to a certain extent, the traffic conflict technique has demonstrated a great potential of using simulation to evaluate safety performance of roadway facilities. The traffic conflict technique attempt is to evaluate the real-time interactions of vehicles to identify and measure the conflicts arisen because of those interactions within the simulated traffic streams and these conflicts reflect the potential of crashes. These surrogate safety measures are more of reflecting higher risk driving condition than an actual representation of reported crash history. Among a number of surrogate safety measures have been established and investigated, Deceleration Rate to Avoid Crash (DRAC) was selected for further investigation in this study. This safety measure considers the interactions between any two consecutive vehicles in the traffic stream and takes the spacing, speed differential, and the response of the following vehicle in the possible crash occurrence. It is expressed as a function of time/space and speed differential between any pair of consecutive vehicles in the traffic stream. It is the required deceleration of the following vehicle to avoid crash because of invasive action of the leading vehicle; either that invasive action is slowdown or that is a lane-changing maneuver. These concepts are illustrated in Figure 4-1 and Figure 4-2 and formulated as Equation 4-1.

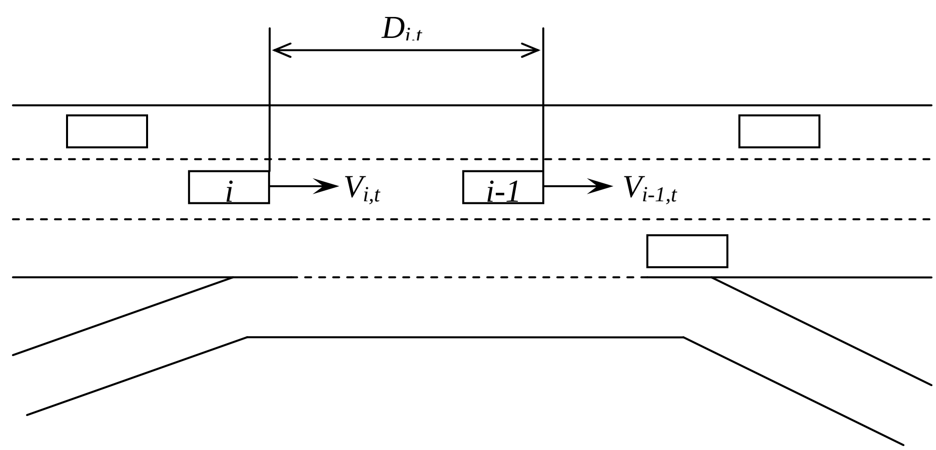



**Fig. 4-1 Conflict between two vehicles caused by slowing-down leading vehicle**

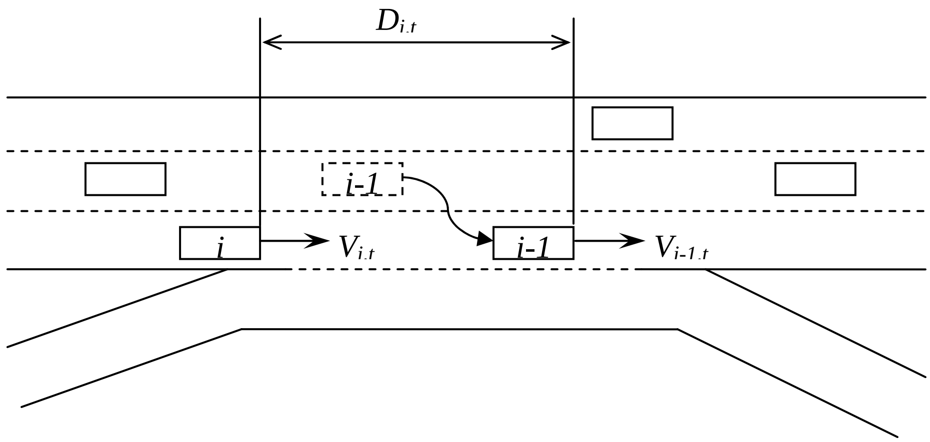



**Fig. 4-2 Conflict between two vehicles caused by lane-changing leading vehicle**

$$DRAC_{i,t+1} = \frac{\left(V_{i,t} - V_{i-1,t}\right)^2}{D_{i,t} - L_{i-1}} \qquad (4-1)$$

Where:

- $DRAC_{i,t+1}$ is Deceleration Rate to Avoid Crash associated with the following (response) vehicle at time step t+1;
- $V_{i,t}$ is speed of the following (response) vehicle at the end of time step t;
- $V_{i-1,t}$ is speed of the leading vehicle at the end of time step t;
- $D_{i,t}$ is the headway distance between the two vehicle at the end of time step t; and,
- $L_{i-1}$ is length of the leading vehicle.

In a study conducted by Cunto and Saccomanno (2007) and their subsequent investigation effort on the issue (Cunto and Saccomanno, 2008), a cash potential index was defined to measure the safety performance and reflect the likelihood of crash occurrence. In another study, Ozbay et al. (2008) included speed as a component of their modified surrogate safety measure and suggested that it would be better indicator of safety performance. By adopting these concepts, a crash potential index was established as below:

$$\begin{cases} \mathrm{CPI_i} = \dfrac{\sum_{t=ti_i}^{t=tf_i}\left(\dfrac{\mathrm{DRAC_{i,t}} - \mathrm{MADR_i}}{\mathrm{MADR_i}}\Delta t.\,V_{i,t}.\,b\right)}{T_i} \\ b = 1 \text{ if } \mathrm{DRAC_{i,t}} > \mathrm{MADR_i} \\ b = 0 \text{ if } \mathrm{DRAC_{i,t}} \leq \mathrm{MADR_i} \end{cases} \qquad (4-2)$$

Where:

- $CPI_i$ is the crash potential index for vehicle i;
- $ti_i$ is the initial simulation time interval for vehicle i;
- $tf_i$ is the final simulation time interval for vehicle i;
- $DRAC_{i,t}$ is the deceleration rate required to avoid crash for vehicle at time t as Equation 4-1;
- $MADR_i$ is the maximum available deceleration rate for vehicle i. This is a unique value, reflecting the braking capability of a particular vehicle in a particular condition of roadway, vehicle, tires etc.;
- $\Delta t$ is the simulation time interval;

- b is a binary state variable which is equal to 1 if at a time t, $DRAC_{i,t}$>$MADR_i$ and it is equal to 0 otherwise; and,
- $T_i$ is the total simulation time for vehicle i (total time vehicle i in the network).

As previously described, MADR is a unique value representing the breaking capability of each vehicle in the traffic stream in a particular condition. Considering the MADR for the entire traffic stream with different vehicles in various conditions, it was suggested to follow a normal distribution with truncated tails (Cunto and Saccomanno, 2007; 2008). There have been different suggestions for the parameters of truncated normal distribution. However, no consistant information was provided. After consulting some TTI's senior researchers, taking their experiences and expertise in to consideration together with suggestions in Cunto and Saccomanno's work (2007, 2008) and Fricke's book (Fricke, 1990), a decision was made to use a truncated normal distribution with the parameters as shown on the table 4-1.

**Table 4-1 Parameters of Truncated Normal Distribution for MADR**

| Parameter | Value | *Note* |
|---|---|---|
| Mean ($ft/s^2$) | 22.54 | *(0.7g)* |
| Standard Deviation ($ft/s^2$) | 1.64 | *($0.5 m/s^2$)* |
| Upper limit ($ft/s^2$) | 38.64 | *(1.2g)* |
| Lower limit ($ft/s^2$) | 12.88 | *(0.4g)* |

*(note: $g = 32.2\ ft/s^2$)*

## 4.2 Selection of simulation model

This section briefly describes the simulation selection process. More information to justify the use of VISSIM, the selected microscopic simulation model for the aforementioned TTI's project (Fitzpatrick et al., 2009), for this specific study is also provided in this section.

A desired traffic simulation model should have capability to replicate traffic operations under various geometric and operational conditions and should be able to provide outputs needed for analyses leading to the predetermined objectives of the study. Any model which considered being appropriate for this study should meet several basic criteria as below:

- The simulation model should simulate traffic operations at a microscopic level. Therefore, only microscopic traffic simulation model can be used.
- The simulation model should have a sophisticated driver behavior model, including car-following and lane-changing, in a wide range of situations. This model should be able to model the intensity and complexity of the weaving section operations.
- The simulation model should be able to keep track of individual vehicles. The simulation model should also have the ability to record and produce outputs with details on the status of each vehicle such as location, speed, acceleration etc., as the vehicle travels through the system.

Among a significant number of currently available microscopic traffic simulation models, three models widely used in both traffic research and engineering practice communities were taken into consideration in this study: VISSIM, CORSIM and PARAMICS. These three candidates were compared based on their characteristics and available features. A quick summary of the major features and characteristics is presented on Table 4-2.

The comparison on Table 4-2 shows that these three simulation models are relatively similar in term of capabilities and features. However, having point-based MOEs is critical in this study; therefore CORSIM was not considered to be appropriate. The candidate simulation models were reduced to only two, which were VISSIM and PARAMICS.

**Table 4-2 Feature Comparisons for Three Candidate Traffic Simulation Models**

| **Features** | **VISSIM** | **CORSIM** | **PARAMICS** |
|---|---|---|---|
| Graphical User Interface | Yes | Yes | Yes |
| Text Editor | Yes | Yes | Yes |
| Developing Tool | Yes – VAP | Yes – RTE | Yes – API |
| Batch Mode | Yes | Yes | Yes |
| Traffic Control | Yield, stop, ramp metering, etc. | Yield, stop, ramp metering, etc. | Yield, stop, ramp metering, etc. |
| O-D Matrix | Yes | Yes | Yes |
| MOE | Point/Link-based | Link-based | Point/Link-based |
| Animation | 2-D & 3-D | 2-D | 2-D & 3-D |
| *Notes:*<br>▪ *VAP: Vehicle Actuated Programming*<br>▪ *RTE: Run-Time Extension*<br>▪ *API: Application Programming Interface* | | | |

Both of these two microscopic simulation models have been successfully used for simulating freeway traffic and weaving section operations. In recent studies, Ozbay et al. (2008) used PARAMICS while Cunto and Saccomanno (2008), Vu et al. (2007), and Roess and Ulerio (2009) used VISSIM.

Because both VISSIM and PARAMICS were qualified from capability and feature standpoints, the final decision was made by the availability and accessibility of the software. By taking this as a key point in the decision making process, VISSIM definitely stood out because it was made available to students on most of the computers in the Zachry Department of Civil Engineering at Texas A&M University. Therefore, VISSIM was selected for this study.

***Characteristics about VISSIM:***

VISSIM is a microscopic, time step and behavior based traffic simulation model developed by PTV Planung Transport Verkehr AG (Germany), distributed and supported in the United States by PTV America. VISSIM has been widely used and regarded as one of the most powerful traffic simulation models available. It has been applied in simulating analyzing all ranges of facilities, from a small intersection to an entire metropolitan area. It has also been used for modeling modes of travel as well as multi-modal system, including cars, trucks, buses, rails, bicyclists and pedestrians. VISSIM uses a psycho-physical driver behavior model developed by Wiedemann in 1974 (Wiedemann, 1974).

### 4.3 Calibration of simulation model

This study used a part of the data used for the TTI's 0-5860 project (Fitzpatrick et al., 2009) and the decision was made to employ VISSIM as well, so the simulation model calibration process which had been done for the aforementioned project by the TTI's research team could be applied to a certain extent, with the consent of the relevant TTI's researchers. The calibration procedure described herein was partly covered under scope of this 0-5860 project.

The entire process of initial test, calibration and final run of the simulation was conducted on "simulation test-beds". A simulation test-bed is a coded network of a transportation facility in which all the geometric configurations and measures, as well as all model parameters remain unchanged; while simulation inputs and routing distribution can be varied. Three simulation test-beds for three different study sites were developed based on actual geometric and traffic data collected as presented in Chapter III "Data collection, reduction and assembly". These field data-based test-beds were used for calibrating the simulation models as well as for the final simulation runs.

The simulation test-beds were developed following a number of major steps as below:

***Step 1:*** Coding the network using the graphical editing tools. Each weaving section was composed of links and connectors to connect those links and create traffic flow from link to link. As described in Chapter III, each study site was divided into five sections labeled from A to E; and the corresponding test-bed was also coded that way. Every section considered as a separate link within the network. The entrance and exit ramps were also coded as separate links and connected to the mainline freeway by connectors. The weaving sections were coded based on geometric data as shown on Table 3-1 and Table 3-2. By considering influence areas as described in section 3-2 and illustrated in Figure 3-14 of Chapter III, these particular parts of the study sites were also included in the coded networks. However, the upstream parts of the mainline freeway and the entrance ramp were not simply coded with lengths of 1,000 feet and 500 feet relatively. Those elements were coded with lengths of at least 2,000 feet to avoid the effects of the activities within the weaving areas as well as any possible occurrence of long queues to the input vehicle distributions. Figure 4-3 provides a clear illustration of the above descriptions.

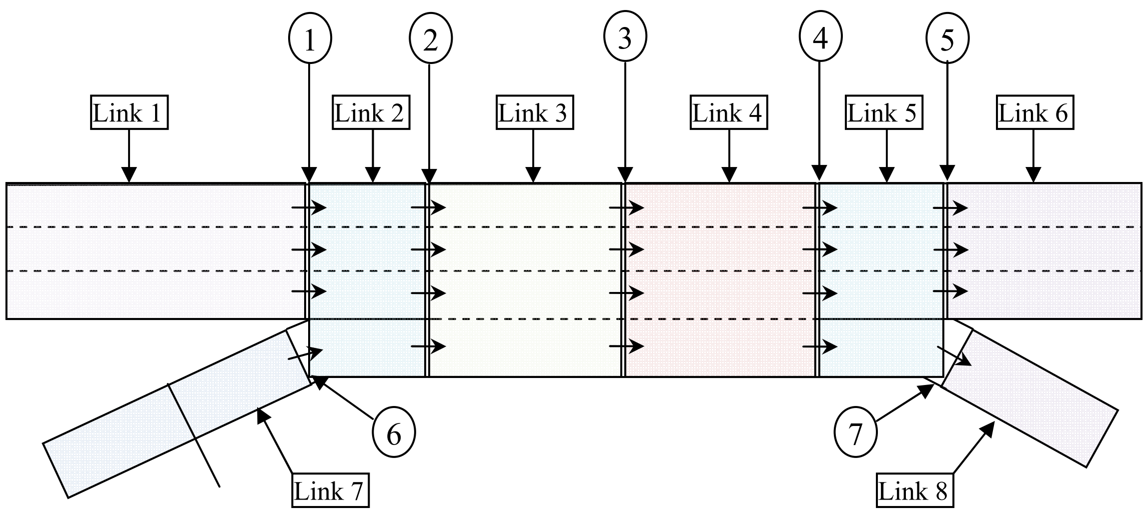



**Fig.4-3 Weaving section elements coded in VISSIM**

***Step 2:*** Defining the base data for simulation. An hour of speed data collected as described in Chapter III were used to define the desired speed distributions. Weidemann 99 was selected as the car following model used in this study. All default parameters were kept and only those parameters found to be significant to the targeted simulation outputs were later adjusted to calibrate the models as presented the next part of this section.

***Step 3:*** Defining routing decisions, vehicle inputs and other simulation model inputs. Routing decisions and vehicle inputs were estimated from the field data collected as described in Chapter III.

***Step 4:*** Defining the simulation outputs and simulation parameters. In this step, the simulation outputs were defined based on the on the necessary data for calibrating the simulation model and running the final simulation. For the purpose of simulation model calibration, the outputs were defined so that the simulated results could be compared to the observed traffic data. The effort of the process was to match the simulated traffic data with those collected from the field. Data collection points were set at positions corresponding to those locations that traffic data were collected. For the final simulation runs, no data collection points were necessary. Each individual vehicle was tracked throughout the weaving area and the information on position, speed, acceleration, and interaction etc., associated with each particular vehicle was recorded. These raw data would later be processed to compute the CPI using Equations 4-1 and 4-2.

After simulation test-beds were created following the above four-step procedure. The objective of the calibration process was to minimize the difference between simulated and the observed data by adjusting and fine-tuning certain parameters of the simulation test-beds. Presenting this in another way, this process was to find a set of *n* parameters:

$$(p_1, p_2, p_3, \ldots, p_n)$$

so that the sum:

$$\sum\left[x_i^{sim}(p_1, p_2, p_3, \ldots, p_n) - x_i^{obs}\right]^2 \qquad (4-3)$$

could be minimal.

In the above expression 4-3:

- $x_i^{sim}$ is the simulated data collected at the collection point *i*; and,
- $x_i^{obs}$ is the observed data collected from the field at the location corresponding to data collection point *i*.

The data used for calibration process included vehicle speeds and accelerations collected at several particular locations where pneumatic tubes were placed as described in Chapter III. The number of lane-changing vehicles observed within each section of the weaving area was also used as a criterion to justify the matching of the simulated weaving area to the actual one.

Behaviors of the traffic stream in the simulation model are significantly influenced by a wide range of parameters, including car-following and lane-changing parameters. An effort of adjusting every parameter of VISSIM to find the best combination of parameter was obviously an overambitious if not impossible target. It could be a very complex and computationally intensive optimization problem. Therefore, instead of trying to seek an absolute solution for this optimization problem:

$$\mathrm{Min}\left\{\sum\left[x_i^{sim}(p_1, p_2, p_3, \ldots, p_n) - x_i^{obs}\right]^2\right\} \qquad (4-4)$$

The calibration effort focused on finding a combination with a limited number of parameters which the following conditions could be satisfied:

$$\frac{\left|x_i^{sim}(p_1, p_2, p_3, \ldots, p_k) - x_i^{obs}\right|}{x_i^{obs}} \leq \varepsilon \qquad (4-5)$$

In the above Equation 4-5, ε is the acceptable error. The allowable error ε was selected to be 0.1 in this study. As being previously discussed, it was not possible to take every parameter into consideration for the calibration process. A limited number of parameters must be selected for being adjusted and fine-tuned to archive the calibration objective as Equation 4-5. The aforementioned TTI's project found that speeds and lane-changes were sensitive to eight VISSIM parameters including six lane-changing and two car-following parameters. In a 2008 study, Cunto and Saccomanno found that their DRAC-based CPI was significantly sensitive to three parameters and two combinations of those three parameters (Cunto and Saccomanno, 2008). Taking all findings of both

TTI's project and Cunto's study, the parameters listed on Table 4-3 were selected for calibration of the VISSIM simulation model.

**Table 4-3 Simulation Input Parameters**

| Factor | Parameter |
|---|---|
| **Base data distributions** | |
| A | Desired speed |
| B | Desired deceleration |
| **Car-following parameters** | |
| C | Min looking ahead distance |
| D | CC0 - Standstill distance |
| E | CC1 - Headway time |
| **Lane changing parameters** | |
| F | Max Deceleration of lane changing vehicle |
| G | Max Deceleration of trailing vehicle |
| H | Accepted Deceleration of lane changing vehicle |
| I | Accepted Deceleration of trailing vehicle |
| J | Safety distance reduction factor |
| K | Max Deceleration for Cooperative braking |

The calibration process involved iterative searches to find combinations of parameters which satisfy the Equation 4-5 for each study site. Seventy five minutes of data collected from each study site as described in Chapter III were used in the calibration process. The first fifteen minutes of data were used for "warm-up" the simulation model, not actually used for matching the observed and simulated data. Only one hour of data, aggregated in fifteen minute intervals were actually used to calibration of VISSIM. Traffic flow rates and routing decision information (calculated from number of lane-changing maneuvers) were used as inputs for simulation model. The findings of the TTI's 5860 project and the Cunto's study allow this study to focus on a small range of parameter values around the parameters found in these two studies. In fact, the simulation process conducted in this study was pretty much a fine-tuning of what had been done for the TTI's project and findings of Cunto's study. In each iterative search step, three simulation runs were performed with three different random seeds. The three-run average values of the simulation outputs including flow rates, mean speeds, lane-changes and mean accelerations, were then compared to those corresponding observed data. The simulation model parameters were adjusted until Equation 4-5 was satisfied.

During data collection and reduction, an operational characteristic within the weaving areas was observed. A review of the videos at all study sites suggested that driver behaviors did not remain the same while travelling along the entire length of the weaving area. While trying to merge or diverge, if the drivers were not able to find sufficient gaps to perform lane-changing maneuvers, they would become more and more aggressive as they got closer to the exit ramp. At the beginning of the weaving area, the drivers were likely to be more relaxed. As travelling downstream, the drivers were willing to accept shorter gaps for both on-ramp and off-ramp vehicles. This characteristic is illustrated in Figure 4-4.

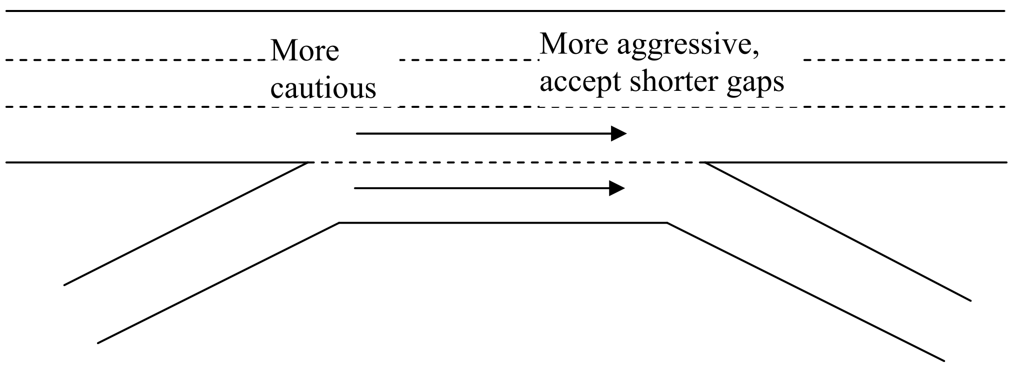



**Fig. 4-4 Progressive driver behaviors**

The observation of progressive driver behaviors and gap acceptance seemed to be logical to the intuitive assumption. This led to the decision of modeling different driver behaviors for different sections within the weaving area, starting with relaxed and unaggressive behaviors at beginning of the weaving area and ending with aggressive driving behaviors at the end of the weaving area. Based on this principle, three study sites were modeled in VISSIM by dividing those sites into four zones corresponding to four sections A, B, C and D, as described in Chapter III and specifying four different driving behavior categories named “Relaxed”, “Normal”, “Moderately Aggressive”, and “Aggressive” for these four zones respectively. These descriptions are illustrated in Figure 4-5.

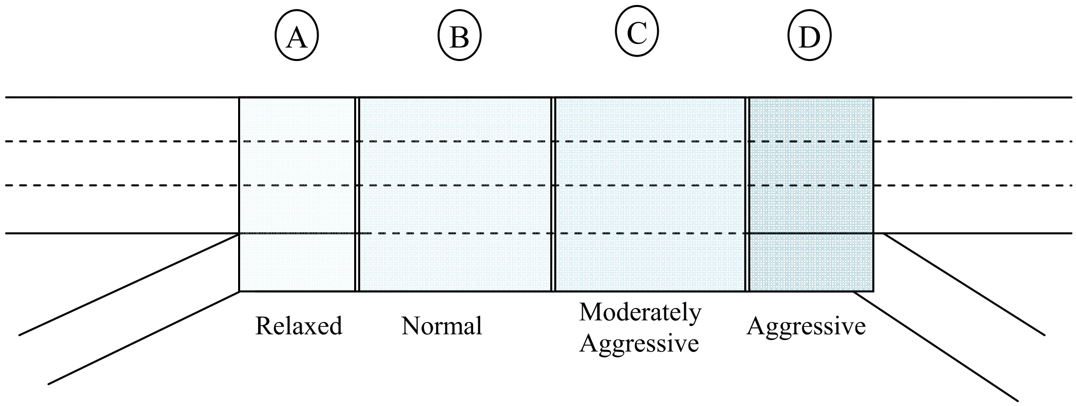



**Fig.4-5 Simulation test-bed with different driving aggressiveness**

The parameters were fine-tuned within a small range around the values found in the TTI's 5860 project until the differences between simulated and the observed data were less than ten per cent as shown in Equation 4-5. The comparisons were made for flow rates, speeds, accelerations and lane-changes; all aggregated in 15-minute intervals. This process was done and repeated for all three study sites until a parameter combination was found to satisfy Equation 4-5. The calibrated VISSIM models for all three study sites were used for final simulation runs as presented in the following section 4.4.

## 4.4 Simulation run and initial processing of simulation outputs

The final simulation runs were performed on the simulation models which had been calibrated as presented in the preceding section. After being calibrated, the simulation models would be reconfigured with new inputs of traffic flow and origindestination (O-D) pattern over time. Traffic flows for the final simulation runs were estimated from data collected by pneumatic tubes and video as presented in Chapter III. The tubes recorded continuous traffic data in the right and auxiliary lanes in three to four consecutive days. However, no tubes were placed in the left lanes of the multilane freeways. Data for those lanes had to be estimated based on both 24-hour continuous tube data and video data. The number of vehicles travelling in the middle lane(s) and median lane were counted from the videos for some hours. Use these counted flow rates in combination with the flow rates collected from pneumatic tubes to estimate the volume distribution factor as the following:

$$k_i = \frac{V_i}{\frac{\sum_{j=1}^{n} V_j}{n}} \qquad (4-6)$$

Where:

- $k_i$ is a factor representing the distribution of traffic volume for lane i;
- $V_i$ is traffic volume for lane i, either from counting on the videos or from tube data; and,
- n is number of lanes.

The volume distribution factor k was used to estimate hourly volumes for all travel lanes as well as the total volume in a 24-hour period as the following Equation 4-7 and Equation 4-8:

$$V_{i,t} = \frac{k_i}{k_1} . V_{1,t} \qquad (4-7)$$

$$V_t = \sum_{i=1}^{n} V_{i,t} \qquad (4-8)$$

Where:

- $V_{i,t}$ is volume estimate for lane i, for $t^{th}$ hour of the day (t ranges from 0 to 23 for a day of 24 hours);
- $k_i$ is volume distribution factor for lane i;
- $V_t$ is total volume for all travel lane, for $t^{th}$ hour of the day (t ranges from 0 to 23 for a day of 24 hours); and,
- n is number of travel lanes.

In the above estimates, $k_i$ was not expected to be constant throughout the 24-hour period. Hourly volumes were not expected to be the same for everyday throughout the year either. Despite this limitation, these estimates, to a certain extent, provided information on the magnitude of traffic activities within the weaving areas from which safety performance indicators might be linked to. Therefore, these estimates were appropriate to be used as inputs for the simulation models.

In addition to the inputs, simulation outputs for the final simulation had to be defined as well. In the final simulation process, output data were not collected at particular points (data collection points). VISSIM was set to tracks and records information of all vehicles at every 0.1 second simulation time step. Based on the required data for identifying the vehicles in conflict and calculating the Crash Potential

Index (described in Section 4.1 of this chapter, Equations 4-1 and 4-2), as well as data necessary for interim steps, the following vehicle record parameters were configured as the simulation outputs. Table 4-4 provides a list with descriptions of the vehicle record parameters.

**Table 4-4 Output Vehicle Record Parameters**

| **Header** | **Vehicle Record Parameter** | **Definition** |
|---|---|---|
| VehNr | Vehicle number | Vehicle number |
| t | Simulation time (s) | Simulation time |
| Link | Link number | Number of the active link |
| x | Link Coordinate (ft) | Link Coordinate [m] at the end of the simulation step |
| LVeh | Leading vehicle | Number of the next (not necessarily relevant) vehicle downstream |
| Head | Headway (ft) | Distance to the next (not necessarily relevant) vehicle downstream |
| Length | Vehicle length (ft) | Vehicle length |
| Lane | Lane number | Number of the active lane |
| DLn | Destination lane | Destination lane number of current lane change |
| V | Speed (mph) | Speed at the end of the simulation step |
| Dv | Speed difference (mph) | Speed relative to the preceding car for the simulation step (>0 = faster) |

Ten simulation runs were performed with ten different random seeds for each study site. Because the simulation required a large amount of computing power, the simulation was set up to run on multiple personal computers in the Civil Engineering computer lab. The VISSIM simulation outputs were written into text files which required additional processing steps before the data could be imported into any statistical analysis software packages for further steps in the workflow. A small program was built to automate this computationally intense process, from reading the data files, step by step, vehicle by vehicle, to importing the data into SAS (SAS, 2004) for further processing and analysis.

After being imported into SAS, the unwanted data would be eliminated. As described in Section 4-3 of this chapter, the simulation test-beds were built with extended upstream entrance ramp and mainline freeway. Therefore, data on all activities occurred beyond the areas of interest (1,000 ft and 500 ft upstream and downstream on the mainline and entrance ramp respectively) were deleted. After importing data into SAS and performing the initial processing of raw data, the data were ready for further processing and analysis as described in Chapter V.

## 4.5 Chapter summary

This chapter describes the principles of traffic conflict technique adopted for this study. This chapter provides details of conflicts between vehicles and the mechanism of crash occurrence and crash avoidance at microscopic level. The formulation of the crash potential index which reflects the level of conflict and therefore potential of crash occurrence is also presented in this chapter.

The selection of an appropriate microscopic simulation software package is also described in this chapter. The popular PTV's VISSIM microscopic simulation model was selected for this study because of its appropriate capabilities and widely availability. The chapter also provides a description about the calibration process carried out for the VISSIM simulation model. The process includes the selection of input parameters and configuration of desired outputs.

The simulation runs and initial data processing carried out are also described in this fourth chapter. Ten simulations with ten different random seeds were conducted. The simulation outputs in VISSIM's raw data format were read reduced and imported into SAS for further processing and analysis.  The next chapter provides details of the data analysis and the results and the analyses carried out.

# CHAPTER V

# DATA ANALYSIS

This chapter includes a description about how the data analysis was carried out and the analysis results. The objective of this process was to analyze and evaluate the correlation between the simulated surrogate safety measures and the actually observed crash data. Pearson correlation coefficient, Spearman rank correlation coefficient and regression model were employed to evaluate the relationship between the simulated data and the field-collected crash ones. The effect of weaving activity within the weaving sections is also examined using Spearman rank correlation coefficient anlaysis. Together with the analyses and the results presented, relevant discussions are also included herewith.

The first section of this chapter provides a brief summary of a weaving analysis for all three study sites using the HCM methodology. This is to preliminarily evaluate the operational conditions at the study sites. Section 5.2 of this chapter describe the data processing and Sections 5.3, 5.4 cover the simple linear analyses and regression modeling respectively. Section 5.5 provides a summary of the analysis to examine the relationship between weaving volume and crashes, number of vehicles in conflict and crash potential index.

## 5.1 Weaving section LOS analysis using HCM method

Traffic operations at three study sites were analyzed and evaluated using the methodology introduced in the 2000 HCM (HCM, 2000). The objective of this analysis is to provide a general idea about the operational conditions of the selected study sites under daily peak traffic conditions.

The 2000 HCM weaving operation analysis procedure is based upon predicting weaving and non-weaving speeds within the weaving areas. These predicted speeds then will subsequently be used to calculate the vehicle density within the weaving area. Level of Service is determined based on this vehicle density. The analysis procedure included in the 2000 HCM requires a lengthy workflow on computation worksheets (HCM, 2000). The calculation is not described in details here. Instead, only the calculation results and a number of key intermediate steps are presented herein.

One hour of traffic data collected from each study site during peak hour were used for this analysis. In addition to geometric configuration data of the weaving sections, traffic data needed include flow rates from the main line freeway and from the entrance ramp, as well as the flow distributions to the mainline freeway and to the exit ramp. Several input parameters such as PHF, Percentage of Truck were estimated from available traffic data collected from each study site. An assumption was made about the driver population as if most of them are regular commuters. A summary of analysis results are presented on the following Table 5-1.

**Table 5-1 Weaving Section LOS Analysis Results**

| Site | Operation Type | Weaving Speed *(mph)* | NonWeaving Speed *(mph)* | Overall Speed *(mph)* | Density *(pc/mile/ln)* | LOS |
|---|---|---|---|---|---|---|
| 1 (SH 288 SB) | Unconstrained | 40.7 | 49.2 | 48.0 | 38.3 | E |
| 2 (IH 45 NB) | Constrained | 41.3 | 62.0 | 55.0 | 33.3 | D |
| 3 (US 67 SB) | Unconstrained | 40.5 | 47.0 | 44.9 | 32.1.9 | D |

## 5.2 Data processing

The simulation outputs, after be initially processed as described in the preceding chapter, were put through further processing and analysis. As presented in Chapter IV, an assumption was made about the distribution of MADR. A random process was made to assign each simulated vehicle an MADR following a truncated normal distribution with parameters listed on Table 4-1. The MADR assignment process was automated and made randomly by employing the function NORMAL() in SAS (SAS, 2004). This SAS function is designed to randomly generate numbers following a standard normal distribution. The values generated were then transformed from a standard normal distribution into normal distribution using the following equation:

$$X = \sigma . Z + \mu \qquad (5-1)$$

Where:

- X is a normal random variable;
- $\sigma$ is standard deviation of the normal distribution; and,
- $\mu$ is mean of the normal distribution.

As previously described in Chapter IV, the VISSIM simulation model tracked and recorded information on every single-vehicle every one tenth of a second. The information recorded was used to compute the DRAC for each vehicle at every simulation time step using Equation 4-1. Because of vehicle interactions, the DRAC varied in time depending on speeds, accelerations, relative positions as well as interactions of the vehicles. The DRAC of each vehicle at any particular time step was then compared with the MADR associated with each vehicle to identify if the vehicle was in conflict with another vehicle in the traffic stream at that particular point of time. Subsequently, Crash Potential Index (CPI) was computed using Equation 4-2. The number of vehicles in conflict with other vehicles identified and CPI calculated were then aggregated in one-hour intervals. This computation process was repeated for all ten individual simulation runs for each study site before the results were averaged for the mean values of the ten simulation runs, also on the one-hour interval basis. Figures 5-1 and 5-2 show 24-hour plots of number of vehicles in conflict and CPI for ten individual simulation runs and the mean values for study site 1. Figure 5-3 shows the observed crashes, mean number of vehicles in conflict and mean CPI after being normalized for this study site. Similarly, Figures 5-4, 5-5 and 5-6 show plots of the results for study site

2. Figures 5-7, 5-8 and 5-9 show plots for study site 3. The graphs show that the fluctuation of the number of crashes is likely to have the same trend as the fluctuation of number of vehicles in conflict and CPI. This means number of crashes is likely to be high when number of vehicles in conflict and CPI are high. However, in some instances, it deviates from this trend. At all three sites, a significant number of crashes occur during night time (from midnight to 6 AM) but the simulated safety performance measures at all three sites tend to be very small. This may be attributed to the inability to reflect the fatigue and poorer driving performance of drivers during this period of time in the simulation model.

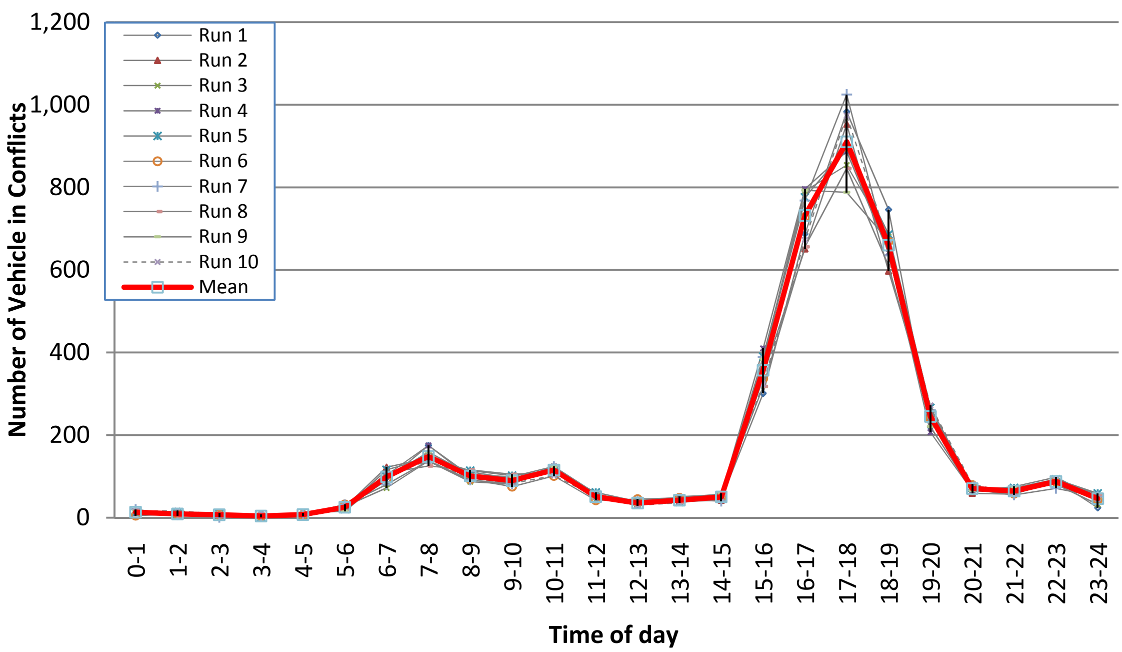



**Fig. 5-1 Number of vehicles in conflict in 24 hours for study site 1 (10 simulation runs)**

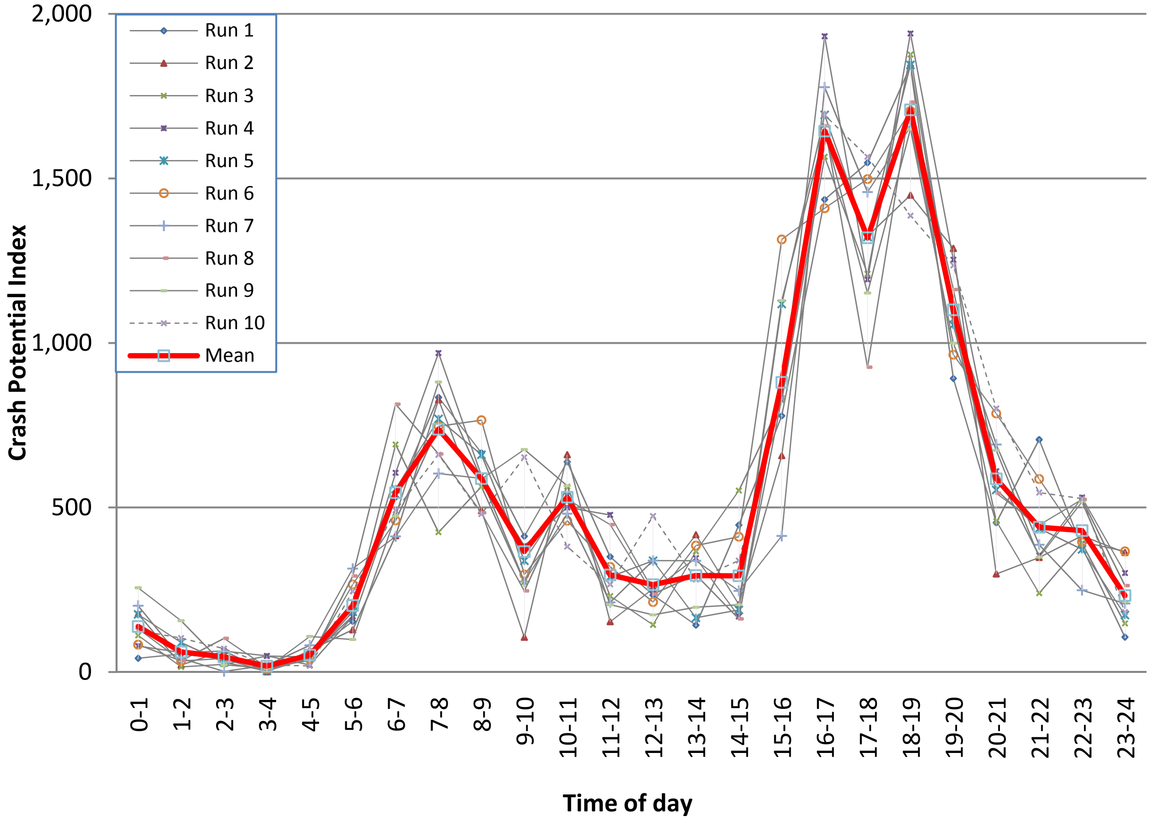
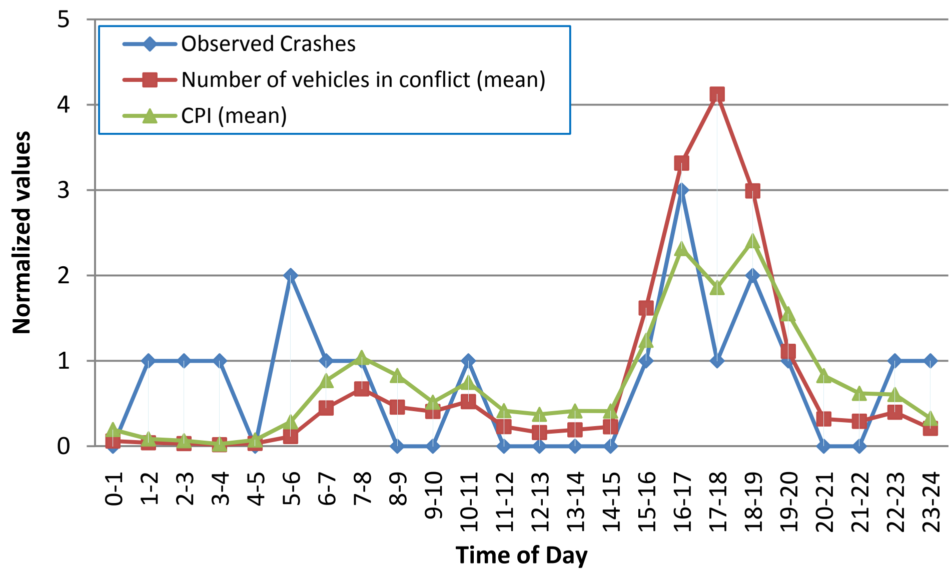



**Fig. 5-2 Crash potential index in 24 hours for study site 1 (10 simulation runs)**

All crashes




Multi-vehicle crashes only

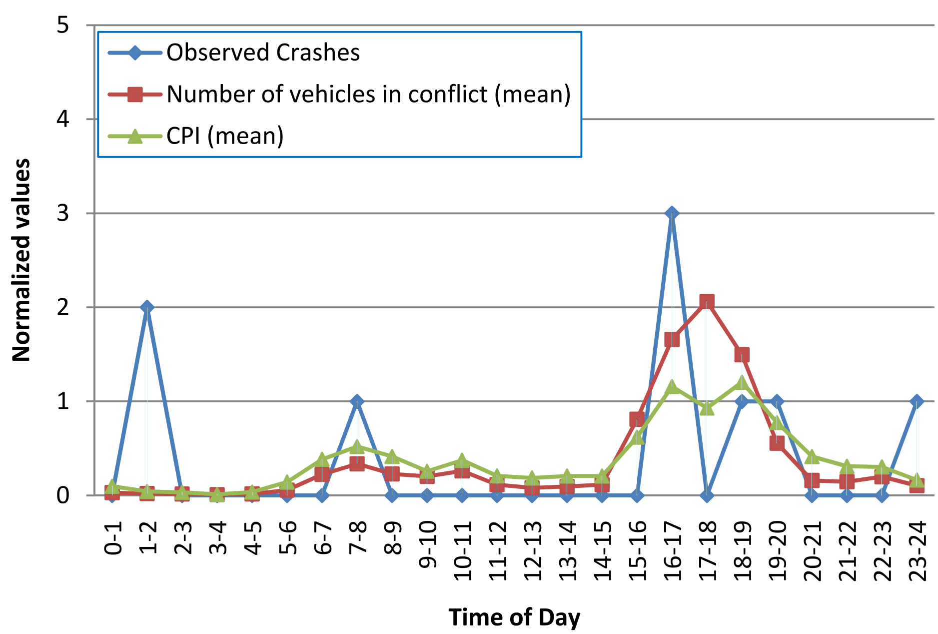



**Fig. 5-3 Observed crashes, normalized number of vehicles in conflict and normalized CPI in 24 hours for study site 1**

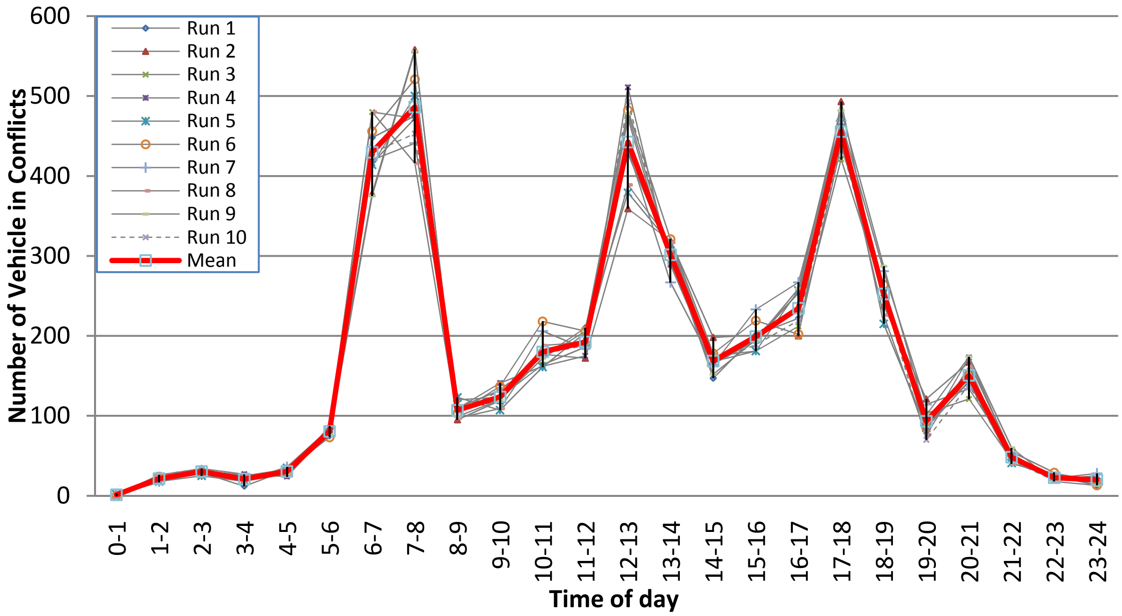



**Fig. 5-4 Number of vehicles in conflict in 24 hours for study site 2 (10 simulation runs)**

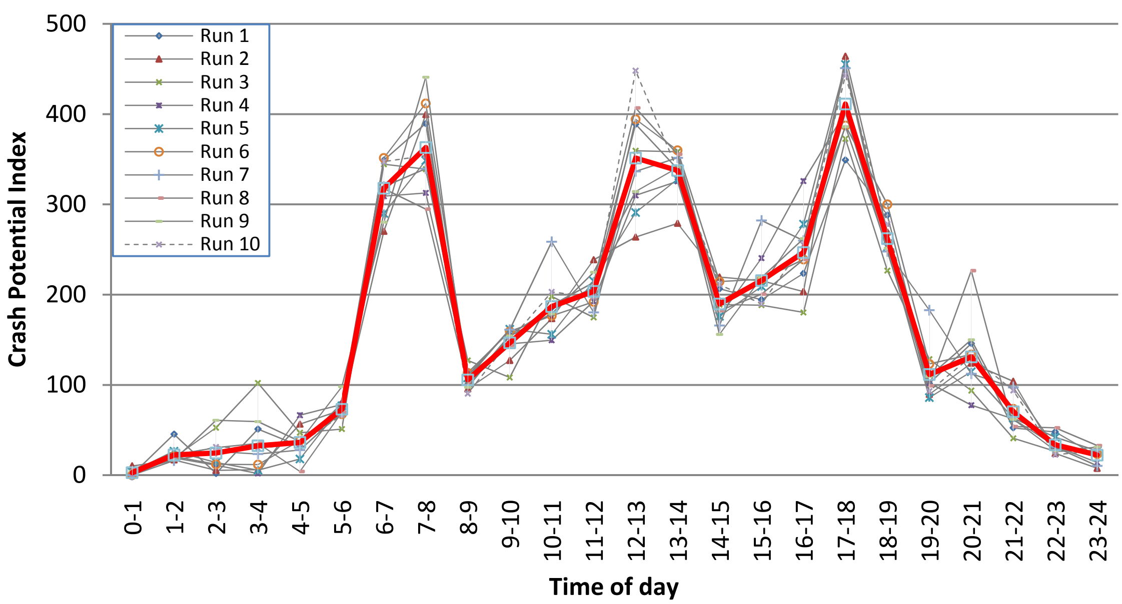
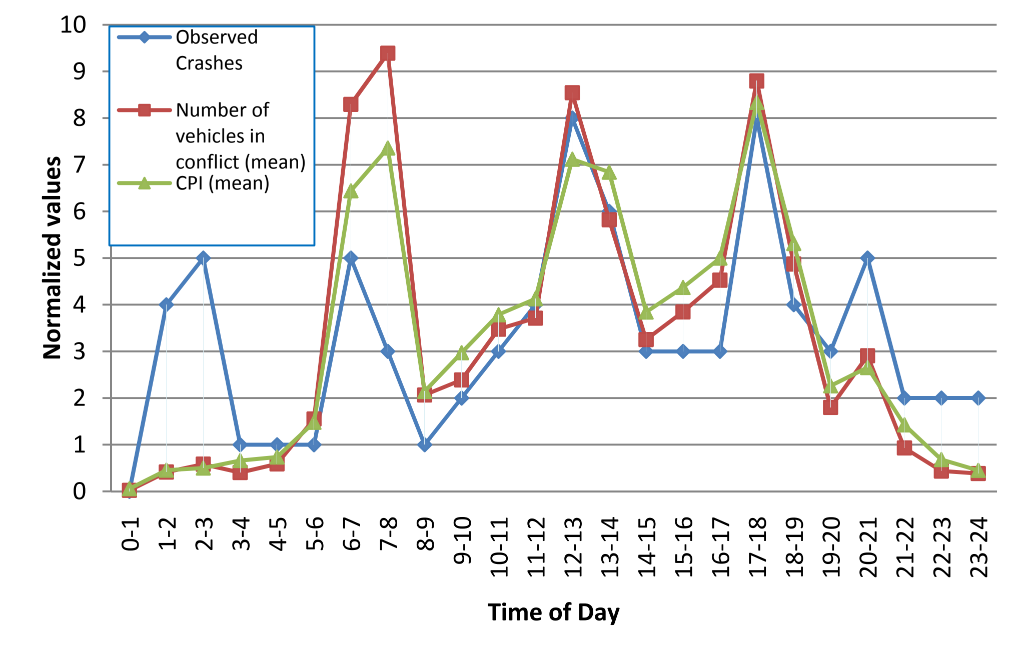



**Fig. 5-5 Crash potential index in 24 hours for study site 2 (10 simulation runs)**

All crashes




Multi-vehicle crashes only

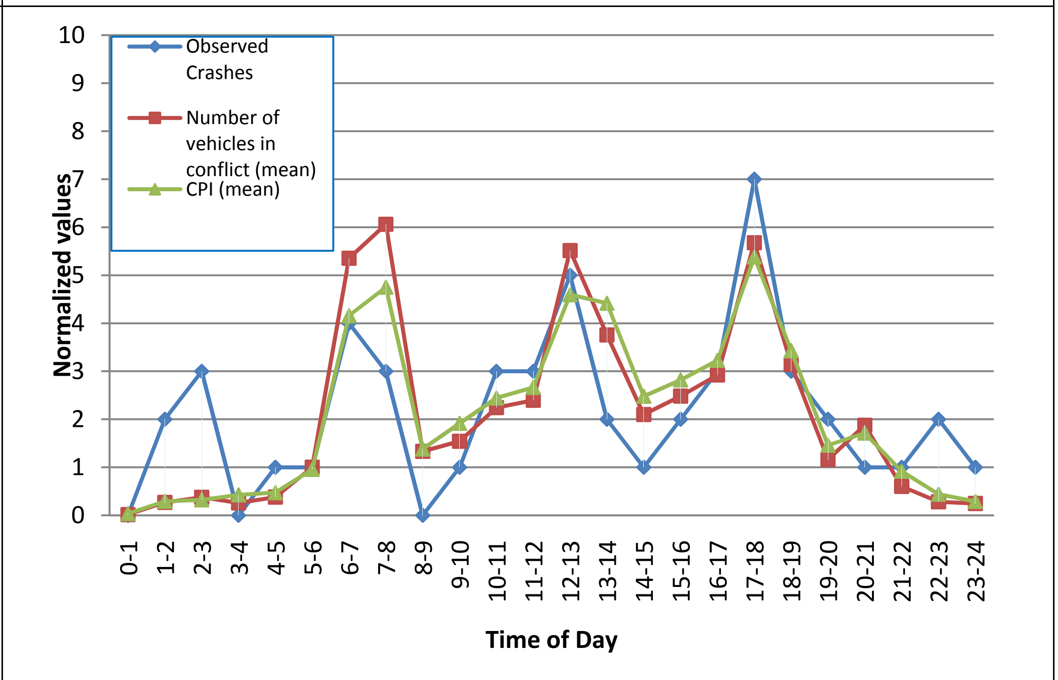



**Fig. 5-6 Observed crashes, normalized number of vehicles in conflict and normalized CPI in 24 hours for study site 2 (10 simulation runs)**

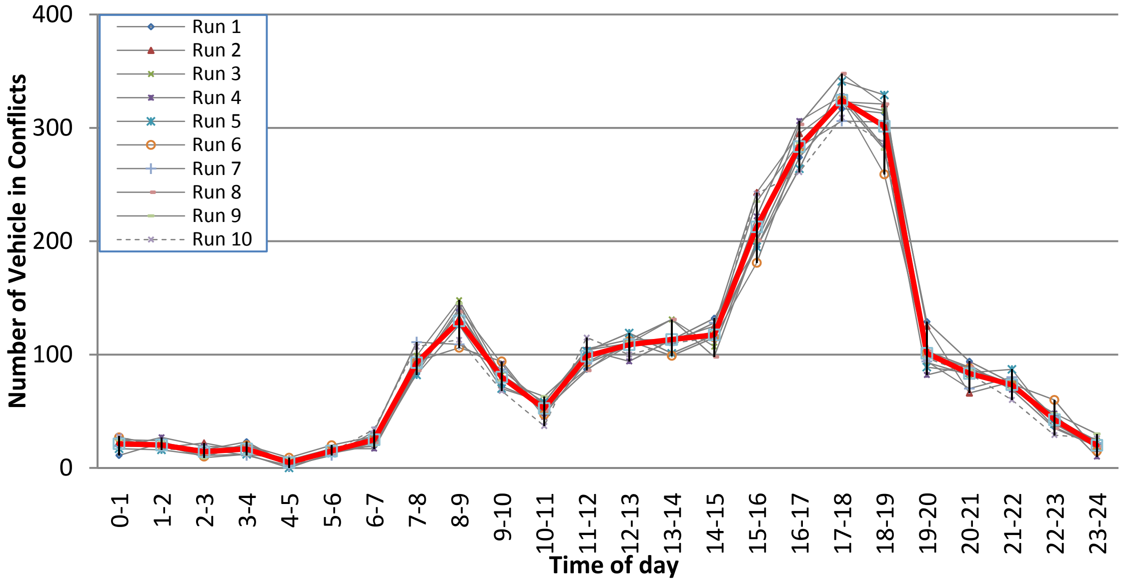



**Fig. 5-7 Number of vehicles in conflict in 24 hours for study site 3 (10 simulation runs)**

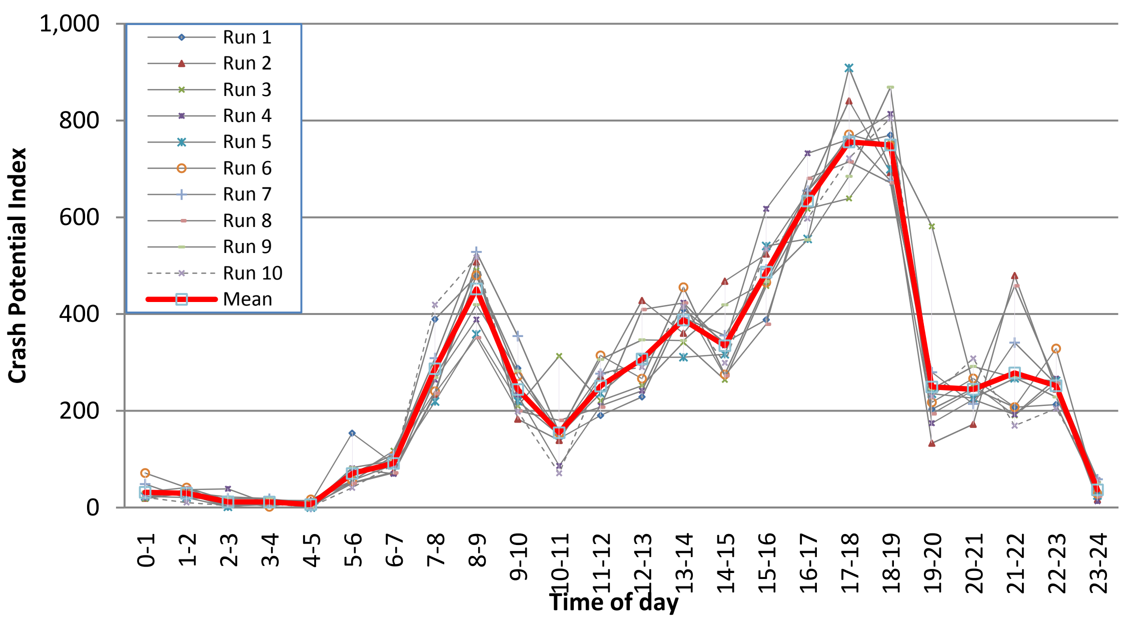



**Fig. 5-8 CPI in 24 hours for study site 3 (10 simulation runs)**

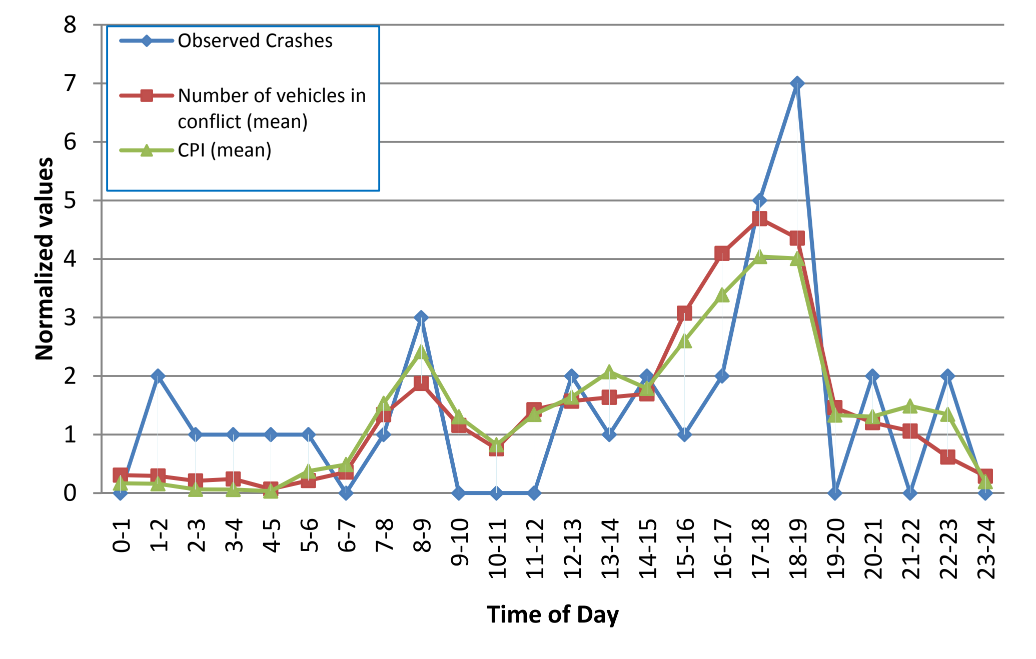
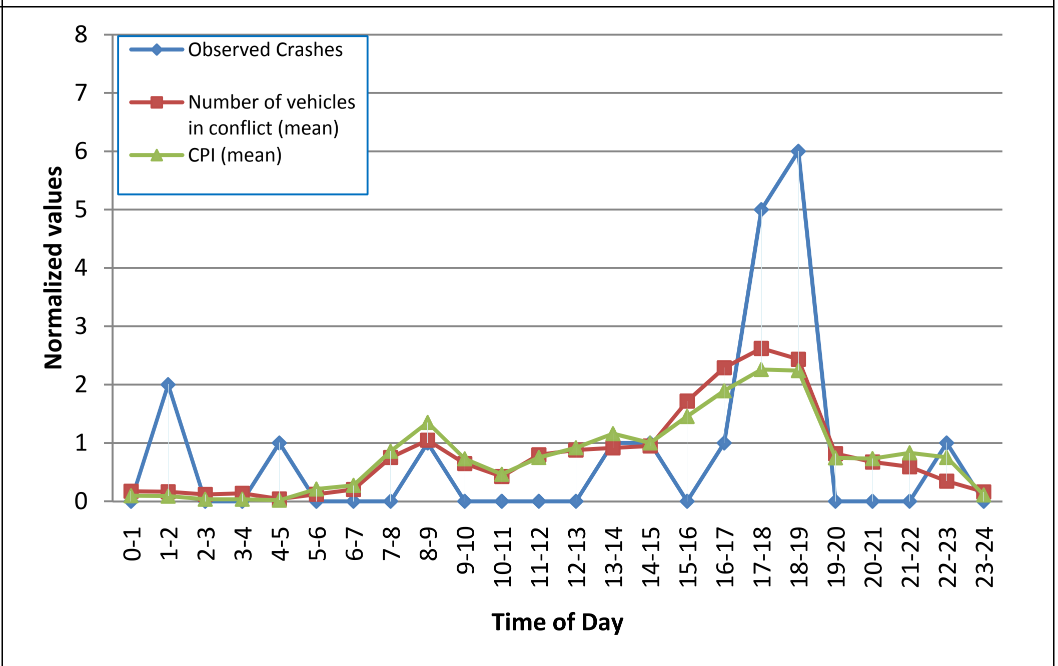
| | |
|---|---|
| All crashes |   |
| Multi-vehicle crashes only |   |

**Fig. 5-9 Observed crashes, normalized number of vehicles in conflict and normalized CPI in 24 hours for study site 3 (10 simulation runs)**

### 5.3 Simple linear correlation between crash and number of vehicles in conflict, crash potential index

One of the key objectives of this study was to find the relationship between the surrogate safety measures, which were selected as number of vehicles in conflict and CPI, and the observed crashes. More specifically, the correlation between the observed crashes and the simulated measures needed to be evaluated. Visually, Figures 5-3, 5-6 and 5-9 show the fluctuations of observed crashes, number of vehicles in conflict and CPI in 24-hour period of time for site 1, site 2 and site 3 respectively. Statistically, Pearson correlation coefficient analysis was used to initially assessing the linear relationship between observed crashes and simulated number of vehicles in conflict as well as the linear relationship between observed crashes and CPI. It could be a premature assumption of a linear relationship between crashes and surrogate measures. Therefore, Pearson correlation coefficient might not be the most appropriate method of linking the data. However, this analysis could provide the first statistically informative look at the data and the linear dependences between observed crashes and the simulated measures being investigated.

Correlation coefficient $\rho$ between Y and X is defined as:

$$\rho = \frac{\sigma_{XY}}{\sigma_X \sigma_Y} \qquad (5-2)$$

Where:

$\sigma_{XY}$ is the covariance between X and Y

$\sigma_X$ and $\sigma_Y$ are the standard deviations of X and Y

Correlation coefficient r for a sample of paired data ($X_i$ and $Y_i$) is calculated as:

$$r = \frac{\sum_{i=1}^{n}(X_i - \overline{X})(Y_i - \overline{Y})}{\sqrt{\sum_{i=1}^{n}(X_i - \overline{X})^2}\,(Y_i - Y)^2} \qquad (5-3)$$

The Spearman rank correlation coefficient uses rankings instead of using the actual data values to calculate the correlation coefficient. By definition, the Spearman rank correlation coefficient ρ is given by:

$$\rho = 1 - \frac{6\sum_{i=1}^{n} D_i^2}{n(n^2 - 1)} \qquad (5-4)$$

Where:

- $D_i$ is the difference in rank of corresponding values. $D_i$=rank($X_i$)-rank($Y_i$); and,
- n is the number of values in each dataset (same for both datasets).

SAS procedure CORR (SAS, 2004) was used to automate the correlation coefficient calculation for all three study sites. A statistics summary and a brief correlation coefficient analysis results for study site 1 are presented on Tables 5-2, 5-3 and 5-4. Figures 5-10 and 5-11 show scatter plots illustrating the relationships between observed crashes and the simulate safety performance measures (the number of vehicles in conflict and CPI).

**Table 5-2 Summary of Statistics for Data in 1-Hour Intervals from Study Site 1**

| Variable | | Number of Observations | Mean | Standard Deviation | Sum | Min | Max |
|---|---|---|---|---|---|---|---|
| Crash | All crashes | 24 | 0.750 | 0.794 | 18 | 0 | 3 |
| | Multi-vehicle crashes only | 24 | 0.375 | 0.770 | 9 | 0 | 3 |
| NoVeh | | 24 | 165.50 | 248.63 | 3,972 | 4 | 910.2 |
| CPI | | 24 | 532.31 | 481.34 | 12,775 | 18.05 | 1,708.01 |

**Table 5-3 Pearson Correlation Coefficients for Data in 1-Hour Intervals from Study Site 1**

| | All crashes | | | Multi-vehicle crashes only | | |
|---|---|---|---|---|---|---|
| | **Crash** | **NoVeh** | **CPI** | **Crash** | **NoVeh** | **CPI** |
| Crash | 1.00 | 0.576 (0.0032) | 0.577 (0.0031) | 1.00 | 0.414 (0.0441) | 0.474 (0.0191) |
| NoVeh | 0.576 (0.0032) | 1.00 | 0.907 (<0.0001) | 0.414 (0.0441) | 1.00 | 0.907 (<0.0001) |
| CPI | 0.577 (0.0031) | 0.907 (<0.0001) | 1.00 | 0.474 (0.0191) | 0.907 (<0.0001) | 1.00 |

*(Note: p-value in the parenthesis)*

**Table 5-4 Spearman Correlation Coefficients for Data in 1-Hour Intervals from Study Site 1**

| | All crashes | | | Multi-vehicle crashes only | | |
|---|---|---|---|---|---|---|
| | **Crash** | **NoVeh** | **CPI** | **Crash** | **NoVeh** | **CPI** |
| Crash | 1.00 | 0.339 (0.105) | 0.291 (0.167) | 1.00 | 0.295 (0.161) | 0.308 (0.143) |
| NoVeh | 0.339 (0.105) | 1.00 | 0.974 (<0.0001) | 0.295 (0.161) | 1.00 | 0.974 (<0.0001) |
| CPI | 0.291 (0.167) | 0.974 (<0.0001) | 1.00 | 0.308 (0.143) | 0.974 (<0.0001) | 1.00 |

*(Note: p-value in the parenthesis)*

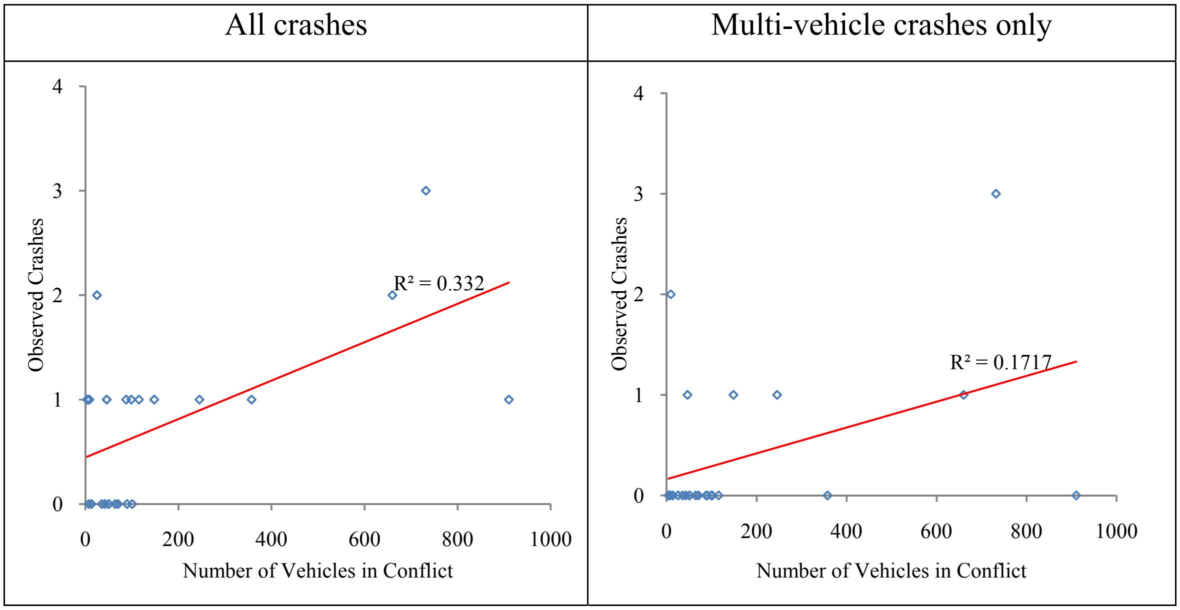



**Fig. 5-10 Scatter plot for observed crashes and number of vehicles in conflict aggregated in 1-hour intervals for study site 1 (2003-2007)**

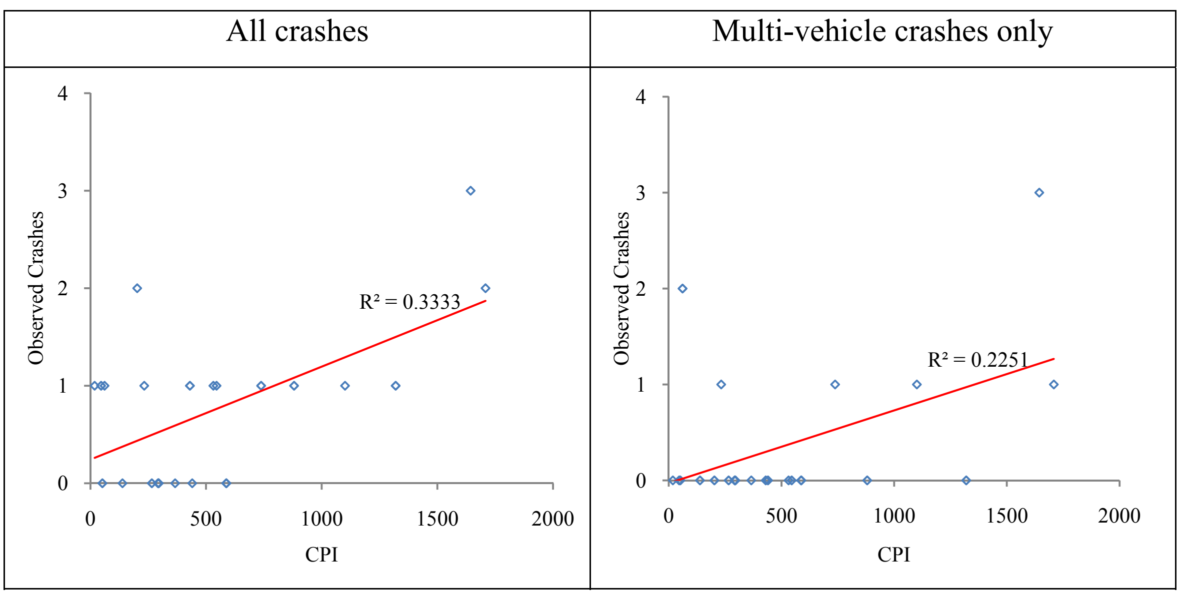



**Fig. 5-11 Scatter plot for observed crashes and CPI aggregated in 1-hour intervals for study site 1 (2003-2007)**

Similarly, a summary of statistics and linear correlation analysis results for data collected from study site 2 are shown on Tables 5-5, 5-6 and 5-7. Scatter plots of observed crashes and surrogate safety measures for this study sites are shown in Figures 5-12 and 5-13.

**Table 5-5 Summary of Statistics for Data in 1-Hour Intervals from Study Site 2**

| Variable | | Number of Observations | Mean | Standard Deviation | Sum | Min | Max |
|---|---|---|---|---|---|---|---|
| Crash | All crashes | 24 | 3.29 | 2.10 | 79 | 0 | 8 |
| | Multi-vehicle crashes only | 24 | 2.13 | 1.65 | 51 | 0 | 7 |
| NoVeh | | 24 | 170.52 | 153.84 | 4,092 | 1.2 | 486.4 |
| CPI | | 24 | 162.42 | 127.21 | 3,898 | 2.78 | 411.25 |

**Table 5-6 Pearson Correlation Coefficients for Data in 1-Hour Intervals from Study Site 2**

| | All crashes | | | Multi-vehicle crashes only | | |
|---|---|---|---|---|---|---|
| | **Crash** | **NoVeh** | **CPI** | **Crash** | **NoVeh** | **CPI** |
| Crash | 1.00 | 0.70 (0.0001) | 0.71 (0.0001) | 1.00 | 0.76 (<0.0001) | 0.74 (<0.0001) |
| NoVeh | 0.70 (0.0001) | 1.00 | 0.97 (<0.0001) | 0.76 (<0.0001) | 1.00 | 0.97 (<0.0001) |
| CPI | 0.71 (0.0001) | 0.97 (<0.0001) | 1.00 | 0.74 (<0.0001) | 0.97 (<0.0001) | 1.00 |

*(Note: p-value in the parenthesis)*

**Table 5-7 Spearman Correlation Coefficients for Data in 1-Hour Intervals from Study Site 2**

| | All crashes | | | Multi-vehicle crashes only | | |
|---|---|---|---|---|---|---|
| | **Crash** | **NoVeh** | **CPI** | **Crash** | **NoVeh** | **CPI** |
| Crash | 1.00 | 0.66 (0.0005) | 0.63 (0.0010) | 1.00 | 0.71 (0.0001) | 0.68 (0.0003) |
| NoVeh | 0.66 (0.0005) | 1.00 | 0.99 (<0.0001) | 0.71 (0.0001) | 1.00 | 0.99 (<0.0001) |
| CPI | 0.63 (0.0010) | 0.99 (<0.0001) | 1.00 | 0.68 (0.0003) | 0.99 (<0.0001) | 1.00 |

*(Note: p-value in the parenthesis)*

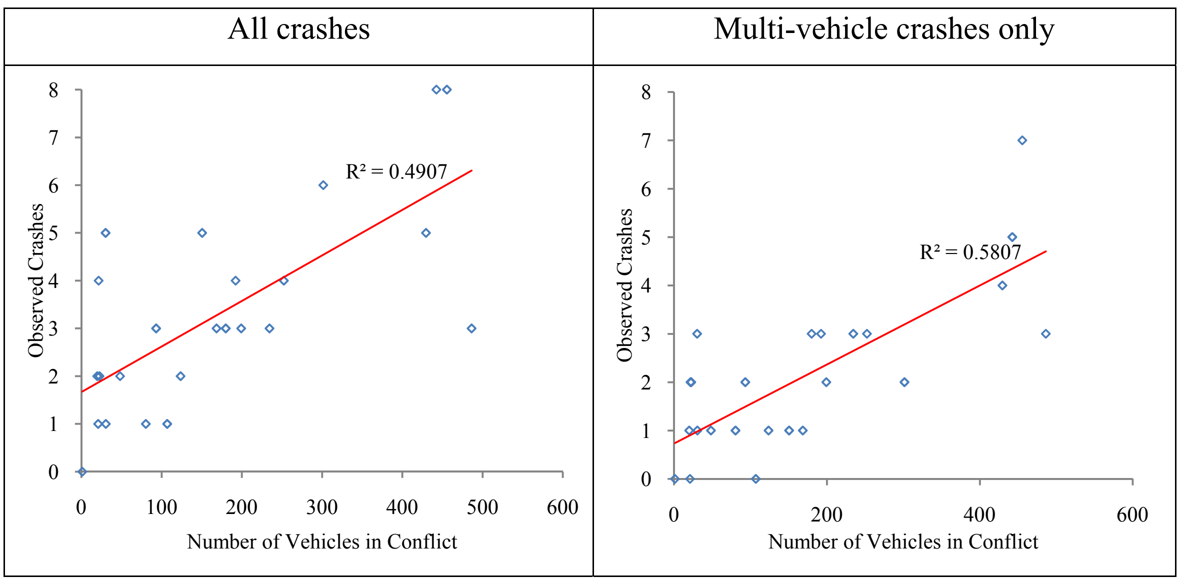



**Fig. 5-12 Scatter plot for observed crashes and number of vehicles in conflict aggregated in 1-hour intervals for study site 2 (2003-2007)**

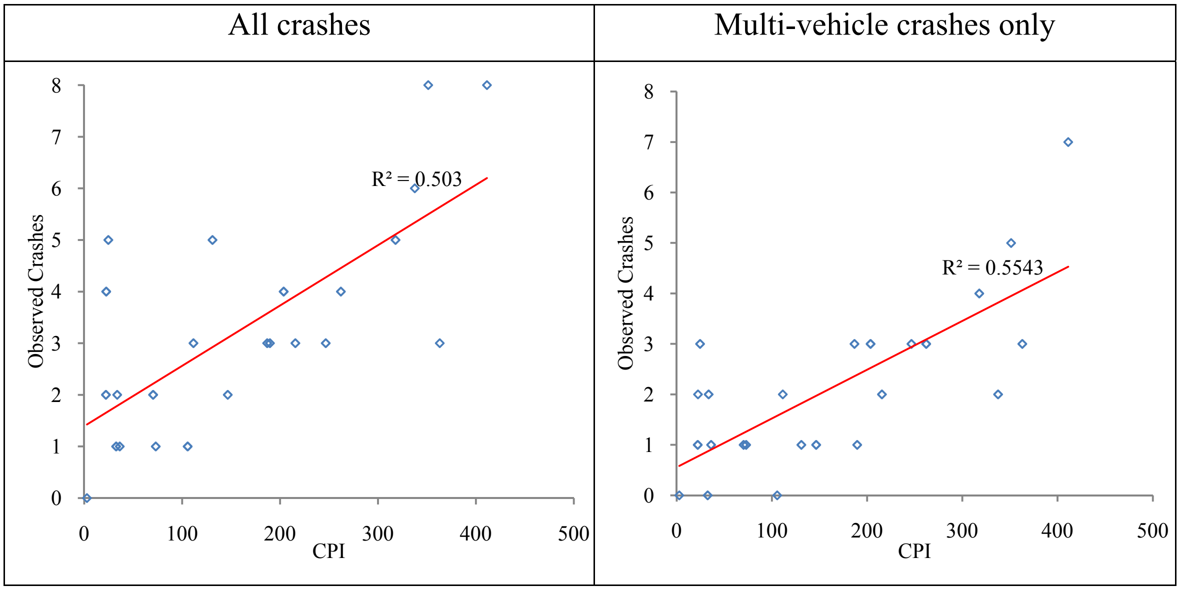



**Fig. 5-13 Scatter plot for observed crashes and CPI aggregated in 1-hour intervals for study site 2 (2003-2007)**

A summary of statistics and linear correlation analysis results for data collected from study site 3 are shown on Tables 5-8, 5-9 and 5-10.

**Table 5-8 Summary of Statistics for Data in 1-Hour Intervals from Study Site 3**

| Variable | | Number of Observations | Mean | Standard Deviation | Sum | Min | Max |
|---|---|---|---|---|---|---|---|
| Crash | All crashes | 24 | 1.42 | 1.69 | 34 | 0 | 7 |
| | Multi-vehicle crashes only | 24 | 0.79 | 1.16 | 19 | 0 | 6 |
| NoVeh | | 24 | 98.05 | 93.44 | 2353 | 4.80 | 324.50 |
| CPI | | 24 | 264.88 | 224.25 | 6357 | 6.31 | 755.64 |

**Table 5-9  Pearson Correlation Coefficients for Data in 1-Hour Intervals from Study Site 3**

| | All crashes | | | Multi-vehicle crashes only | | |
|---|---|---|---|---|---|---|
| | **Crash** | **NoVeh** | **CPI** | **Crash** | **NoVeh** | **CPI** |
| Crash | 1.00 | 0.71 (0.0001) | 0.72 (<0.0001) | 1.00 | 0.69 (0.0002) | 0.67 (0.0004) |
| NoVeh | 0.71 (0.0001) | 1.00 | 0.97 (<0.0001) | 0.69 (0.0002) | 1.00 | 0.97 (<0.0001) |
| CPI | 0.72 (<0.0001) | 0.97 (<0.0001) | 1.00 | 0.67 (0.0004) | 0.97 (<0.0001) | 1.00 |

*(Note: p-value in the parenthesis)*

**Table 5-10  Spearman Correlation Coefficients for Data in 1-Hour Intervals from Study Site 3**

| | All crashes | | | Multi-vehicle crashes only | | |
|---|---|---|---|---|---|---|
| | **Crash** | **NoVeh** | **CPI** | **Crash** | **NoVeh** | **CPI** |
| Crash | 1.00 | 0.47 (0.02) | 0.51 (0.01) | 1.00 | 0.40 (0.0525) | 0.43 (0.0383) |
| NoVeh | 0.47 (0.02) | 1.00 | 0.95 (<0.0001) | 0.40 (0.0525) | 1.00 | 0.95 (<0.0001) |
| CPI | 0.51 (0.01) | 0.95 (<0.0001) | 1.00 | 0.43 (0.0383) | 0.95 (<0.0001) | 1.00 |

*(Note: p-value in the parenthesis)*

Figure 5-14 and Figure 5-15 are scatter plots of data collected from site 3 aggregated in 1–hour intervals. These plots illustrate the relationship between observed crashes and the two simulated safety measures being investigated in this study which are number of vehicles in conflict and CPI.

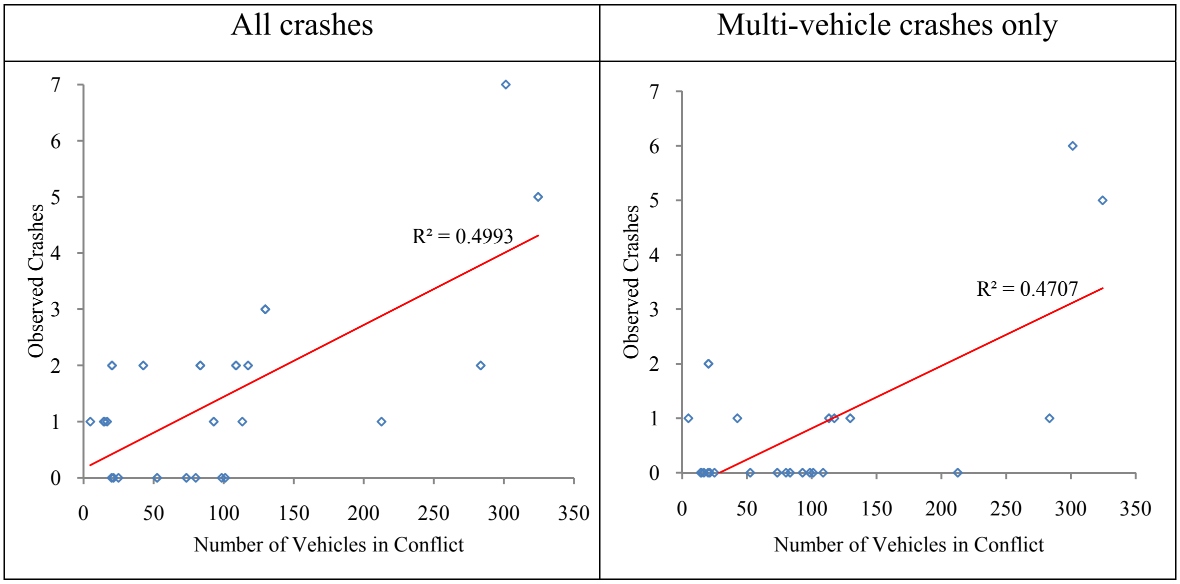



**Fig. 5-14 Scatter plot for observed crashes and number of vehicles in conflict aggregated in 1-hour intervals for study site 3 (2003-2007)**

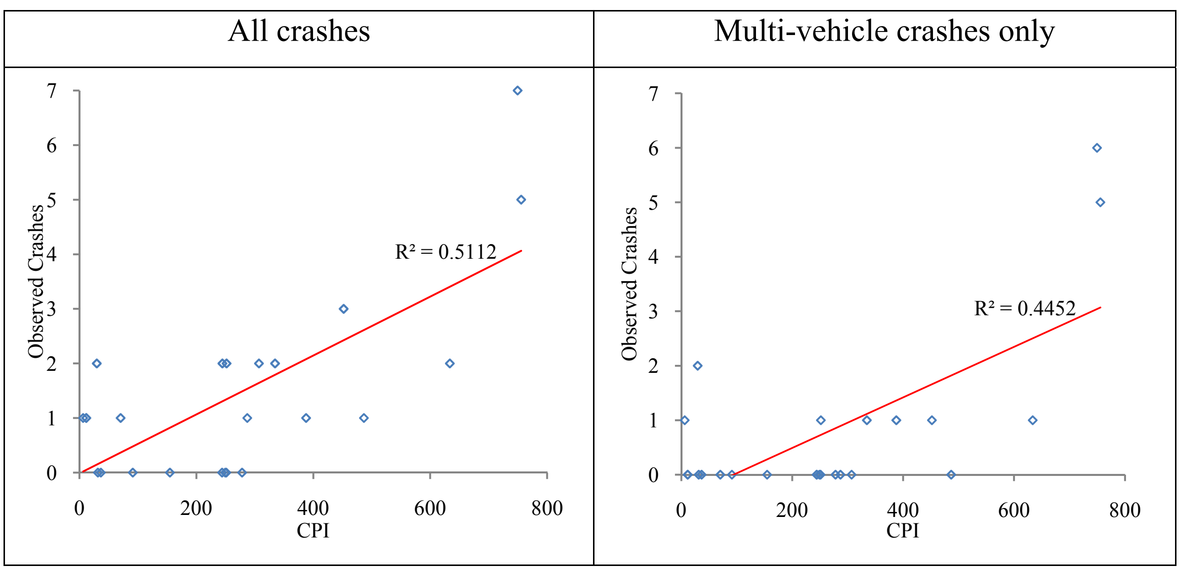



**Fig. 5-15 Scatter plot for observed crashes and CPI aggregated in 1-hour intervals for study site 3 (2003-2007)**

All analysis results previously presented are based on datasets with all data aggregated in 1-hour intervals. Table 5-11 is the Spearman rank correlation coefficient analysis using part of data shown on Tables 3-6 to 3-8. These data were aggregated based on time of day which previously defined and described in Chapter III as: Night (from 0 to 6 AM), Morning peak (from 6 to 9 AM), Off-peak (9 to 15, Evening peak (from 15 to 18) and Evening off-peak (from 18 to midnight). Scatter plots of these data are also shown in Figures 5-16 and 5-17. Similarly, the analysis results for data collected from study site 2 and study sites 3 are shown on Table 5-12 and Table 5-13 respectively. Figures 5-18, 5-19 and Figures 5-20, 5-21 show the scatter plots illustrating the relationships between the average observed crashes per hour and the simulated safety performance measures.

**Table 5-11 Spearman Correlation Coefficients for Data Aggregated by Time of Day from Study Site 1**

| | All crashes | | | Multi-vehicle crashes only | | |
|---|---|---|---|---|---|---|
| | **Crash** | **NoVeh** | **CPI** | **Crash** | **NoVeh** | **CPI** |
| Crash | 1.00 | 0.56 (0.32) | 0.56 (0.32) | 1.00 | 0.82 (0.09) | 0.82 (0.09) |
| NoVeh | 0.56 (0.32) | 1.00 | 1.00 (<0.0001) | 0.82 (0.09) | 1.00 | 1.00 (<0.0001) |
| CPI | 0.56 (0.32) | 1.00 (<0.0001) | 1.00 | 0.82 (0.09) | 1.00 (<0.0001) | 1.00 |

*(Note: p-value in the parenthesis)*

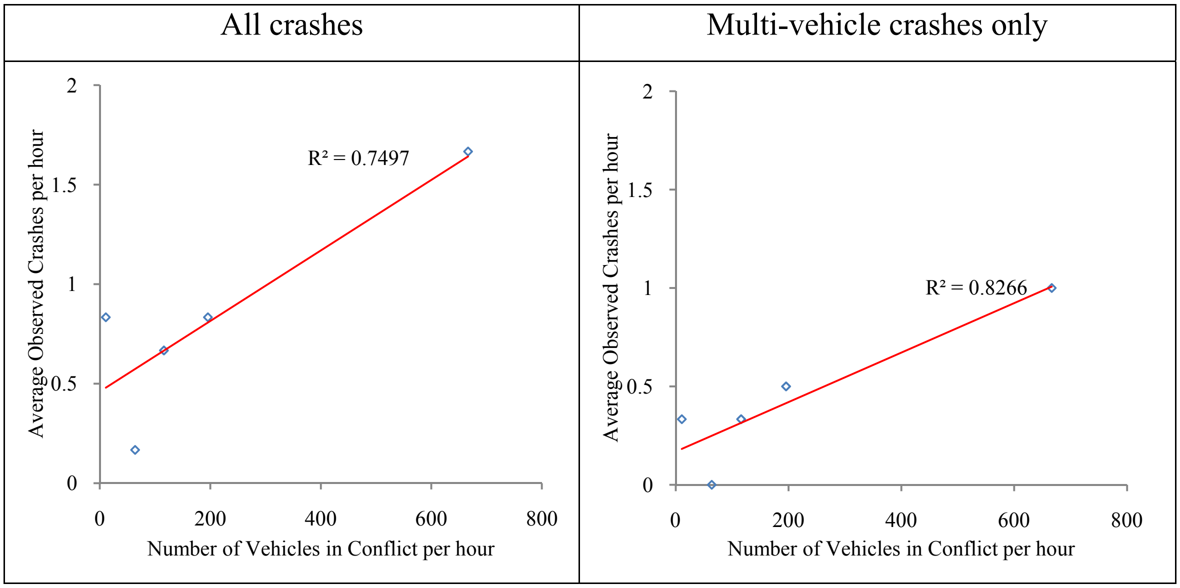



**Fig. 5-16 Scatter plot for observed crashes and number of vehicles in conflict aggregated by time of day for study site 1 (2003-2007)**

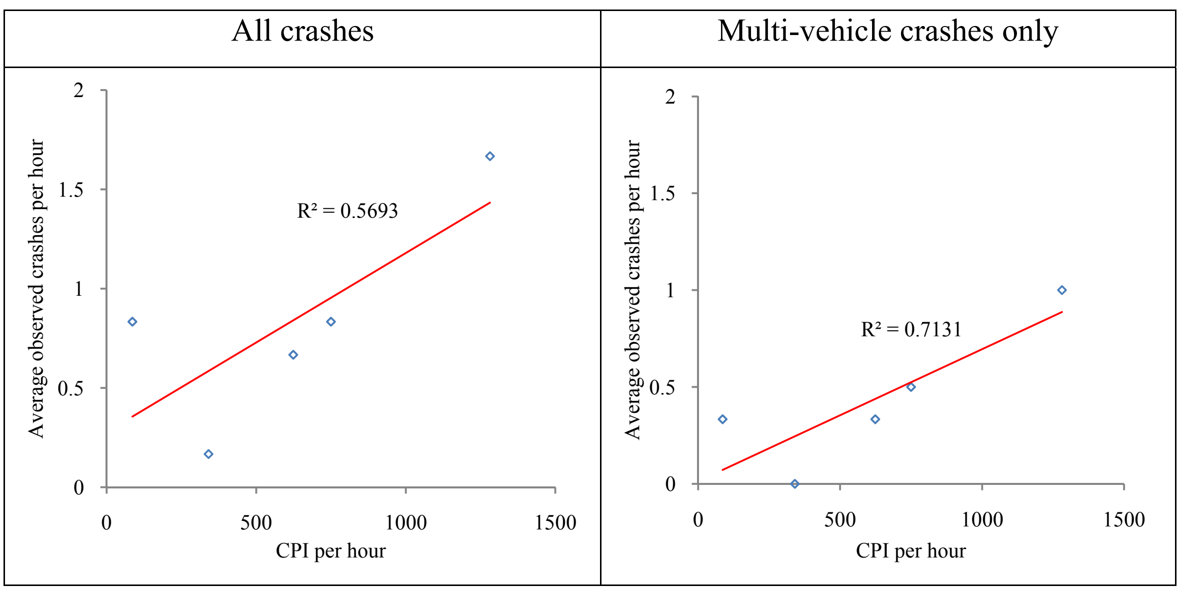



**Fig. 5-17 Scatter plot for observed crashes and CPI aggregated by time of day for study site 1 (2003-2007)**

**Table 5-12 Spearman Correlation Coefficients for Data Aggregated by Time of Day from Study Site 2**

| | All crashes | | | Multi-vehicle crashes only | | |
|---|---|---|---|---|---|---|
| | **Crash** | **NoVeh** | **CPI** | **Crash** | **NoVeh** | **CPI** |
| Crash | 1.00 | 0.56 (0.32) | 0.82 (0.09) | 1.00 | 0.70 (0.19) | 0.90 (0.04) |
| NoVeh | 0.56 (0.32) | 1.00 | 0.90 (0.04) | 0.70 (0.19) | 1.00 | 0.90 (0.04) |
| CPI | 0.82 (0.09) | 0.90 (0.04) | 1.00 | 0.90 (0.04) | 0.90 (0.04) | 1.00 |

*(Note: p-value in the parenthesis)*

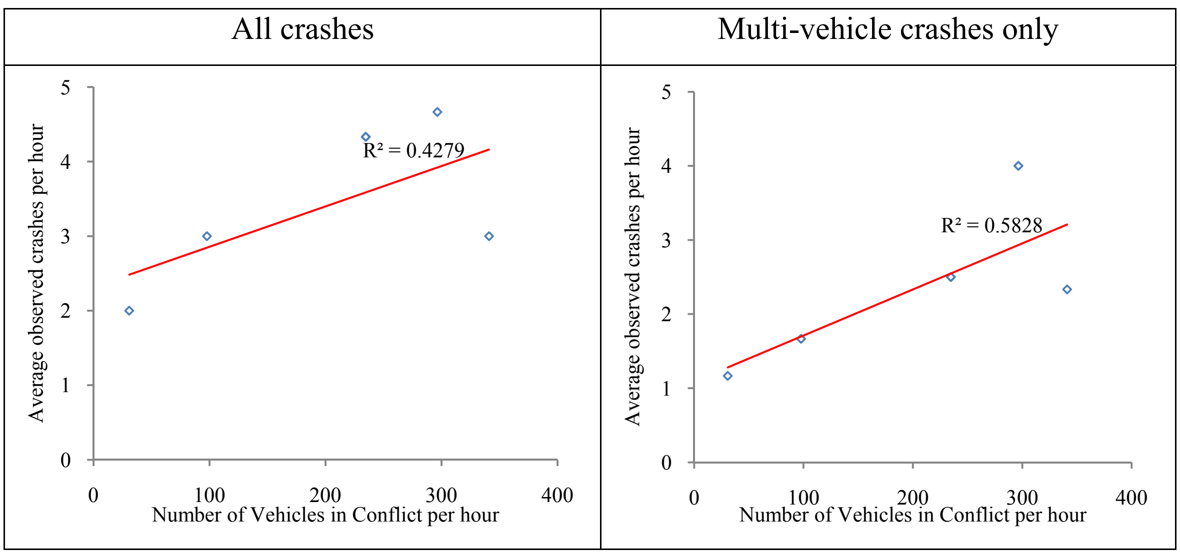



**Fig. 5-18 Scatter plot for observed crashes and number of vehicles in conflict aggregated by time of day for study site 2 (2003-2007)**

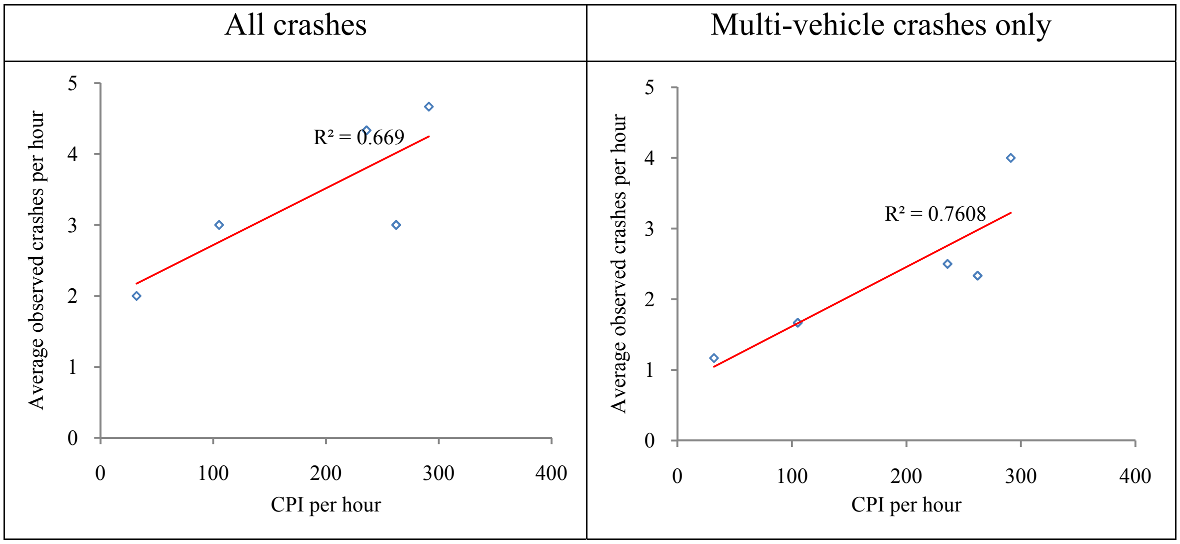



**Fig. 5-19 Scatter plot for observed crashes and CPI aggregated by time of day for study site 2 (2003-2007)**

**Table 5-13 Spearman Correlation Coefficients for Data Aggregated by Time of Day from Study Site 3**

| | All crashes | | | Multi-vehicle crashes only | | |
|---|---|---|---|---|---|---|
| | **Crash** | **NoVeh** | **CPI** | **Crash** | **NoVeh** | **CPI** |
| Crash | 1.00 | 0.70 (0.19) | 0.70 (0.19) | 1.00 | 0.67 (0.22) | 0.67 (0.22) |
| NoVeh | 0.70 (0.19) | 1.00 | 1.00 (<0.0001) | 0.67 (0.22) | 1.00 | 1.0 (<0.0001) |
| CPI | 0.70 (0.19) | 1.00 (<0.0001) | 1.00 | 0.67 (0.22) | 1.0 (<0.0001) | 1.00 |

*(Note: p-value in the parenthesis)*

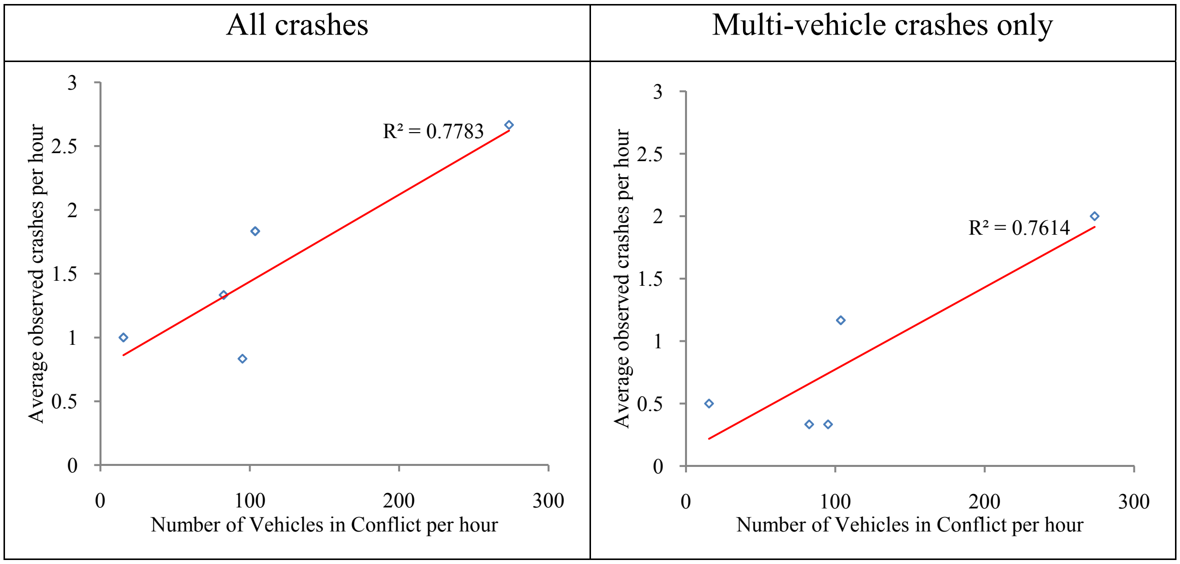



**Fig. 5-20 Scatter plot for observed crashes and number of vehicles in conflict aggregated by time of day for study site 3 (2003-2007)**

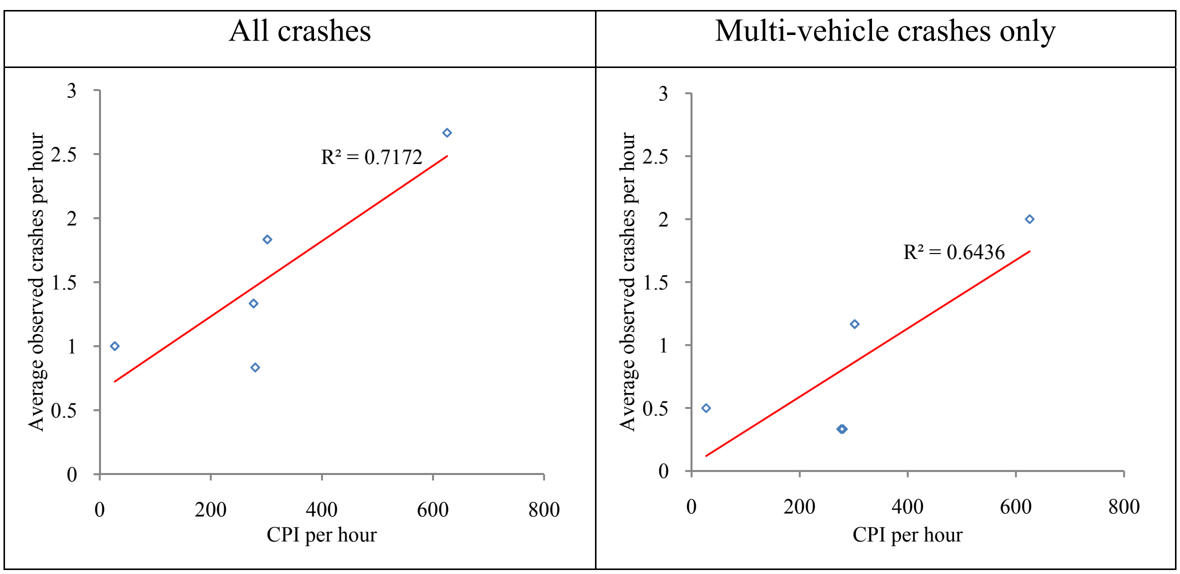



**Fig. 5-21 Scatter plot for observed crashes and CPI aggregated by time of day for study site 3 (2003-2007)**

The analyses with Pearson correlation coefficients revealed that crashes and number of vehicles in conflicts and crashes and CPI for site 2 and site 3 were not highly correlated with correlation coefficients of 0.70, 0.71 and 0.71, 0.72 for data set with all crashes included; and 0.76, 0.74 and 0.69, 0.67 for the data set with only multi-vehicle crashes respectively. Meanwhile, the results for study site 1 did not lead to the same level of correlation found for the other two study sites. With correlation coefficients of 0.576, 0.577 and 0.414, 0.472 for two data sets for, the data would not be deemed as strongly correlated. The results for all three sites were likely to suggest that CPI and NoVeh might have similar correlation levels to the observed crashes because they all resulted in similar correlation coefficients in all three cases. However, the differences were very small and drawing a conclusion solely from this analysis would be inappropriate and unwarranted. The aberrant results from site 1 could be attributed to the small number of crashes observed on this study site over five year period. Only eighteen crashes on this site were reported in five years and included in this analysis. However, because of this small number of crashes to be distributed over twenty four hours of time, there was hardly a crash distribution that could be compared to the simulated surrogate measures.

Comparisons between the results from two different data sets suggested better correlation for the data set with only crashes of multi-vehicle involvement in one case (site 2), similar correlations in another case (site 3) and worse correlation in the third case (site 1). Intuitively, the data set with single-vehicle crashes excluded would have better results. However, the results are obviously mixed. Again, this could be the small number of crashes observed. The small number of crashes could also be the reason why Spearman rank correlation test did not work well for the data set aggregated in 1-hour intervals. The test resulted in very small correlation coefficients for the relationship between crashes and surrogate safety measures (NoVeh, CPI). The small number of crashes distributing throughout 24 one-hour intervals resulted in many intervals with 0 or only 1 crash. Those intervals would have the same rank and therefore, the rank-based test did not work properly.

However, for the data set aggregated based on time of day, Spearman rank correlation coefficients are relatively high in most cases, especially for the data set with multi-vehicle crashes only. These indicate good linear relationship between average crashes of all severity levels per hour and the surrogate safety measures. The results can be translated as number of crashes (or average number of crashes per hour) is likely to be high when the simulated safety measures are high and vice versa. However, the p-values of the test are large; exceeding the 0.05 level in most cases. Therefore, the results are not reliable at this level of confidence. This issue might be caused by the small sample sizes of the data. The scatter plots and the relatively high R-square values of the linear regression lines also suggest linear relationships between crashes and the surrogate safety measures.

## 5.4 Statistical modeling

After the initial examination of the data by using correlation coefficient and rank correlation coefficient, a step further was taken, using regression modeling to explore the links of Crash to NoVeh and Crash to CPI. It is important to note that the regression models described below were not estimated to predict crashes, but just to examine the relationship between crashes and the number of vehicles in conflict and the CPI.

Crash data are obviously discrete and non-negative. This characteristic makes Poisson and Negative Binomial appropriate for modeling crash data. Past studies have found crash data likely to be over-dispersed (the variance is greater that the mean). Although Negative Binomial have been recommended to be better model for crash data (Milton and Mannering, 1998; Lord et al., 2005), both Poisson and Negative Binomial were used to test their fit with the data. The testing revealed that Poisson worked better for the data in this study and it was selected for further investigation.

The regression model development was an effort to link each of the simulated surrogate measures to the observed crashes through a link function other than a simple linear one. One of the key elements in the regression model development is selecting a functional form to link crashes to the relevant variable(s). For this study, the following functional form was selected:

$$\mu_i = \beta_0 . L . e^{\beta_1 x_1} \qquad (5-4)$$

Where:

- $\mu_i$ is the estimated number of crashes occurring within $i^{th}$ hour of the day (ranges from 0 to 23);
- $x_i$ is the explanatory variable, it could be either NoVeh or CPI;
- $\beta_0$ is the intercept, one of the regression model parameters to be estimated;
- L is the weaving section length (shown on Table 3-10), used as an offset of the model; and,
- $\beta_1$ is one of the model to be estimated. It is the coefficient for either NoVeh or CPI.

SAS procedure GENMOD (SAS, 2004) was employed to automate the coefficient estimation and goodness of fit assessment. The analysis used both crash data sets, one with all crashes included (single-vehicle and multi-vehicle crashes), and the other one without single-vehicle crashes. Results are shown on Tables 5-14 to 5-19 of this chapter. Goodness of fit assessment criteria and model parameter estimates for site 1 are shown on Table 5-14 and Table 5-15. Similarly, modeling results for site 2 and site 3 are shown on Table 5-16, Table 5-17 and Table 5-18, Table 5-19, respectively.

**Table 5-14 Regression Model Parameter Estimates for Observed Crash-Number of Simulated Vehicles in Conflict for Study Site 1**

| **Goodness of Fit Criteria** | | | | | | |
|---|---|---|---|---|---|---|
| **Criterion** | **All crashes** | | | **Multi-vehicle crashes only** | | |
| | **DF** | **Value** | **Value/DF** | **DF** | **Value** | **Value/DF** |
| Deviance | 22 | 17.61 | 0.80 | 22 | 22.60 | 1.03 |
| Pearson Chi-Square | 22 | 14.24 | 0.65 | 22 | 29.05 | 1.32 |
| Log Likelihood | | -20.74 | | | -15.61 | |

| **Parameter Estimates** | | | | | | |
|---|---|---|---|---|---|---|
| **Parameter** | **All crashes** | | | **Multi-vehicle crashes only** | | |
| | **Estimated Value** | **Test Statistics** | **p-value** | **Estimated Value** | **Test Statistics** | **p-value** |
| Intercept | -0.1263 | 15.51 | <0.0001 | -2.118 | 19.40 | <0.0001 |
| NoVeh | 0.0017 | 5.81 | 0.0160 | 0.0022 | 5.28 | 0.0216 |

**Table 5-15 Regression Model Parameter Estimates for Observed Crash-CPI for Study Site 1**

| **Goodness of Fit Criteria** | | | | | | |
|---|---|---|---|---|---|---|
| **Criterion** | **All crashes** | | | **Multi-vehicle crashes only** | | |
| | **DF** | **Value** | **Value/DF** | **DF** | **Value** | **Value/DF** |
| Deviance | 22 | 17.11 | 0.778 | 22 | 20.61 | 0.937 |
| Pearson Chi-Square | 22 | 14.89 | 0.677 | 22 | 36.83 | 1.674 |
| Log Likelihood | | -20.485 | | | -14.62 | |

| **Parameter Estimates** | | | | | | |
|---|---|---|---|---|---|---|
| **Parameter** | **All crashes** | | | **Multi-vehicle crashes only** | | |
| | **Estimated Value** | **Test Statistics** | **p-value** | **Estimated Value** | **Test Statistics** | **p-value** |
| Intercept | -1.533 | 14.27 | 0.0002 | -2.658 | 17.02 | <0.0001 |
| CPI | 0.001 | 5.98 | 0.0145 | 0.0015 | 6.88 | 0.0087 |

**Table 5-16 Regression Model Parameter Estimates for Observed Crash-Number of Simulated Vehicles in Conflict for Study Site 2**

| **Goodness of Fit Criteria** | | | | | | |
|---|---|---|---|---|---|---|
| **Criterion** | **All crashes** | | | **Multi-vehicle crashes only** | | |
| | **DF** | **Value** | **Value/DF** | **DF** | **Value** | **Value/DF** |
| Deviance | 22 | 18.28 | 0.83 | 22 | 15.67 | 0.71 |
| Pearson Chi-Square | 22 | 16.85 | 0.77 | 22 | 13.10 | 0.60 |
| Log Likelihood | | 21.98 | | | -4.92 | |

| **Parameter Estimates** | | | | | | |
|---|---|---|---|---|---|---|
| **Parameter** | **All crashes** | | | **Multi-vehicle crashes only** | | |
| | **Estimated Value** | **Test Statistics** | **p-value** | **Estimated Value** | **Test Statistics** | **p-value** |
| Intercept | -0.422 | 4.72 | 0.0299 | -1.047 | 16.80 | <0.0001 |
| NoVeh | 0.0026 | 14.43 | 0.0001 | 0.0033 | 15.92 | <0.0001 |

**Table 5-17 Regression Model Parameter Estimates for Observed Crash-CPI for Study Site 2**

| **Goodness of Fit Criteria** | | | | | | |
|---|---|---|---|---|---|---|
| **Criterion** | **All crashes** | | | **Multi-vehicle crashes only** | | |
| | **DF** | **Value** | **Value/DF** | **DF** | **Value** | **Value/DF** |
| Deviance | 22 | 17.23 | 0.78 | 22 | 15.39 | 0.70 |
| Pearson Chi-Square | 22 | 17.18 | 0.78 | 22 | 13.50 | 0.61 |
| Log Likelihood | | 22.50 | | | -4.78 | |

| **Parameter Estimates** | | | | | | |
|---|---|---|---|---|---|---|
| **Parameter** | **All crashes** | | | **Multi-vehicle crashes only** | | |
| | **Estimated Value** | **Test Statistics** | **p-value** | **Estimated Value** | **Test Statistics** | **p-value** |
| Intercept | -0.543 | 6.07 | 0.0138 | -1.193 | 16.70 | <0.0001 |
| CPI | 0.0034 | 14.81 | 0.0001 | 0.0043 | 15.28 | <0.0001 |

**Table 5-18 Regression Model Parameter Estimates for Observed Crash-Number of Simulated Vehicles in Conflict for Study Site 3**

**Goodness of Fit Criteria**

| Criterion | All crashes | | | Multi-vehicle crashes only | | |
|---|---|---|---|---|---|---|
| | DF | Value | Value/DF | DF | Value | Value/DF |
| Deviance | 22 | 24.43 | 1.11 | 22 | 24.43 | 1.11 |
| Pearson Chi-Square | 22 | 19.65 | 0.89 | 22 | 32.32 | 1.47 |
| Log Likelihood | | -12.93 | | | -11.03 | |

**Parameter Estimates**

| Parameter | All crashes | | | Multi-vehicle crashes only | | |
|---|---|---|---|---|---|---|
| | Estimated Value | Test Statistics | p-value | Estimated Value | Test Statistics | p-value |
| Intercept | -1.135 | 13.43 | 0.0002 | -2.41 | 22.04 | <0.0001 |
| NoVeh | 0.0067 | 20.14 | <0.0001 | 0.0102 | 23.63 | <0.0001 |

**Table 5-19 Regression Model Parameter Estimates for Observed Crash-CPI for Study Site 3**

**Goodness of Fit Criteria**

| Criterion | All crashes | | | Multi-vehicle crashes only | | |
|---|---|---|---|---|---|---|
| | DF | Value | Value/DF | DF | Value | Value/DF |
| Deviance | 22 | 22.36 | 1.02 | 22 | 22.86 | 1.04 |
| Pearson Chi-Square | 22 | 18.48 | 0.84 | 22 | 41.86 | 1.90 |
| Log Likelihood | | -11.90 | | | -10.24 | |

**Parameter Estimates**

| Parameter | All crashes | | | Multi-vehicle crashes only | | |
|---|---|---|---|---|---|---|
| | Estimated Value | Test Statistics | p-value | Estimated Value | Test Statistics | p-value |
| Intercept | -1.359 | 14.94 | 0.0001 | -2.78 | 22.12 | <0.0001 |
| CPI | 0.0031 | 21.16 | <0.0001 | 0.0048 | 23.56 | <0.0001 |

The modeling outcomes indicate that the data fit relatively well with Poisson models, some of the results for study site 1 and, the results for study site 2 show signs of under-dispersion (Deviance/DF below 1) to some certain extents. In most cases, the results show that the model coefficients for both Noveh and CPI were estimated to be positive values and the estimates are statistically significant at 5 per cent level for all three study sites. The goodness of fit is also supported by both Deviance and Pearson Chi-square, of which the Value/DF ratios are found to be around one. In some instances, these ratios are greater or smaller than one. However, those do not drastically deviate from one, point where is considered to be the perfect fit.

These results from both Pearson correlation coefficient analysis and regression modeling suggest that links between observed crashes to the simulated NoVeh and CPI exist and such links could be observed and identified statistically. The analysis revealed decent simple linear correlations between observed and simulated data in two of the three cases but links based on Poisson model were found to be stronger and more consistent in all cases, even when no perfect fit was found. These findings support what have been found in previous studies, even relatively limited, that simulation could possibly be used as an alternative to actual crash history to evaluate safety performance of weaving area as well as other transportation facilities in general.

## 5.5. Examining the effect of weaving volume

As an intuitive assumption, magnitude of weaving activity within the weaving area should be strongly correlated to number of vehicles in conflict (NoVeh) and crash potential index (CPI). It could probably be linked to number of actual crashes. The relationships between weaving volume and actual crashes, NoVeh, and CPI were examined using the Pearson correlation and Spearman rank correlation analyses. Table 5-20 and Table 5-21 show the correlation analyses results for data collected from study site 1 aggregated in 1-hour intervals. Scatter plots of the weaving volume and crashes, NoVeh, and CPI are shown in Figures 5-22 and 5-23. Similar results for study sites 2 and study site 3 are summarized on Tables 5-22, 5-23, and Table 5-24, 5-25 respectively. Figures 5-24, 5-25 and Figures 5-26, 5-27 show scatter plots for data from study site 2 and study site 2 respectively.

**Table 5-20 Pearson Correlation Coefficients between Weaving Volume and Crashes, NoVeh, and CPI for Study Site 1**

| | All crashes | Multi-vehicle crashes only |
|---|---|---|
| Weaving volume - Crashes | 0.34 (0.10) | 0.23 (0.28) |
| Weaving volume - NoVeh | 0.93 (<0.0001) | |
| Weaving volume - CPI | 0.87 (<0.0001) | |

*(Note: p-value in the parenthesis)*

**Table 5-21 Spearman Correlation Coefficients between Weaving Volume and Crashes, NoVeh, and CPI for Study Site 1**

| | All crashes | Multi-vehicle crashes only |
|---|---|---|
| Weaving volume - Crashes | 0.10 (0.65) | 0.23 (0.28) |
| Weaving volume - NoVeh | 0.82 (<0.0001) | |
| Weaving volume - CPI | 0.87 (<0.0001) | |

*(Note: p-value in the parenthesis)*

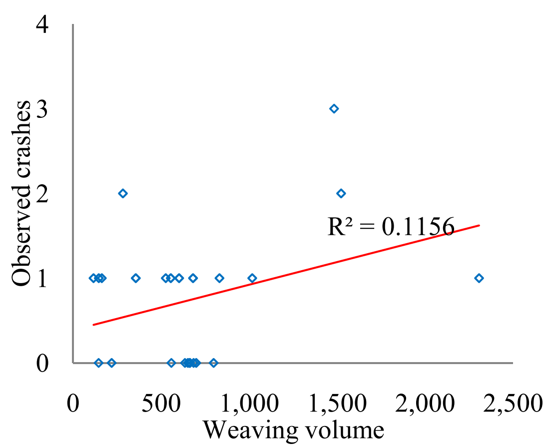
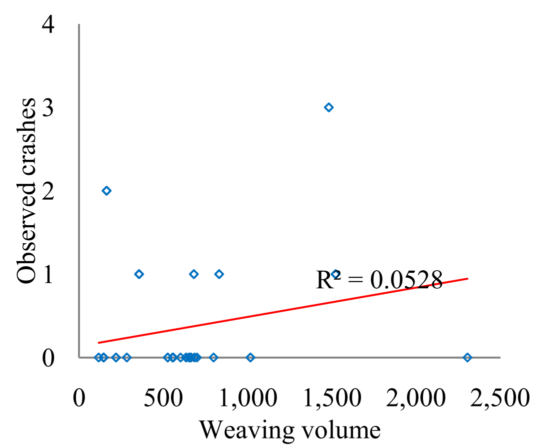
| All crashes | Multi-vehicle crashes only |
|---|---|
|  |  |





**Fig. 5-22 Scatter plot for observed crashes and weaving volume aggregated in 1-hour intervals for study site 1 (2003-2007)**

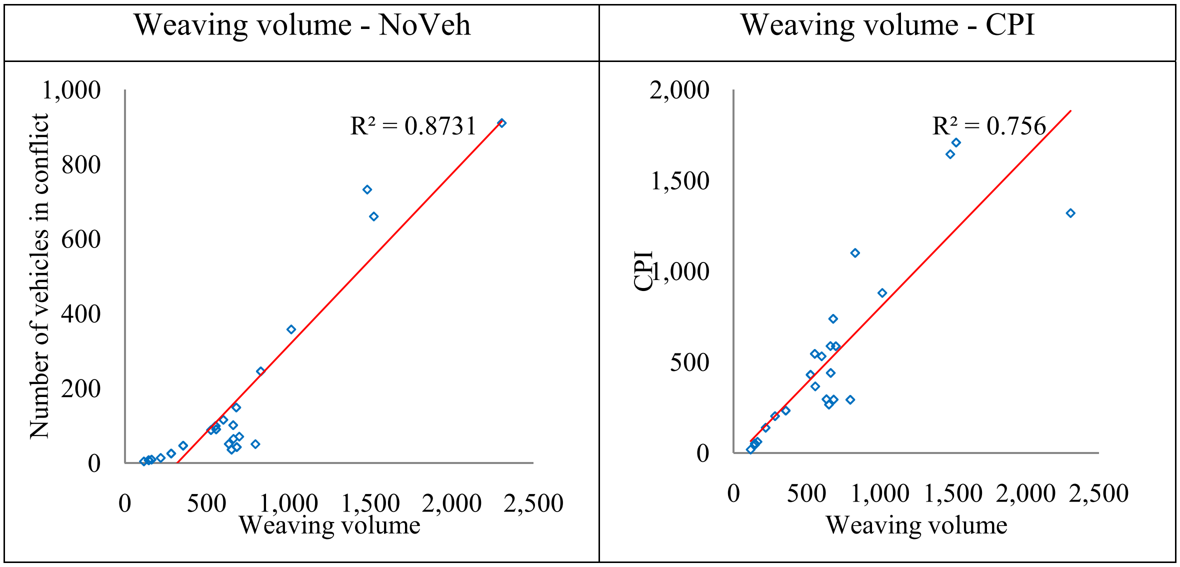



**Fig. 5-23 Scatter plot for number of vehicles in conflict, CPI and weaving volume aggregated in 1-hour intervals for study site 1**

**Table 5-22 Pearson Correlation Coefficients between Weaving Volume and Crashes, NoVeh, and CPI for Study Site 2**

| | All crashes | Multi-vehicle crashes only |
|---|---|---|
| Weaving volume - Crashes | 0.34 (0.11) | 0.45 (0.03) |
| Weaving volume - NoVeh | 0.78 (<0.0001) | |
| Weaving volume - CPI | 0.82 (<0.0001) | |

*(Note: p-value in the parenthesis)*

**Table 5-23 Spearman Correlation Coefficients between Weaving Volume and Crashes, NoVeh, and CPI for Study Site 2**

| | All crashes | Multi-vehicle crashes only |
|---|---|---|
| Weaving volume - Crashes | 0.30 (0.16) | 0.46 (0.03) |
| Weaving volume - NoVeh | 0.85 (<0.0001) | |
| Weaving volume - CPI | 0.85 (<0.0001) | |

*(Note: p-value in the parenthesis)*

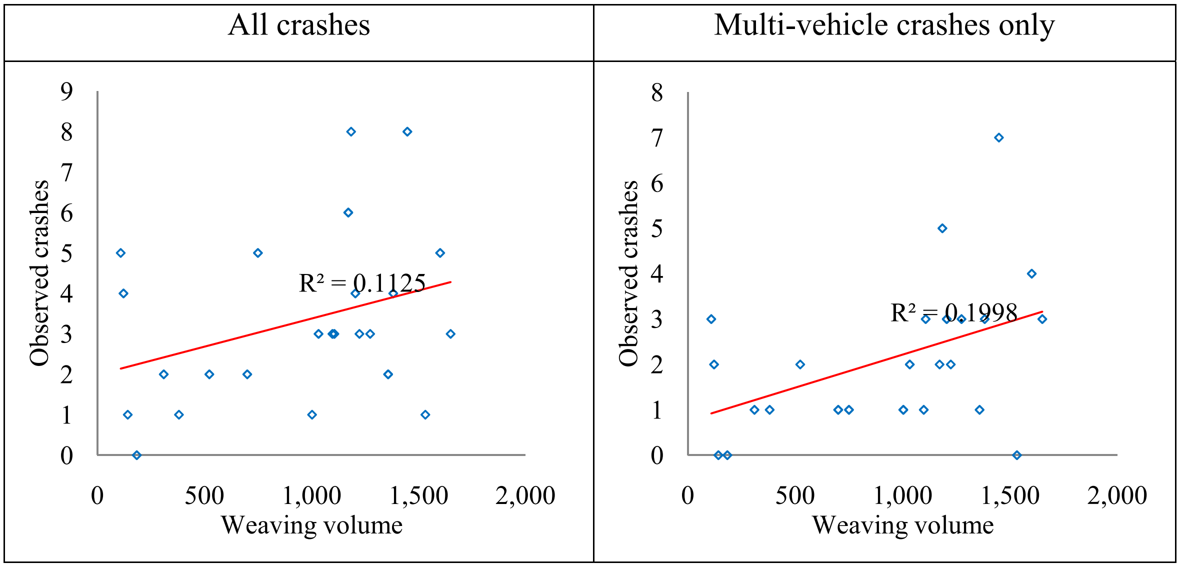



**Fig. 5-24 Scatter plot for observed crashes and weaving volume aggregated in 1-hour intervals for study site 2 (2003-2007)**

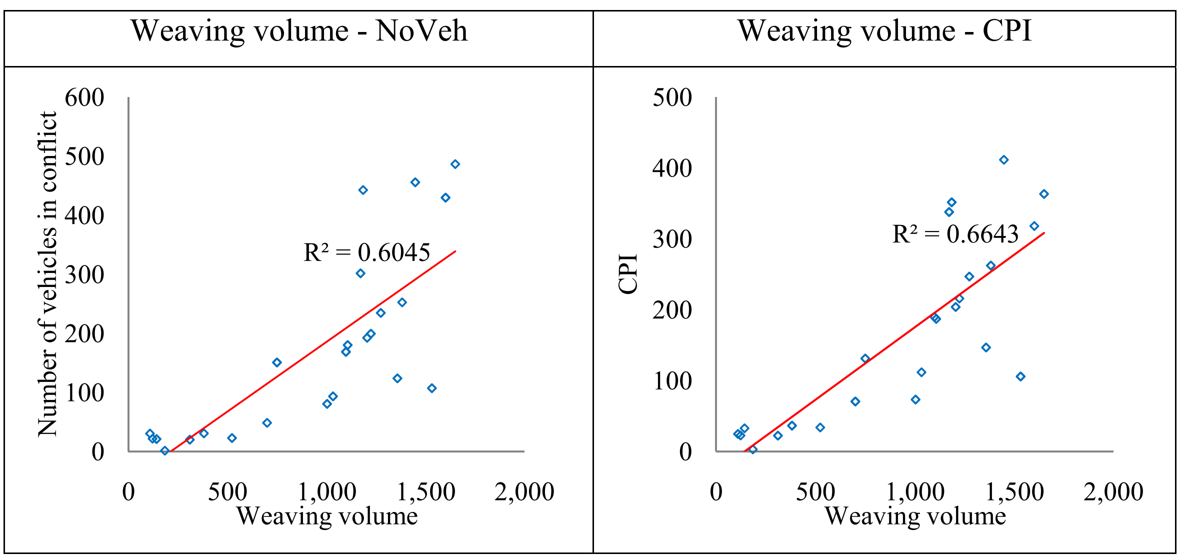



**Fig. 5-25 Scatter plot for number of vehicles in conflict, CPI and weaving volume aggregated in 1-hour intervals for study site 2**

**Table 5-24 Pearson Correlation Coefficients between Weaving Volume and Crashes, NoVeh, and CPI for Study Site 3**

| | All crashes | Multi-vehicle crashes only |
|---|---|---|
| Weaving volume - Crashes | 0.51 (0.01) | 0.47 (0.02) |
| Weaving volume - NoVeh | 0.93 (<0.0001) | |
| Weaving volume - CPI | 0.94 (<0.0001) | |

*(Note: p-value in the parenthesis)*

**Table 5-25 Spearman Correlation Coefficients between Weaving Volume and Crashes, NoVeh, and CPI for Study Site 3**

| | All crashes | Multi-vehicle crashes only |
|---|---|---|
| Weaving volume - Crashes | 0.38 (0.06) | 0.32 (0.13) |
| Weaving volume - NoVeh | 0.96 (<0.0001) | |
| Weaving volume - CPI | 0.93 (<0.0001) | |

*(Note: p-value in the parenthesis)*

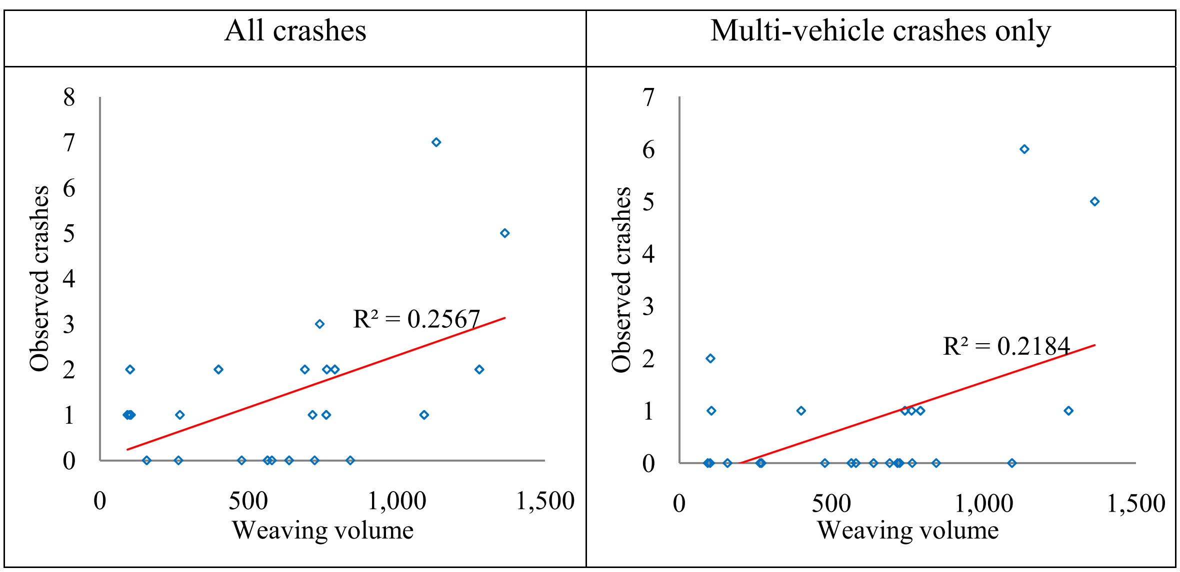



**Fig. 5-26 Scatter plot for observed crashes and weaving volume aggregated in 1-hour intervals for study site 3 (2003-2007)**

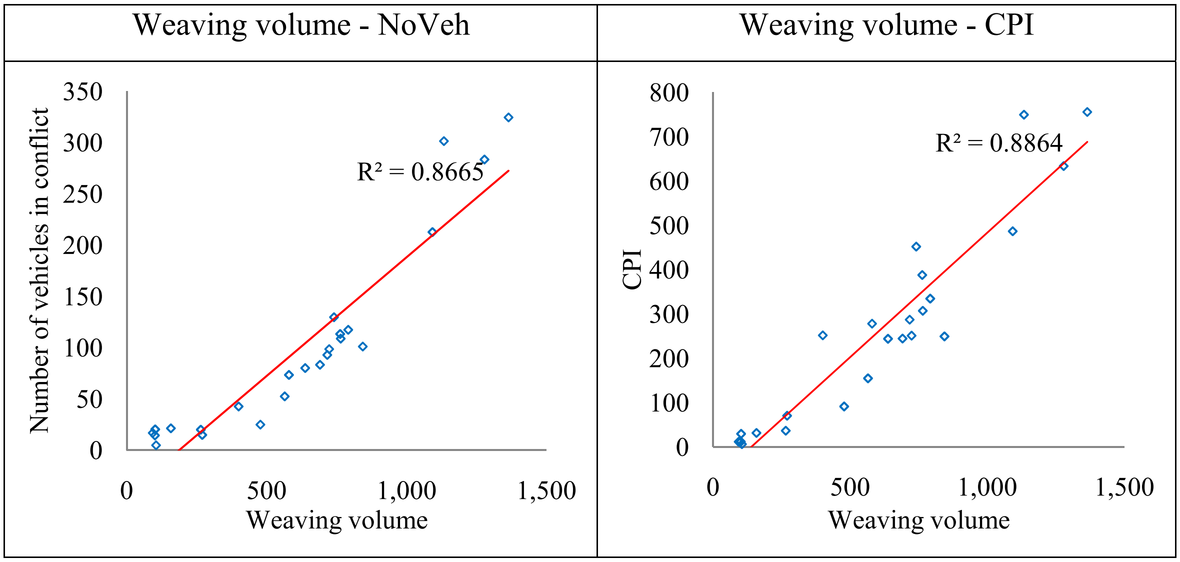



**Fig. 5-27 Scatter plot for number of vehicles in conflict, CPI and weaving volume aggregated in 1-hour intervals for study site 3**

The correlation analysis results provide clear statistical evidence to support the intuitive assumption about strong linear relationships between the magnitude of weaving activity (weaving volume) and the simulated safety measures. This linear link is also reinforced by the scatter plots and the high R-squared values for the linear regression lines. However, the results do not lead to establishing a linear link between weaving volume and observed crashes. The cause of crash might be attributed to many different factors, and magnitude of weaving activity is just one of them; and obviously, further study on this issue is needed.

The safety performance of weaving areas was found to have a strong relationship with weaving lengths as previous studies (Cirillo, 1970; Pulugurtha and Bhatt, 2008). In general, crashes (or its variations like crash rate) were found to decrease as the weaving length increases and vice versa. Under the scope of this study, this issue was not examined. More extensive studies should be done to investigate the effect of weaving length to safety performance, both actual crash history and simulated safety measures, and to verify what has been found.

## 5.6 Chapter summary

This chapter covers the extensive exploration and analysis of data with the objective of finding the relationship between observed crash data and the simulated surrogate safety performance measures. The simulated data used for the analysis were the average of ten simulation runs with ten different random seeds. The analyses carried out include Pearson correlation coefficient, Spearman rank correlation coefficient, and Poisson regression modeling. Pearson correlation coefficient was used for data aggregated in 1-hour intervals. Spearman rank correlation coefficient was employed for both datasets aggregated in 1-hour intervals and aggregated based on time of day (night, morning peak, off-peak, evening peak, and evening off-peak). Pearson correlation coefficient suggested decent linear relationships between observed crashes and number of vehicles in conflict and between observed crashes and CPI for two of the three study sites with the correlation coefficients ranging from 0.67 to 0.76. However, the linear correlation is weak at one study site with the correlation coefficients in the 0.4 and 0.5 ranges. Spearman rank correlation coefficient analysis did not work for the data aggregated in 1-hour intervals probably because of the small number of crashes distributing throughout 24 one-hour intervals. This resulted in many intervals with same rank. This rank-based test was also employed to examine the data aggregated in a different way based on time of day. The analysis results provided evidence about linear relationships between observed crashes (average number of crashes per hour) and NoVeh and CPI. In addition to Pearson and Spearman rank correlation coefficients, the regression modeling also revealed that the observed crashes may relate to the simulated safety measures following a Poisson function.

The analyses presented in this chapter also revealed strong linear relationships between the magnitude of weaving activity and the simulated safety measures, which are number of vehicles in conflict and crash potential index. The effect of weaving length on safety was not examined in this study. Even though, a small number of previous studies have revealed this relationship.

In this chapter, the operational conditions of the three study sites were also analyzed using the weaving analysis methodology in the 2000 HCM. The results showed reasonable LOS for all three weaving areas in normal daily peak condition.

The next chapter covers the conclusions drawn from this study and recommendations for future studies.

# CHAPTER VI

# CONCLUSIONS AND RECOMMENDATIONS

The concept adopted and explored, supported by the analyses performed in this study led to the conclusions and recommendations included in this chapter. The main objective of this study was to investigate the possibility of applying the traffic conflict technique and microscopic simulation to evaluate safety performance of weaving area. Guiding by this objective, this study was an effort of examining the concept as well as the microscopic mechanism of traffic conflict technique, adopting the technique and applying for weaving area, running the simulation, computing, analyzing and validating the results with actual crash data. Based on the findings of this study, the following conclusions could be made:

- Microscopic simulation models are not designed to simulate conflicts or crashes. In simulation, vehicles never hit each other. Therefore, it is not possible to directly evaluate safety using simulation. However, by adopting the traffic conflict technique, microscopic simulation could possibly be used for evaluating safety performance of weaving areas, in addition to intersections and freeway segments as previously investigated, to a certain extent, by a number of studies (Fazio el al., 1993; Cunto and Saccomanno, 2007, 2008; Cunto et al., 2009; Ozbay et al., 2008).

- Surrogate safety measures based on deceleration rates to avoid crash could have links to crash data. These surrogate measures could also reflect the pattern of crash history. Therefore, they could possibly be used as an alternative to crash history as traffic safety indicators.
- The connection between the surrogate safety measures and the real-life conflict in the traffic stream has not been well understood. Validation of surrogate measures is still based on real-world crash data and crash database can never fully estimate the actual risks of the roadway. At this point, surrogate safety measures are still far way from becoming a reliable alternative for crash data.
- The differences between mechanisms of rear-end and side-swipe crashes have not been clearly addressed, investigated, distinguished and modeled in this simulation-based technique. The current knowledge on this issue is not sufficient.
- The technique only considers a conflict between two vehicles interacting with each other in the traffic stream. It does not address the issue of single-vehicle crashes while this type of crash is definitely not rare.
- Although, no strong, consistent and reliable link between conflicts estimated by simulation and actual crash data has been found, possibly because simulating conflicts is not what the current simulation models are designed to do, microscopic simulation still has potential applications in traffic safety study.

- A complete simulation-based safety evaluation method has not been developed because of lack of comprehensive knowledge on the issue from a broader view. No such method could provide an answer on how safe or unsafe a transportation facility is just by analyzing geometric and traffic data. However, with the finding of this study, once again the applicability of simulation in safety study has been reinforced. Even with its current limitation, this technique can definitely be used to evaluate safety performances of different designs of new facilities or it can also be applied for choosing the best option among different improvement strategies of currently existing facilities.

Based on the conclusions drawn in this chapter, several recommendations could be made:

- At this point, microscopic simulation-based surrogate safety measures should only be considered as an alternative to crash history as safety indicators of transportation facilities. It is not well understood enough to replace crash data. Further analysis should be done with larger data sets for better understanding on this issue.
- This technique should be utilized in the design process to evaluate safety performances because no crash history information exists. Even if the modeling and CMF-based method is used, the simulation based technique

could still provide more information and firmer ground for decision making process.

- Crash history has its own political power. Crashes and fatalities have more impacts on public perception and policy makers than any other alternatives even if those surrogates are equally good or even better from an engineering and technical standpoint. Therefore, these need to be taken into consideration and used appropriately.
- Mechanism of single-vehicle crashes should be addressed and investigated in future studies.
- Side-swipe crashes need to be studied more and better understood in future efforts.
- The combination of crashes, volumes and exposures could have links to the surrogate safety performance measures. The length of weaving areas was found to have an effect on safety in a small number of previous studies. It could also have an effect on theses simulated measures. These issues need to be further investigated in future studies.
- The ultimate target in studying simulation-based safety evaluation technique should be a complete method that could be used for comprehensive safety evaluation.

**REFERENCES**

AASHTO., 2004. A Policy on Geometric Design of Highways and Streets. American Associatation of State Highway and Transportation Officials, Washington D.C.

Batenhorst, R.A., Gerken, J.G., 2000. Operational analysis of terminating freeway auxiliary lanes with one-lane and two-lane exit ramps: a case study. In: Proceeding of the Mid-Continent Transportation Symposium, Ames, IA.

Borchardt, D. W., Ballard, A. J., Stokton, W. R., 1984. An Analysis of Urban Freeway Operation and Modifications. Federal Highway Administration. Report No. FHWATX-84-46+210-12, Washington D.C.

Cassidy, M., Skabardonis, A., May, A. D., 1989. Operation of Major Freeway Weaving Sections: Recent Empirical Evidence. Transport. Res. Rec. 1225, 61-72.

Cassidy, M., May, A. D., 1991. Proposed Analytical Technique for Estimating Capacity and Level of Service of Major Freeway Weaving Sections. Transport. Res. Rec. 1320, 99-109.

Cirillo, J. A., 1970. The Relationship of Accidents to Length of Speed-change Lanes and Weaving Areas on Interstate Highways. Highway Res. Rec. 312, 17-32.

Chen, H. Liu, P., Lu, J. Behzadi, B., 2009. Evaluating the Safety Impacts of the Number and Arrangement of Lanes on Freeway Exit Ramps. Accid. Anal. Prev. 41 (3), 543-551.

Cunto, F., Saccomanno, F., 2007. Micro-level Traffic Simulation Method for Assessing Crash Potential at Intersections. In: Proceedings of the 86$^{th}$ Annual Meeting of the Transportation Research Board, Washington D.C.

Cunto, F., Saccomanno, F., 2008. Calibration and Validation of Simulated Vehicle Safety Performance at Signalized Intersections. Accid. Anal. Prev. 40 (3), 1171-1179.

Cunto, F., Duong, D., Saccomanno, F., 2009. Comparision of Simulated Freeway Safety Performance with Observed Crashes. Transport. Res. Rec. 2103, 88-97.

Darzentas, J., Cooper, D. F., Storr, P. A., McDowell, M. R.C., 1980. Simulation of Road Traffic Conflicts at T-Junctions. Simulation 34 (May), 155-164.

Fazio, J., Rouphail, N. M., 1986. Freeway Weaving Sections: Comparison and Refinement of Design and Operations Analysis Procedures. Transport. Res. Rec. 1091, 101-109.

Fazio, J., Holden, J., Rouphail, N. M., 1993. Use of Freeway Conflict Rates as an Alternative to Crash Rates in Weaving Section Safety Analyses. Transport. Res. Rec. 1401, 61-69.

Fisher, R. L., 1949. Weaving Traffic. In: Proceedings of Highway Research Board 28, 364-369.

Fitzpatrick, K., Porter R. J., Pesti G., Chu C. L., Park E. S., Le T. Q., 2009. Guidelines for Spacing Between Ramps. Draft Report, FHWA/TX-10/0-5860-1. Texas Transportation Institute, College Station, TX.

Fricke, B. L., 1990. Traffic Accident Reconstruction, Vol. 2. Northwestern University Traffic Institute, Evanston, IL.

Gettman, D., Head, L., 2003. Surrogate Safety Measures from Traffic Simulation Models. Transport. Res. Rec. 1840, 104-115.

Glad, R. W., Milton, J. C., Olson, D. K., 2001. Weaving Analysis and Performance: The Washington State Case Study. Washington State Department of Transportation. Report No. WA-RD 515.1, Olympia, WA.

Golob, T. F., Recker, W. W., Alvarez, V. M., 2004. Safety Aspect of Freeway Weaving Sections. Transport. Res. Part A (38), 35-51.

HCM., 1950. Highway Capacity Manual. U.S Goverment Printing Office, Washington D.C.

HCM., 1965. Highway Capacity Manual. Highway Research Board, Washington D.C.

HCM., 1985. Highway Capacity Manual. Transportation Research Board, Washington D.C.

HCM., 2000. Highway Capacity Manual. Transportation Research Board, Washington D.C.

Leisch, J. E., 1979. New Technique for Design and Analysis of Weaving Sections on Freeway. ITE Journal 49 (3), 26-29.

Leisch, J. E., 1983. Completion of Procedures for Analysis and Design of Weaving Sections. Federal Highway Administration, Washington D.C.

Lord, D., Manar, A., Vizioli, A., 2005. Modeling Crash-Flow-Density and Crash-Flow-V/C Ratio Relationships for Rural and Urban Freeway Segments. Accid. Anal. Prev. 37 (1), 185-199.

Lunenfeld, H., 1993. Human Factors Associated with Interchange Design Features. Transport. Res. Rec. 1385, 84-89.

May, A. D., 1990. Traffic Flow Fundamentals. Prenitice-Hall, Upper Saddle River, NJ.

Milton, J., Mannering, F., 1998. The Relationship Among Highway Geometrics, Traffic Related Elements and Motor Vehicle Accident Frequencies. Transportation 25, 395-413.

Ostrom, B., Leiman, L., May, A. D., 1993. Suggested Procedures for Analyzing Freeway Weaving Sections. Transport. Res. Rec. 1398, 42-48.

Ozbay, K., Yang, H., Bartin, B., Mudigonda, S., 2008. Derivation and Validation of New Simulation-based Surrogate Safety Measure. Transport. Res. Rec. 2083, 105-113.

Perkins, S., Harris, J., 1967. Criteria for Traffic Conflict Characteristics. General Motor Corporation. Report No. GMF 632, Warren, MI.

Pignataro, L. J., McShane, W. R., Roess, R. P., Lee, B., Crowley, K. W., 1975a. Weaving Areas: Design and Analysis. NCHRP Report 159. Transportation Research Board, Washington D.C.

Pignataro, L. J., McShane, W. R., Roess, R. P., Lee, B., Crowley, K. W., 1975b. Recommended Procedure for Weave Area Operation and Design (Abridgment). Transport. Res. Rec. 538, 16-19.

Pulugurtha, S. S., Bhatt, J., 2008. Evaluation of Weaving Sections Through Safety Assessments. ITE 2008 Technical Conference and Exhibit, Miami, FL.

Reilly, W., Kell, J., Johnson, J., 1984. Weaving Analysis Procedures for the New Highway Capacity Manual. JHK & Associates, Tucson, AZ.

Roess, R. P., McShane, W. R., Pignataro, L. J., 1974. Configuration, Design, and Analysis of Weaving Sections. Transport. Res. Rec. 489, 1-12.

Roess, R. P.. 1987. Development of Weaving Area Analysis Procedures for the 1985 Highway Capacity Manual. Transport. Res. Rec. 1112, 17-22.

Roess, R. P., Ulerio, J. M., 2000. Weaving Area Analysis in Year 2000 Highway Capacity Manual. Transport. Res. Rec. 1710, 145-153.

Roess, R.P., Ulerio, J. M., Prassas, E. S., Schoen, J., Vandehey, M., Reilly, W., Kittelson, W. , 2008. Analysis of Freeway Weaving Sections. National Cooperative Highway Research Program, Transportation Research Board. Final Report for NCHRP Project 3-75. Washington D.C.

Roess, R., Ulerio, J., 2009. Capacity of Freeway Weaving Segments. In: Proceedings of the 88th Annual Meeting of the Transportation Research Board 2009, Washington D.C.

Rosenbaum, M. J., Pinnell, C., Kemper, W. J., 1982. Synthesis of Safety Research Related to Traffic Control and Roadway Elements. Federal Highway Administration. Report No. FHWA-TS-82-232, Washington D.C.

Sarhan, M., Hassan, Y., Abd El Halim, A. O., 2008. Safety Performance of Freeway Sections and Relation to Length of Speed-change Lanes, Can. Journal of Civil Eng. 35 (5), 531-541.

SAS Institute., 2004. SAS/ETS 9.1 User's Guide. SAS Institute Publishing. Cary, NC.

Sayed, T., Brown, G., Navin, F., 1994. Simulation of Traffic Conflicts at Unsignalized Intersections TSC-Sim. Accid. Anal. Prev. 26 (5), 593-607.

Skabardonis, A., Cassidy, M., May, A. D., Cohen, S., 1989. Application of Simulation to Evaluate the Operation of Major Freeway Weaving Sections. Transport. Res. Rec. 1225, 91-98.

PTV AG., 2009. VISSIM User's Manual. Planung Transport Verkehr AG. Germany.

Vu, T., Roess, R., Ulerio, J., Prassas, E., 2007. Simulation of a Weaving Section. In: Preceedings of the 86th Annual Meeting of the Transportation Research Board 2007, Washington D.C.

Wiedemann, R., 1974. Simulation of Traffic Flow. Institute for Transportation Science, University of Karlsruhe, Germany.

Wynn, F. H., 1946. Weaving Practices on One-way Highways. Bureau of Highway Traffic, Yale University, New Haven, CT.

# VITA

Thanh Quang Le received his Bachelor of Engineering degree from Hanoi University of Communications and Transport, Vietnam, in May 2003. He entered the Civil Engineering program at Texas A&M University in September 2007 and received his Master of Science degree in December 2009. His research interests include traffic safety, simulation and traffic operation.

Mr. Le may be reached at Zachry Department of Civil Engineering, Texas A&M University, College Station, TX 77843. His email is quangthanhgt@gmail.com